UNIVERSIDAD POLITECNICA DE VALENCIA

DEPARTAMENTO DE INGENIERIA HIDRAULICA Y MEDIO AMBIENTE

Analysis of monitoring network: spatial and temporal reduction and their consequences on the ecological classification in the scope of Water Frame Directive

Tesis doctoral

Abramić Andrej

Directores

Dr. Julio González del Río Rams

Dr. María Nieves Martínez Alzamora

Valencia-2010

Acknowledgements

I am indebted to a large number of people who have provided assistance and support throughout this PhD but first and foremost, I would like to express my sincere gratitude to my supervisors, Dr. Julio González del Río and Dr. Nieves Alzamora for their invaluable guidance and advice and for making this research such an experience.

I am grateful to Professor González del Río for accepting me as a PhD student from the very beginning. His wide knowledge and his logical way of thinking have been of great value for me. I am also indebted to Professor Alzamora, who played an active role in my research, encouraged me enormously and inspired me with interest in applied statistics. Her understanding, encouraging and personal guidance have provided a good basis for the present thesis.

I am also very grateful to Carmen González from the SAIH unit at the Confederación Hidrográfica del Júcar, for her very important contribution in providing with all the pluviometric data I needed for this work and for her closeness and friendship.

I would like to deeply acknowledge the help of the following colleagues and friends from the IIAMA to whom I am indebted for their encouragement and support in the production of this thesis: Xelo, Pedro, Jose, José Antonio, Lidia, Miguel, Jaime, Germà, Jorge, Eli, Diego and Noemi. Gracias a tod@s!

I would like to express my gratitude to Professor Rafael García Bartual, my friend Rafa, for his sensible counsel, his thoughtful conversations and fruitful discussions during the very hard times, which contributed to my thinking and studies along the way.

I would have never finished this PhD without a close group of friends by my side. They have helped me out over the course of my PhD by offering advice or support.

My final and most heartfelt acknowledgement goes to my family; this work is dedicated to Leo, my son, who has been my inspiration since he was born; to my parents, Srđan and Jasna, whose encouragement and support have been essential, even though the distance and separation; and to Carmen, always next to me.

Abstract

The Water Framework Directive (WFD) is the most substantial piece of water legislation ever produced by the European Commission, and will provide the major driver for achieving sustainable management of water in Member States for many years to come. It requires that all inland and coastal waters within defined river basin districts must reach at least good status by 2015 and defines how this should be achieved through the establishment of environmental objectives and ecological targets for surface waters.

The WFD has established a system of biological (*phytoplankton, macroalgae, marine angiosperms and macroinvertaebrate)* hydromorphological and physiochemical indicators that determine the ecological quality of coastal water bodies. This study is focused on the coastal monitoring network which analyzes phytoplankton biomass through chlorophyll *a* (a surrogate indicator), established in the scope of the WFD. The objective is to provide an ecological classification for Valencian coastal water bodies.

Since the coastal monitoring network was activated in August 2005, this study manages historical data for three years. Ecological classification quality is assessed using 90th percentile valules (chlorophyll *a)* and ecological class thresholds intercalibrated between the Member States of the Mediterranean Geographical Intercalibration Group (MED-GIG).

The calculation of the 90th percentile value of a sample presents some difficulties since the empirical CDF is represented by a ladder chart, which results in undefined situations. There are different criteria for calculating sample percentiles in these cases. In fact, most statistical software provides us with different results since they use criteria that are similar but not equal. These differences are caused by the lack of a universally accepted procedure. Because of this, the possibility exists that the ecological classification is inadequate. It is suggested that this issue should be discussed during the intercalibration process for MED-GIG and that common criteria to be used by all member states should be established.

The results of the exploratory study for the temporal reduction show unpredicted consequences in the ecological quality classification for the Valencian coastal waters. Integrating pluviometric, wave, salinity and chlorophyll *a* series analyzed for the three different climate conditions that were found for the Valencian coastal waters, it was realized that the trophic balance of the ecosystems depends on anthropogenic influence and environmental factors that are determined by climate seasonality and irregularity. Three types of coastal water body ecosystems have been determined: stable, unstable and intermediate. The results demonstrate that survey frequency can be reduced for both stable and unstable systems; even though temporal reduction is only recommended for stable water body ecosystems.
Additionally, in the exploratory study of possible reductions in the number of stations visited, the comparative study of the 90th percentile has confirmed that it was possible to reduce the number of stations with no change in ecological classification.

A methodology has been developed to select the optimal subset of sampling stations to be surveyed. The reason for this is the fact that the ecological balance of coastal waters is subject to strong instability, and a simple descriptive study would lose its usefulness if the

characteristics of the area were to change. The proposed methodology could be used routinely by scientists.

The particular study of the empirical CDF of the data collected for each water body has confirmed that it is desirable that the empirical CDF in the full sample does not differ substantially from the empirical CDF in the subsample in order to attain the greatest similarity between the respective 90^{th} percentiles. Also, it would be interesting that the subsample includes the greatest values of the original sample. The method for selecting the stations to be surveyed has been approached from the perspective of an inter-observer variability problem. Different k-combinations of sampling stations have been approached as different observers which provide monthly measurements of chlorophyll *a* concentration in the area. In order to compare the concordance between two different observers, two measures have been considered: Euclidean distance and intra-correlation coefficient. To ease the interpretation of the Euclidean distance between observers, multidimensional scaling has been used.

To detect possible changes in unchecked sampling stations, we suggest estimating the concentrations of chlorophyll *a* in these stations using regression models. The virtual values obtained this way should be verified with occasional revisions to detect possible changes.

Resumen

La Directiva Marco Europea del Agua (WFD en sus siglas en inglés) es la norma legislativa más importante que ha establecido la Comisión Europea, en materia de aguas, y constituye un importante avance para lograr una gestión sostenible del agua, en los países miembros, a medio y largo plazo. Esta norma exige que todas las aguas, continentales y costeras, alcancen en el 2015 un status bueno o muy bueno e indica los pasos a seguir para conseguirlo, mediante el establecimiento de objetivos ecológicos y medioambientales en las mismas. Para concretar estos objetivos, la DMA ha establecido un sistema de indicadores biológicos (fitoplancton, macroalgas, angiospermas marinos y macroinvertebrados), hidro-morfológicos y fisicoquímicos que determinan la calidad ecológica de las diferentes masas de agua.

Para proporcionar una clasificación ecológica de las masas de agua costeras en la Comunidad Valencia existe una red de monitoreo costera que analiza la biomasa de fitoplancton, de forma indirecta a partir de la clorofila *a*. La clasificación ecológica de las aguas costeras se ha realizado siguiendo las indicaciones del Grupo Geográfico de Intercalibración del Mediterráneo (MED-GIG). En las mismas, se establece la calidad de las aguas costeras en base al valor del percentil 90 del total de mediciones de clorofila *a* recogidas en sucesivas campañas mensuales durante un período de 5 años.

La red de monitoreo se puso en marcha en agosto del 2005 y el estudio se cerro en agosto del 2008, y por lo tanto se manejan datos históricos de tres años. Las clasificaciones ecológicas obtenidas con estos datos son todavía provisionales, ya que la DMA exige un período de 5 años para obtener una clasificación.
Se ha considerado la posibilidad de reducir el número de campañas realizadas a lo largo del año y la posibilidad de reducir el número de estaciones revisadas en cada campaña y que consecuencias lleva esta reducción por estado ecológico.

Los resultados del estudio exploratorio para la reducción temporal, muestran que una reducción no controlada en el número de campañas realizadas a lo largo del año en las aguas costeras valencianas, puede producir alteraciones impredecibles en su clasificación ecológica. Un estudio conjunto de las series pluviométricas, de oleaje, salinidad y clorofila a, para los tres climas marítimos presentes en las aguas costeras valencianas, ha permitido observar que el equilibrio trófico de los ecosistemas, en esta zona, depende tanto de la influencia antropogénica como de factores medioambientales determinados por la estacionalidad e irregularidad del clima. En base a ello, se han definido tres tipos de ecosistemas de masas de aguas costeras: estable, inestable e intermedio. El estudio particularizado de cada tipo indica que existe la posibilidad de reducir la frecuencia del monitoreo para los sistemas estables e inestables; aunque se recomienda únicamente para ecosistemas estables.

Los resultados del estudio comparativo del valor del percentil 90 muestral, para la reducción espacial han proporcionado resultados prometedores, ya que en la mayor parte de las masas de agua analizadas existía la posibilidad de realizar diferentes reducciones en el número de estaciones visitadas en cada campaña, sin que ello comporte importantes alteraciones en el valor del percentil, y, consecuentemente, en la clasificación ecológica de la calidad de las aguas costeras.

Evidentemente, la identificación las estaciones de muestreo que deberían ser incluidas en la campaña de muestreo, no podía realizarse únicamente en base a la simple desviación entre los percentiles 90, ya que cualquier variación en los datos podría llevarnos a consecuencias contradictorias. Es por ello que se han estudiado, las características que deberían reunir las estaciones analizadas para que la desviación del percentil 90 de los datos recogidos no se desviara significativamente del percentil 90 correspondiente a la muestra completa. El problema de la selección de estaciones, ha sido abordado desde la perspectiva de un problema de concordancia entre observadores. Se han identificado las diferentes combinaciones de las estaciones de muestreo de una masa de agua, como diferentes observadores que miden la concentración de clorofila *a* en la zona, en las diferentes campañas. Se ha considerado como observador de referencia al observador que incluye todas las estaciones disponibles en la masa de agua y se han localizado los observadores que presentaban una mayor concordancia con el observador de referencia.

Dado que el equilibrio ecológico en las masas de agua costeras está sujeto a una fuerte inestabilidad, y las conclusiones de estudio podrían perder validez si se producen importantes alteraciones en la zona; se han estimado, adicionalmente, con el fin de detectar posibles cambios en las estaciones de muestreo no inspeccionadas, modelos de regresión que permiten obtener estimaciones mensuales de las concentraciones de clorofila *a*, en las estaciones no revisadas, a partir de las mediciones obtenidas en las estaciones revisadas. La realización de revisiones puntuales en estas estaciones y la comparación entre valores virtuales y observados sería un comprobante de que no existen alteraciones importantes en la masa de agua.

RESUM

La Directiva Marco Europea de l'Aigua (WFD en les seues sigles en anglès) és la norma legislativa més important que ha establert la Comissió Europea, en matèria d'aigües, i constituïx un important avanç per a aconseguir una gestió sostenible de l'aigua, en els països membres, a mitjà i llarg termini. Esta norma exigeix que totes les aigües, continentals i costaneres, aconsegueixen en el 2015 un estatus bo o molt bo i indica els passos a seguir per a aconseguir-ho, per mitjà de l'establiment d'objectius ecològics i mediambientals en les mateixes. Per a concretar estos objectius, la DMA hi ha establit un sistema d'indicadors biològics (fitoplàncton, macroalgues, angiospermes marins i macroinvertebrats), hidro-morfològics i fisicoquímics que determinen la qualitat ecològica de les diferents masses d'aigua.

Per a proporcionar una classificació ecològica de les masses d'aigua costaneres en la Comunitat València hi ha una xarxa de monitoratge costanera que analitza la biomassa de fitoplàncton, de forma indirecta a partir de la clorofil·la *a*. La classificació ecològica de les aigües costaneres s'ha realitzat seguint les indicacions del Grup Geogràfic d'Intercalibratge del Mediterrani (MED-GIG). En les mateixes, s'estableix la qualitat de les aigües costaneres basant-se en el valor del percentil 90 del total de mesuraments de clorofil·la *a* arreplegades en successives campanyes mensuals durant un període de 5 anys.

La xarxa de monitoratge es va posar en marxa a l'agost del 2005 i l'estudi es turó a l'agost del 2008, i per tant es manegen dades històriques de tres anys. Les classificacions ecològiques obtingudes amb estes dades són encara provisionals, ja que la DMA exigeix un període de 5 anys per a obtindre una classificació.
S'ha considerat la possibilitat de reduir el nombre de campanyes realitzades al llarg de l'any i la possibilitat de reduir el nombre d'estacions revisades en cada campanya i que conseqüències porta esta reducció per estat ecològic.

Els resultats de l'estudi explorador per a la reducció temporal, mostren que una reducció no controlada en el nombre de campanyes realitzades al llarg de l'any en les aigües costaneres valencianes, pot produir alteracions impredictibles en la seua classificació ecològica. Un estudi conjunt de les sèries pluviomètriques, d'onatge, salinitat i clorofil·la a, per als tres climes marítims presentes en les aigües costaneres valencianes, ha permès observar que l'equilibri tròfic dels ecosistemes, en esta zona, depèn tant de la influència antropogènica com de factors mediambientals determinats per l'estacionalitat i irregularitat del clima. Basant-se en això, s'han definit tres tipus d'ecosistemes de masses d'aigües costaneres: estable, inestable i intermedi. L'estudi particularitzat de cada tipus indica que hi ha la possibilitat de reduir la freqüència del monitoratge per als sistemes estables i inestables; encara que es recomana únicament per a ecosistemes estables.

Els resultats de l'estudi comparatiu del valor del percentil 90 mostral, per a la reducció espacial han proporcionat resultats prometedors, ja que en la major part de les masses d'aigua analitzades existia la possibilitat de realitzar diferents reduccions en el nombre d'estacions visitades en cada campanya, sense que això comporte importants alteracions en el valor del percentil, i, conseqüentment, en la classificació ecològica de la qualitat de les aigües costaneres.

Evidentment, la identificació de les estacions de mostratge que haurien de ser incloses en la campanya de mostratge, no podia realitzar-se únicament basant-se en la simple desviació entre els percentils 90, ja que qualsevol variació en les dades podria emportar-nos a conseqüències contradictòries. És per això que s'han estudiat, les característiques que haurien de reunir les estacions analitzades perquè la desviació del percentil 90 de les dades arreplegades no es desviés significativament del percentil 90 corresponent a la mostra completa. El problema de la selecció d'estacions, ha sigut abordat des de la perspectiva d'un problema de concordança entre observadors. S'han identificat les diferents combinacions de les estacions de mostratge d'una massa d'aigua, com a diferents observadors que mesuren la concentració de clorofil·la *a* en la zona, en les diferents campanyes. S'ha considerat com a observador de referència a l'observador que inclou totes les estacions disponibles en la massa d'aigua i s'han localitzat els observadors que presentaven una major concordança amb l'observador de referència.

Atès que l'equilibri ecològic en les masses d'aigua costaneres està subjecte a una forta inestabilitat, i les conclusions d'estudi podrien perdre validesa si es produeixen importants alteracions en la zona; s'han estimat, addicionalment, a fi de detectar possibles canvis en les estacions de mostratge no inspeccionades, models de regressió que permeten obtindre estimacions mensuals de les concentracions de clorofil·la *a*, en les estacions no revisades, a partir dels mesuraments obtinguts en les estacions revisades. La realització de revisions puntuals en estes estacions i la comparació entre valors virtuals i observats seria un comprovant que no hi ha alteracions importants en la massa d'aigua.

1.Introduction

1.1.WATER FRAME DIRECTIVE

On 23 October 2000, the "Directive 2000/60/EC of the European Parliament and of the Council establishing a framework for the Community action in the field of water policy" or, in short, the EU Water Framework Directive (or even shorter the WFD) was finally adopted. The Directive was published in the Official Journal (OJ L 327) on 22 December 2000 and entered into force the same day.

The Water Framework Directive is the most substantial piece of water legislation ever produced by the European Commission, and will provide the major driver for achieving sustainable management of water in Member States for many years to come.
It requires that all inland and coastal waters within defined river basin districts must reach at least good status by 2015 and defines how this should be achieved through the establishment of environmental objectives and ecological targets for surface waters. The result will be a healthy water environment achieved by taking due account of environmental, economic and social considerations.

WFD introduce a new perspective politics to manage inland and costal waters for the Member States (MS) of the European Union. It establishes a river basin as a main unit of water resource management. This division on river basin to manage water resources is established due to the fact that natural limits have to predominate over other administrative division. This new politics also establish an essential need of basin planning process, economics analysis of water use, need of public information and consultation etc. to manage adaptation of WFD in the scope of the European Community.

Water is used by industry to make products, assist industrial processes and generate power. It is used by the wider public to support many domestic, agricultural and recreational uses. Water should also sustain healthy environments and a range of wildlife. To strike a balance between environmental, social and economic interests, it is necessary to initialize river basin planning process.
The river basin planning process determines all types of water bodies: rivers, lakes, estuaries, coastal waters and groundwaters and will:

- *describe the current condition of water environment;*
- *identify where current or historic activities are affecting the ecological quality of the water environment;*
- *detail the actions required to ensure to maintain quality of waters that already meet WFD standards;*
- *set out the actions needed to deliver environmental improvements.*

In coastal areas, the WFD is in force up to one nautical mile from the territorial baseline of a Member State for a Good Ecological Status and up to 12 nautical miles for a Good

Chemical Status. To facilitate the assessment of the coastal water bodies, different systems are used for the ecological and chemical status.

For the ecological status, no absolute standards are set for the whole community, because of the ecological variability. For every water type, national classification schemes for the quality elements, phytoplankton, macroalgae, marine angiosperms and macroinvertaebrate are set. The WFD classification scheme includes five status categories: High, Good, Moderate, Poor and Bad.

Classification process combines the biological, hydro-morphological and physico-chemical elements. High status is defined as biological, physico-chemical and morphological conditions that are associated by low or zero of human pressure. Assessment of quality is based on the extent of deviation from reference condition:

'High status'- no deviation

'Good status'-slight deviation

'Moderate status'- moderate deviation

The National classification schemes of the Member states have to be harmonized in the European Union context. The purpose of the intercalibration is not to harmonize the assessment systems, but only their results. The descriptions of the reference situations and the classification schemes may differ, but the outcomes of the different schemes have to be comparable. The intercalibration will aim at the "high-good" and the "good-moderate" boundaries. The difference between the qualifications good and moderate is essential in the WFD, because it decides if measures are required or not.
The chemical status of the water is tested by means of a list of over thirty priority substances (prioritized by the European Commission). The Ecological Quality Standards (EQS) are differentiated for inland surface waters (rivers and lakes) and other surface waters (transitional, coastal and territorial waters). Two types of EQS are set; Annual average concentrations and maximum allowable concentrations, for protection against long-term and for short-term, direct and acute ecotoxic effects, respectively.

The Water Framework Directive requires that an integrated monitoring programme be established within each river basin district. These monitoring programs will, in many cases, be extensions or modifications of existing programs and will enable collection of the physical, chemical and biological data necessary to assess the status of surface and groundwater bodies in each river basin district.
Finally, the plan and program of the WFD to manage and put in order integrated natural cycle of water in the synergy with a human use and activity and to introduce the basic environmental features includes:

- The concept of river basin management is introduced to all Member States through the establishment of river basin districts as the basic management units(Article 3)
- Although its prime aims are environmental, the Directive embraces all three principles of sustainable development. Environmental, economic and social needs must all be taken into account when river basin management plans are being developed (Article 9).
- The river basin management plans will not allow further deterioration to existing water quality (Article 4).

- The polluter pays principle is incorporated through a review of measures for charging for water use, including full environmental cost recovery (Article 9).
- Public participation and the involvement of stakeholders are key requirements of the river basin management planning process (Article 14).

1.2.Monitoring in the scope of WFD

1.2.1.Monitoring

Monitoring programs are designed to assess water quality and, within results of survey, to inform about the state of water to management authorities. The general framework for any monitoring program deals with the definition of objectives, priorities and process of optimization, implementation of quality control and management of monitoring success.

As the general objective of monitoring, specified in the WFD, is to verify compliance with water quality objectives or to establish the reasons for noncompliance so that appropriate measures can be activated (J. G. Ferreira. 2007), the WFD demands to MS's to establish a surveillance monitoring programs to allow the water quality assessment. Due the water body ecological status, it is necessary to activate more intensive quality control. For the water bodies identified as being at risk of failing to meet environmental objectivities, the WFD requires to establish operational or investigation monitoring program.
The monitoring process optimization is covered by dynamic investigation, which deals with achieving objectivities and improving the efficiency by reducing sampling in time and space. The results of analysis during the monitoring process, sometimes quite simple, identify lack or abundance of information that can be corrected by an appropriate redesign of the network.
The change of priorities in the monitoring program, according to the management issues, is possible but, it is necessary to revise in detail in the scope of defined objectives. This includes analysis of the anthropogenic and non-anthropogenic pressures, susceptibility (freshwater influence, tidal mixing and other environmental conditions) and the actual state of ecological system.
The efficiency of monitoring networks can be optimized using statistic or mathematical models, with condition of sufficient data information and using robust methods. Spatial data analysis, using geographical information system tools, can underline problems or advantages of any monitoring network to be improved.

The implementation of quality control is an essential issue of environmental monitoring. Data quality is an important consideration for any environmental network survey to ensure that objectives are met and conclusions are not misled by inaccurate of data (J. G. Ferreira. 2007). Errors inevitably occur both in the process of sampling and in the analysis of water samples.
The aim of an appropriate quality assurance procedure is to quantify and control these errors. The quality assurance procedures may take the form of standardization of sampling and analytical methods, which are developed or in the process of development.

Finally, the success of each monitoring plan must be assessed with an appropriate management that can provide a mechanism for the adaptations or corrections to improve survey process.

1.2.2.Types of monitoring in the scope of WFD

The Water Framework Directive requires that an integrated monitoring program be established within each river basin district. This program should be operational, at the latest, by 22 December 2006 and must be in accordance with requirements of Annex V.

The Annex V of WFD describes that for coastal waters it is necessary to establish monitoring programs for biological elements(composition, abundance and biomass of phytoplankton; other aquatic flora and benthic invertebrate fauna) hydromorphological condition (morphological conditions, tidal regime), general elements (transparency, thermal conditions, oxygenation conditions, salinity and nutrient conditions) and finally, for the specific pollutants (priority and other substances discharged in significant quantities in the water body).

The monitoring network shall be designed to provide coherent and comprehensive summary of ecological status. It will permit classification of water bodies with normative defined into WFD document and established by intercalibration process.

The Directive describes three different types of monitoring strategies: surveillance, operational and investigative monitoring. They all have different aims in terms of detection or mapping of the environmental status.

Surveillance Monitoring

The WFD states: "For each period to which a river basin management applies surveillance, monitoring shall be established. The objective of the surveillance monitoring is to provide information for:

- Supplementing and validating the assessment of the likelihood that transitional or coastal waters will fail to meet the environmental quality objectives;
- The efficient and effective design of future monitoring programs;
- The assessment of long-term changes in natural conditions in order to distinguish between non-natural and natural alterations in the ecosystem;
- The assessment of long-term changes resulting from widespread anthropogenic activity.

The results of surveillance monitoring shall be reviewed and used in combination with the impact assessment to determine or adjust requirements for current and other monitoring programs.

On the basis of these results, the risk of failing to meet WFD environmental objectives shall be evaluated in the surveyed water bodies and an operational monitoring program established. Before implementing operational programs and to ascertain the causes of a water body failing to achieve the environmental objectives, investigative monitoring shall be considered, which may provide insight into reasons for any unknown excess."

The minimum monitoring frequencies indicated in Annex V of the WFD may not be adequate or realistic for coastal waters. A minimum of 6 months frequency, indicated for phytoplankton, has a low confidence level and with 2 campaigns during the year is really difficult to determinate composition disturbance or blooms frequency.

The spatial resolution will be determined on the basis of the water bodies defined for each system (Ferreira et al.2006) with at least one station per water body.

Operational monitoring

Operational monitoring will be carried out for all those water bodies identified as being at risk of failing their environmental objectives. This means that whereas surveillance monitoring is broader in the scope, operational monitoring will generally focus on subset of elements, e.g. primary and secondary eutrophication symptoms in the case of nutrient related problems (Bricker et al.2003).

Implementation of operational monitoring is based on the result of surveillance monitoring, the pressures on a water body (priority substances discharges) or both. Design of operational monitoring depends on the type of the pressure (point or diffuse source or hydromorphological changes) and the sample stations should be selected to assess the magnitude and the impact of ecological disturbance. Water bodies at risk to fail the environmental objectivities should be monitor for those quality elements which are indicative of the pressures.
Water Frame Directive do not specify the frequency of the operational monitoring. It states that the monitoring frequency shall be chosen to achieve an acceptable level of confidence and precision, but campaign intervals do not exceed those specified for surveillance monitoring.

Investigative monitoring

The Water Frame Directive specifies three cases when investigative monitoring is required:
Where the reason for any exceedances is unknown,
Where surveillance monitoring indicates that water body is not likely to be achieved ecological objectivities and operational monitoring has not already been established,
To ascertain the magnitude and impacts of accidental pollution.

The results of monitoring should be used as a base for the establishment of the specific measures for the achievement of the environmental objectivities.
This kind of survey will be designed to the specific case or problem being investigated. It could be more intensive, in terms of monitoring frequencies, and focused on relevant quality elements, on the particular water bodies or parts of water bodies.
This kind of monitoring might also include alarm or early warning monitoring, for example, for protection against accidental pollution.

1.2.3.Mediterranean networks

The WFD requires that MS's compares the results of monitoring and classification systems in each eco-region and due that decision it establishes Geographical Intercalibration Groups (GIG) on the basis of surface water body types. The Mediterranean GIG includes seven Member States: Spain, France, Italy, Slovenia, Greece, Cyprus and Malta plus Croatia as an accession country. During this process, it is established the data base on the CIRCA where is possible to examine historical data and monitoring networks.
Most national MS's monitoring programs started before the implementing the WFD in the scope of investigation, eutrophication assessment or national surveys projects. These historical data have been used for the identification of coastal water bodies, comparing national surveys methods between MS's and finally, to assess ecological classification system in Intercalibration process. The examination of historical data highlighted the huge heterogeneity between national monitoring programs for different aspects:

- Which part of the coastal waters is surveyed? – (inshore, nearshore or both)
- How are situated water sample stations in the area? - (transects sited vertical to the coast line, scattered, single station)
- Number of the sample stations (single, transect of three or two stations, etc…)
- Sampling method (various water column sampling points or just the superficial sampling)
- Distance to the pressure (continental loads, ports, waste water sewers, submarine outfalls….)
- Campaigns frequency
- Analyzing methods (probe-in situ, laboratory methods or both)

Figure 1: Different type of monitoring networks (three station transect vertical to the coast line, scattered situation)

Beyond all these differences, WFD does not require standardization of the surveys methods, but intercalibration of the results and classification procedures. The heterogeneity of the monitoring networks will make the intercalibration exercise more complex and the task more difficult to achieve.
As Valencian survey network was redesigned for the WFD needs, the other MS's monitoring networks are also updated and modified for the new task. The survey of coastal waters is a dynamic process. Consequently, monitoring networks are modified,

due the environmental needs, legislation, identified water bodies, continental pressure and eventually, because of conceptual model modification.

1.3.INTERCALIBRATION

The Annex V of Directive 2000/60/EC describes a process to ensure the comparability among Member States of biological monitoring results, which is an essential part of the ecological status classification. This requires that MS's compare the results of monitoring and classification systems in each eco-region of the European Community through an intercalibration network, to assess consistency and comparability of the national methods with WFD and other Member States.
The intercalibration process starts with the European Commission (EC) decision on August 2005 for the establishment of a register of sites to create a network in accordance with Directive 2000/60/EC of the European Parliament. In this document, intercalibration exercise divides Member States into Geographical Intercalibration Groups (GIG) on the basis of surface water body types.
GIGs are established for the European Rivers, Lakes and Transitional and Coastal waters. Transitional and Coastal waters of European Union are divided into three geographical sections:
Baltic , North- East Atlantic and Mediterranean (list of associated countries Table 1)

Baltic GIG	North-East Atlantic GIG	Mediterranean GIG
Denmark	Belgium	Cyprus
Estonia	Denmark	France
Finland	France	Greece
Germany	Germany	Italy
Latvia	Ireland	Malta
Lithuania	Netherlands	Slovenia
Poland	Portugal	Spain
Sweden	Spain	
	Sweden	
	United Kingdom	

Table 1: Geographical Intercalibration Groups

1.3.2.Mediterranean GIG

The Mediterranean GIG includes seven Member States, Spain, France, Italy, Slovenia, Greece, Cyprus and Malta plus Croatia as an accession country. It is divided into 4 subgroups, based on the system of organisms which are used to assess ecological classification for the coastal waters:

a.Phytoplankton
b.Benthic macroinvertebrate fauna
c.Macroalgae
*d.*Angiosperms: *Posidonia oceanica*

INTERCALIBRATION 2005 (reported by 2006)

All subgroups rapidly agreed on the IC types on base of substratum and depth:

Type Name of Type Substratum Depth
CW - M1 Rocky shallow coast rocky shallow
CW - M2 Rocky deep coast rocky deep
CW - M3Sedimentary shallow coast sedimentary shallow
CW - M4Sedimentary deep coast sedimentary deep

Table 2: IC types on the basis of substratum and depth

During the first meeting of the Phytoplankton subgroup it was agreed that complex issues of phytoplankton composition of various species can be bridged in the first phase of IC using chlorophyll *a,* which is consider as a good abundance and biomass indicator.
Phytoplankton subgroup has revised the Type distinction and, analyzing data, find necessary to define water bodies coastal types that are significant for phytoplankton, which has to be a previous step to intercalibrate the reference condition and to set common boundaries:

Reference Condition

Reference conditions will be different according to different water types.

Two (or more) types of coastal waters, significant for phytoplankton, are under development, based on water column stability, as a criterion for identification of coastal types. Then, for each type, sites with reference conditions, high and good status (H/G sites selected for intercalibration based on expert judgement and data on pressures) will be considered.

Setting boundaries

No clear correlation between pressures and chlorophyll *a* was found, so class boundaries could not be defined by discontinuities in the pressure vs. chlorophyll *a* diagram.

Procedure for setting the H/G boundary has been agreed upon. The boundary between high and good will be set according to the 90th percentile for chl a data, of all the selected sites per type.

(Intercalibration Report – Coastal GIGs Technical summary September 2006)

INTERCALIBRATION 2006 (reported by the beginning of 2007)

In order to define a new system for the typology of coastal water masses in the Mediterranean basin at CIRCA (an extranet tool, developed under the European Commission) it is established a data base with the following data:

MS	N° Sites	N° Records	Period	Freq (d)	Profile	Temp Sal data
France	3	2366		7	not available	available
Italy	11	2541	2001-2004	15	available	available
Valencia	46	858	1991-1992	15	not available	available
Slovenia	2	332	1997-2004	30	available	available
Cyprus	48	158	2005	60	not available	not available
Catalonia	17	69	2004	90	available	available
Croatia	19	1784	2000-2004	120	not available	available
Balearic	64	128	2005	365	not available	not available

Table 3: Available data on the CIRCA server

Data base has allowed development of a new typology which approach is based on continental influence level, using stability parameter- derived from temperature and salinity values in the water column. The subgroup adopts surface density as proxy indicator for static stability where temperature and salinity are involved in circulation and mixing dynamic for coastal marine system (Russo et al., 2006, in press).
On the basis of surface density values three major water types have been defined:

	Type I	Type II	Type III
σ t (density)	<25	25<d<27	>27

Table 4: MED GIG defined density values for the coastal typology

The same three water types are defined below as salinity classes due that seasonal fluctuation is marginal in various parts of Mediterranean basin:

	Type I	Type II	Type III
salinity	<34.5	34.5<d<37.5	>37.5

Table 5: MED GIG defined salinity values for the coastal typology

The three different water types, in an ecological perspective, can be described as follows:

Type I coastal sites highly influenced by freshwater inputs

Type II coastal sites not directly affected by freshwater inputs

Type III coastal sites not affected by freshwater inputs

After establishing three different water types for the Mediterranean basin, further distinction has been suggested and approved by the MSs:
First is the splitting of the coastal water type III in two different sub-basins; the Eastern and the Western Mediterranean, according to the different trophic conditions.

Second is subdivision of type II which includes two subtypes; type II-A (moderately influenced by freshwater input, as was determinate before) and Type II-B (coastal Mediterranean waters influenced by the Atlantic Ocean)
The types presence in the different MSs are shown in the following Table 6:

Types description			**France**	**Spain**	**taly**	**Slovenia**	**Croatia**	**Greece**	**Cyprus**
Type I	Highly influenced by freshwater input		**X**		**X**				
Type II	A	Moderately influenced by freshwater input (continent influence)	**X**	**X**	**X**	**X**			
	B	Influenced by Atlantic waters		**X**					
Type III	W	Not influenced by freshwater input	**X**	**X**	**X**		**X**		
	E	Not influenced by freshwater input						**X**	**X**

Table 6: Water types identified for Mediterranean coastal waters

The part of the IC focused on the reference condition at this time was not intercalibrated, mainly because every MSs proposed its own reference conditions based on their phytoplankton experts knowledge. This was not applicable, but all MSs still retain unchanged that the reference condition will be different for the three intercalibrated water types.
The examination of data provided by MedGIG MS's highlighted the huge heterogeneity, mainly due to the different monitoring schemes.
National methods adopted, mostly, three kinds of metrics: percentile90th, annual geometric mean and average. Depending on the MS, the metrics were calculated using only surface data or water-column integrated data, covering different period.

Phytoplankton experts from MSs decided to adopt a final agreement based on 90th percentile on raw data. In order to make data values more comparable, MSs which have been using geo-mean or average, decided to translate or recalculate their own values using the same metric as other countries.

As it was mentioned before, the MSs adopt a final agreement about using P90th percentile on the row data, with at least monthly sampling frequency in the surface layer. Harmonization of boundaries and EQR values were intercalibrated for the type II (only for II-A, type II-B is problem to intercalibrate as the coastal waters influenced by Atlantic, at MEDGIG exist just at the south of Spain) and type III (for Eastern and Western Mediterranean). Type I boundaries for ecological assessment was not intercalibrated due to the lack of data (only Italy has provided sufficient data). All MSs approved that the boundaries will be applied when 5 years data are available.

Type II

Type	MS	REFERENCE	H/G	EQR	G/M	EQR
T2 - A	Slovenia France Spain Italy	1.9	2.4	*0.80*	3.6	*0.53*

Type III

Type	MS	REFERENCE	H/G	EQR	G/M	EQR
T3 – Western Med	France Spain Italy	0.9	1.1	*0.80*	1.8	*0.50*
T3 – Eastern Med	Cyprus Greece	0.08	0.1	*0.80*	0.40	*0.20*

Table 7: Proposed EQR values and chlorophyll *a* thresholds (reference condition, high/good and good/moderate) for the ecological classification from the MS

In summer 2007, an annex of technical report for the EC was released. It described National Methods included in IC. In this document, national methods of monitoring and harmonization are detailed explained due to the intercalibration exercise for Slovenia, France, Greece and Spain which contains different methods for Catalunya, Valencia, Andalucia and the Balearic Islands.
This study was performed when IC has not been finished yet, during the long standstill phase. So P90 is the metric used for this analysis, with boundaries of ecological classes for the type II-A and type III established at the last MedGIG meeting, without reference condition reported and after adopted by the European Commission.

Type	**Ecological Quality Ratios**		**Values (µg/l, 90%ile)**	
	High-Good boundary	***Good-Moderate boundary***	***High-Good boundary***	***Good-Moderate boundary***
Type IIA	0.80	0.53	2.4	3.6
Type IIIW	0.80	0.50	1.1	1.8
Type IIIE	0.80	0.20	0.1	0.4

Table 8: Intercalibrated values for MED-GIG adopted by European Commission

The annex "DESCRIPTION OF NATIONAL METHODS INCLUDED IN THE INTERCALIBRATION" divides into Inshore and Nearshore waters the methodology for the Spanish coastal waters.

As Catalonia maintains survey on Nearshore and Inshore waters, comparing the values is reached common agreement among the Mediterranean Spanish Regions. The common Nearshore agreed boundaries corresponding to around 50% of those for Catalan and Valencian Inshore waters.

The problem about how to intercalibrate the Nearshore and Inshore surveys probably will be an issue of the discussion for the next Intercalibration meetings. This part of intercalibration should contain a serious analysis for different types of coastal waters (I, II, and III) in the Mediterranean Basin. These analysis should include historical data of biological surveys (chlorophyll a, phytoplankton...) of various years and also should integrate environment data as pluviometry, anemometry, sea currents and geomorphological information.
As this part of IC exercise didn't started yet, in this study the ratio of 50% will be used for the relation of Nearshore-Inshore for the type II-A and type III with values:

Type II-A; High/Good limit 4,8 mg/m^3 and Good/Moderate 7,2 mg/m^3

Type III; High/Good limit 2,2 mg/m^3 and Good/Moderate 3,6 mg/m^3

1.3.3.Coastal waters Eutrophication

The word “eutrophication” has its roots in two Greek words: “eu” which means “well” and “trope” which means “nourishment”. Eutrophication has been recognized as a problem in freshwater systems for many years but, only in the past three decades, concern has grown over the widespread occurrence of eutrophic conditions in transitional, coastal and marine systems.

Due to the complexity of the phenomena, the lack of consistent data sets, the lack of a harmonized approach to assess eutrophication, the severity and extent of the problem had not been adequately characterized at national levels and not harmonized on a European scale in the past.

Now, when the different process and effects are well know and documented, marine eutrophication is considered as one of the mayor threats for the health of coastal ecosystems (Cloern, 2001; Conley et al., 2002; Ronnberg and Bonsdorff, 2004).
Eutrophication is addressed in several EU policies such as Urban Waste Water Treatment (UWWT; 91/271/EEC), Nitrates (91/676/EEC) and Water Framework Directives (WFD; 2000/60/EC). A number of international conventions address the lack of a harmonized approach to the issue of the eutrophication in marine waters, including OSPAR (North East Atlantic) and HELCOM (Baltic Sea).

The EC Urban Waste Water Treatment Directive 91/271/EEC defines eutrophication cause as “the enrichment of water by nutrients, especially nitrogen and/or phosphorus” and the EC Nitrates Directive 91/676/EEC nutrient enrichment is generally defined by nitrogen compounds. The difference between these definitions about the main cause of acceleration /incrementation of organic enrichment can be found at the focus of Nitrate Directive, which, perhaps unsurprisingly, rests on losses of nitrogen from agriculture (Jesper at al.,2005).

The implementation of WFD, which treats with eutrophication indirectly by Good/Moderate boundary generated a need of common understanding and stronger coordination between the directives which deals with Eutrophication. On November 2005, a common implement strategy for WFD was published as a consensus agreed by water directors of EC.

The document: TOWARDS A GUIDANCE DOCUMENT ON EUTROPHICATION ASSESSMENT IN THE CONTEXT OF EUROPEAN WATER POLICIES defines:
Eutrophication is the accelerated production of organic matter, particularly algae and higher forms of plant life, in a water body usually caused by an increase in the amount of nutrients being discharged to the water body. As a result of accelerated algal production (primary impact), a variety of impacts may occur, including nuisance and toxic algal blooms, depleted dissolved oxygen, and loss of submerged aquatic vegetation (secondary impacts), undesirable disturbance of the balance of organisms present in the water, and deterioration of the quality of the water concerned.

The term “eutrophic” is used in the mentioned guidance to refer to the situation, when the natural trophic status is out of balance because of anthropogenic interventions. A pressure (in this case nutrient enrichment) causes an adverse change in biological quality elements (e.g. ‘composition, abundance and biomass of phytoplankton’). This in turn might cause indirect effects on physicochemical quality elements (e.g. transparency, oxygenation conditions), and other biota (e.g. macro-invertebrates). Water bodies that fail to achieve Good Ecological Status due to these effects of human induced nutrient enrichment can be considered to be “eutrophic” due to the process of eutrophication.

On a worldwide oceanic scale, the prevailing limiting conditions vary slightly. For open oceanographic conditions, nitrogen rather than phosphorus is generally assumed to be the production limiting factor. In transitional and coastal waters, anthropogenic nitrogen enrichment is also the most important cause of eutrophication. Nitrogen is generally considered to limit primary productivity in most world oceans and seas, but numerous studies have suggested that the Mediterranean Sea could be an exception. Bertland et al (1980) noted that the Mediterranean Sea is different from most of the other major ocean basins. He stated that phosphorus seems to play an important role in phytoplankton limitation. Several studies, carried out in the western basin, shown that phosphorus is either the principal limiting nutrient or is the limiting nutrient for the part of the year (e.g. Fiala et al 1976). The result of the studies performed for the Adriatic Sea also indicated that phosphorus is a limiting nutrient in the growth of phytoplankton (Pojed and Kveder 1977, Vukadin and Stojadinski 1976). The studies realized by Becacos- Kontos (1977) suggested that the limiting factor in the Aegean Sea were sometimes phosphorus and sometimes phosphorus and nitrogen. Similar results are obtained by Bonin (1989) at the Israel coast and Bertland (1987) at the Levantine basin. Azov (1986) found the correlation between chlorophyll *a* and PO_4^{3-} content at two locations off Haifa.

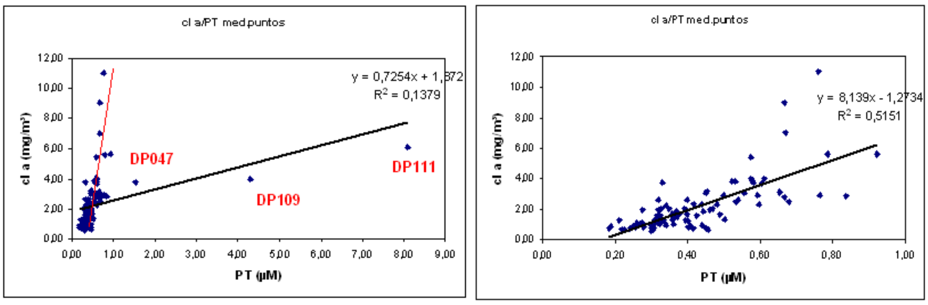


Figure 2: Results of correlation analysis for the chlorophyll *a* and total phosphorus presented on the ISMS07. Analysis clearly separated the results sampled at the areas directly influenced by urban waste waters and correlate chlorophyll *a* and total phosphorus by R^2=0,52

For the Valencian coasts various correlation analyses were performed (2007). There is a relationship for the average of chlorophyll *a* and the total phosphorus concentrations at the stations established for the Water Frame Directive. (those results were presented by the author of this study at ISMS07, 16)

Marine eutrophication is mainly an inshore problem that affects lagoons, harbors, estuaries and coastal areas adjacent to river mouths of highly populated river basins and/or that receive sewage from coastal cities. This creates particular situations that differ in their production characteristics from those of offshore waters. Due this, it is important to survey coastal waters trophic state in the near shore and in shore zone.

1.4.Monitoring strategies

During the research process, establishing the strategy about how to analyze the coastal monitoring network efficiency and the possible reduction, numerous studies and published reports that deal with the coastal monitoring issue were revised.

As reviewed in the WFD Guidance documents, various reports of other initiatives and efforts for the coastal and marine waters are also considered. According to reviewed documentations, it is identified differences in the monitoring programs established by European directives or initiatives. The WFD main difficulty arises from the concept that it appears to have been developed from a freshwater perspective and then, extended to the more variable estuaries and coasts (de Jong at al.2006).

The technical design of monitoring networks is related to the determination of: (i) monitoring sites; (ii) monitoring frequencies; (iii) variables to be sampled; (iv) duration of sampling (L.M. Nunes at al. 2006). At this study, the first two aspects will be discussed, the third one refers to the phytoplankton biomass surrogated by chlorophyll *a* and the last one is specific for the other type of monitoring.

Minimal frequency of twice a year established for the phytoplankton biomass survey at the WFD is inconsistent, in terms of natural spatial and temporal variability, management actions and decision making (de Jong at al.2006).

For instance, the United Nations Programme for the Assessment and Control of Pollution in the Mediterranean region (MED POL) demands a minimum of four campaigns (analyzing chlorophyll *a* and phytoplankton) per year and recommends monthly or bimonthly frequency. The eutrophication monitoring program proposes three transects (minimum three stations) per site, sampling superficial and vertical profiles to support Trophical Index for marine systems (TRIX) (Vollenweider et al. 1998). TRIX is based on chlorophyll *a*, oxygen saturation, total nitrogen and total phosphorus to characterize the trophic state of coastal marine waters.

The Black Sea Commission (formed by six coastal Black Sea states under the similar UN action initiative) for the Black Sea Monitoring and Assessment Programme (BSIMAP) proposes equivalent monitoring frequencies. This programme accepted common principles and platform for the elaboration of regionally coordinated monitoring and assessment system (COAST, WFD and OSPAR-JAMP). This programme assesses the eutrophication by trend analysis in nutrient levels but also recommends analyzing biota (chlorophyll a and phytoplankton) as response on the anthropic pressure. For the biota is demanded a minimum of four campaigns per year as for the MED POL.

The Convention for the Protection of the Marine Environment of the North-East Atlantic (OSPAR - 1992) , signed by fifteen Governments of the western coasts and catchments of Europe, together with the European community, has worked to identify threats to the marine environment and has organized, across its maritime area, programs and measures to ensure effective national action to combat them.

The OSPAR launched an Agreement on the Eutrophication Monitoring Programme (2005) to manage eutrophication by establishing sub-regions of the maritime area with different levels of priority:

1. Non-problem areas with regard to eutrophication
   - The monitoring programme has the function of detecting changes in the eutrophication status or confirming the status of particular areas as non-problem areas. Monitoring effort should be limited to a limited number of parameters and a limited frequency of measurements.

2. Problem areas with regard to eutrophication
   - The monitoring programme should focus on long-term trends in nutrient concentration and on a selection of related eutrophication effect parameters. A larger number of parameters and a higher sampling frequency should be considered comparing to the non-problem areas, in order to satisfy statistical requirements. The spatial coverage should also be more focused than for non-problem areas. Monitoring should continue until the non-problem area status is achieved.

3. Potential problem areas with regard to eutrophication
   - With regard to their unknown status, potential problem areas with regard to eutrophication should be monitored as if they were problem areas, for a trial period not exceeding five years. This should enable the area to be reclassified as either a problem area or a non-problem area with regard to eutrophication.

For the Non-Problem areas only nutrient enrichment monitoring is demanded (once per three years). Potential problem and Problem areas have to be minimal annually sampled; chlorophyll *a* and phytoplankton (during the algal growing season) and nutrients (during winter when algal growth is at a minimum).

The Joint Assessment and Monitoring Programme (JUMP) is an OSPAR Commission strategy which supports and produces a series of thematic assessments. JUMP "Eutrophication Monitoring Guideline: Chlorophyll *a* in Water" points that: *Chlorophyll a concentrations vary substantially during the growth season and may vary considerably from year to year as a consequence of many factors (e.g. meteorological/ hydrographic conditions). Thus, it may be difficult to establish temporal trends. In order to establish temporal trends, sampling needs to cover the entire growth season. This leads to the possibility of assessing a mean value for the spring season and a mean value for the entire growth season.*

JUMP Eutrophication Monitoring Guidelines are revised on the demand of the OSPAR commission by the International Council for the Exploration of the Sea (ICES). ICES response on the "required frequency of sampling" was that it cannot provide any additional guidance. *Considering the short generation time of phytoplankton, a very high sampling frequency is needed to cover the succession and development of the phytoplankton communities. Sampling frequency must be often enough to resolve bloom events at minimum, sampling should be weekly or biweekly, but this may need to be more frequent depending on local conditions and the species being observed.*

HELCOM is the governing body of the "Convention on the Protection of the Marine Environment of the Baltic Sea Area" - more usually known as the Helsinki Convention.

HELCOM. This institution works to protect the marine environment of the Baltic Sea from all sources of pollution through intergovernmental co-operation. COMBINE is the Manual for the Marine Monitoring, which should be applied also for coastal waters monitoring. Programme for monitoring of eutrophication and its effects (part C of COMBINE) divides sampling stations into:

1. Mapping stations with minimum annual frequency that analyze:
   -Nutrients before the onset of the phytoplankton growth period
   -Oxygen/hydrogen sulphide and nutrient conditions in the near bottom waters
   -Zoobenthos

2. High frequency sampling stations with minimum monthly sampling (but weekly in the vegetative period) that analyze in the “cruise” stations:
   -Temperature and salinity
   -Oxygen/hydrogen sulphide
   -Nutrients
   -Chlorophyll-a
   -Phytoplankton

   High frequency stations are also sampled by Ship-of-opportunity and Automatic fixed stations that provide significantly dense frequencies of data collecting.
   Ship-of–opportunity- provides unattended recording and sampling on ferries and other commercial ships with regular schedules. It gives a possibility to collect data with high temporal and spatial resolution in the surface layer of the sea, with large spatial extent. These kinds of measurements supply important information, especially for the real time monitoring, and early warning system of, e.g. toxic algal blooms, and can also serve as reference and calibration for satellite images.
   Automatic fixed stations: They make possible to collect high frequency data on temperature, salinity, oxygen, light attenuation and current speed/direction.

COMBINE monitoring strategy for the chlorophyll *a* is based on the same principles as described above, in the OSPAR “Eutrophication Monitoring Guideline: Chlorophyll *a* in Water”.

The document provided by Joint Research Center (JRC) “Monitoring strategies for phytoplankton in the Baltic Sea coastal waters”, the monitoring systems for the Baltic UE countries (Baltic – GIG) is reviewed. Monitoring frequencies are based on the HELCOM program for the eutrophication. Phytoplankton monitoring is generally conducted in the summer period, but the sampling frequency is also more intense during summer months (Heiskanen at al. 2005). On Figure 3, the number of samples per month and the number of samples per station for the countries of Baltic – GIG are represented.

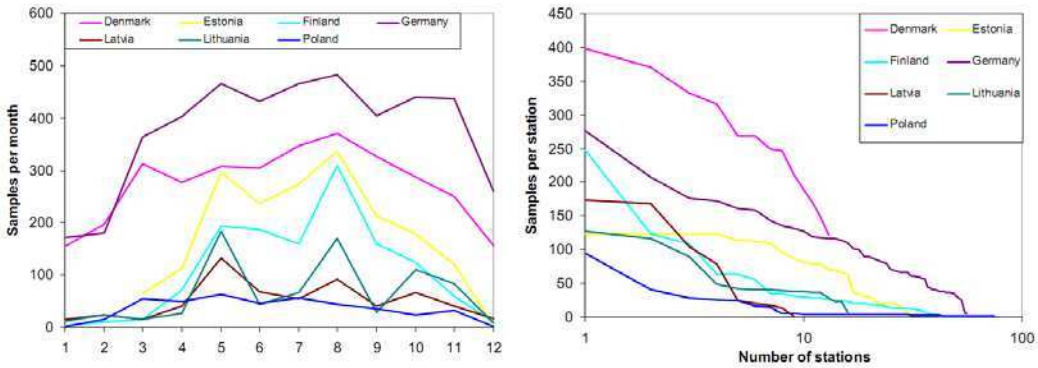


Figure 3: Number of samples per month for the different countries providing data; number of samples per station for the different countries providing data. For better illustration at second diagram of the differences the X-scale is logarithmic.

The amount of work involved for the high frequency monitoring, also by its specialist nature, results in high cost per sample and it is very time consuming (Dubelaar at al. 2004). As a result, neither qualitative nor quantitative surveys generally sample phytoplankton biomass adequately (Smayda at al.1998, Barreta at al.1998). Dubelaar at al. 2004 compares phytoplankton biomass dynamic with music – CD:

Besides the numerical accuracy: "the more bits the better", the other basic principle is that the sampling frequency should be at least double the highest significant frequency in the music in order to obtain good reproduction. If the sampling frequency is lowered too much, an increasingly large part of the signal fluctuations (system dynamics) remains unnoticed, with correspondingly impaired reproduction, which may lead to, in cases of ecosystem monitoring, questionable ecological interpretations.

The high frequency monitoring described above (with minimum frequency of 5 or 6 times per week) is very expensive and hardly approved by water managers and stakeholders. As technical field measurements, involving equipment and manpower, are generally associated with high cost and prolonged analysis time, modeling is mostly preferred (Batzias at al.2007)

There are numerous studies using the calibrated models (systems for the early warning) based on previous observations, scientific understanding, statistics, neural networks etc. that can forecast future environmental conditions. By these models it is possible to predict algal blooms and intense monitoring activities for a particular period and a particular area.

Despite a general acknowledgement in the scientific community on the benefits of using monitoring and modeling jointly, it has not been a common practice in the European monitoring programs so far. In the research communities, it is generally accepted that monitoring and modeling are interlinked activities (Hojberg and Refsgaard at al 2007). Models may potentially be used at many stages in the WFD cycle to support a variety of the tasks that must be carried out (Hattermann and Kundzewicz 2007).

Hojberg and Refsgaard at al. 2007 proposes (and describes in various examples) the use of models to support the monitoring requirements in the Water Frame Directive:

1. Quality Assurance of Monitoring Data - specifically mentioned as an important issue in the WFD guidance document on monitoring (EC 2003). Quality assurance is confined to sampling and laboratory procedures, and the minimum

requirement to monitoring data is that they should not contradict with simple well known basic principles introduced in the confident model.

2. Interpolation and Extrapolation in Time and Space – monitoring data are discrete in time and space and some interpolation technique is, therefore, required to transform these discrete points into a temporally and spatially continuous image of variables. This technique can be used (if reference site do not exist) to extrapolate the reference conditions (WFD explicitly mentions modeling as a suitable method).This challenge is well discussed by Nielesen et al.(2003) and Wasson et al.(2003)

3. Conceptual Model – establishing conceptual model/understanding, plays a key role in the WFD guidance of monitoring (EC 2003). The importance of the conceptual model is obvious, since the future management strategies are based on a conceptual understanding.

4. Assess Effects of Anthropogenic Activities – Monitoring data are used to identify trends in the ecological status and to assess whether an implemented programme of measures has had the expected effects. In this regard, the natural climate variability induces variability in the variables to be analyzed. Models should be able to explain some of the natural variability and, thereby, enhance the signal from the anthropogenic activity. The key idea is to make a statistical test of trend or shift by analyzing the residuals between model simulation and observation. In this way, the residual does not include the part of the natural variability, but includes the signal from the anthropogenic activities.

5. Design Monitoring programme – For a design of monitoring programme we have to assess the balance between the information content and the cost of the survey. It is also important to assess relation between the measurement effort (number of locations and measurement frequency) and the reduction of the uncertainty. A basic principle in the application of models-based monitoring design is the link of model simulations to optimization schemes, by which the best monitoring strategy under specified constraints can be identified. By this technique, alternative monitoring designs could be improved.

De Jonge et al. (2006) suggests reviewing national monitoring programs in order to integrate water quality monitoring and biological monitoring and change from “station oriented monitoring” to “basin or system oriented monitoring” in combination with specific “cause–effect” studies for highly dynamic coastal systems. His study also describes a spatial reduction of Dutch monitoring programme in 1992, reducing 35% of sampling stations (from 400 to 260). The monitoring program was reviewed by the government agency and considered “costly and data rich, but information poor”. Mainly, the budget-driven reduction of the monitoring program has indirectly been simulated by the way authorities interpreted their task. The monitoring system and data were analyzed mainly statistically within a wider environmental context and with large gaps. Due this, De Jonge suggests that developments in the monitoring programs respond to the current pressures, rather than assuming a responsibility of the safeguard of environment. Although monitoring programs need to be (re)evaluated regularly with acknowledge, there is insufficient funding to measure and monitor everything and there is need to achieve cost-effective monitoring and thus, rely on surrogates for the detecting of change.

Hunt at al. (2007) states that reductions in the number of sampling stations is less detrimental to the quality of data for annual decision making than reductions in the number of surveys. To evaluate alternative monitoring program design he used a method that accounts for degree of redundancy in the current sampling design, both spatially and temporally. He examined redundancy by correlation of data obtained by reduced monitoring strategies VS data obtained by initial monitoring design.

1.5.Aims of the study

The aim of this study is to analyze reduction for the Valencian coastal water monitoring network. This study is focused on the effects on the ecological classification of water bodies due to survey reduction within the scope of the Water Frame Directive.
This study will use historical data sampled and analyzed during the period from August 2005 to July 2008. The Water Frame Directive demands a minimum of five years of data to establish ecological classification for water bodies, which was not available at the moment this study was being carried out. Because of this, provisional ecological classifications were established (for Valencian coastal water bodies) using three-year data in order to allow an analysis of the consequences of monitoring reduction.

The reduction analysis of the coastal monitoring network will be realized in two aspects:

1. Temporal reduction – reducing monthly campaigns to bimonthly and trimestral frequency. The relationship between monthly surveys and reduced frequency monitoring campaigns will be analyzed.

2. Spatial reduction –results obtained by using a reduced number of sampling stations will be analyzed.

For both studies (temporal and spatial), the consequences that produce the reduction will be analyzed, compared to the results obtained using the reference data:

- Modification to ecological classification for water bodies that is generated by reduction
- Value discrepancies that are obtained using data derived by different reduction alternatives

Separately, it will be analyzed if it is possible to apply new reduction policy to the environmental network (temporal and spatial) that can guarantee reliable results.

Finally, once the possibility to realize a reliable alternative for the monitoring network is established, a new methodology will be developed parallel to this that will enable:

- Selection based on the statistical analysis of the reduction policy (temporal or spatial)
- Awareness of errors, risks, the level of reliability and accuracy of the new (reduced) monitoring network

This methodology should be applied, when five years of data become available, to design a new future monitoring network that will guarantee that significant information will not be lost and that the results generated will be the same as (or as similar as possible to) the results that are provided by the present monitoring network.

The alternative monitoring design that will be proposed by this study will be in total accordance with the policy and demands established by the Water Frame Directive. For each proposed monitoring network modification (temporal or spatial reduction) methods will be determined that can identify if the reduction decision was incorrect due to an extreme change in the ecosystem.

2.Study area

2.1.Mediterranean Sea

The name Mediterranean derives from the Latin word mediterraneus, meaning "in the middle of earth" (medius, "middle" + terra, "land, earth"). This is either due to the sea being surrounded by land (especially compared to the Atlantic Ocean) or that it was at the centre of the known world. The Mediterranean Sea has been known by a number of alternative names throughout human history. For example the Romans commonly called it Mare Nostrum (Latin, "Our Sea").
The Mediterranean Sea is an intercontinental sea positioned between Europe to the north, Africa to the south and Asia to the east, covering an approximate area of 2.5 million km². The longitude from the East to West is 3860 Km and maximum with North-South is about 1600 Km.
The Mediterranean Sea is connected to the Atlantic Ocean by the Strait of Gibraltar on the west and to the Sea of Marmara and the Black Sea, by the Dardanelles and the Bosporus respectively, on the east. The Sea of Marmara is often considered a part of the Mediterranean Sea, whereas the Black Sea is generally not. The 163 km long man-made Suez Canal in the southeast connects the Mediterranean Sea to the Red Sea.
Large islands in the Mediterranean include Cyprus, Crete, Euboea, Rhodes, Lesbos, Chios, Kefalonia, Corfu, Naxos and Andros in the eastern Mediterranean; Sardinia, Corsica, Sicily, Cres, Krk, Brač, Hvar, Pag, Korčula and Malta in the central Mediterranean; and Ibiza, Majorca and Menorca (the Balearic Islands) in the western Mediterranean.

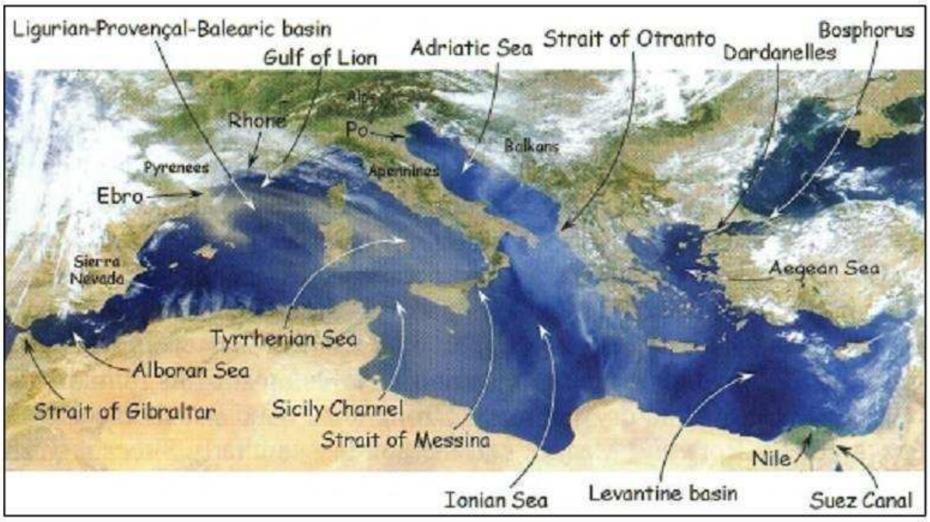


Figure 1: Mediterranean sea basins

The Mediterranean Sea is the largest semi-enclosed basin in Europe which is divided from Strait of Messina to Sicily Chanel into two big basins:

1.West Basin that includes sub-basins :

- Alboran Sea lying between Spain on the north and Morocco and Algeria on the south,
- Ligurian –Provincial –Balearic basin limited by two continental tight platforms with depression between Levante Coast and Balearic Islands
- Tyrrhenian sea in front of Occidental coast of Italy

2.East basin that Mediterranean sea sub-basins :

- Ionian Sea as a central Mediterranean sub-basin
- Levantine Basin, southeastern part of the Mediterranean Sea
- Aegean sea basin which is situated between the coasts of Greece and Turkey
- Adriatic sea basin which is situated between Italy and the Balkan Peninsula

The Mediterranean Sea has an average depth of 1.370 meters and the deepest recorded point is 5,267 meters in the Calypso Deep in the Ionian Sea.

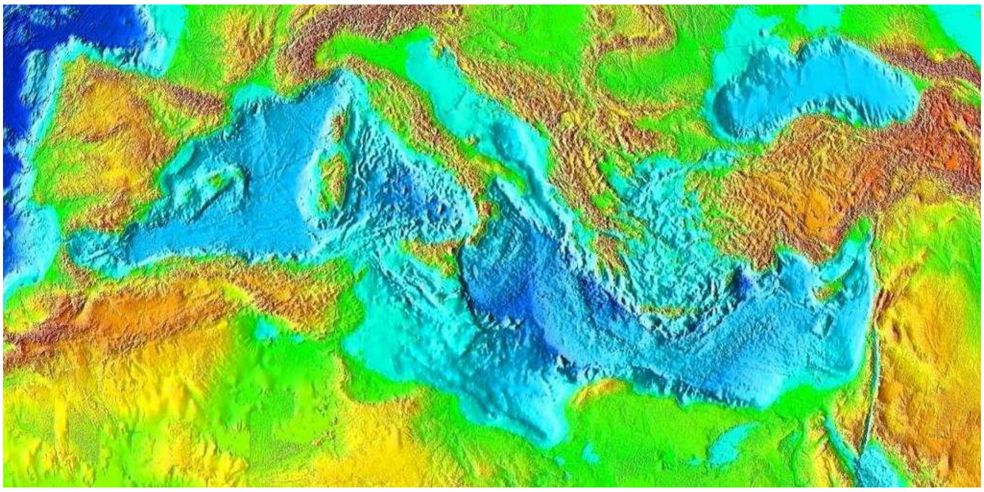


Figure 2: Mediterranean Sea surface, altimetry screenshot from NASA

Tides are very limited (very low amplitudes as it is possible to observe on diagram bellow Figure 3) as a result of the narrow connection with the Atlantic Ocean.
The highest amplitudes are detected in the Golf of Gabes and the noticeable tide levels can be noticed only at the Eastern part, at North Adriatic and around the regions close to Gibraltar.

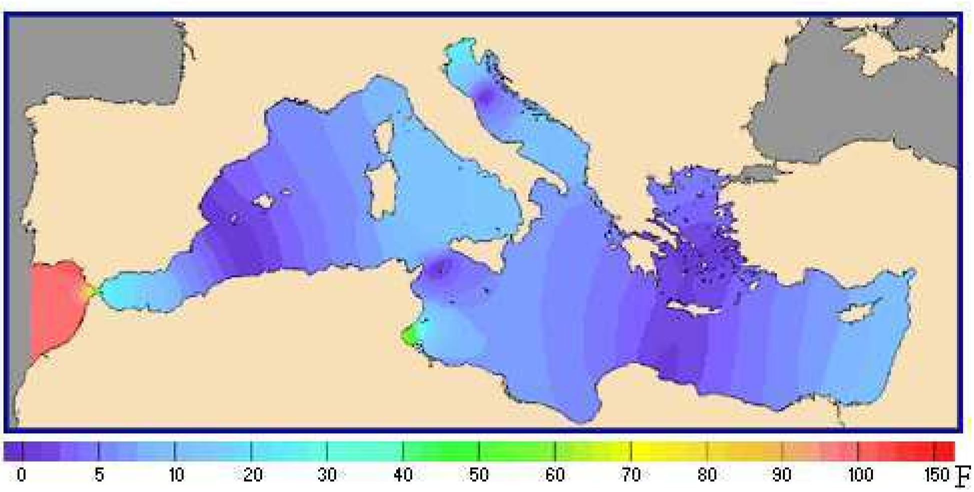


Figure 3: Legos via AVISO/Altimetry

The superficial water temperature varies during the year seasons. In summer is between 21 º to 30º C and during winter between 10 º to 15ºC. On the sea depth from 100 till 200 meters the temperature is stable with 13º C.

Evaporation greatly exceeds precipitation and river runoff in the Mediterranean, affecting the water circulation within the basin. The quantity of fresh water flowing into the Mediterranean from rivers is only one-third of the amount lost through evaporation. Evaporation is especially high in its eastern half, causing the water level to decrease and salinity to increase eastward. This imbalance causes a pressure gradient which draws relatively cool, low-salinity water from the Atlantic across the basin (Figure 4).

The average salinity concentration is about 38‰ and increase on the East part of Mediterranean basin due to the level of the evaporation up to 39‰.

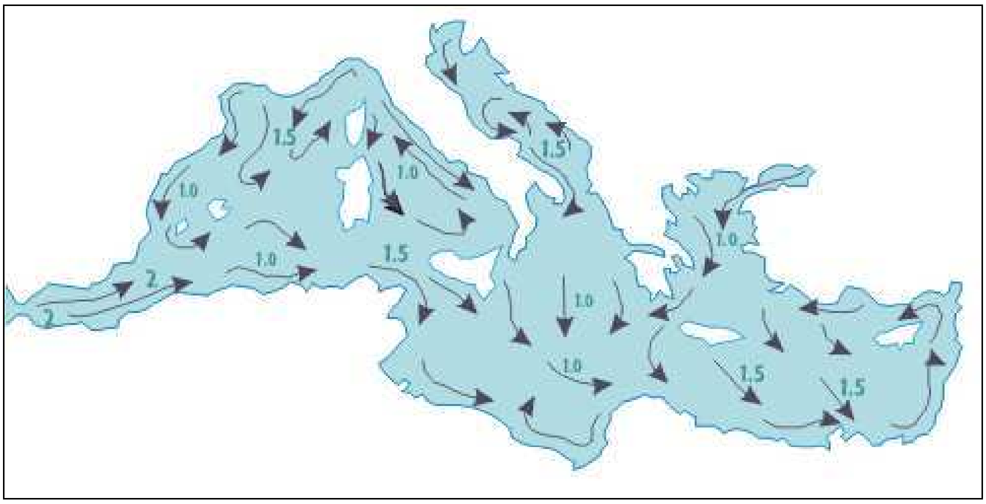


Figure 4: Predominant currents for June

The northern margin receives freshwater from big rivers as Ebro, Rhône, Po but the largest Mediterranean river is the Nile (southeastern sector Egypt). Fluvial waters deliver into the sea large amounts of anthropogenic substances, which, together with atmospheric deposition, are the main sources of pollutants.

Beyond the fluvial inputs (69 rivers that contribute with a 283 km^3 per year), Mediterranean Sea has a water deficit as previous mentioned because of evaporation and exchange with Atlantic waters in the Gibraltar. At the strait of Gibraltar negative water balance is produced by the cause of superficial oceanic water entrance. Entrance of Atlantic superficial waters is estimated by one million of m^3/s, while around 95% of deep water ran out from the Mediterranean Sea. Atlantic superficial waters that enter in then Mediterranean Sea are poor with nutrients, although that have higher concentrations compared by the same level Mediterranean waters (Rodríguez at al., 1982). Due to that imbalance, the Nitrogen as Phosphorus load for the Mediterranean basin is considerable lower in comparison with the Atlantic Ocean. Mediterranean is a oligotrophic sea and superficial as deep waters are considerable more poor with a nutrients than nearby Atlantic zone. The concentration of nutrients is decreasing from the West to East.

In the Mediterranean Sea, nutrients are important tracers of biological cycles, new production, natural and anthropogenic inputs and transfer processes (Bethoux et al., 1999). These low nutrient levels generally cannot support a large biomass. Besides the oligotrophic properties of the open Mediterranean Sea, recent observations of high nutrient levels in the Gulf of Lions and the northern Adriatic Sea have raised the question of how the increasing anthropogenic inputs from land and river discharge and from rainfall may influence the nutrient concentrations and lead to a critical eutrophication (Denis-Karafistan at al., 1998).

According to the European Environment Agency (1999) the main human activities and pressures identified are:

(1) population growth;
(2) tourism;
(3) agriculture;
(4) fishing and aquaculture;
(5) industry;
(6) maritime traffic;
(7) discharge from sewage outfalls; and
(8) discharge via rivers.

Domestic, industrial and agricultural activities are considered to be the three main pollution sources for the Mediterranean basin which loads nutrient concentrations can lead to eutrophication.

2.2.Coastal waters of Valencia

2.2.1.Introduction

The Valencian Community (officially *Comunitat Valenciana* or *Comunidad Valenciana* in Valencian and Spanish language, respectively) is an Autonomous region located in central and south-eastern Spain. It is divided into three provinces, from south to north: Alicante, Valencia and Castellón. It covers 23,259 km² of land with 5.02 million inhabitants (2008 data).
Mediterranean clime is dominant for the Valencian Community. It is possible to distinguish three sub-areas:

- Coastal area between Cataluña and Marina County is characterized by typical Mediterranean clime.
- In the inland part of the Community, it is possible to sense continental clime influence proceeding from inside of the peninsula. It is distinguished by lower temperatures and higher level of precipitations.
- The area at the South of Alicante is characterized by higher temperatures and low precipitation level.

2.2.2.Hydrographic basins

There are only two major rivers: Segura river, in Alicante province (whose source is in Andalusia) and Júcar river, in Valencia province (whose source is in Castile-La Mancha); both are subjected to very intense human regulation for cities, industries and agricultural consumption. The River Turia is the third largest river with a source in the Aragón, Teruel province. Due to the agricultural usage, both rivers in the area have a low river-flow and often they are almost dry during the summer.

The territory is crossed by three big river (hydrographical demarcations):

- The Jucar hydrographical demarcation is extended through the major part of the Valencian Community.
- The Segura hydrographical demarcation is situated at the South.
- The Ebro hydrographical demarcation is crossing the North part of the Valencian Community and has a contact with coastal waters in Cataluña.

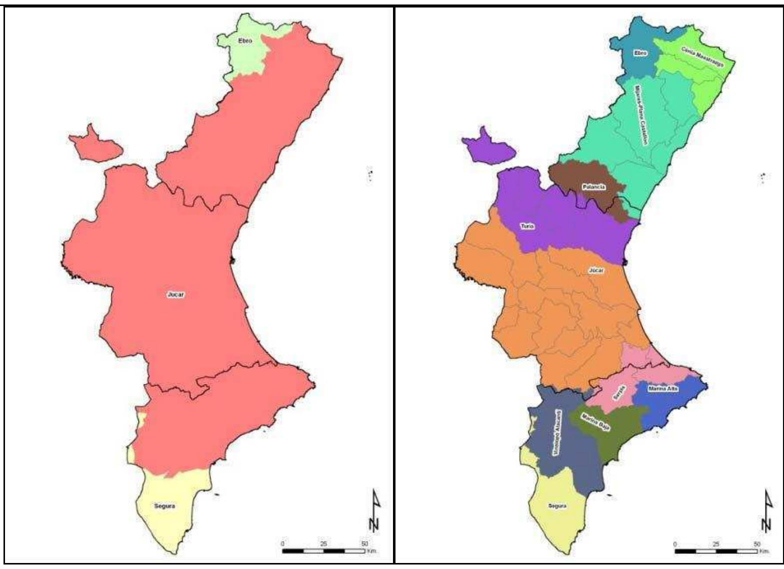


Figure 5: Basins and sub-basin of Valencian community

The most important rivers covered by the Jucar hydrographical demarcation are: Cenia, Mijares, Palancia, Turia, Júcar, Serpis and Vinalopo. All rivers in the Jucar basin flow into the coastal waters of Valencia.
Segura river is crossing the south part of the Valencian Community and flows into the Mediterranean sea in the province of Alicante, at Guardamar del Segura.

2.2.3.Coastal waters

The *Comunidad Valenciana* has 437 km of Mediterranean coastline crossing three provinces, Castellón, Valencia and Alicante. Valencian coastal waters include a few small islands like Columbretes and Tabarca. Columbretes are volcanic islands, consist of small islands and rocks that are spread on the 2500 ha. They are depopulated and protected by state environmental law. Tabarca Island is located 11 miles from Alicante. It is habited during all year and during the summer the population increases as for the tourism. Tabarca is declared as a Marine reserve and it is protected by state environmental law.

Valencian coast includes various geomorphologic elements as beaches, dune chains and cliff coasts. Terrestrials systems feeding the coastal waters with sediments that mostly include clay, sand and gravel. This sedimentary material is dispersed all around the coast line due to the marine currents. The most dominating current for the Valencian coast line has a direction North-South that makes a river Ebro (which flow into Mediterranean Sea, but on the south of Cataluña) an important sediments source.

The type of substratum determines the two principal coastal ecosystems: sandy coast with a soft seabed and cliff coasts with a rocky bottom. Due to the instability of the sandy beaches (marine erosion), artificial marine protection structures are installed all along the coast.
Geomorphologically, the coast is divided by San Antonio cape on the North and South. At the North coast, sandy or gravel beaches dominate, with short parts of rocky coasts. The South part is predominated by cliff and rocky coasts and, at the south of Santa Pola cape, sandy beaches prevail.

Circulation in the West Mediterranean sea is determinate by temperature or salinity differential (thermohaline circulation, thermo- referring to temperature and -haline referring to salt content, factors which together establish the density of sea water), though in the onshore, it could be modified by winds, waves or other environment factors.

In general, Superficial Mediterranean circulation makes a cyclonic (counterclockwise in Northern Hemisphere) spin. Atlantic Ocean waters that enter in the Straits of Gibraltar are spread along the North African coast and in the central Mediterranean basin (Figure 4). The currents pass within Balearic Island and Sardinia, cover the South French coast and the Gulf of Lion and return to South tracing the Spanish littoral.

Observing the superficial currents in the scale of the Valencian coast, general direction is North-South. Therefore, currents direction could be changed (in the low-scale) due the natural defenses or anthropic influence (marine structures).

Mediterranean sea is classified as a micro-tidal sea, with astronomic tide (caused by attractions of sun and moon) variation less than half meter (maximum values are around the 50cm, Figure 3). On the Figure 3 it is possible to notice that the Valencian littoral has almost lowest astronomic tides in the Mediterranean basin. For the Valencian coast, the metrological tides are more significant (caused by local meteorological conditions) and they can exceed values of 1m variations. This type of tides does not generate sea currents or make significant influence on the existing ones, but they are important because of their potential influence on the coastal morphologic for the sediment coasts (sea erosions, inundations, etc.).

The waves in the Mediterranean basin are an important hydrodynamic factor as generator of the local currents (littoral zone) and the transport of sediments. The wave influence on the Valencian coastal waters is determined by two fundamental geographical areas divided by Sant Antonio cape. North area is generally affected by waves proceeding from North-East and East (the area is determinate by the Balearic Islands and the importance of components vary due the coastal fetch exposition) and the South area is exposed from North-East until the South-West due the coastal line orientation.

Dominant winds in the area come from the North-East (Gregal), East (Levante) and the West winds from inland peninsula. At the South of Sant Antonio cape, dominant winds come from the South-West and, due to the geomorphological properties of the inland, the winds from the West are not that significant. During most part of the summer, the breeze SE (Garbi) regime for the Valencian gulf coast dominates.

2.2.4.Coastal water bodies

Water Frame Directive guidance document n.º2 , "Identification of water bodies" define coastal waters as:
"*Every point of which is at a distance of one nautical mile on the seaward side from the nearest point of the baseline from which the breadth of territorial waters is measured, extending where appropriate up to the outer limit of transitional waters*".

The essence and one of the first steps in the implementation of an environmental policy of WFD is to identify the water bodies. The identification of the water bodies have to be in accordance with river basins and the w. b. should be coherent sub-units. In other words, the coastal water bodies should be identified and assigned to the nearest or most appropriate river basin district.

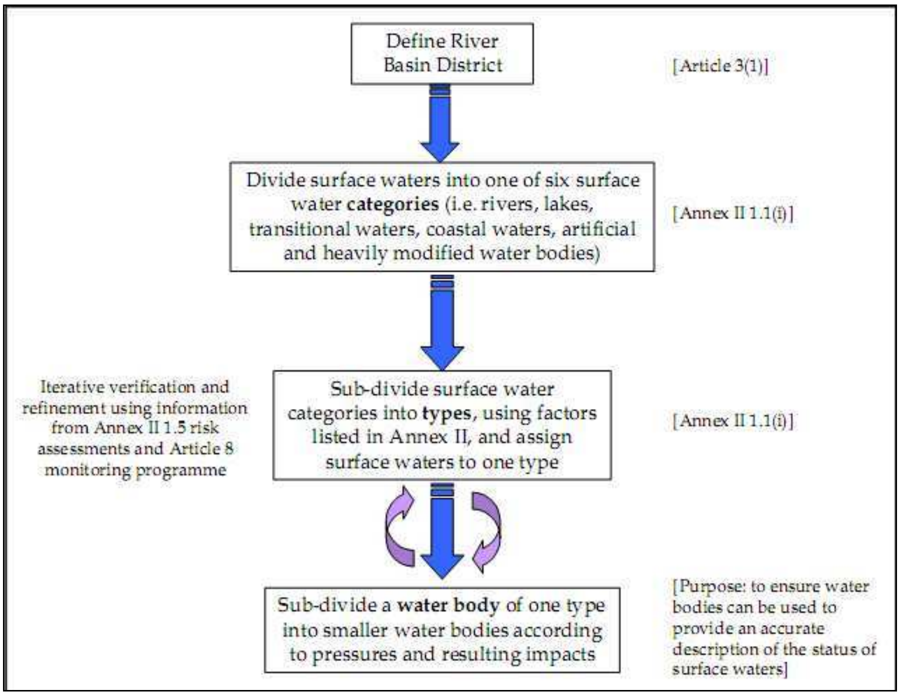


Figure 6: Summary of suggested hierarchical approach to the identification of surface water bodies.

All coastal waters in Valencia belongs to two river district Jucar and Segura (Figure 7). As the Jucar district (hydrographical demarcation) is spread on more than 80% of coast line, for the identification and division of the water bodies are used the Jucar sub-basins.

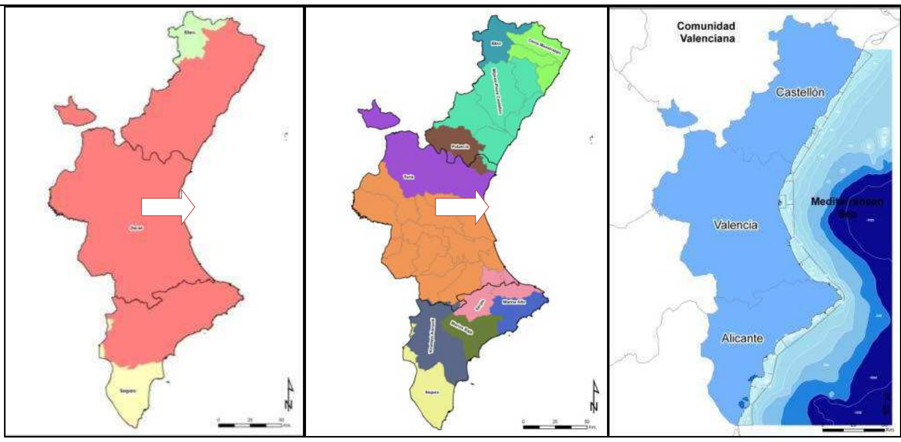


Figure 7 : Basins and sub-basins identification and division of coastal water bodies

When assigning a stretch of coastal water to a River Basin District the objective is to ensure that coastal waters are assigned to the closest possible or the most appropriate natural management unit and to minimize any unnecessary splitting of coastal stretches.

Apart from sub-basins division, coastal morphology and geographical futures are used to establish limits between water bodies (as Sierra de Irta, Cullera Cape…).

Finally, the 437 km of coast line is divided on 18 coastal water bodies and 6 water bodies that are significantly influenced by the commercial ports (Annex 2). The water bodies that have significant port influence are indicated on the Table 1.

M.A.	LIMITES	TIPO	Area (Ha2)
0041	Puerto de Castellon	Puerto	2068
006	Puerto de Sagunto	Puerto	3533
0081	Puerto de Valencia	Puerto	5438
0101	Puerto de Gandia	Puerto	441
0102	Puerto de Denia	Puerto	260
0161	Puerto de Alicante	Puerto	814

Table 1:Coastal water bodies significantly influenced by ports

The water body 0081 (Puerto de Valencia) covers the biggest area (5 438 Ha^2) as the most important commercial port in Valencia region. The only commercial port managed by local government (Puerto de Denia) covers the smallest area as 260 Ha^2.

The division from the rest of coastal waters is applied because in these areas the littoral is significantly modified. This morphological change probably will result that all port influenced areas will be identified as Heavily Modified Water Bodies. Beyond that, in the low level of renovation environment (Mediterranean ports where the water renovation is mainly driven by local wind fields and micro-tidal forces are negligible) can be excepted reduced water quality.

From the 18 coastal water bodies (Table 2 debajo de) the biggest area covers w.b. 010 (26 833^2) that is extending from Gandia until San Antonio Cape. The smallest area covers (excluding the port influenced w.b.) water body 012 (3 136 Ha^2) that is situated between Punta de Moraira y Peñon de Ifach (Calpe). Water bodies on the North until the Sant Antonio cape, have a sand (sedimentary) sea bad substratum, excluding the w.b. 002 (Sierra de Irta). Water body 002 is determinate by Irta Mountain at Jucar north sub-basin, where *rocky substrate* dominates the *sea* bottom. All this water bodies have a low sea bottom gradient and they are identified as shallow waters.

Passing the Sant Antonio cape, the shallow waters are substituted with rocky deep coast (rocky substrate) water bodies and there are extending until w.b. 015. Water body 015 is identified as a shallow with transition substratum (rocky bottom to sand seafloor). South behind w.b. 014, at littoral of Valencian Community dominates the shallow waters with a low slope gradient of seafloor. Water bodies 016 and 018 are identified as shallow waters with sand sea floor and 017 and 019 as shallow waters with mixed sea bed properties (as was detected for the w.b. 015)

During the intercalibration process (1.3, page 9) water types for the Mediterranean Sea are defined on the surface density values or equivalent salinity (Table 4 and Table 5, part Intercalibration page 11). To identify water body types (due to influence of fresh water inputs) annual mean salinity values of inshore samplings are used. There are fourteen water bodies designated as water bodies type II-A (Not directly affected by freshwater inputs) and 10 water bodies classified as type III (Not affected by freshwater inputs).

On Figure 6, it is easily observed that the area affected by freshwater inputs is located on the north coast, bordering Catalan coastal waters and on the south San Antonio Cape. Part of the coast with water bodies designated as a type III on the south is bordering Murcia coastal waters.

M.A.	LIMITES	TIPO	Area (Ha2)
001	Limite CV-Sierra de Irta	Costera	13007
002	Sierra de Irta	Costera	4427
003	Sierra de Irta-Cabo de Oropesa	Costera	10553
004	Cabo de Oropesa-Burriana	Costera	14100
005	Burriana-Canet	Costera	12228
007	Costa norte de Valencia	Costera	15227
008	Puerto de Valencia-Cabo de Cullera	Costera	19731
009	Cabo Cullera-Puerto de Gandia	Costera	17063
010	Puerto de Gandia-Cabo de San Antonio	Costera	26833
011	Cabo San Antonio-Punta de Moraira	Costera	5699
012	Punta de Moraira-Peñon de Ifach	Costera	3136
013	Peñon de Ifach-Punta de les Caletes	Costera	8944
014	Punta de les Caletes-Barranco de Aguas de Busot	Costera	14714
015	Barranco de Aguas de Busot-Cabo Huertas	Costera	7555
016	Cabo Huertas-Santa Pola	Costera	13400
017	Santa Pola-Guardamar del Segura	Costera	14404
018	Guardamar del Segura-Cabo Cervera	Costera	10862
019	Cabo Cervera-Limite CV	Costera	13848

Table 2: Coastal water bodies in littoral of Valencia (not influenced by ports)

This study is focused on the coastal water bodies from the table above, but it is added coastal water body 006 which is classified as port influenced. The exception is made with water body 006 (Puerto de Sagunto) because it covers the area significantly larger than port installations. Water body has a very reduced size (if we compare with 18 w.b. listed above) and it is clearly influenced by Sagunto Port. This study did not find adequate to differentiate this water body from the rest of water bodies of Valencia Gulf.

Figure 8: Valencian coastal water bodies (more detailed maps at Annex 2)

3.Methods and source of data

3.1.Monitoring networks that provides data for the Valencian coastal waters

3.1.1.Coastal waters monitoring network

Water Frame Directive demands that Member States establish the monitoring networks, to obtain for each river basin and their waters a general and complete vision of their ecological status. For coastal waters a survey is demanded so that water bodies could provide ecological status and related ecological potential.

Monitoring information is needed for the following reasons:

- Classification of status of all water bodies.
- To support risk assessment procedures.
- Design of future monitoring programs.
- Assessment of long-term changes whose causes are both natural and anthropogenic.
- Assessment of compliance with standards and objectives.
- Estimation of pollution load transfers across international boundaries or into seas.
- Assessing the efficacy of measures applied to water bodies designated as at risk.
- Ascertaining formerly unidentified reasons for failure to achieve environmental objectives.
- Assessing the impact of accidental pollution.
- Use in inter-calibration exercises.

In order to obtain the required information, Member States have to establish an integrated monitoring programme that embraces the physical, chemical and biological data, needed to assess the status of coastal water bodies.

The monitoring network that surveys 437km of Valencian coast started to be functional in August 2005. By network (Annex 2) are surveyed 18 coastal and 6 port influenced water bodies (Castellon, Port of Sagunto, Valencia, Gandia, Denia and Alicante). As it was agreed on the first Intercalibration exercise meeting, the survey is maintained by monthly campaigns; therefore, the WFD demands a minimum of three months frequency. The survey network, at each monitoring station takes an inshore water sample. Each water sample is taken at 10cm depth, in the surf zone, but before the waves collapsing area (to avoid the sediment resuspension that can affect the bio-chemic water quality). Taking the water samples at the inshore, makes the survey more efficient and easy, without needs to use a vessel.
The monitoring networks analyze:

- Environmental data which are taken in situ at the moment of taking water sample: temperature, wind speed and direction, cloudiness and sea condition
- Biological data: phytoplankton recounts and composition analysis, concentration of chlorophyll *a*
- Physic-chemical data: Salinity, pH values and concentration of nutrients

Sample stations locations (Annex 2) are chosen according to their continental waters influence (rivers, gullies…) and of well-know anthropogenic impacts (sea outfalls, geomorphologic changes of the coast…). The first survey was a pilot campaign in August 2004. It provided first information about the Valencian coastal waters. During the pilot campaign 134 water samples were analyzed, sampled all through 437km of coast line. After that first campaign, it was decided to reduce the number of sample station because, at that moment, the information was copious and overlapping. For the next campaign the number of sampling stations was bisected on 70 (the minimum number of samples during the period august 2005-july 2008) and was changing in function of survey need, investigation and requirements to implement WFD. According to the diagram below, after October 2005, the number of sampling stations has been increased and, in July 2008, it reached its maximum with 108 sample stations.

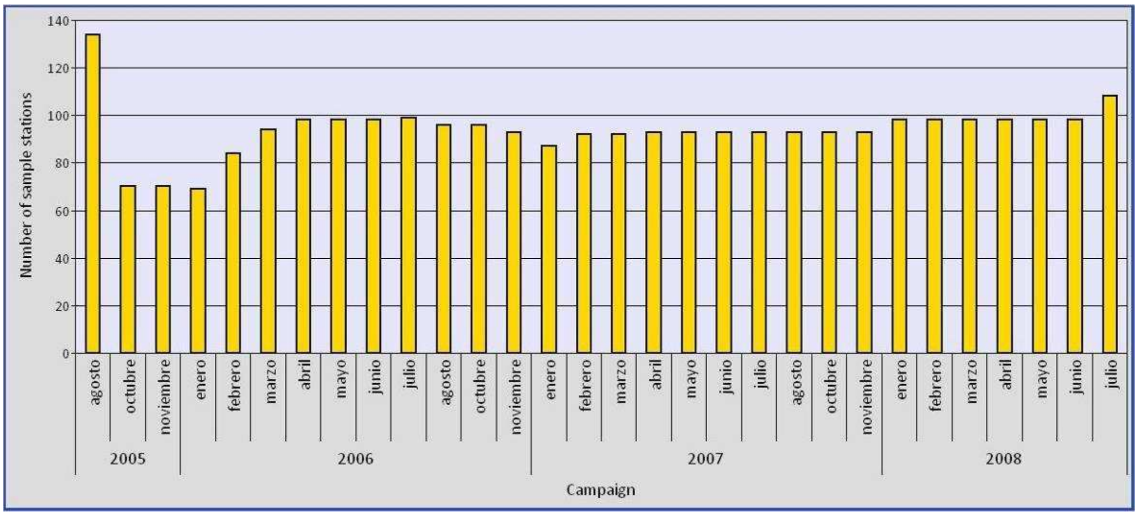


Figure 1: Number of sampling stations for each campaign at the period August 2005 till July 2008

During these three years more than 2500 samples have been analyzed:

Water body 001

-156 water samples
-5 active sample stations, 4 active from the pilot campaign, August 2005
-1 station active since February 2006
-Historical data from the 10 sample stations

Water body 002

-147 water samples
-5 active sample stations , 1 active from the pilot campaign
-4 stations active since February 2006
-Historical data from the 6 sample stations

Water body 003

-102 water samples
-4 active sample stations, 3 active from the pilot campaign
-1 station active since June 2006
-Historical data from 8 sample stations

Water body 004
- -107 water samples
- -5 active sample stations, 3 active from the pilot campaign
- -2 stations active since January 2008
- -Historical data from 11 sample stations

Water body 0041
- -34 water samples
- -4 active sample stations, 1 active from the pilot campaign
- -3 stations active since July 2008
- -Historical data from 4 sample stations

Water body 005
- -168 water samples
- -5 active sample stations, 5 active from the pilot campaign
- -Historical data from 10 sample stations

Water body 006
- -85 water samples
- -5 active sample stations, 2 active from the pilot campaign
- -3 stations active since January 2008
- -Historical data from 5 sample stations

Water body 007
- -118 water samples
- -6 active sample stations, 1 active from the pilot campaign
- -3 stations active since January 2007 and 1 station since July 2006
- -Historical data from 7 sample stations

Water body 0081
- -66 water samples
- -4 active sample stations, 2 active from the pilot campaign
- -2 stations active since July 2008
- -Historical data from 5 sample stations

Water body 008
- -149 water samples
- -5 active sample stations, 4 active from the pilot campaign
- -1 station active since April 2007
- -Historical data from 5 sample stations

Water body 009
- -166 water samples
- -5 active sample stations, 5 active from the pilot campaign
- -Historical data from 8 sample stations

Water body 010
- -187 water samples
- -6 active sample stations, 6 active from the pilot campaign
- -Historical data from 12 sample stations

Water body 0101
- -34 water samples

-4 active sample stations, 1 active from the pilot campaign
-3 stations active since July 2008
-Historical data from 4 sample stations

Water body 0102
-No data
-2 active sample stations, active since July 2008

Water body 011
-166 water samples
-6 active sample stations, 3 active from the pilot campaign
-3 stations active since March 2006
-Historical data from 16 sample stations

Water body 012
-112 water samples
-5 active sample stations, 3 active from the pilot campaign
-2 stations active since January 2008
-Historical data from 5 sample stations

Water body 013
-104 water samples
-4 active sample stations, 3 active from the pilot campaign
-1 station active since June 2006
-Historical data from 10 sample stations

Water body 014
-128 water samples
-4 active sample stations, 4 active from the pilot campaign
-Historical data from 7 sample stations

Water body 015
-34 water samples
-4 active sample stations, 1 active from the pilot campaign
-3 station active since July 2008
-Historical data from 6 sample stations

Water body 016
-98 water samples
-4 active sample stations, 2 active from a pilot campaign
-2 station active since February 2007
-Historical data from 6 sample stations

Water body 0161
-62 water samples
-4 active sample stations, 2 active from a pilot campaign
-2 station active since July 2008
-Historical data from 5 sample stations

Water body 017
-118 water samples
-4 active sample stations, 3 active from a pilot campaign
-1 station active since February 2007

-Historical data from 5 sample stations

Water body 018

-65 water samples
-4 active sample stations, 1 active from a pilot campaign
-2 station active since February 2007 and 1 since July 2007
-Historical data from 5 sample stations

Water body 019

-133 water samples
-4 active sample stations, 4 active from a pilot campaign
-Historical data from 9 sample stations

More detailed data are presented at Annex.2

3.1.2.Pluviometric Jucar river basin network

In order to achieve a modern and efficient management of the hydrological resources and its operation systems, the General Directorate for Hydraulic Works created a program for the establishment of an Automatic Hydrological information System (SAIH Sistema Automático de Información Hidrológica). The SAIH is a real-time information system, designed and firstly implemented in 1989 to the Jucar basin. It is created to facilitate the decision making in the management of the hydrological resources and the forecast of floods. The SAIH captures in real-time the hydrological and hydraulic (infrastructures situation) data and other basic meteorological data. It transmits all data to the corresponding command centre, where they are analyzed and applied to the solution of management of the water, in both normal circumstances (exploitation) and in emergency situations (droughts and floods).
The SAIH captures the data by means of different sensor devices and transmits them to a communication network. At the beginning, its communications system was by radio but, currently, that system is being replaced by means of a network VSAT (satellite system), between the concentration points and the center of process.

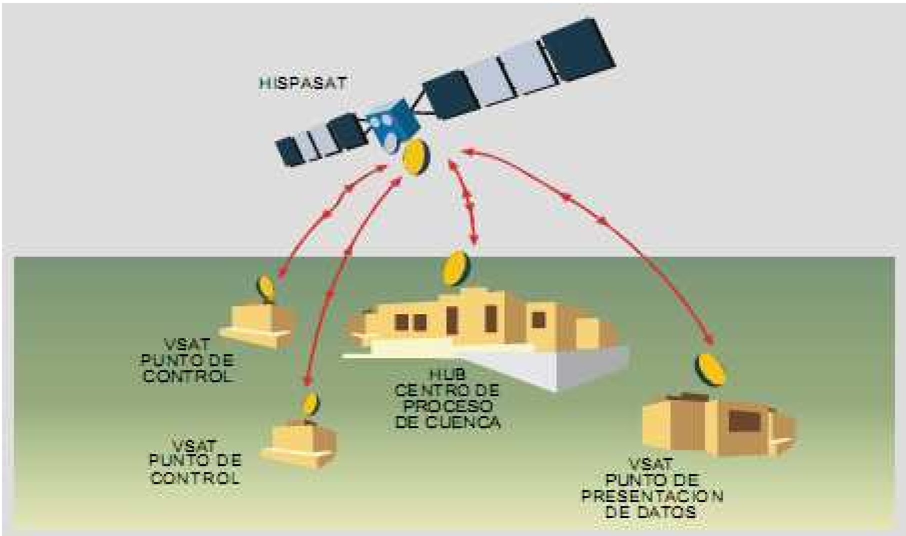


Figure 2: Jucar basin communication system that transmit data for the Jucar basin (from the stations to the control center)

Jucar basin SAIH consists of 137 stations that provide different type of data, as mentioned above. For this study, pluviometric data are used. Those data are obtained from selected pluviometers that are located as near as possible to the coast line with the condition that basin river flow is not interrupted by dam or any anthropogenic process or obstacle.

3.1.3.Wave data networks

Territorial waters of Valencia are surveyed by various measurement networks. These networks are designed to obtain, in real time, detailed information about the physical features (waves, tides, temperature, wind, etc) of the Spanish territorial waters. There are four networks with different objectives: deep waters, coasts, current meters and tide gauges. For the needs of this study, data obtained by Coastal and the Deep Sea networks are used. Both networks (Figure 3) consist of wave buoys - used to measure the movement of the water surface as a wave train. The wave train is analyzed to determine statistics like the significant wave height and period, and wave direction.
The Deep Sea Network is based on 13 Seawatch and 3 WaveScan buoy stations which are moored in the open sea far away from the coast line. Instruments are located at points with depths between 200 and 800 m and measure atmospheric and oceanographic parameters. Measurements are transmited every hour to Puertos del Estado via satellite and directly posted to this web page.

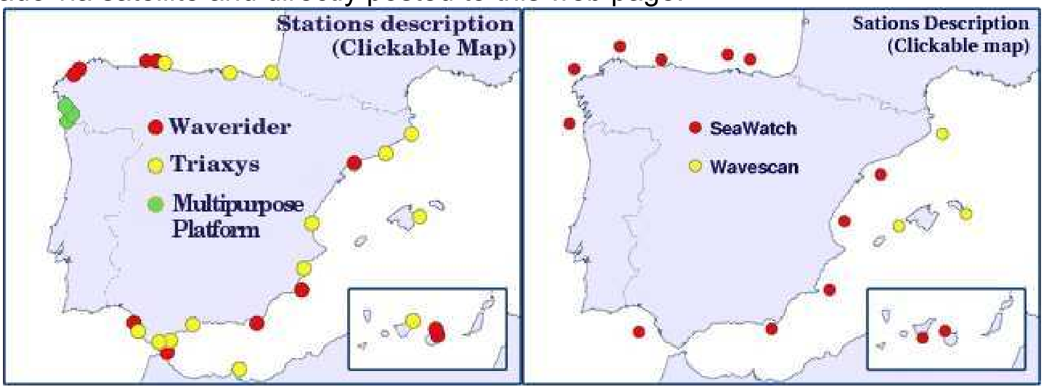


Figure 3: Deep Sea Network and Coastal Network at the Spanish territorial waters (www.puerto.es)

The Coastal Network, belonging to Puertos del Estado, is providing real time data in some specific points located at shallow waters. The main objective of the measurements is to complement those of the Deep Sea Network at those locations of special interest for the port operations or wave modeling validation. The buoys employed are scalar and directional Waverider and directional Triaxys. Coastal buoys are moored in swallow waters, in the vicinity of the port facilities (less than 100 meters depth). In most of the cases, measurements are affected by the shape of the coast line and the sea ground. This means that they obtain information that is only representative of the local conditions.

Figure 4: Wave buoys different types used to obtain data for the Spanish territorial waters (www.puerto.es)

At the coastal waters of Valencia are located three wave buoys, two of them are part of Coastal Network and one from the Deep Sea Network. For this study, WANA data set is used, generated and distributed by Puertos del Estado. The WANA data set consists of predictions using the wave model WAM (WAMDI, 1988). These predictions are done in a grid with a cell size of 0.125° x 0.125° (15 Km). The obtained information of the model in each point is the wave directional spectrum. The wave directional spectrum is used to obtain basic information, such as the significant height (Hs), period (Tp), mean and directions *(Gunther et al., 1991; Gómez & Carretero et al, 1997).*

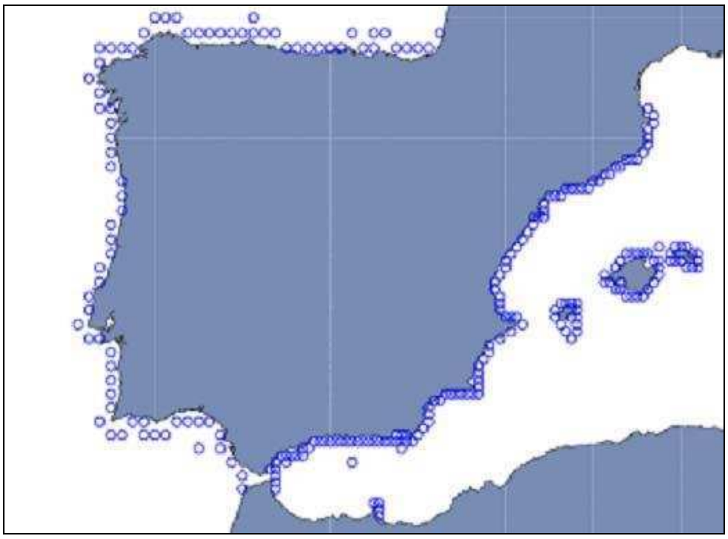


Figure 5: Wana points obtained by the WAM model

For the Valencian coast also exists the SeaMAR data, but in this moment this “virtual network” is inactive and data series are only completed until year 2001. WANA set covers the period from 1996 and provides data currently. For each annual point resulting reports are posted on the web site (www.puerto.es) where data are processed and collected for the meteorological year (December to November). Results are presented in the plots and statistics.

3.2.Analytical methods

For this study are used data obtained by:

-Coastal waters monitoring network established for the WFD
-Pluviometric Jucar river basin network
-Wave buoy networks established by Puertos de Estado

Each network provided data that were analyzed and used to obtain the results of this study.

3.2.1.Coastal waters monitoring network established for the WFD

From the Coastal waters monitoring network established to provide information for the WFD are used chlorophyll a and salinity data. As it mentioned before the water samples are taken in the 437 Km of the Valencian coast line. Each water sample are taken in the two liters high density polyethylene bottle. After water sampling, all samples are temporally stored at the low temperature in the portable refrigerator, during the day of the campaign. This temporally storing never exceed 12 hours, before water samples arrived in the laboratory. When water samples arrive in the laboratory, they are stored into different aliquots due the different storing method (APHA 1988). 800 ml are filtered by cellulose acetate membrane to obtain the nutrient and salinity concentrations. The cellulose acetate membrane are stored on the -20 ºC and after are used to obtain the chlorophyll a concentrations.
Salinity concentrations are obtained by Portasal 8410A salinometer, which is previously calibrated with standard sea water. The results have the ± 0,01 salinity unit precision.
To obtain chlorophyll a concentrations, are used trichromatic method based on the spectroscopy (APHA 1998). Filters are introduced to 6 ml 90% acetone with 1% calcium carbonate. Optic density is determinate by different bands 630, 647 and 664 nm, to obtain pigment content. Chlorophyll a concentrations are calculated according to the equations of Jeffrey and Humphrey (at al.1975).

3.2.2.Pluviometric Jucar river basin network

Pluviometric data were obtained by Jucar river basin Automatic Hydrological information System (SAIH). The SAIH system has 137 stations that transmitting the data via VSAT satellite system. For this study are used only pluviometric data obtained by rain gages. A rain gauge (also known as a udometer or a pluviometer) is used to gather and measure the amount of precipitation over a set period of time. The rain gages used by SAIH is the tipping bucket that consists of a large copper cylinder set into the ground. At the top of the cylinder is a funnel that collects and channels the precipitation. The precipitation falls onto one of two small buckets or levers which are balanced in same manner as a scale. After an amount of precipitation equal to 0.2 mm falls, the lever tips and an electrical signal is sent to the recorder. As the all rain gages are part of SAIH, the data are transmitted to the Jucar Confederation in the real time.

3.2.3.Wave buoy networks established by Puertos de Estado

Traditionally the description of the wave conditions at a certain location and time is characterized by some parameters that define *the sea state*. Nowadays the sea state can be assessed through buoys instruments (scalar and directional).

In case of Puertos del Estado buoys, every hour they record instant elevations of sea level for a time interval in which the sea state can be considered to be constant. This duration (e.g. 20 minutes) depends on the buoy model and will be enough to obtain a representative sample of the wave conditions in that hour. Then the time series of instant elevations is processed using standard methods (zero crossing and spectral analysis) to determine the parameters that characterize the sea state in that hour: Hs,Tm,Tp, etc. Today, all this process is usually done on board the buoy and later the outputs are transmitted to land. Though, the scalar buoys of the Coastal Network still transmit the measured sample of elevations to a station on land where the analysis is performed.

The data used for this study is WANA set, which are the daily wave forecast output from the third generation Wave Model, WAM (Gunther et al.,1991).

The WAM-model solves the wave transport equation explicitly without any presumptions on the shape of the wave spectrum. It represents the physics of wave evolution in accordance with our current knowledge for the full set of degrees of freedom of a 2D wave spectrum. The wave directional spectrum is used to obtain basic information such as the significant height (Hs), period (Tp), mean and directions *(Gunther et al., 1991; Gómez & Carretero, 1997).* The results of the WANA set are published in the form of the annual reports (significant heights diagrams and wave roses) on the web site of Puertos de Estado that are used for this study.

3.3.Provisional ecological classifications

The use of chlorophyll *a* might provisionally set an ecological quality status for the water bodies after three years monthly sampling campaigns. There is a serious lack of data in almost all water bodies influenced by the ports activities until now. From the very beginning, these water bodies have been surveyed by one or maximum two sample stations. The more adequate survey started on July 2008, when at least 4 sample stations were situated in each port influenced water body. Finally, for the coastal waters of the Valencian Community, it is also established a minimum of four sample stations per water body.
For these three years, more than 2500 samples have been analyzed (details at chapter 3.1.1 Coastal waters monitoring network)

During the Intercalibration exercise, all Member States decided that it is necessary a minimum of 5 years historical data for any ecological classification. In this study, for each water body (where lack of data has not been identified) four calculated ecological quality statuses are established. None of them are official but they are all established in the purpose of analyzing the spatial and the temporal reduction of the monitoring network. Three years time interval is divided into three one-year periods (August 2005/July 2006, August 2006/July 2007, August 2007/July 2008) which contains 12 sampling campaigns. The ecological classification for each water body is established for each one-year period and also for the whole period of three years (August 2005-July 2008).

Subsequently, the 90th percentile of chlorophyll *a* values is calculated for each water body and for all periods. The quality status is established on the basis of chlorophyll *a* 90th percentile value, using the ecological classes thresholds (High/Good and Good/Moderate), published according to the last document of the European Commission (COMMISSION DECISION, establishing, pursuant to Directive 2000/60/EC, the values of the Member State monitoring system classifications as a result of the intercalibration exercise). As the reference condition values for type II-A and type III are not distinct in this document, the Ecological Quality Ratio (EQR) was not calculated.
Analyzing the number of sampled data for the coastal water bodies 015 and 018, it is possible to establish the provisional ecological status only for the three year period. Finally, due to the lack of data, this classification certainly provides a biased idea of eutrophication status and distorted information of ecological systems.

During the process of calculating 90th percentile it was highlighted that different statistical software use different algorithm to obtain quantile values. This means that the 90th percentile will not be the identical value, if it is calculated by different statistical software (SPSS, StatGraphic, S-Plus, Excel…).
There is no a universally accepted definition of a percentile from sample data, as Hyndman and Fan claimed in 1996 on "Sample Quantiles in Statistical Packages". In this study, they discuss nine different definitions of percentile and related algorithms.

To establish a provisional ecological quality status, 90th percentile for each water body was calculated using the definition that is implanted in the spreadsheet Excel® and the statistical packages SPSS® and STATGRAPHICS®..

3.3.1.Provisional ecological classification status August 2005- July 2008

The 90th percentile values are calculated for the 19 water bodies, excluding the port influenced ones because of lack of data.
In the Table 1 below, it is possible to observe:

- Campaigns period
- Water body code
- Type of water body in accordance to the continental influence level
- 90th percentile of chlorophyll *a* (mg/m^3) calculated by SPSS□
- Provisional ecological status calculated in function of 90th percentile values calculated by SPSS□
- 90th percentile of chlorophyll *a* (µg/l) calculated by EXCEL algorithm
- Provisional ecological status calculated in function of 90th percentile values calculated by EXCEL algorithm
- Discrepancy of 90th percentiles calculated by different software (SPSS□ 90th percentile - EXCEL P90th percentile)
- Number of the samples for each water body that are used to obtain 90th percentile

PERIOD	WATER BODY	TYPE	P90 CHL a SPSS(µg/l)	E.C.S. SPSS	P90 CHL a EXCEL(µg/l)	E.C.S. EXCEL	p90 SPSS-p90 EXCEL	N. of samples
2005/08-2008/07	001	II A	3,54	HIGH	3,49	HIGH	0,05	156
2005/08-2008/07	002	II A	3,50	HIGH	3,42	HIGH	0,08	147
2005/08-2008/07	003	II A	2,18	HIGH	2,16	HIGH	0,02	102
2005/08-2008/07	004	II A	3,72	HIGH	3,58	HIGH	0,13	107
2005/08-2008/07	005	II A	5,69	GOOD	5,67	GOOD	0,02	168
2005/08-2008/07	006	II A	4,73	HIGH	4,60	HIGH	0,13	85
2005/08-2008/07	007	II A	5,13	GOOD	5,10	GOOD	0,03	117
2005/08-2008/07	008	II A	7,96	MODERATE	7,92	MODERATE	0,04	149
2005/08-2008/07	009	II A	8,22	MODERATE	7,87	MODERATE	0,36	166
2005/08-2008/07	010	II A	5,91	GOOD	5,67	GOOD	0,25	187
2005/08-2008/07	011	III	2,26	GOOD	2,24	GOOD	0,03	166
2005/08-2008/07	012	III	2,27	GOOD	2,25	GOOD	0,02	112
2005/08-2008/07	013	III	3,61	MODERATE	3,39	GOOD	0,22	104
2005/08-2008/07	014	III	1,79	HIGH	1,76	HIGH	0,03	128
2005/08-2008/07	015	III	4,76	MODERATE	4,45	MODERATE	0,31	34
2005/08-2008/07	016	III	6,26	MODERATE	6,23	MODERATE	0,03	98
2005/08-2008/07	017	III	5,46	MODERATE	5,40	MODERATE	0,06	118
2005/08-2008/07	018	III	3,32	GOOD	3,17	GOOD	0,16	65
2005/08-2008/07	019	III	3,10	GOOD	2,98	GOOD	0,12	133

Table 1: 90th percentile and ecological classifications obtained by SPPS and EXCEL for the three year period

To establish a provisional ecological status as described in chapter “Intercalibration exercise”, the ratio of 50% is used for the relation of Nearshore-Inshore for type II-A and type III with values:

Type II-A; High/Good limit 4,8 mg/m^3 and Good/Moderate 7,2 mg/m^3
Type III; High/Good limit 2,2 mg/m^3 and Good/Moderate 3,6 mg/m^3

At the North part of the Valencian Coastal waters, which are recognized as influenced by freshwater (type II-A), five water bodies are associated with High, three with Good and two by Moderate ecological status. Water bodies situated at the South of San Antonio cape are identified as not influenced by fresh water (type III) and it seems that, according to these results, they are exposed by higher anthropogenic perturbation, or high freshwater discharges during the sampling period. Just one water body is classified with High ecological status, four with Good and other four with Moderate ecological class. This classification is established on the basis of 90th percentile calculated by SPSS and the qualitative results provided by Excel demonstrate discrepancy at water body 013. The 90th percentile computed by SPSS 3.61 µg/l classifies water body as Moderate and the same percentile calculated by EXCEL provides value 3.39 µg/l and the same water body is classified as Good.

The quantitative discrepancy between 90th percentiles calculated by two different algorithms (presented in the 8th column of Table 1) may have deep consequences on ecological classification as it is demonstrated on the 013 water body example. The values dissimilarity of the same statistical metric (in this case 90th percentile) allows a manipulation of the ecological results and may mislead conclusions about the ecosystems.
Finally, it could be concluded that there is a strong need to assess this issue by the institution that manage the implementation of the Water Frame Directive (European Commission, ECOSTAT…) and it should be a theme of discussion at the next meeting of phytoplankton group in Intercalibration.

3.3.2.Provisional ecological classification status for one year periods

Ecological classification statuses, as described above, are established for each water body for one year periods.

August 2005-July 2006

PERIOD	WATER BODY	TYPE	P90 CHL a SPSS(μg/l)	E.C.S. SPSS	P90 CHL a EXCEL(μg/l)	E.C.S. EXCEL	p90 SPSS-p90 EXCEL	N. of samples
2005/08-2006/07	001	II A	3,36	HIGH	3,03	HIGH	0,33	49
2005/08-2006/07	002	II A	3,33	HIGH	3,12	HIGH	0,21	35
2005/08-2006/07	003	II A	2,19	HIGH	2,16	HIGH	0,02	30
2005/08-2006/07	004	II A	6,98	GOOD	6,50	GOOD	0,47	32
2005/08-2006/07	005	II A	4,80	HIGH	4,52	HIGH	0,28	54
2005/08-2006/07	006	II A	5,37	GOOD	4,44	HIGH	0,93	20
2005/08-2006/07	007	II A			3,90	HIGH		10
2005/08-2006/07	008	II A	9,19	MODERATE	7,99	MODERATE	1,21	44
2005/08-2006/07	009	II A	5,05	GOOD	4,60	HIGH	0,45	53
2005/08-2006/07	010	II A	6,02	GOOD	5,93	GOOD	0,09	57
2005/08-2006/07	011	III	1,56	HIGH	1,50	HIGH	0,06	50
2005/08-2006/07	012	III	2,97	GOOD	1,97	HIGH	0,99	34
2005/08-2006/07	013	III	1,66	HIGH	1,30	HIGH	0,35	30
2005/08-2006/07	014	III	1,63	HIGH	1,54	HIGH	0,09	41
2005/08-2006/07	015	III			3,84	MODERATE		11
2005/08-2006/07	016	III	5,98	MODERATE	5,33	MODERATE	0,65	22
2005/08-2006/07	017	III	3,11	GOOD	2,90	GOOD	0,21	34
2005/08-2006/07	018	III			2,20	GOOD		13
2005/08-2006/07	019	III	3,39	GOOD	3,00	GOOD	0,38	44

Table 2: 90th percentile and ecological classifications obtained by SPPS and EXCEL for the annual period August 2005-July 2006

As it was mentioned before, the water bodies 007, 015 and 018 have a lack of data (for the period of August 2005- July 2006) with less than 20 samples per water body. The statistical program SPSS denied to calculate a 90th percentile values for the data series with less than 20 samples, though for the Excel this was not representing a problem. So, for the first period of data we have a qualitative and quantitative discrepancy of results and different number of ecological classifications.

For the type II-A using SPSS, four water bodies are qualified as High (001,002,003 and 005), four as Good (004,006,009 and 010), one as Moderate quality status (008) and one water body is detected lack of data to be ecologically qualified (007).

For the type III, using also SPSS, three water bodies are classified as high (011,013 and 014), three (012, 017 and 019) as Good, one as Moderate and two as water bodies with lack of data (015 and 018).

Using the Excel algorithm, the results are different for the water bodies 006, 009 and 012, which are classified as a High quality with discrepancy 0,93, 0,45 and 0,99 µg/l, respectively. The water body 007 is classified as High, 015 as Moderate and 018 as Good quality, on the base of 10, 11 and 13 water samples.

August 2006-July 2007

PERIOD	WATER BODY	TYPE	P90 CHL a SPSS(µg/l)	E.C.S. SPSS	P90 CHL a EXCEL(µg/l)	E.C.S. EXCEL	p90 SPSS-p90 EXCEL	N. of samples
2006/08-2007/07	001	II A	3,69	HIGH	3,65	HIGH	0,04	56
2006/08-2007/07	002	II A	3,38	HIGH	3,34	HIGH	0,04	59
2006/08-2007/07	003	II A	2,49	HIGH	2,32	HIGH	0,17	34
2006/08-2007/07	004	II A	3,33	HIGH	3,25	HIGH	0,07	34
2006/08-2007/07	005	II A	5,74	GOOD	5,44	GOOD	0,30	58
2006/08-2007/07	006	II A	4,41	HIGH	4,12	HIGH	0,29	21
2006/08-2007/07	007	II A	5,11	GOOD	4,95	GOOD	0,16	43
2006/08-2007/07	008	II A	5,35	GOOD	5,16	GOOD	0,18	50
2006/08-2007/07	009	II A	5,51	GOOD	5,05	GOOD	0,46	59
2006/08-2007/07	010	II A	4,51	HIGH	4,49	HIGH	0,03	66
2006/08-2007/07	011	III	2,69	GOOD	2,69	GOOD	0,00	60
2006/08-2007/07	012	III	3,67	MODERATE	2,79	GOOD	0,88	36
2006/08-2007/07	013	III	4,40	MODERATE	3,65	MODERATE	0,75	36
2006/08-2007/07	014	III	1,82	HIGH	1,65	HIGH	0,17	46
2006/08-2007/07	015	III			2,27	GOOD		12
2006/08-2007/07	016	III	4,59	MODERATE	3,78	MODERATE	0,81	35
2006/08-2007/07	017	III	3,80	MODERATE	3,75	MODERATE	0,06	40
2006/08-2007/07	018	III	3,98	MODERATE	3,18	GOOD	0,80	23
2006/08-2007/07	019	III	3,25	GOOD	3,10	GOOD	0,15	47

Table 3: 90th percentile and ecological classifications obtained by SPPS and EXCEL for the annual period August 2006-July 2007

For the period of August 2006 and July 2007, just at one water body is detected lack of data and SPSS established 18 ecological statuses. For the same period, at the Northern part of the Valencian coast (type II-A) six water bodies are recognized as High (001,002,003,004,006 and 010) and four as Good ecological quality (005,007,008 and 009). Two water bodies of the Southern of San Antonio cape are classified as Good (011 and 019), one as a High (014) and five as Moderate (012,013,016,017 and 018).
Using the Excel algorithm, there is no qualitative discrepancy for the water bodies type II-A (maxim 90th percentile discrepancy is 0,46 µg/l) and for the type III, there are two water bodies classified as Good (012 and 018) which disagrees with previous classification.
Water body 015, which disposes of 12 samples for this period, is classified as a w.b. with Good ecological quality.

August 2007-July 2008

PERIOD	WATER BODY	TYPE	P90 CHL a SPSS(µg/l)	E.C.S. SPSS	P90 CHL a EXCEL(µg/l)	E.C.S. EXCEL	p90 SPSS-p90 EXCEL	N. of samples
2007/08-2008/07	001	II A	3,43	HIGH	3,21	HIGH	0,22	51
2007/08-2008/07	002	II A	4,49	HIGH	3,75	HIGH	0,74	53
2007/08-2008/07	003	II A	2,19	HIGH	2,05	HIGH	0,14	38
2007/08-2008/07	004	II A	3,44	HIGH	3,24	HIGH	0,20	41
2007/08-2008/07	005	II A	6,98	GOOD	6,65	GOOD	0,33	55
2007/08-2008/07	006	II A	4,80	GOOD	4,70	HIGH	0,10	39
2007/08-2008/07	007	II A	6,06	GOOD	5,82	GOOD	0,23	64
2007/08-2008/07	008	II A	11,81	MODERATE	11,27	MODERATE	0,54	55
2007/08-2008/07	009	II A	10,38	MODERATE	10,13	MODERATE	0,24	54
2007/08-2008/07	010	II A	8,02	MODERATE	7,58	MODERATE	0,43	64
2007/08-2008/07	011	III	2,00	HIGH	1,87	HIGH	0,13	56
2007/08-2008/07	012	III	2,17	HIGH	2,11	HIGH	0,06	42
2007/08-2008/07	013	III	4,04	MODERATE	3,84	MODERATE	0,20	38
2007/08-2008/07	014	III	1,83	HIGH	1,81	HIGH	0,02	41
2007/08-2008/07	015	III			4,60	MODERATE		11
2007/08-2008/07	016	III	7,16	MODERATE	7,13	MODERATE	0,03	41
2007/08-2008/07	017	III	9,56	MODERATE	7,90	MODERATE	1,66	44
2007/08-2008/07	018	III	3,80	MODERATE	3,43	GOOD	0,37	29
2007/08-2008/07	019	III	2,46	GOOD	2,38	GOOD	0,08	42

Table 4: 90th percentile and ecological classifications obtained by SPPS and EXCEL for the annual period August 2007-July 2008

The lack of data is only detected for water body (015) which disables calculation of 90th percentile with SPSS.
Four water bodies type II-A are classified as High (001,002,003 and 004), three as Good (005,006 and 007) and three as Moderate (008,009,010) for the last one year period. For the type III, three water bodies are classified as High, four as Moderate and just one as Good.
Qualitative discrepancy is presented for water bodies 006 (High) and 018 (Good) using Excel algorithm and water body 015 is classified as Moderate.

3.4.Temporal reduction

At the first IC meeting of the Mediterranean basin (MED-GIG), it was commonly agreed that the monitoring of phytoplankton (a complex biological indicator surrogated by chlorophyll *a*) has to be increased to a monthly frequency. This frequency of monthly campaigns results in elevated financial expenses and human efforts. After three years of monthly monitoring campaigns, it is possible to analyze historical data for the Valencian coast in the scope of reducing the frequency of the campaign. The goal of this analysis is to determine the consequences, in the ecological classification of the coastal water bodies, produced by a reduction in the campaign frequency. This analysis simulated reduced campaigns and compared the quality classifications corresponding to the reduced campaigns with that obtained for the complete dataset. Analysis of temporal reduction is separated into three parts:

-Effects of a reduction in the frequency of campaigns on the ecological classification
-Seasonal pattern regularity
-Instability in the ecological classification

3.4.1.Effects of a reduction in the frequency of campaigns on the ecological classification

In the study of the effects produced by reduction campaigns frequency, on the water body classification, we have considered the possibility of bimonthly campaigns or trimestral campaigns. In the case of bimonthly campaigns, we have studied the possibility that the campaigns will be conducted in even months or uneven months and for trimestral campaigns we have studied the three possible combinations as stated in the Table 5.

Campaigns						
	Monthly	Bimonthly Campaigns		Trimestral Campaigns		
N	original set of data	1. set	2.set	1.set	2.set	3.set
1	August	August	-	August	-	-
2	September	-	September	-	September	-
3	October	October	-	-	-	October
4	November	-	November	November	-	-
5	December	December	-	-	December	-
6	January	-	January	-	-	January
7	February	February	-	February	-	-
8	March	-	March	-	March	-
9	April	April	-	-	-	April
10	May	-	Jul	May	-	-
11	June	June	-	-	June	-
12	July	-	July	-	-	July

Table 5: Temporal reduction scheme, two sets of bimonthly and three sets of trimestral campaigns

During the period (August 2005/July 2008) more than 2300 sample were analyzed and the results were stored. The study of the impact on the ecological classification of water bodies has been developed for annual periods (August 2005/July 2006, August 2006/July 2007 and August 2007/July 2008) and for the full term of three years (August 2005/July 2008). For each period, was obtained from the full samples, the subsamples corresponding to the five bimonthly or trimestral campaigns listed in the table. Therefore, it has been analyzed 20 different sub-samples for each body of water.

After selecting the 20 subsamples from each full sample, it has been calculated, the 90th percentiles for the subsamples and it has been obtained the ecological classifications of the water body, based on the percentiles obtained. Finally, it has carried out a comparison between the ecological classification of the water body, obtained from the original sample and the ecological classifications, obtained from the different subsamples considered. This allowed us to study the possible changes that occur in the ecological classification of each water body, due to a reduction campaigns frequency.

The results obtained for each water body are as follows:

1. The number of samples used to obtain 90th percentile values for each water body (for the five reductions in the frequency of campaigns and for the four periods considered).
2. The ecological classifications established using the five types of campaigns with decreased frequency and the ecological classifications established using the complete data (for each water body and for the four periods)
3. The percentage and number of quality results, obtained using the subsamples, that either preserve, upgrade or degrade ecological classifications (for the five reductions in the frequency of campaigns and for the four periods considered)
4. Mean absolute percentage error (MAPE) in the value of the 90th percentile (for the five reductions in the frequency of campaigns and for the four periods considered)

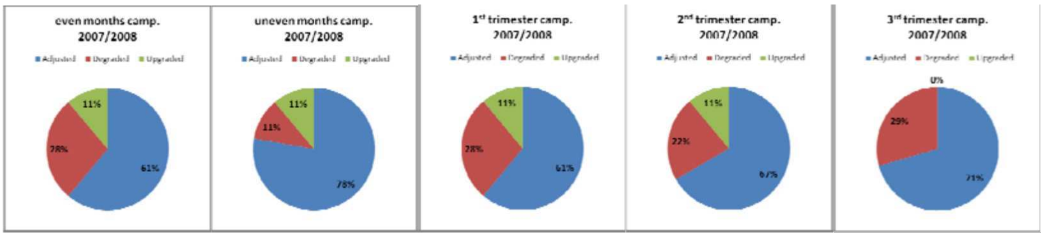


Figure 6: Quality results of bimonthly and trimestral campaigns (August 2007 – July 2008)

3.4.2.Seasonal pattern

The Valencian coast is divided into three parts with different maritime climate conditions. Maritime climates are identified based on coastal morphology and corresponding fetch. For each part, maritime climate conditions, wave, rain, salinity and chlorophyll *a* time series are analyzed.

Wave and pluviometric series are used to identify their influence on the trophic state of Valencian coastal waters. Beside that, climate seasonality and possible annual pattern for the wave, rain, salinity and chlorophyll *a* are also analyzed.

For each climate, annual wave roses are analyzed, which provide frequency, intensity and direction of wave events during the four one-year periods. Wave roses are analyzed for 2005, 2006, 2007 and 2008. Like the wave roses, wave time series are also presented in annual diagrams for the same periods. The analyzed wave data is published on the Puertos del Estado web page (http://www.puertos.es) as a part of the WANA data set network.
Rain data is provided by the Jucar Basin Authorities and the Automatic Hydrological Information System (SAIH). This data is provided as daily time series for more than 100 stations that measure precipitation in the Jucar basin. For the purpose of this study, 15-day accumulation data is generated from the daily series. By doing this, each sampling campaign date has a related 15-day accumulation (presented in diagrams), with the purpose of determining rainfall influence on coastal waters. The pluviometric data used for this study contains the amount of daily rain in addition to the quantity of rainfall (accumulation) in the last 15 days.
Further, salinity concentrations were also analyzed for one water body from each maritime climate. The analyzed water bodies are surveyed by five sampling stations. Because of this, five different salinity values are analyzed for each campaign. Campaign average salinity values are calculated in order to enable analysis using pluviometric and wave data (only one value per campaign).
Finally, campaign average chlorophyll *a* values series were analyzed. For these values, analysis was performed on interannual seasonality and what effect hydrodynamic conditions (waves) and continental influence (mostly determined by rain conditions) have on the water body's trophic condition.

3.4.3.Instability in the ecological classification

Instability in the ecological classification of a water body, after a temporary reduction occurs more easily when the 90th percentile of the chlorophyll measurements made in the system is located near the threshold that separates the ecological status High / Good or the threshold that separates the ecological status Good / Moderate. The classification is generally similar for most of the alternatives for reducing the frequency of campaigns when the 90th percentile is not located near a threshold or is sufficiently high to ensure that the water body belongs unequivocally to Moderate ecological status.

On this basis, we propose a simple technique, based on the average and the sample standard deviation of chlorophyll measurements taken at a sampling station, to know if it would be reasonable to propose a reduction in the frequency of the campaigns in some sampling station.

This technique is based on the following approach:

- We assume that measurements of chlorophyll a in a sampling station are approximately normally distributed

- We call *buffer distance average value* - average distance between the measurements of chlorophyll a in the whole sample and the nearest border (High / Good or Good / Moderate).

- After the adjustment of a normal distribution, to the set of measurements taken on a sampling station, we call *indicator value* to the distance between the 90th percentile of the adjusted normal distribution and the mean of this distribution.

- Whenever the sum of sample average and sample standard deviation (84th percentile), of measurements included in the full sample, are above the threshold Good / Moderate, although the sample is reduced, it is unlikely that the 90th percentile of a subsample for a temporal reduction, transferred this threshold and we consider that the system has a *regular classification*, regardless of the variability of the variable.

- Otherwise, if the *indicator value* is less than the *buffer distance average value*, although the sample is reduced, it is unlikely that the 90th percentile of a subsample for a temporal reduction, transferred the threshold and we can consider that the system also has a *regular classification*. But, if the *indicator value* exceeds the *buffer distance average value,* there is increased the probability that the 90th percentile of a subsample for a temporal reduction transferred the threshold and we consider that the system has an *up-and-down classification*.

It is not expected that a reduction in the frequencies of the campaigns, produces changes in the ecological classification of the system, whether this system has a regular classification, but it is easy to produce changes if the system has an up-and-down classification.

3.5.Spatial reduction

The *spatial reduction*, or reduction in the number of sampling stations reviewed, has been approached from the perspective of a problem of inter-observer variability. From the standpoint of *spatial reduction*, it has been considered that the different k-observers of sampling stations of a water body, characterize different *observers* that provide different monthly measurements of *chlorophyll a* concentration in the area. Therefore, our aim has been to quantify the concordance between the *reference observer,* which takes measurements from *all the sampling stations* currently active in a water body and the *reduced observers* which take measurements from the different *subsets* that can be formed from the currently active sampling stations .

The study of spatial reduction is also separated into three parts:

- Effects of a reduction in the number of sampling stations reviewed on the ecological classification
- Selecting the best observers
- Post-reduction management

3.5.1.Effects of a reduction in the number of sampling stations reviewed on the ecological classification

For the spatial reduction analysis, it has been developed a previous exploratory study, as it has been done on the case of the temporal reduction. On this preliminary study it has been examined, in an exhaustive manner, all the possible reductions in the number of reviewed sampling stations.

To analyze the consequences for the ecological classification due the spatial reduction, first step was to determinate all possible station combinations. Water bodies that are included into this analysis have four, five or six active sampling stations. To determinate number of possible combinations for each water body is used this mathematical equation:

$$C_n^k = \frac{n!}{k!(n-k)!}$$

where n is the number of active sampling stations and k is the number of selected sampling stations

By this equation is calculated number of possible combination for the each water body (depending on the number of active sampling stations) and presented into Table 6:

w.b. number of active s. stations	number of s. stations in the combination	number of combinations	all possible combinations
4	4	1	**15**
	3	4	
	2	6	
	1	4	
5	5	1	**31**
	4	5	
	3	10	
	2	10	
	1	5	
6	6	1	**63**
	5	6	
	4	15	
	3	20	
	2	15	
	1	6	

Table 6: Number of possible k-observers for a w. b. surveyed by four, five and six sampling stations

With SPSS® statistical package we have calculated the 90th percentile of measurements taken in the sampling stations included in each combination for the following periods:

•August 2005 – July 2008 (whole data set – three year period)
•August 2005 – July 2006 (first annual period)
•August 2006 – July 2007 (second annual period)
•August 2007 – July 2008 (third annual period)

In order to,

- Know the global effect of *reducing the number of sampling stations*, on the numerical value of 90 percentile and the ecological classification, in each water body

- Compare the global effect of *reducing the number of sampling stations*, on the numerical value of 90 percentile and the ecological classification for the four periods considered, in each water body.

- Know the effect of the *number of stations unrevised* on the numerical value of 90 percentile and the ecological classification system in each water body

- Compare the deviations in estimating the 90th percentile and know the correspondence in the ecological classification between the *reference observer* and each *limited observer*

we have obtained the following results,

1. Calculate percentage and number of observers that established unchanged ecological classification for each water body and for all analyzed periods separately (diagrams and tables)
2. Calculate percentage and number of observers that upgrade or degrade ecological classification (established by complete data) for each water body and for all analyzed periods separately (diagrams and tables)

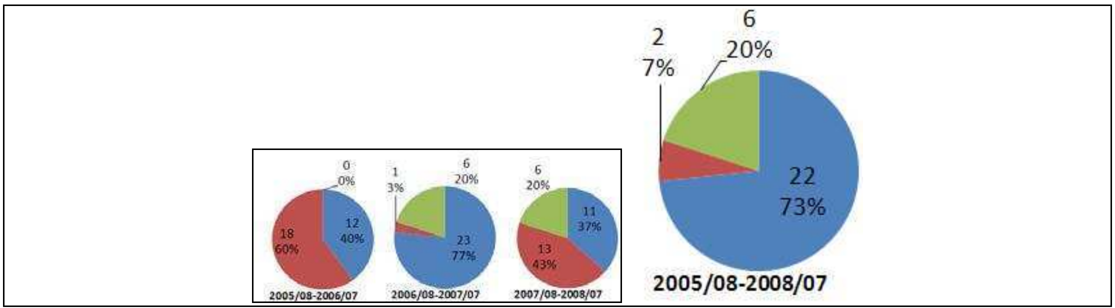


Figure 7: example of number and % (four periods) of k-observers that preserve, upgrade or degrade ecological classification established by complete data

3. Calculate percentage of the combinations, with same number of sampling stations, which establish unchanged ecological classifications (for each water body). This analysis is preformed for the all four periods. The results are presented as percentage for the observers that preserve ecological classifications for the all four periods and just for the three year period.

N. of excluded s. stations	N. of comb. realized for all 4 periods	N. of comb. that do not modify eco. class. (4 periods)	% of comb. that do not modify eco. class. (4 periods)	N. of comb. realized for 3 years period	N. of comb. that do not modify eco. class. (3 years period)	% of comb. that do not modify eco. class. (3 years period)
1	5	5	100,00%	5	5	100%
2	10	5	50,00%	10	10	100%
3	10	6	60,00%	10	9	90%
4	2(from5)	1	50,00%	5	4	80%

Table 7: example of table with number and percentage of k-observers (grouped by number of excluded stations) that do not modify ecological classification

4. For each analyzed period (and for all water bodies) is calculate the 90th percentile absolute average biases (produced by 90th percentiles calculated by k-observers). Average absolute biases are represented separating all periods and integrating the three annual periods.

average absolute bias 2005-2006	average absolute bias 2006-2007	average absolute bias 2007-2008	annual average absolute bias	average absolute bias 2005-2008
0,64	0,61	1,00	0,75	0,47

Table 8: example of table with interannual average absolute biases; annual average absolute bias, average absolute bias for the three year period

5. For each observer is calculate 90th percentile bias and if the obtained ecological classification is upgraded, degraded or unchanged. This analysis is preformed

for all four periods. Results are represented in the annex “Descriptive statistic results”.

6. For each observer is calculated 90^{th} percentile average absolute bias (comparing with the 90^{th} percentiles obtained by complete data) for the all four periods. This value is used to ascendingly order k-observers in the presented tables and detect combination with lowest error values.
7. For each observer is determinate if ecological classifications were unchanged (compared by the ecological classifications obtained by complete data) for the all four periods

w.b.	k-observer	- n stations	average absolute error (4 periods)	Corresponding eco. class. (4 periods)	Bias 90 PCTL (3 years)	eco. class. (3 years)
019	DP125 DP127 DP129	1	0,155	yes	0,000	corresp.
019	DP129 DP127	2	0,235	yes	-0,036	corresp.
019	DP125 DP131	2	0,273	no	0,318	corresp.
019	DP125 DP127 DP131	1	0,275	yes	0,610	corresp.
019	DP129 DP127 DP131	1	0,308	no	-0,108	corresp.
019	DP131 DP127	2	0,338	no	0,472	corresp.
019	DP127	3	0,368	yes	0,728	corresp.
019	DP125 DP129 DP131	1	0,455	no	-0,219	corresp.
019	DP125 DP129	2	0,498	no	-0,153	corresp.
019	DP125 DP127	2	0,550	yes	0,687	corresp.
019	DP125	3	0,590	no	0,610	corresp.
019	DP129 DP131	2	0,900	no	-0,720	degrade
019	DP131	3	1,213	no	-0,985	degrade
019	DP129	3	2,925	no	-0,512	degrade

Table 9: Example of table with 90th percentile average errors calculated by reduced sampling stations combination and ecological classification adjustment four periods and three year period.

3.5.2.Selecting the best observers

If the preliminary study indicates that it is possible to realize a reduction in the number of sampling stations revised in a water body, the next stage will be to find characteristics that identify to the good observers and to develop a methodology to locate the best observers for each system.

In the preliminary study about the effects of a reduction in the number of sampling stations reviewed on the ecological classification, for each observer is calculated 90^{th} percentile average absolute bias (comparing with the 90^{th} percentiles obtained by complete data) for the three-year period. The comparison between the 90th percentile of the limited observers, and the 90th percentile of the reference observer, allow us preselecting the observers with less bias and to discard the observers with more bias.

The order established between the pre-selected observers, based on the 90th percentile of the empirical cdf, may be easily modified with small changes in the data. The reason is that this order is based only on the value of the 90th percentile of the empirical cdf in a relatively small sample. That is why, we have tried to find a more stable ordination, based on the proximity between the monthly averages of chlorophyll a measurements collected at the sampling stations checked every month.

Since observers are characterized by sampling stations, from which they receive the information, the first step to know the properties of an observer will be to study the distributions of chlorophyll a measurements collected at each sampling station. The box-whisker plots help us to better understand these distributions, especially to know the extreme values that will impact directly on the 90th percentiles.

Box-and-whisker Plots is a graphical method commonly used to compare the frequency distribution of several samples. It is also very useful to detect extreme values. A box is drawn extending from the lower quartile of the sample to the upper quartile. This is the interval covered by the middle 50% of the data values when sorted from smallest to largest. A line is drawn at the median (the middle value). Whiskers are drawn from the edges of the box to the largest and smallest data values, unless there are values unusually far away from the box. Outside points, which are points more than 1.5 times the interquartile range (box width) above or below the box, are indicated by point symbols. If outside points are present, the whiskers are drawn to the largest and smallest data values which are not outside points.

Additionally, Quantile Plots, specially adapted, allow us to compare the 90th percentiles of the measurements collected at each station and knowing its proximity to the thresholds that separate the ecological status. A *Quantile Plot* represents the empirical CDF of a data set and Statgraphics allows overlapping graphics to compare the frequency distributions of two or more datasets.

If we want the information provided by a limited observer, based on a limited number of sampling stations, on the ecological status of a body of water is consistent with the information provided by the reference observer, with all the sampling stations that are currently active, it is desirable that the empirical cumulative distribution function *(empirical CDF)* of *chlorophyll a* concentration, in the full sample, does not differ substantially from the empirical CDF of the subsample, specially in the higher percentiles. Additionally, it would be also interesting that the subsample includes the greatest values of the original sample. Thus, the difference between the 90^{th} percentile of the data analyzed by a limited observer and the 90^{th} percentile of the data analyzed by the reference observer will be minimal.

It is not too difficult to identify the sampling stations in which are obtained the greatest measurements and include them in the subset of sampling stations to be reviewed. However, it is not so easy to identify a subset of sampling stations that, in addition to contain these values, meets the requirement that the dataset corresponding to measurements taken, on them, have an empirical CDF similar to the empirical CDF of the original sample.

The first difficulty, which arises, is how to synthesize the information about the status of a water body, in a given month, obtained by a particular *observer,* from the measurements of *chlorophyll a*, collected in the subset of sampling stations that characterize to this *observer*.

The arithmetic mean is a statistic which is heavily influenced by extreme values. This property is generally not desirable in an estimator; but can be useful in our problem. Indeed, a statistic, sensitive to outliers, will allow us to know if the sampling stations that characterize an observer, contain the extreme values of the original dataset.

Moreover if, month to month, the averages of measurements made in the sampling stations that characterize an observer, are similar to the averages of measurements collected at the sampling stations that characterize the reference observer, this will be an indicator of proximity between the empirical CDF. of measurements collected in the group of stations that characterize the observer and the empirical CDF of the original dataset.

The second issue to determine was choosing the proximity or distance measures to use to compare the concordance between the monthly averages of two different observers. There have been considered two measures: Interclass correlation coefficient (ICC) and Euclidean distance.

If there is any correlation between the monthly averages of measurements made by an observer limited and the monthly averages of measurements made by the observer of reference, this indicates that the limited observer perceives the ups and downs that occur in the levels of chlorophyll a in the system. On the other hand, if the Euclidean distance between the monthly averages is small, since the averages are sensitive to extreme values, this indicates, first, that the observer included in their measurements, outliers, and second, that the information collected is balanced not providing an alarmist view of the alterations to be found in the system.

The ICC, that enables us to obtain a measure of the degree of consistency or agreement between the observers, is based on the decomposition of the total variability of the measurements into two components: variability due to differences between observers and variability due to differences between measurements for each observer. The ICC is then defined as the proportion of the total variability due to the variability of the measurements.

The ICC is calculated under a model alpha 2-factor and mixed effects, and seeking absolute agreement in which to establish an agreement between two variables should vary not only in parallel but their means and variances must be equal. We calculated the lower and upper limits of the ICC with a confidence interval of 95%. Furthermore, using a qualitative scale usual in this type of study in which the agreement was considered very good (0.90), good (0.80 to 0.90), moderate (0.60 to 0.80) and poor (0-.60) [7]

The Euclidean distance between two observers is the square root of the sum of squared differences between the monthly averages of the limited observer and the reference observer. To facilitate the interpretation of the results obtained with the Euclidean distance, it has been used the technique called Multidimensional Scaling.

The MDS is a technique of multivariate analysis that allows us to represent the proximity between a set of objects as distances in a space of low dimensionality. The spatial representation of the neighbourhood is so that if two objects are valued as much alike they are within walking distances one from the other and vice versa. The axes are placed ensuring that each expresses the maximum variability in the distances contained in the input array. The representation facilitates the interpretation of the neighbourhood to show them visually instead of numerically. It is also possible to observe significant subsets of objects or interpreting the dimensions. It has been used the Proxscal algorithm, starting by placing the objects in the configuration all at the same distance of each other and taking one iteration to improve this high-dimensional configuration, followed by a dimension-reduction operation to obtain two dimensions, mmaximum number of iterations 100, stops iterating when the difference in consecutive stress values is less than 0.0001

3.5.3.Post reduction Management

As seen in the above section, the concordance of a limited observer and the reference observer can be evaluated using the ICC and the Euclidean distance between the monthly average of the measurements taken by a limited observer and the reference observer. A high value of the ICC and a low value of the Euclidean distance guarantee that the limited observer can provide a complete and unbiased view of the changes experienced by the chlorophyll *a* in this system. The final selection between the good reduced observer should be made taking into account other factors such as effectiveness, cost, etc…

One of the criteria that can help take the final decision is to evaluate the ability to obtain virtual values of the measurements of chlorophyll *a* in the unrevised stations, from measurements taken at the stations visited, using regression models,
We believe that, there are factors that can alter, significantly, the ecological stability of the areas where the unrevised sampling stations are located. To detect these possible changes in the levels of chlorophyll *a*, it is suggested to estimate the concentrations of chlorophyll *a* in these unrevised sampling stations from chlorophyll concentrations observed in stations reviewed using regression models. These virtual values of chlorophyll *a* concentration in the unchecked sampling stations should be contrasted with the values observed during occasional revisions of these stations. In the event that, in subsequent revisions, the values observed in any of these stations stray from the virtual values, more frequent monitoring of the sampling station should be considered, as well as a possible review of the optimal combination of stations that are to be reviewed regularly.

If, in the post-management of the spatial reduction, we take occasional measurements in the unrevised stations, and we found repeatedly, significant discrepancies between virtual and observed values, this will indicate us that there have been significant anthropogenic changes in the specific area where this station is located. This may have important implications for the spatial reduction, because maybe the “selected observer” is no longer the most suitable to provide a complete and stable view of the area and we should consider performing a new selection.

Therefore, the selection of the sampling stations to be reviewed should not be a permanent selection and should be able to be modified if circumstances indicate that it is convenient.

In order to know the ability of obtain virtual values of the measurements taken in the unchecked sampling stations and select the best observer between the good observers identified, we calculate the Pearson correlation coefficients between the sampling stations actually activated, and adjust a regression model to obtain virtual values for each sampling station. The selection of the best independent predictors is made by stepwise method, considering α=0’05 level of significance for the incorporation of a variable and α=0’10 level of significance for the elimination of a variable. The F statistic is used to include variables and the t-test is used as the criteria to remove them. The statistic used for testing the goodness of fit of the final estimated model is de adjusted R Square.

The value of the adjusted R Squared help us to know if it is possible to do a post-management of spatial reduction, and which observer is the most appropriate among the good observers identified.

Because there is a relatively high percentage of missing data at some stations, and since the objective of this section is simply to indicate how one would obtain virtual values for unrevised stations, the missing data has been completed using the method of k-nearest neighbors. With this method, the missing data has been replaced in some stations, with values estimated from other monthly observations with similar values in the rest of the sampling stations.

Euclidian distance has been used as a measure of distance between monthly observations. An important element to consider is the number of neighbors from which the estimate is made. If it is too small, the effect may be a high variance in the estimate. On the other hand, if the imputation is made from an excessively large number of neighbors, the effect may be the introduction of bias into the estimation of information from individuals too far away. We have obtained the estimated values for missing data as the average of the values measured in the three nearest neighbors. It has been discarded observations with missing values in more than two seasons.

3.6.Calculating the 90th percentile

The percentile value and graphing was done using the spreadsheet Excel and the statistical packages SPSS and STATGRAPHICS. It drew our attention, the fact that the values obtained did not coincide in some cases, being sometimes significant deviations. Moreover, even within a single package, sometimes, the numeric value of the percentiles did not coincide with the values used to draw the Box-and-whisker Plot, the Quantile Plot or the Quantile-Quantile Plot. Since this could affect to the conclusions of the study, we have sought the causes of these deviations and tried to find references to other researchers on this topic (Hydman at al.1996).

Let $(x_1,.......,x_n)$ represent the observed values in a sample of size n from a homogeneous population with distribution function F(x) and let $x_{(1)} \le x_{(1)} .... \le x_{(n)}$ the order statistics of the samples value placed in ascending order.

The empirical distribution function is usually considered as a step function of $\Re$ in [0,1], denoted by F_n ,which takes the values:

$$F_n(x_{(i)}) = \begin{cases} 0 & x < x_{(1)} \\ i/n & x_{(i)} \le x < x_{(i+1)} \\ 1 & x \ge x_{(n)} \end{cases}$$

As the sample size tends to infinity $F_n(x)$ will converge in probability to $F(x)$, for all x. In this sense, $F_n(x)$ can be considered a good estimator of $F(x)$. However, when the sample is small, $F_n(x)$ is the distribution function of a discrete distribution concentrated on n points and $F_n(x)$ is not a reasonable estimation of $F(x)$.

If $F(x)$ is estimated with the previous expression, the probability that the variable assumes a value greater than the highest sample value observed will be equal to zero. Therefore, we will be underestimating the value of $F(x)$. In order to reduce this error, it has been proposed several alternative estimators of F(x). Weibull (1939) and Gumbel (1939) proposed to define $F_n(x)$ as follows,

$$F_n(x_{(i)}) = \begin{cases} 1/(n+1) & x < x_{(1)} \\ i/(n+1) & x_{(i)} \le x < x_{(i+1)} \\ 1/(n+1) & x \ge x_{(n)} \end{cases}$$

Thus, the values observed divide the range of the variable into n+1 regions, each with probability 1/(n+1). In particular,

$$P(X < x_{(1)}) = 1/(n+1) \qquad P(X > x_{(n)}) = 1/(n+1)$$

Gumbel (1939) also proposed use the following estimator,

$$F_n\left(x_{(i)}\right)=\begin{cases}0 & x<x_{(1)}\\ (i-1)/(n-1) & x_{(i)}\leq x<x_{(i+1)}\\ 1 & x\geq x_{(n)}\end{cases}$$

Thus, the observed values divide the range of the variable into n-1 regions, each one with probability 1/(n-1).

If p is a given number in the range 0 <p <1, a quantile of order p is a point, ζ_p, such that

$$P\left(X<\zeta_p\right)\leq p \qquad and \qquad P\left(X\leq\zeta_p\right)\geq p$$

There are several methods for estimating the quantile, of order p, for a distribution, based on a rounding or a interpolation scheme. Most of the statistical packages or spreadsheets use the interpolation scheme in order to obtain the sample quantile, Q_p, and the expressions to obtain it may be written as follows,

$$Q_p=x_{(k)}+\gamma\left(x_{(k+1)}-x_{(k)}\right)$$

where k is a value that meets that $F_n\left(x_{(k)}\right)\leq p<F_n\left(x_{(k+1)}\right)$

Although the formula will change, depending on the estimator of the distribution function considered, always it will be true that,

$$\frac{F_n\left(x_{(k+1)}\right)-F_n\left(x_{(k)}\right)}{x_{(k+1)}-x_{(k)}}=\frac{p-F_n\left(x_{(k)}\right)}{Q_p-x_{(k)}}$$

and, therefore, we have,

$$Q_p=x_{(k)}+\frac{p-F_n\left(x_{(k)}\right)}{F_n\left(x_{(k)}\right)-F_n\left(x_{(k)}\right)}\left(x_{(k+1)}-x_{(k)}\right)$$

Thus,

$$\gamma=\frac{p-F_n\left(x_{(k)}\right)}{F_n\left(x_{(k+1)}\right)-F_n\left(x_{(k)}\right)}$$

Some statistical software packages, such as SPSS®, consider, by default, that $F_n\left(x_{(k)}\right)=k/(n+1)$ and therefore

$$\frac{k}{(n+1)}\leq p<\frac{k+1}{(n+1)}\Rightarrow k\leq(n+1)p\leq k+1\Rightarrow k=\lfloor(n+1)p\rfloor=\lfloor np+p\rfloor$$

where $\lfloor u\rfloor$ denotes the largest integer not greater than u.

If we obtain $h = (n+1)p$ and then split h into its integer component, k, and decimal component, d, such that $h = k + d$, we have

$$\gamma = \frac{p - k/(n+1)}{1/(n+1)} = (n+1)p - k = h - \lfloor h \rfloor = d$$

Some spreadsheets, such as Excel®, consider, by default, that $F_n(x_{(k)}) = (k-1)/(n-1)$ and therefore

$$\frac{k-1}{(n-1)} \leq p < \frac{k}{(n-1)} \Rightarrow k \leq (n-1)p + 1 \leq k + 1 \Rightarrow k = \lfloor (n-1)p + 1 \rfloor = \lfloor np + (1-p) \rfloor$$

If we obtain $h = (n-1)p + 1$ and then split h into its integer component, k, and decimal component, d, such that $h = k + d$, we have

$$\gamma = \frac{p - (k-1)/(n-1)}{1/(n-1)} = (n-1)p - (k-1) = (n-1)p + 1 - \lfloor (n-1)p + 1 \rfloor = d$$

In our case, quality of coastal waters, is established based on the 90th percentile and obviously the method used by Excel will provide us with a value lower than, or equal to, the value obtained by the first method, since (1-p) will be lower than p and the interpolated value corresponding to $x_{\lfloor np+(1-p) \rfloor}$ will be lower than or equal to the interpolated value corresponding to $x_{\lfloor np+p \rfloor}$. Consequently, the classification of the water body will be better if we obtain the 90th percentile using Excel. Therefore, we suggest that the Mediterranean Geographical Intercalibration Group (MED-GIG) treat this issue and provide guidance on this point. This will avoid confusion and will ensure comparable results.

4.Temporal reduction analysis

4.1.Temporal reduction data analysis - exploratory study

The reduction analysis is performed using three-year historical data from the Valencian coastal monitoring network. This analysis consists of bimonthly and trimestral campaign simulations (data reduction, covering all possibilities) and a percentile value calculation to establish ecological classification status, comparing them with the results calculated using the complete data set of qualitative and quantitative values and determining the consequences of reduction.

Bimonthly campaigns:

-Even months – August, October, December, February, April, June

-Uneven months – September, November, January, March, May, July

Trimestral campaigns:

-1st set – August, November, February, May,

-2nd set – September, December, March, June

-3rd set – October, January, April, July

4.1.1.Three-year period data analysis

To simulate a reduced bimonthly frequency survey, campaigns are divided into even/uneven months.

The first data set of three year historical data erases all campaigns that are provided during even months, starting in August 2005 and ending in June 2008. Starting from number 2343, the data set is reduced to 1244 water samples. The second data set excludes the 18 uneven-month campaigns and the data set is reduced to 1099 data samples.

The establishment of trimestral campaigns causes the original data set to suffer more seriously from reduction of data. The first trimestral data set, which begins in August 2005 and ends in May 2008, amounts to 840 water samples. The second data set begins in September 2005 and the last campaign was completed in June 2008, comprising 745 data samples. Finally, the third set of trimestral campaigns takes 758 data samples into account, the first campaign was performed in October 2005 and the last in July 2008.

In Table 1, it is possible to observe a number of water samples that are used to calculate 90th percentile and to establish an ecological classification status for each water body.

WATER BODY	No. of samples monthly campaigns	No. of samples even months	No. of samples uneven months	No. of samples 1st Set (trimestral campaigns)	No. of samples 2nd Set (trimestral campaigns)	No. of samples 3rd Set (trimestral campaigns)
001	**156**	85	71	55	49	52
002	**147**	76	71	51	51	45
003	**102**	55	47	37	31	34
004	**107**	59	48	38	35	34
005	**168**	88	80	60	53	55
006	**85**	46	39	31	26	28
007	**118**	61	57	40	39	39
008	**149**	78	71	52	48	49
009	**166**	88	78	59	54	53
010	**187**	96	91	69	57	61
011	**166**	88	78	61	52	53
012	**112**	61	51	40	35	37
013	**104**	56	48	37	33	34
014	**128**	68	60	46	39	43
015	**34**	18	16	12	11	11
016	**98**	53	45	38	31	29
017	**118**	63	55	42	38	38
018	**65**	35	30	25	21	19
019	**133**	70	63	47	42	44
Σ	**2343**	**1244**	**1099**	**840**	**745**	**758**

Table 1: Number of samples for each water body; monthly, bimonthly and trimestral campaigns (August 2005 – July 2008)

The next step was to calculate the 90th percentile to establish the ecological classification for the new reduced datasets. The same methodology has been used (limits for High/Good and Good/Moderate boundaries) as in the previous chapter, where a provisional ecological classification status for each water body with an original data set has been established.

The consequences of the bimonthly or trimestral reductions are comparable using the quantitative and qualitative (ecological classification) results represented in Table 2. Table 3 represents the comparison of ecological classification to the original data set and error percentage of 90th percentile for each body of water. Finally, mean absolute percentage error is provided at the foot of the table (MAPE - mean is used for summarization across the absolute series values). The MAPE values represent changes of the 90th percentiles - quantitative values (established by reduced datasets) for the whole reduced campaign.

The ecological classification status is calculated for each body of water using only the even month campaigns. Of 19 statuses, 14 of them are equal to the originally established ones, which provides a 73,7% correspondence. The five erroneous classifications present 3 (15,8%) degradations (006, 018 and 019) and 2 (10,5%) upgrades (008 and 009) of ecological classification calculated by reduced dataset. Comparing the quantitative values, the majority of 90th percentiles are elevated compared to the original ones. The MAPE for the bimonthly campaigns is 14,5% and percentage error varies from -43% to 16,8% per water body.

78.9% of uneven month campaigns are unchanged, with 15 of 19 calculated ecological classes remaining unchanged. All 4 (21,0%) erroneous classifications upgraded (007, 011, 012 and 013) the previously calculated classifications. In contrast, mostly all of the calculated 90th percentile values for the uneven month campaigns are inferior to the percentiles calculated with the whole data set. The MAPE indicates the highest error rate (18,2%) of quantitative values for all reduced datasets.

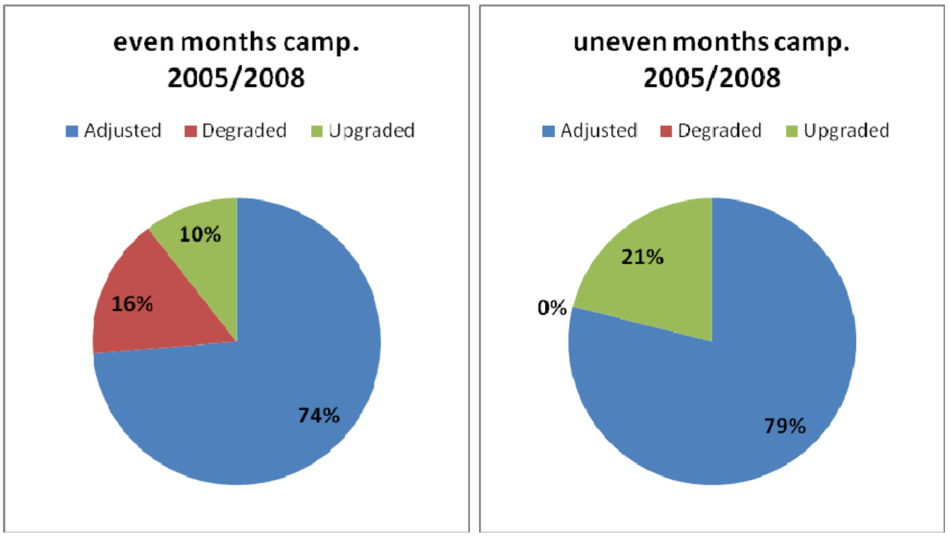


Figure 1: Quality results of bimonthly campaigns (August 2005 – July 2008)

By establishing ecological statuses with trimestral campaigns, the results adapt similarly or slightly better than those calculated using a bimonthly reduction set. The first and third trimestral data sets (starting in August and October, respectively) have 15 unchanged statuses (78,95%) and the second one (starting in September) 17 (89,47%) has just 2 incorrect statuses (007 has an upgraded and 004 a degraded ecological classification). The number of upgraded and degraded classifications is identical for the first and third reduction data sets; 2 (10,5%) and 2 (10,5%). For the first trimestral campaigns, the ecological status of water bodies 006 and 012 upgraded while that of 011 and 015 degraded. The MAPE values for the trimestral data sets are 15,4%, 13,6% and 17,5% respectively (Table 3), with the lowest value for the second trimestral data set which has the highest amount of unchanged ecological statuses. The error percentage for the quantitative values varies from -98,8% to 50,4% for the trimestral data campaigns.

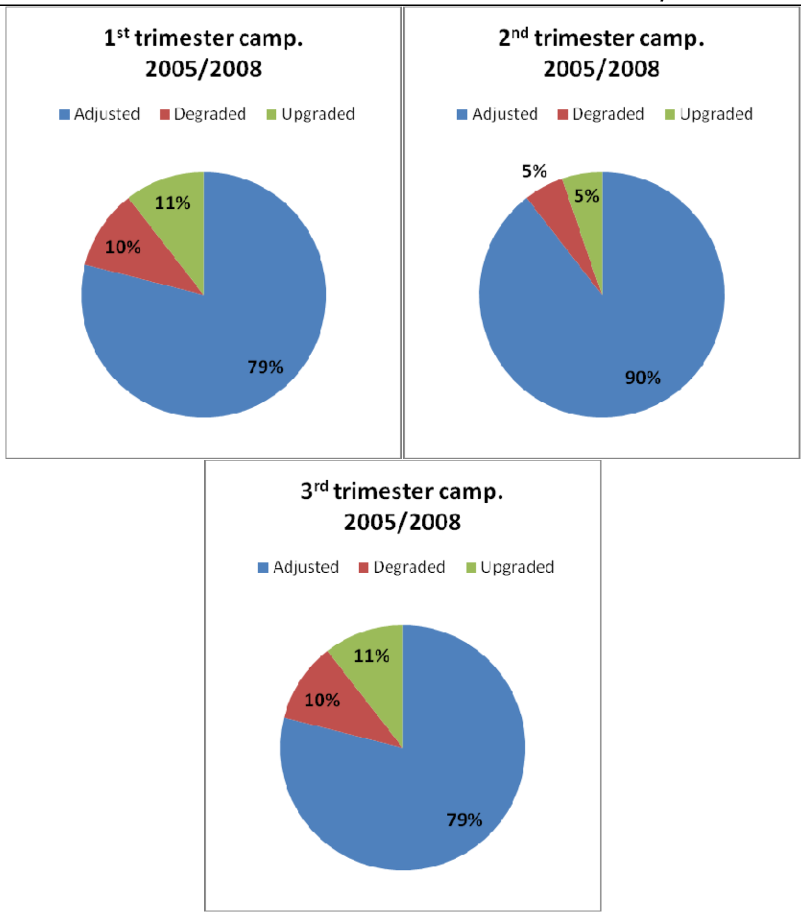


Figure 2: Quality results of trimestral campaigns (August 2005 – July 2008)

WATER BODY	TYPE OF WATER BODY	90th PTCL Chl a monthly campaigns	Classification monthly campaigns	90th PTCL Chl a even months	Classification even months	90th PTCL Chl a uneven months	Classification uneven months	90th PTCL Chl a 1st Set (trimestral campaigns)	Classification 1st Set (trimestral campaigns)	90th PTCL Chl a 2nd Set (trimestral campaigns)	Classification 2nd Set (trimestral campaigns)	90th PTCL Chl a 3rd Set (trimestral campaigns)	Classification 3rd Set (trimestral campaigns)
001	II A	3,54	HIGH	3,89	HIGH	2,79	HIGH	4,22	HIGH	3,40	HIGH	3,49	HIGH
002	II A	3,50	HIGH	3,65	HIGH	3,03	HIGH	3,49	HIGH	3,81	HIGH	3,10	HIGH
003	II A	2,18	HIGH	2,54	HIGH	1,80	HIGH	2,03	HIGH	2,44	HIGH	2,29	HIGH
004	II A	3,72	HIGH	4,56	HIGH	3,08	HIGH	3,71	HIGH	5,00	GOOD	3,42	HIGH
005	II A	5,69	GOOD	5,92	GOOD	5,02	GOOD	5,02	GOOD	5,17	GOOD	6,38	GOOD
006	II A	4,73	HIGH	5,07	GOOD	4,11	HIGH	5,60	GOOD	4,28	HIGH	4,83	GOOD
007	II A	5,13	GOOD	6,06	GOOD	4,59	HIGH	5,10	GOOD	4,50	HIGH	6,04	GOOD
008	II A	7,96	MODERATE	7,10	GOOD	12,16	MODERATE	7,80	MODERATE	10,06	MODERATE	8,28	MODERATE
009	II A	8,22	MODERATE	6,83	GOOD	8,98	MODERATE	9,55	MODERATE	8,11	MODERATE	7,08	GOOD
010	II A	5,91	GOOD	6,04	GOOD	5,15	GOOD	5,28	GOOD	6,05	GOOD	5,98	GOOD
011	III	2,26	GOOD	2,47	GOOD	1,72	HIGH	1,63	HIGH	2,49	GOOD	2,38	GOOD
012	III	2,27	GOOD	3,25	GOOD	1,86	HIGH	3,63	MODERATE	2,24	GOOD	2,66	GOOD
013	III	3,61	MODERATE	4,37	MODERATE	3,14	GOOD	4,20	MODERATE	4,77	MODERATE	1,79	HIGH
014	III	1,79	HIGH	1,94	HIGH	1,43	HIGH	1,41	HIGH	1,81	HIGH	1,85	HIGH
015	III	4,76	MODERATE	5,96	MODERATE	4,70	MODERATE	3,51	GOOD	5,27	MODERATE	9,46	MODERATE
016	III	6,26	MODERATE	5,99	MODERATE	7,35	MODERATE	5,31	MODERATE	7,52	MODERATE	6,42	MODERATE
017	III	5,46	MODERATE	5,88	MODERATE	4,82	MODERATE	3,68	MODERATE	6,48	MODERATE	7,23	MODERATE
018	III	3,32	GOOD	4,12	MODERATE	2,21	GOOD	3,55	GOOD	2,55	GOOD	4,43	MODERATE
019	III	3,10	GOOD	3,75	MODERATE	2,26	GOOD	3,15	GOOD	2,42	GOOD	3,51	GOOD

Table 2: Ecological classifications calculated by monthly, bimonthly and trimestral campaigns for the period August 2005 – July 2008

w.b.	Even months		Uneven months		1st set (trimestral camp.)		2nd set (trimestral camp.)		3rd set (trimestral camp.)		w.b.MAPE
	unchanged class	error %	unchanged class	error %	unchanged class	error %	unchanged class	error %	unchanged class	error %	
001	0	-9,88%	0	21,34%	0	-19,14%	0	4,01%	0	1,36%	11,15%
002	0	-4,20%	0	13,38%	0	0,23%	0	-8,81%	0	11,38%	7,60%
003	0	-16,28%	0	17,56%	0	6,83%	0	-11,97%	0	-5,00%	11,53%
004	0	-22,65%	0	17,16%	0	0,11%	-1	-34,48%	0	8,15%	16,51%
005	0	-3,90%	0	11,75%	0	11,84%	0	9,26%	0	-12,03%	9,76%
006	-1	-7,30%	0	13,07%	-1	-18,53%	0	9,45%	-1	-2,24%	10,12%
007	0	-18,21%	1	10,53%	0	0,64%	1	12,28%	0	-17,74%	11,88%
008	1	10,83%	0	-52,71%	0	2,06%	0	-26,33%	0	-4,02%	19,19%
009	1	16,93%	0	-9,24%	0	-16,12%	0	1,45%	1	13,89%	11,52%
010	0	-2,18%	0	12,99%	0	10,72%	0	-2,27%	0	-1,08%	5,85%
011	0	-9,32%	1	24,17%	1	27,88%	0	-9,81%	0	-5,08%	15,25%
012	0	-43,02%	1	17,92%	-1	-59,62%	0	1,19%	0	-16,95%	27,74%
013	0	-20,91%	1	13,02%	0	-16,45%	0	-32,13%	2	50,42%	26,59%
014	0	-8,38%	0	20,44%	0	21,38%	0	-1,06%	0	-3,18%	10,89%
015	0	-25,13%	0	1,34%	1	26,20%	0	-10,76%	0	-98,82%	32,45%
016	0	4,25%	0	-17,39%	0	15,10%	0	-20,20%	0	-2,59%	11,90%
017	0	-7,82%	0	11,60%	0	32,55%	0	-18,78%	0	-32,53%	20,66%
018	-1	-23,96%	0	33,41%	0	-6,86%	0	23,30%	-1	-33,35%	24,18%
019	-1	-21,06%	0	27,23%	0	-1,48%	0	21,90%	0	-13,23%	16,98%
Campaign MAPE		14,54%		18,22%		15,46%		13,65%		17,53%	

Table 3: Changes in classification, 90th percentile error (%) and MAPE for each w.b.; MAPE for each reduced dataset; period August 2005-July 2008

4.1.2.One-year periods, reduced campaigns

August 2005-July 2006

The same analysis is performed on a one year period, the difference being that the one-year period comprises less data and the reduction is more critical.

WATER BODY	No.of samples monthly campaigns	No.of samples even months	No.of samples uneven months	No.of samples 1st Set (trimestral campaigns)	No.of samples 2nd Set (trimestral campaigns)	No.of samples 3rd Set (trimestral campaigns)
001	**49**	27	22	18	13	18
002	**35**	18	17	12	11	12
003	**30**	17	13	11	**8**	11
004	**32**	17	15	13	**8**	11
005	**55**	30	25	20	15	20
006	**25**	15	10	11	**6**	8
007	**10**	**5**	**5**	**3**	**3**	**4**
008	**44**	24	20	16	12	16
009	**53**	29	24	20	14	19
010	**57**	29	28	22	15	20
011	**50**	27	23	18	15	17
012	**34**	20	14	13	**9**	12
013	**30**	18	12	12	**7**	11
014	**41**	23	18	15	10	16
015	**11**	**6**	**5**	**4**	**3**	**4**
016	**22**	13	**9**	10	**4**	**8**
017	**34**	19	15	13	**9**	12
018	**13**	**8**	**5**	**6**	**3**	**4**
019	**44**	24	20	16	12	16
Σ	**669**	**369**	**300**	**253**	**177**	**239**

Table 4: Number of samples for each water body; monthly, bimonthly and trimestral campaigns (August 2005 – July 2006)

It is visible from the above table that the original data set numbers 669 water samples. In emulating the bimonthly campaigns, that number is reduced to 369 (even months) or 300 (uneven months), and to emulate trimestral frequency the number is cut to 253, 177 and 239 water samples.
Water bodies 007, 015 and 018 are identified as bodies with insufficient data (10, 11 and water samples) and so are excluded from the analysis.
With a reduction of data to simulate lower frequency surveys, bodies of water with less than 10 water samples (in the reduced dataset) are also excluded from the analysis. This excludes water bodies 003, 004, 006, 012, 013, 016, and 017 from the analysis of the second trimester (starting in September) reduction, 006 and 016 from the third trimestral reduction and 016 from the uneven month campaigns.
The quality and quantity results of the one year reduction analysis are represented in Table 5. Table 6 represents changes to the quality results set and 90th percentile error percentage for each water bodies.
For the even month campaigns, ecological classification statuses are established for all 16 possible water bodies. The number of established classes which remain unchanged in comparison to the calculation using the complete data set is 12 (75.0%)

and the remaining 4 (25,0%) ecological statuses are degraded (esablished for 004, 005, 009 and 010). The MAPE is 19,2% which indicates the lowest error percentage of quantitative values for all reduced datasets (Table 4).
Uneven month reduction establishes 15 ecological classes and 10 of them remain unchanged with a percentage of 66,67% . The 5 (33,3%) changed established classes represent 2 upgrades (009 and 010) and 3 degradations (011, 013 and 019) of ecological classification. The MAPE value is considerably higher (44,0%) in comparison to the even month values and discrepancies with the results vary from between 46,4% to -226,0%.

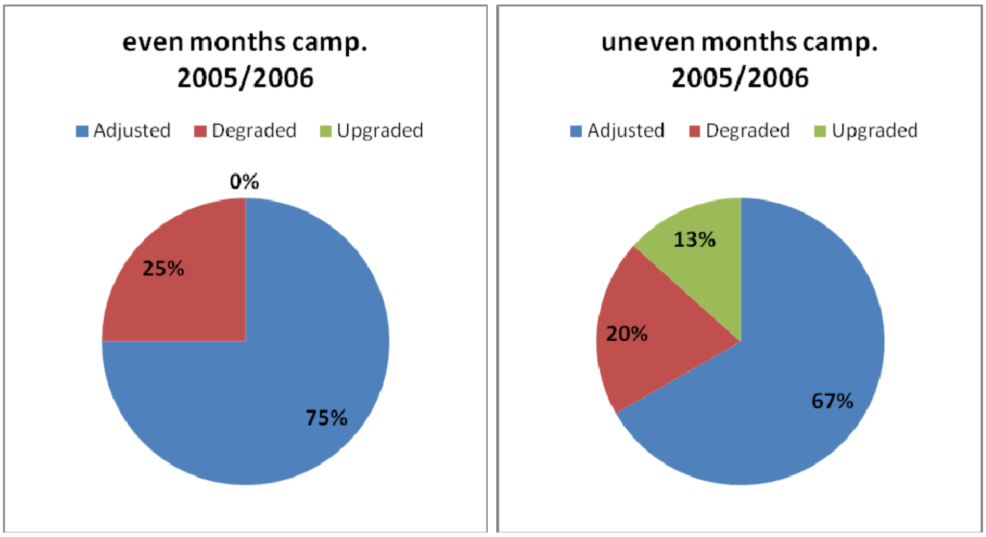


Figure 3: Quality results of bimonthly campaigns (August 2005 – July 2006)

Of the trimestral campaigns, only the first set (starting in August) establishes ecological status for 16 water bodies, with 50% of them changed. Six changed statuses are degraded classifications (001. 005, 011, 012, 013 and 017) and two of them are upgraded (009 and 019) classifications. The MAPE campaign value is 56,0%, which indicates a large discrepancy with the quantitative results, whose error varies between 37,5% and -229,38%.
The second trimestral set, which starts in September, establishes just 9 of 16 classes, and 5 classes (55,5%) remain unchanged. Four (44,4%) changed classes are degraded (005, 009, 010 and 019) in ecological status. Despite the low ecological classification, the MAPE has a 50,0% value with variation ranging from 4,3% to -234,5%.
The trimestral data set which starts with campaigns in October consists of 14 classified water bodies and 11 classifications (78,5%) remain unchanged. In the last reduced dataset, 2 water bodies are degraded (004 and 009) and one is upgraded (012) in ecological classification. The MAPE value is 27,7%, which is a considerable level of quantitative error.

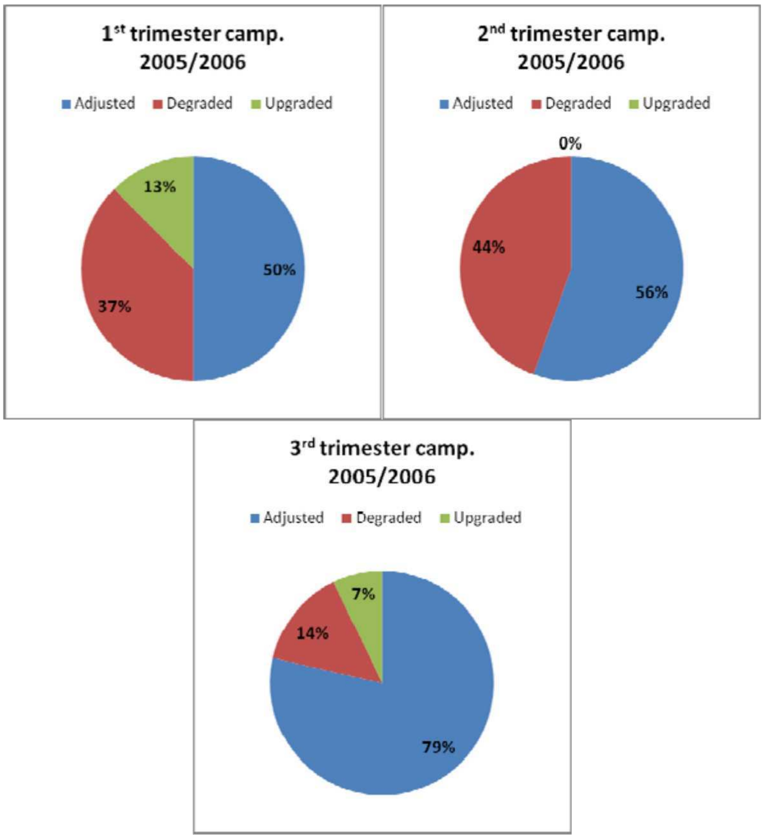


Figure 4: Quality results of trimestral campaigns (August 2005 – July 2006)

WATER BODY	TYPE OF WATER BODY	90th PTCL Chl a monthly campaigns	Classification monthly campaigns	90th PTCL Chl a even months	Classification even months	90th PTCL Chl a uneven months	Classification uneven months	90th PTCL Chl a 1st Set (trimestral campaigns)	Classification 1st Set (trimestral campaigns)	90th PTCL Chl a 2nd Set (trimestral campaigns)	Classification 2nd Set (trimestral campaigns)	90th PTCL Chl a 3rd Set (trimestral campaigns)	Classification 3rd Set (trimestral campaigns)
001	II A	3,36	HIGH	4,42	HIGH	2,19	HIGH	4,91	GOOD	3,83	HIGH	2,96	HIGH
002	II A	3,33	HIGH	4,29	HIGH	2,37	HIGH	4,57	HIGH	3,93	HIGH	2,98	HIGH
003	II A	2,19	HIGH	2,44	HIGH	1,40	HIGH	2,87	HIGH	-	-	1,53	HIGH
004	II A	6,98	GOOD	7,34	MODERATE	7,13	GOOD	6,51	GOOD	-	-	9,41	MODERATE
005	II A	4,80	HIGH	5,60	GOOD	3,54	HIGH	5,01	GOOD	7,28	MODERATE	4,22	HIGH
006	II A	5,37	GOOD	5,75	GOOD	4,94	GOOD	5,79	GOOD	-	-	-	-
007	II A	-	-	-	-	-	-	-	-	-	-	-	-
008	II A	9,19	MODERATE	9,19	MODERATE	21,12	MODERATE	22,17	MODERATE	8,79	MODERATE	15,05	MODERATE
009	II A	5,05	GOOD	8,98	MODERATE	3,88	HIGH	3,31	HIGH	8,62	MODERATE	8,98	MODERATE
010	II A	6,02	GOOD	8,96	MODERATE	3,22	HIGH	6,10	GOOD	7,30	MODERATE	5,80	GOOD
011	III	1,56	HIGH	1,41	HIGH	2,48	GOOD	2,24	GOOD	2,09	HIGH	1,12	HIGH
012	III	2,97	GOOD	3,46	GOOD	3,42	GOOD	9,77	MODERATE	-	-	1,30	HIGH
013	III	1,66	HIGH	1,30	HIGH	5,40	MODERATE	5,40	MODERATE	-	-	0,84	HIGH
014	III	1,63	HIGH	1,59	HIGH	1,70	HIGH	1,97	HIGH	1,61	HIGH	1,73	HIGH
015	III	-	-	-	-	-	-	-	-	-	-	-	-
016	III	5,98	MODERATE	5,71	MODERATE	-	-	7,62	MODERATE	-	-	-	-
017	III	3,11	GOOD	3,31	GOOD	3,26	GOOD	4,16	MODERATE			3,18	GOOD
018	III	-	-	-	-	-	-	-	-	-	-	-	-
019	III	3,39	GOOD	2,74	GOOD	3,87	MODERATE	2,12	HIGH	11,32	MODERATE	3,42	GOOD

Table 5: Ecological classifications calculated by monthly, bimonthly and trimestral campaigns for the period August 2005 – July 2006

	Even months		Uneven months		1st set (trimestral camp.)		2nd set (trimestral camp.)		3rd set (trimestral camp.)		w.b. MAPE
w.b.	**unchanged class**	**error %**	**unchanged class**	**error %**	**unchanged class**	**error %**	**unchanged class**	**error %**	**unchanged class**	**error %**	
001	0	-31,43%	0	34,82%	-1	-46,10%	0	-14,11%	0	11,93%	27,68%
002	0	-28,80%	0	28,95%	0	-37,27%	0	-18,02%	0	10,60%	24,73%
003	0	-11,66%	0	36,08%	0	-31,32%			0	29,86%	27,23%
004	-1	-5,20%	0	-2,16%	0	6,72%			-1	-34,78%	12,22%
005	-1	-16,69%	0	26,30%	-1	-4,38%	-2	-51,77%	0	12,09%	22,25%
006	0	-7,12%	0	8,07%	0	-7,86%					7,68%
007											
008	0	0,00%	0	-129,84%	0	-141,19%	0	4,32%	0	-63,71%	67,81%
009	-1	-77,96%	1	23,11%	1	34,46%	-1	-70,83%	-1	-77,96%	56,86%
010	-1	-48,79%	1	46,61%	0	-1,35%	-1	-21,29%	0	3,69%	24,34%
011	0	9,86%	-1	-58,64%	-1	-43,47%	0	-33,93%	0	28,17%	34,81%
012	0	-16,53%	0	-15,35%	-1	-229,38%			1	56,12%	79,34%
013	0	21,26%	-2	-225,97%	-2	-225,97%			0	49,15%	130,59%
014	0	2,09%	0	-4,55%	0	-20,76%	0	1,04%	0	-6,20%	6,93%
015											
016	0	4,58%			0	-27,33%					15,96%
017	0	-6,60%	0	-5,06%	-1	-33,98%			0	-2,54%	12,05%
018											
019	0	19,20%	-1	-14,18%	1	37,46%	-1	-234,48%	0	-1,15%	61,29%
Campaign MAPE		19,24%		43,98%		58,06%		49,98%		27,71%	

Table 6: Changes in classification, 90th percentile error (%) and MAPE for each w.b; MAPE for each reduced dataset; period August 2005-July 2006

August 2006-July 2007

During the second year of monthly campaigns, 815 water samples were taken that are considered in the reduction analysis. Bimonthly sets are reduced to a nearly equal quantity of data (408 and 407) and trimestral campaigns consist of 274, 262 and 279 samples respectively.

WATER BODY	No. of samples monthly campaigns	No. of samples even months	No. of samples uneven months	No. of samples 1st Set (trimestral campaigns)	No. of samples 2nd Set (trimestral campaigns)	No. of samples 3rd Set (trimestral campaigns)
001	**56**	30	26	18	19	19
002	**59**	30	29	20	20	19
003	**34**	18	16	12	10	12
004	**34**	18	16	12	12	10
005	**58**	28	30	20	18	20
006	**21**	10	11	**6**	**7**	**8**
007	**43**	20	23	13	13	17
008	**50**	24	26	16	16	18
009	**59**	30	29	19	20	20
010	**66**	32	34	23	19	24
011	**60**	32	28	22	18	20
012	**36**	18	18	12	12	12
013	**36**	18	18	12	12	12
014	**46**	22	24	16	15	15
015	**12**	**6**	**6**	**4**	**4**	**4**
016	**35**	18	17	12	12	11
017	**40**	20	20	13	13	14
018	**23**	11	12	**8**	**7**	**8**
019	**47**	23	24	16	15	16
Σ	**815**	**408**	**407**	**274**	**262**	**279**

Table 7: Number of samples for each water body; monthly, bimonthly and trimestral campaigns(August 2006 – July 2007)

Observing the amount of water sample data, it is clear that water body 007 is better monitored in the second year of the survey with 43 samples that make it possible to analyze the consequences of reduction for all reduced datasets. The survey is also more intensive for water body 018, but the number of samples does not allow a trimestral reduction analysis. Water body 015 is still poorly surveyed with only 12 water samples. The number of samples from water body 006 did not increase during the second year, because of which it is possible to simulate only bimonthly campaigns. Simulating bimonthly survey frequency, uneven and even month campaign simulations are performed for each water body, except for water body 015 (18 classifications out of 19). The trimestral campaign simulation results in 16 ecological classification statuses per analysis excluding water body 006, 015 and 018.
Table 8 represents the water body classification with 90th percentiles for the monthly, bimonthly and trimestral campaign frequencies. Changes in ecological classification and the percentage of error is represented in Table 9.

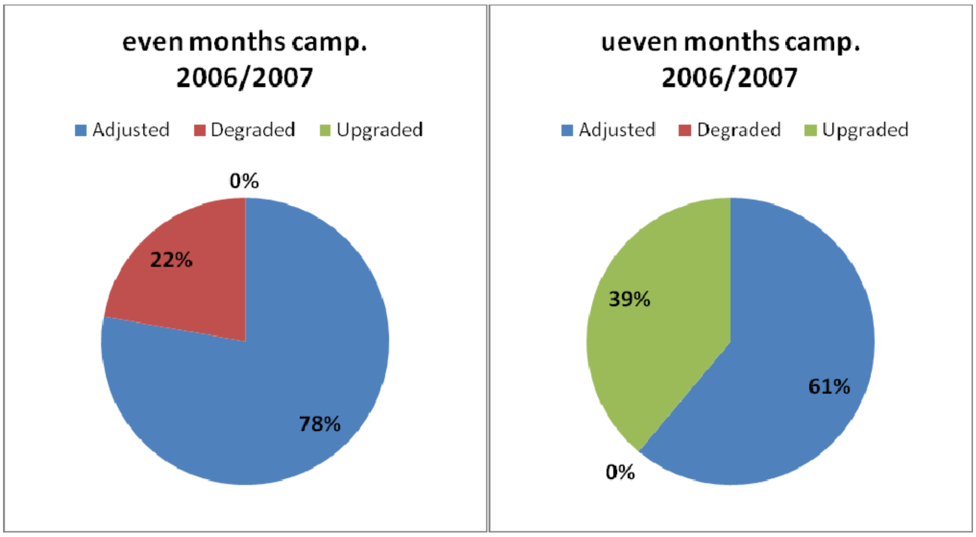


Figure 5: Quality results of bimonthly campaigns (August 2006 – July 2007)

In this annual period, the results of temporal bimonthly reduction are better 14 (77,7%) for even months, with 4 (22,2%) statuses differing from the original ones, all of which are degraded (006, 010, 014 and 019) ecological classes. The campaign MAPE value is a low 18,7% and the water body error varies from 5,39% to -62,66%.

Uneven month reductions count 11 (61,1%) unchanged and 7 (38,89%) upgraded (007, 009, 011, 012, 013, 018 and 019) ecological statuses. The campaign MAPE value is considerably higher 32,0% as a percentage of changed ecological statuses in comparison with the results of the even campaign.

The results for the 1st set of the trimestral campaigns shows that 8 (50,0%) are unchanged, and of the other half of the classifications 2 (12,5%) are degraded (010 and 019) and 6 (37,5%) are upgraded (005, 007, 011, 012, 016 and 017) compared to their original ecological status. The MAPE, considering the qualitative results, is unexpectedly low (22,8%) and the 90th percentile error varies between 54,5% and -31,7%.

The second data set has slightly better results with 9 corresponding classes (56,2%), with 4 (25,0%) degraded (008, 009, 010 and 014) and 3 upgraded (005, 007, 019) ecological classes. The MAPE value 32,8% is the highest of all five reduced datasets.

Third trimestral set analysis consists of more than 13 corresponding classes (81,2%) and 1 (6,2%) degraded (014) and 2 (12,4%) upgraded (009 and 013) statuses. The campaign MAPE value is unexpectedly high 29,8%, which is not comparable with the qualitative results of the reduced dataset.

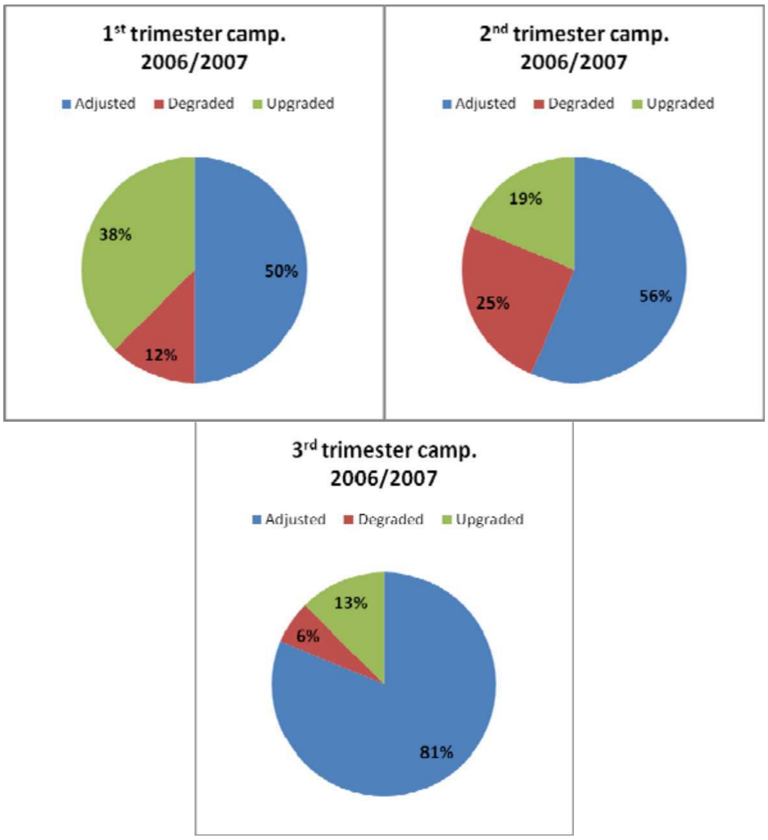


Figure 6: Quality results of trimestral campaigns (August 2006 – July 2007)

WATER BODY	TYPE OF WATER BODY	90th PTCL Chl a monthly campaigns	Classification monthly campaigns	90th PTCL Chl a even months	Classification even months	90th PTCL Chl a uneven months	Classification uneven months	90th PTCL Chl a 1st Set (trimestral campaigns)	Classification 1st Set (trimestral campaigns)	90th PTCL Chl a 2nd Set (trimestral campaigns)	Classification 2nd Set (trimestral campaigns)	90th PTCL Chl a 3rd Set (trimestral campaigns)	Classification 3rd Set (trimestral campaigns)
001	II A	3,69	HIGH	3,49	HIGH	4,66	HIGH	3,94	HIGH	3,64	HIGH	4,34	HIGH
002	II A	3,38	HIGH	3,38	HIGH	3,59	HIGH	2,69	HIGH	3,83	HIGH	3,08	HIGH
003	II A	2,49	HIGH	2,84	HIGH	1,77	HIGH	1,13	HIGH	3,17	HIGH	2,72	HIGH
004	II A	3,33	HIGH	3,35	HIGH	3,16	HIGH	3,48	HIGH	2,22	HIGH	3,48	HIGH
005	II A	5,74	GOOD	6,02	GOOD	5,58	GOOD	4,32	HIGH	4,07	HIGH	6,97	GOOD
006	II A	4,41	HIGH	6,89	GOOD	3,33	HIGH	-	-	-	-	-	-
007	II A	5,11	GOOD	5,43	GOOD	3,39	HIGH	4,61	HIGH	4,70	HIGH	5,84	GOOD
008	II A	5,35	GOOD	5,57	GOOD	5,01	GOOD	5,87	GOOD	7,31	MODERATE	4,81	GOOD
009	II A	5,51	GOOD	5,54	GOOD	4,33	HIGH	4,93	GOOD	7,64	MODERATE	4,50	HIGH
010	II A	4,51	HIGH	5,27	GOOD	3,23	HIGH	4,94	GOOD	5,51	GOOD	4,25	HIGH
011	III	2,69	GOOD	2,92	GOOD	1,18	HIGH	1,29	HIGH	3,12	GOOD	2,92	GOOD
012	III	3,67	MODERATE	4,66	MODERATE	1,67	HIGH	1,85	HIGH	5,11	MODERATE	6,12	MODERATE
013	III	4,40	MODERATE	7,15	MODERATE	1,52	HIGH	4,80	MODERATE	10,06	MODERATE	1,45	HIGH
014	III	1,82	HIGH	2,95	GOOD	1,25	HIGH	1,26	HIGH	2,56	GOOD	2,50	GOOD
015	III	-	-	-	-	-	-	-	-	-	-	-	-
016	III	4,59	MODERATE	4,07	MODERATE	6,50	MODERATE	3,21	GOOD	6,40	MODERATE	6,10	MODERATE
017	III	3,80	MODERATE	3,80	MODERATE	6,68	MODERATE	3,40	GOOD	3,67	MODERATE	9,40	MODERATE
018	III	3,98	MODERATE	4,59	MODERATE	2,67	GOOD	-	-	-	-	-	-
019	III	3,25	GOOD	4,59	MODERATE	2,11	HIGH	4,29	MODERATE	1,70	HIGH	3,48	GOOD

Table 8: Ecological classifications calculated by monthly, bimonthly and trimestral campaigns for the period August 2006 – July 2007

w.b.	Even months		Uneven months		1st set (trimestral camp.)		2nd set (trimestral camp.)		3rd set (trimestral camp.)		w.b. MAPE
	unchanged class	error %	unchanged class	error %	unchanged class	error %	unchanged class	error %	unchanged class	error %	
001	0	5,39%	0	-26,19%	0	-6,83%	0	1,33%	0	-17,65%	11,48%
002	0	0,15%	0	-6,21%	0	20,36%	0	-13,40%	0	8,88%	9,80%
003	0	-14,06%	0	28,76%	0	54,46%	0	-27,31%	0	-9,28%	26,77%
004	0	-0,66%	0	4,84%	0	-4,54%	0	33,26%	0	-4,75%	9,61%
005	0	-4,84%	0	2,82%	1	24,68%	1	29,12%	0	-21,38%	16,57%
006	-1	-56,35%	0	24,36%							40,36%
007	0	-6,42%	1	33,57%	1	9,64%	1	7,87%	0	-14,38%	14,38%
008	0	-4,17%	0	6,23%	0	-9,74%	-1	-36,69%	0	10,10%	13,39%
009	0	-0,49%	1	21,42%	0	10,53%	-1	-38,69%	1	18,26%	17,88%
010	-1	-16,86%	0	28,43%	-1	-9,46%	-1	-22,09%	0	5,83%	16,53%
011	0	-8,40%	1	56,25%	1	52,08%	0	-16,13%	0	-8,36%	28,25%
012	0	-27,00%	2	54,45%	2	49,59%	0	-39,28%	0	-66,83%	47,43%
013	0	-62,66%	2	65,51%	0	-9,05%	0	-128,69%	2	67,14%	66,61%
014	-1	-62,27%	0	31,24%	0	30,97%	-1	-40,81%	-1	-37,73%	40,61%
>z 015											
016	0	11,25%	0	-41,63%	1	30,03%	0	-39,58%	0	-32,96%	31,09%
017	0	0,00%	0	-75,60%	1	10,54%	0	3,55%	0	-147,17%	47,37%
018	0	-15,27%	1	32,92%							24,10%
019	-1	-41,12%	1	35,31%	-1	-31,71%	1	47,70%	0	-6,91%	32,55%
Campaign MAPE		18,74%		31,98%		22,76%		32,85%		29,85%	

Table 9: Changes in classification, 90th percentile error (%) and MAPE for each w.b; MAPE for each reduced dataset; period August 2006-Julay 2007

August 2007-July 2008

In the third year of survey 858 water samples are taken from coastal water bodies. The number of samples increased in water bodies 006, 018 and it is now possible to perform a trimestral temporal reduction analysis (with the exception of 018 for the third trimestral data set which consists of 7 water samples). 015 is still poorly monitored with only 11 water samples for this period. Even campaigns consist of 466 samples and uneven campaigns of 392 samples, trimestral campaigns number 312, 306 and 240 water samples respectively.

WATER BODY	No. of samples monthly campaigns	No. of samples even months	No. of samples uneven months	No. of samples 1st Set (trimestral campaigns)	No. of samples 2nd Set (trimestral campaigns)	No. of samples 3rd Set (trimestral campaigns)
001	**51**	28	23	19	17	15
002	**53**	28	25	19	20	14
003	**38**	20	18	14	13	11
004	**41**	24	17	13	15	13
005	**55**	30	25	20	20	15
006	**39**	21	18	14	13	12
007	**64**	35	29	23	23	18
008	**55**	30	25	20	20	15
009	**54**	29	25	20	20	14
010	**64**	35	29	24	23	17
011	**56**	29	27	21	19	16
012	**42**	23	19	15	14	13
013	**38**	20	18	13	14	11
014	**41**	23	18	15	14	12
015	**11**	**6**	**5**	**4**	**4**	**3**
016	**41**	22	19	16	15	10
017	**44**	24	20	16	16	12
018	**29**	16	13	11	11	**7**
019	**42**	23	19	15	15	12
Σ	**858**	**466**	**392**	**312**	**306**	**240**

Table 10: Number of samples for each water body; monthly, bimonthly and trimestral campaigns (August 2007 – July 2008)

The quality and quantity results for the reduction analysis for the period from August 2007-July 2008 are represented in Table 11, and changes in class, error percentage and MAPE values in Table 12.
Even months contain 11 (61,1%) unchanged classes of 18, with 5 (27,7%) degraded (001, 007, 011, 012 and 019) and 2 (11,1%) upgraded (006 and 010) as compared to their original status. The MAPE values for the third year are slightly lower, and the even month reduction data set has a MAPE value of 31,0% .
The number of unchanged classes is higher for uneven month campaigns 14 (77,7%) and counts 2 (11,1%) upgraded (018 and 019) and 2 (11,1%) degraded (002 and 005) statuses. The quantitative results also are better correlated with a 23,4% MAPE value.

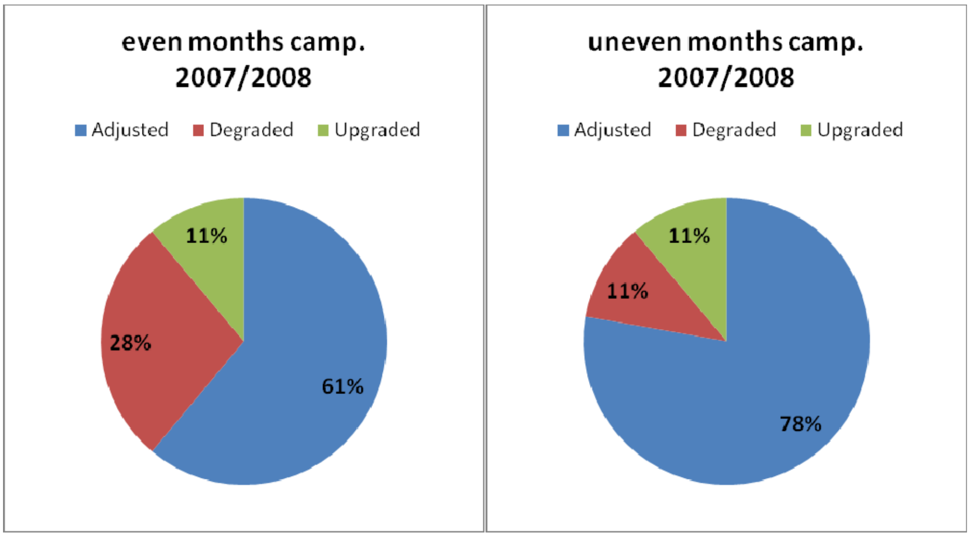


Figure 7: Quality results of bimonthly campaigns (August 2007 – July 2008)

The trimestral reduction analysis ended with 11 (61,1%), 12 (66,7%) and 12 (70,6%) unchanged ecological statuses respectively. Only for the third data set were 17 ecological statuses established due to the lack of data for body of water 018 water body. All 5 (29,4%) changed classes in the third data set represent degradations (002, 004, 005, 012 and 019), 5 (27,7%) classes are degraded (001, 005, 007, 011 and 012) and 2 (11,1%) upgraded (006 and 018) in the first set, 4 (22,2%) degraded (002, 006, 007 and 011) and 2 (11,1%) upgraded (018 and 019) in the second set. The MAPE values are 24,7%, 43,5%, and 26,9% respectively for the three trimestral data sets.

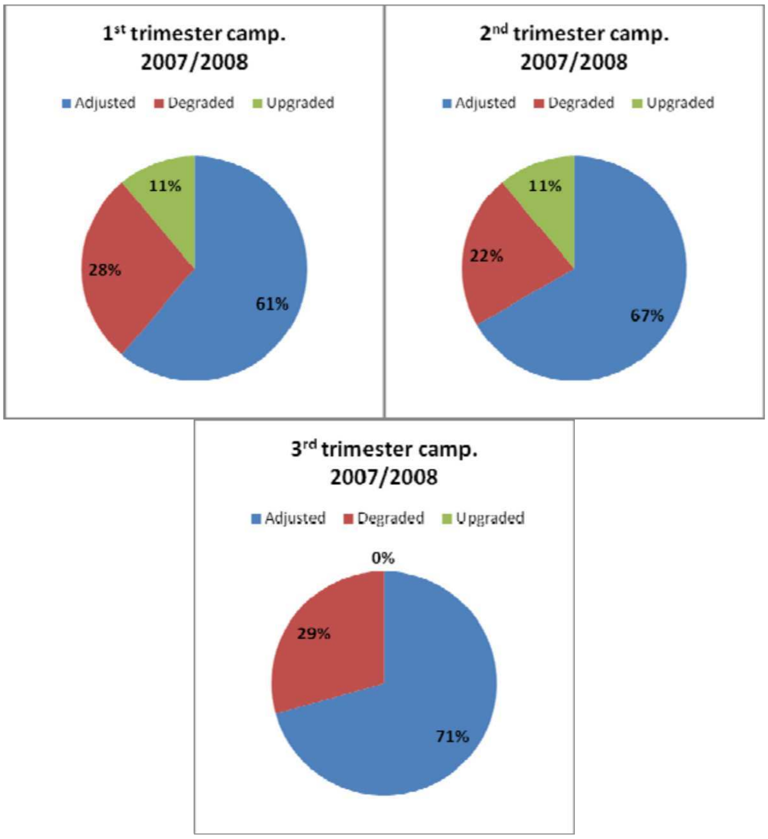


Figure 8: Quality results of trimestral campaigns (August 2007 – July 2008)

WATER BODY	TYPE OF BODY OF WATER	90th PTCL Chl a monthly campaigns	Classification monthly campaigns	90th PTCL Chl a even months	Classification even months	90th PTCL Chl a uneven months	Classification uneven months	90th PTCL Chl a 1st Set (trimestral campaigns)	Classification 1st Set (trimestral campaigns)	90th PTCL Chl a 2nd Set (trimestral campaigns)	Classification 2nd Set (trimestral campaigns)	90th PTCL Chl a 3rd Set (trimestral campaigns)	Classification 3rd Set (trimestral campaigns)
001	II A	3,43	HIGH	6,01	GOOD	2,56	HIGH	6,00	GOOD	3,80	HIGH	3,58	HIGH
002	II A	4,49	HIGH	3,93	HIGH	6,04	GOOD	3,82	HIGH	5,21	GOOD	5,11	GOOD
003	II A	2,19	HIGH	2,52	HIGH	1,97	HIGH	3,10	HIGH	2,50	HIGH	2,02	HIGH
004	II A	3,44	HIGH	4,58	HIGH	3,17	HIGH	3,46	HIGH	4,36	HIGH	4,80	GOOD
005	II A	6,98	GOOD	6,75	GOOD	7,61	MODERATE	7,23	MODERATE	5,96	GOOD	7,90	MODERATE
006	II A	4,80	GOOD	4,78	HIGH	6,76	GOOD	4,55	HIGH	14,43	MODERATE	5,74	GOOD
007	II A	6,06	GOOD	8,29	MODERATE	4,98	GOOD	8,47	MODERATE	8,39	MODERATE	6,13	GOOD
008	II A	11,81	MODERATE	7,87	MODERATE	19,12	MODERATE	12,84	MODERATE	18,23	MODERATE	8,49	MODERATE
009	II A	10,38	MODERATE	10,33	MODERATE	10,54	MODERATE	13,68	MODERATE	9,56	MODERATE	7,35	MODERATE
010	II A	8,02	MODERATE	6,72	GOOD	8,74	MODERATE	8,20	MODERATE	7,38	MODERATE	11,05	MODERATE
011	III	2,00	HIGH	2,45	GOOD	1,80	HIGH	3,35	GOOD	2,45	GOOD	1,98	HIGH
012	III	2,17	HIGH	2,70	GOOD	1,87	HIGH	3,60	MODERATE	1,99	HIGH	4,46	MODERATE
013	III	4,04	MODERATE	5,16	MODERATE	4,28	MODERATE	4,99	MODERATE	4,03	MODERATE	6,68	MODERATE
014	III	1,83	HIGH	1,95	HIGH	1,48	HIGH	1,78	HIGH	1,88	HIGH	2,12	HIGH
015	III	-	-	-	-	-	-	-	-	-	-	-	-
016	III	7,16	MODERATE	6,94	MODERATE	7,73	MODERATE	7,21	MODERATE	8,14	MODERATE	7,07	MODERATE
017	III	9,56	MODERATE	23,80	MODERATE	5,75	MODERATE	5,66	MODERATE	36,93	MODERATE	10,09	MODERATE
018	III	3,80	MODERATE	5,76	MODERATE	1,90	HIGH	3,56	GOOD	2,75	GOOD	-	-
019	III	2,46	GOOD	3,61	MODERATE	1,07	HIGH	2,76	GOOD	1,70	HIGH	4,13	MODERATE

Table 11: Ecological classifications calculated by monthly, bimonthly and trimestral campaigns for the period August 2007 – July 2008

w.b.	Even months		Uneven months		1st set (trimestral camp.)		2nd set (trimestral camp.)		3rd set (trimestral camp.)		w.b.MAPE
	unchanged class	**error %**	**unchanged class**	**error %**	**unchanged class**	**error %**	**unchanged class**	**error %**	**unchanged class**	**error %**	
001	-1	-75,48%	0	25,39%	-1	-75,13%	0	-10,80%	0	-4,38%	38,24%
002	0	12,47%	-1	-34,55%	0	14,96%	-1	-15,94%	-1	-13,65%	18,31%
003	0	-14,78%	0	9,99%	0	-41,20%	0	-14,23%	0	7,85%	17,61%
004	0	-32,99%	0	7,85%	0	-0,64%	0	-26,69%	-1	-39,53%	21,54%
005	0	3,28%	-1	-9,06%	-1	-3,68%	0	14,64%	-1	-13,27%	8,79%
006	1	0,50%	0	-40,73%	1	5,21%	-1	-200,63%	0	-19,54%	53,32%
007	-1	-36,98%	0	17,75%	-1	-39,82%	-1	-38,63%	0	-1,17%	26,87%
008	0	33,33%	0	-61,91%	0	-8,70%	0	-54,42%	0	28,08%	37,29%
009	0	0,43%	0	-1,63%	0	-31,81%	0	7,87%	0	29,16%	14,18%
010	1	16,21%	0	-9,05%	0	-2,31%	0	7,95%	0	-37,89%	14,68%
011	-1	-22,68%	0	9,96%	-1	-67,85%	-1	-22,68%	0	0,70%	24,78%
012	-1	-24,65%	0	13,67%	-2	-66,39%	0	8,13%	-2	-105,82%	43,73%
013	0	-27,69%	0	-5,96%	0	-23,53%	0	0,27%	0	-65,40%	24,57%
014	0	-6,57%	0	18,95%	0	2,74%	0	-2,68%	0	-15,94%	9,38%
015											
016	0	3,09%	0	-7,93%	0	-0,73%	0	-13,63%	0	1,33%	5,34%
017	0	-148,90%	0	39,81%	0	40,77%	0	-286,26%	0	-5,52%	104,25%
018	0	-51,45%	2	50,00%	1	6,37%	1	27,58%			33,85%
019	-1	-47,01%	1	56,45%	0	-12,33%	1	30,73%	-1	-68,05%	42,91%
Campaign MAPE		31,03%		23,37%		24,68%		43,54%		26,90%	

Table 12: Changes in classification, 90th percentile error (%) and MAPE for each w.b; MAPE for each reduced dataset; period August 2007-July 2008

4.1.3. Discussion

- As mentioned above, the results of the analysis of temporal reduction show unpredicted consequences. The percentages of the unchanged ecological classes, as expected, are higher for the three year period analysis due the number of water samples. The percentage of unchanged classifications (for the three year period) does not descend below 74%, which can be considered a fairly good result, but the changed classification cannot be predicted and could have unexpected consequences for any of the water bodies.
- The one year periods have more serious data reduction and the number of data in the reduced sets (especially for the trimestral campaigns) is more critical. Due to this, the bimonthly reduction sets have slightly better results, unchanged percentages descend to 61% and the trimestral campaigns analysis to 50%. These percentages present a high level of uncertainty and totally exclude the bimonthly and trimestral frequency campaign option for the one year period body of water classification.
- Finally the WFD demands a five year period of survey data to establish ecological status for the coastal bodies of water. Also, the directive requires a survey with a minimum three-month frequency for chlorophyll a data. The percentage of uncertainty (for bimonthly or trimestral campaigns) will not significantly decrease for the five year period (comparing for the three year results), because it does not depend on the quantity of samples (when it exceeds a certain threshold), but rather depends on the trend, regularity and patterns of the environmental parameters. These analysis results question the minimum survey frequency established by WFD and supports one of the first decisions of the MEDGIG experts (at the first intercalibration reunion) to monthly monitor coastal waters in the scope of ecological classification.

- The same analysis for all four periods is performed by EXCEL algorithm, the results are slightly different (due the dissimilar 90^{th} percentile values) and have similar unpredicted consequences. Due to this, results are not presented, because it is preferable to avoid redundant analysis which cannot contribute any new significant conclusion.

4.2.Seasonal pattern regularity

The idea of monitoring a network of possible temporal reduction which does not affect ecological classification was based on the regularity of the chlorophyll a temporal (annual) pattern.
In the Valencian littoral, phytoplankton biomass proliferation is basically conditioned by two factors:

- Continental influence loads
- Hydrodynamic conditions

Autumn rainy periods, usually accompanied by sea storms, and summer drought periods are considered to be the seasonal climate characteristics of the Valencia region. During the rainy periods, the coastal waters are significantly affected by ecological stressors (increased loads of freshwater with nutrients – even high chlorophyll a concentrations – and other pollutants discharged from the land) that affect the biomass dynamic and can disturb eutrophic balance. Through the ever more frequent drought periods, terrestrial runoffs are entirely absent, which reduces the inflow of nutrients and pressure from ecological stressors.
During the rainy period it is (usually) easy to observe the rain/plume effect, which consists of inshore waters (especially type II-A) having a decreased salinity due to the freshwater continental influence discharged from rivers, gullies, groundwaters, etc. Usually, the decreasing salinity is accompanied by an increase in nutrient loads which are taken up quickly by phytoplankton (and bacteria), entering into the food chain as particulate organic matter, leading to organic enrichment (Furnas 1997). Finally, the organic enrichment is detected by higher concentrations of phytoplankton biomass (chlorophyll *a).*
Although this effect in its conceptual schema is correlated with the intensity of rain events, in reality this process is more complex.
First, the rain/plume effect is affected by various anthropogenic processes such as land use together with irrigation, water accumulation at dam reservoirs, water gates, redirection of river flows etc. All of these processes reduce (or totally suspend) freshwater terrestrial runoffs, or simply retain inflow of the catchment area (basin) to the sea.
Besides this, coastal hydrodynamics is also an essential background driver, currents and waves are essential to determine the flushing or retention of freshwater, as well as corresponding nutrient concentrations. During sea storms (accompanied by elevated wave height), oceanic waters arrive with high salinity and accelerate mixing process of freshwater and seawater.
Due to this, if rainy periods coincide with sea storms, the salinity values can increase instead of reducing in concentration as expected. This accelerated mixing process can also reduce the nutrient concentrations that affect the phytoplankton biomass, and chlorophyll a values remain low.
However, sea storms can also affect biomass concentrations in the opposite way. Elevated wave amplitude can affect thermocline stability and can provoke a kind of "upwelling" process that can enrich superficial waters with nutrients. Elevated wave trough can penetrate the thin thermocline layer and provoke the "upwelling" flow of nutrient-rich deep water which then mixes with superficial layers. Superficial and shallow water enriched with nutrients responds in a proliferation of phytoplankton and hence higher chlorophyll a values. This process can be easily detected through elevated chlorophyll a concentrations shortly after sea storms during drought periods.

Local winds develop superficial currents that can influence biomass dynamics and spatial distribution so that the area of continental water influence directly depends on the direction of the wind.

4.2.1.Valencian Community seasonal patterns

To identify climatic patterns, historical data from the Jucar Basin pluviometry network and various buoys that collect wave height data are used. The pluviometric data is presented as a time series in the diagram as values of 15 days of accumulated rain and daily rain events.

The wave data is presented as:

- Annual wave roses that show the long-term distribution of wave height and direction
- Daily events diagram in annual format (temporal series of significant wave heights)

Finally, all environmental parameters (waves and rain) and salinity are superimposed to describe their mutual interaction and their influence on phytoplankton biomass (chlorophyll a). To this end, the campaign date function of the rain series has been modified. Rainfall is presented as values of 15 days accumulated rain for the campaign sampling date. Using this methodology, the amount of rainfall (in the coastal water sub-basin) for the past 15 days, before the water sampling, is identified. The 15-day “buffer” is chosen as proxy for:

- Time that fallen rain needs to drain from the basin into the water flow (river, gully…)
- Time from drainage to inflow into coastal waters
- Time that phytoplankton biomass needs to react to (freshwater) nutrient pressure

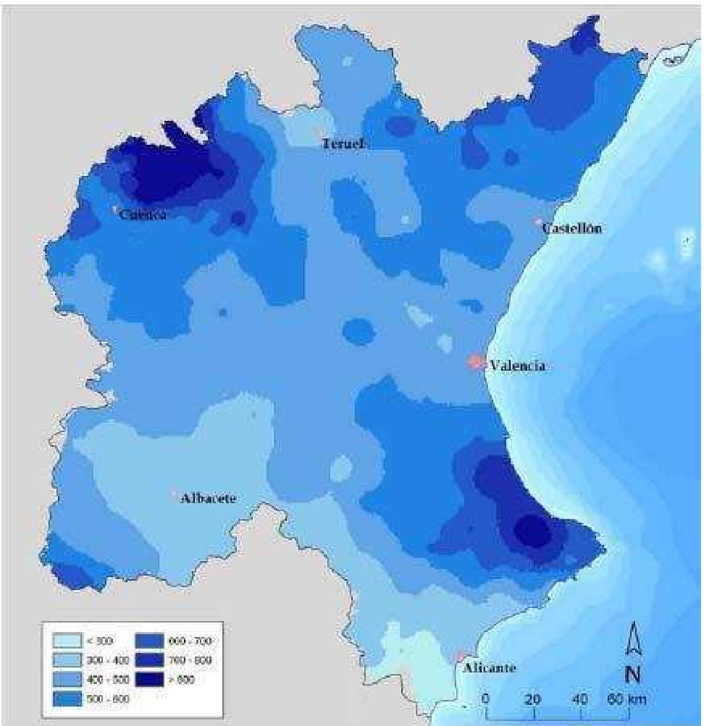


Figure 1: Average annual precipitations (l/m^2 per year) for the Jucar basin

Since wave influence on the coastal waters was analyzed in this part, it is important to identify the presence of various maritime climate conditions for the Valencian littoral.

Maritime climate conditions are determined by the coastal morphology (orientation of the coast line, Balearic Islands and other coastal barriers, seabed slope…) and the corresponding fetch (length of open water superficies) which determines the size of produced waves.
Three different maritime climate regimes are identified for the Valencian coastal waters:

1. Vinaroz – Cabo de Oropesa-including water bodies:
   001,002,003

2. Cabo de Oropesa - Saint Antonio Cape-including water bodies:
   004,005,006,007,008,009,010

3. Saint Antonio Cape - Sant Pedro del Pinatar -including water bodies:
   011,012,013,014,015,016,017,018,019

Therefore this part of study analyzes abiotic parameters (rain, waves and salinity) and phytoplankton biomass (chlorophyll *a*) interaction and pattern regularity for the three maritime climates mentioned above.

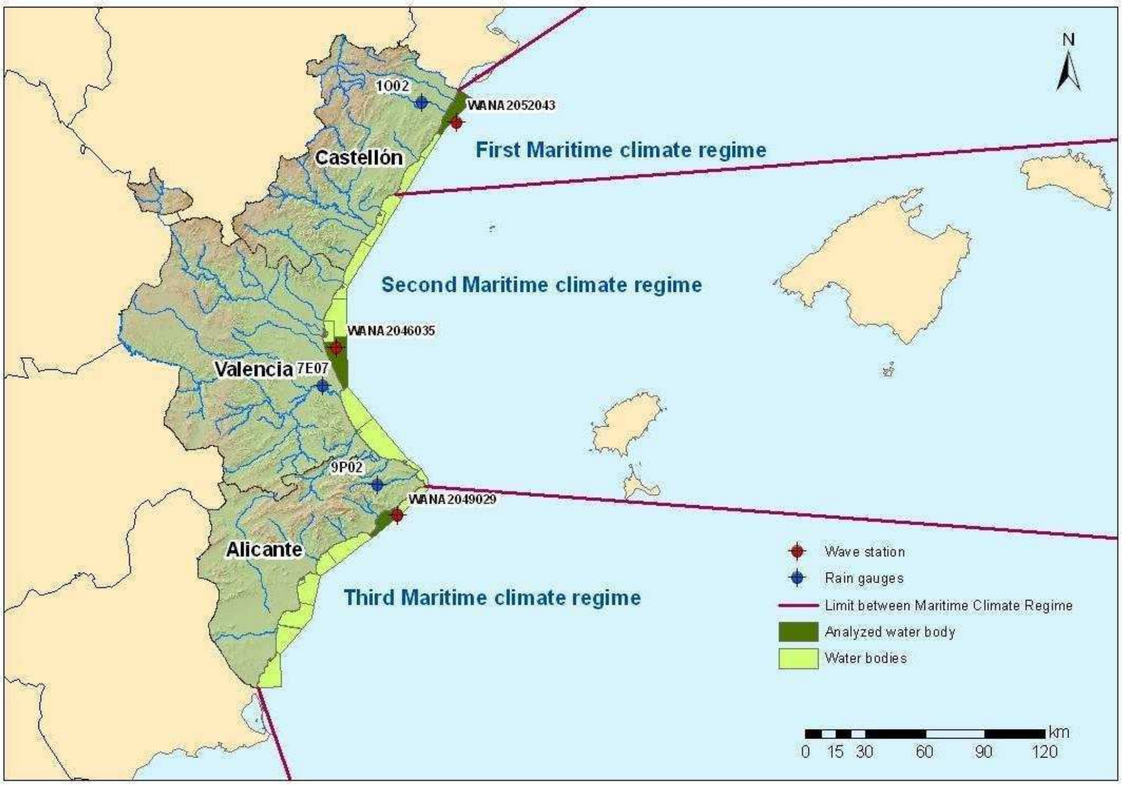


Figure 2: Three different maritime climate regimes identified for the Valencian coastal water

1st Maritime climate regime (Vinaroz –Cabo de Oropesa)

The first sea climate regime for the North Coast of Valencia is determined by a NE-SW extension of the littoral sheltered on the North by the delta of the river Ebro. The coastal fetch is determined by the Catalan coast on the north, by the French and Italian coast and Corsica on the northeast and by Sardinia on the east. The fetch length from southwest to southeast is determined by the Valencian coast and the Balearic Islands.

<u>Wave series</u>

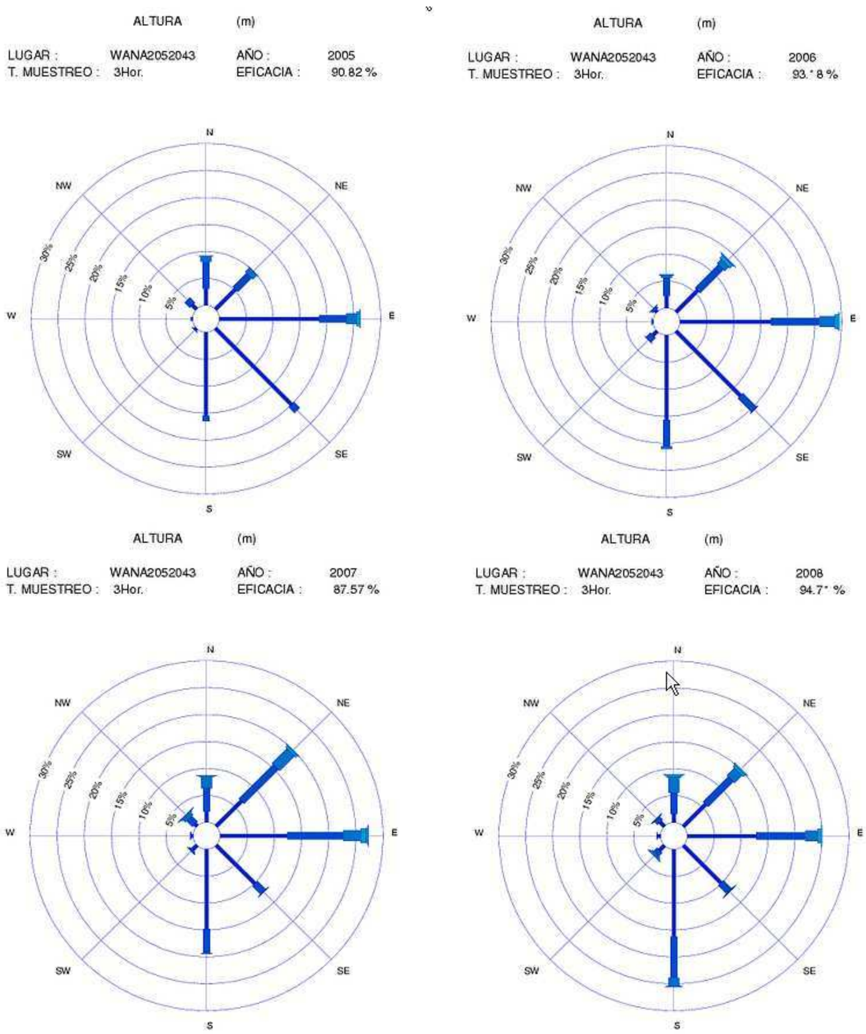


Figure 3: Wave roses; annual evolution for 2005, 2006, 2007 and 2008; data from point WANA 2052043

The point WANA 2052043 detected waves proceeding from the east 26% of the time in 2005 with a maximum height of approximately 2.5m. The waves from the southeastern component affect the coast 21% of the time and those from the southern component about 16% of the time. The percentage of time affecting the coast by southern and southeastern components is significant, but the wave height values (as expected) are considerably

lower. The North-East component affects the coast a low percentage of time but the wave heights are significant.
Sea storms with significant wave heights (proceeding from the northeast and east) are detected in the end of February and in November of 2005. From the end of April until the end of October, a "calm period" is detected with one single event (September) that exceeds 1m of significant wave height.

In 2006, 3-meter-high waves are detected during a sea storm in December and important events are sensed by the buoy in the January, February, May and the beginning of November. During 2006, waves proceeding from the east were detected almost 30% of the year, from the southeast 21% of the year, and from the south 21%. The highest waves (3m) are detected for the northeastern component, encompassing 13% of the year. The southern and southeastern components affect the coast more than 40% of the year but with wave heights that are significantly lower. The frequency and intensity of sea storms are also significantly higher than those detected for 2005 due to the fact that a long "calm period" is not detected.

2007 has three important wave components; the eastern component (27%) with waves exceeding 3 meters, northeastern component (20%) and southern component affecting the coast line 19.5% of the year. Waves proceeding from the northeast have relevant values of detected significant heights. The significant height values detected for the southern component are low and not comparable with wave heights proceeding from the northeast and east. The southeastern component for 2007 has no relevant percentage of time affecting the coast (less than 12%) and significant heights do not exceed relevant values for this study. Sea storms that exceed 2 meter wave height are detected in January, March, August, October and December. The end of April until the end of August cannot be identified as a "calm period", but it is significant that no sea events were detected exceeding 1.5m significant wave height.

Waves proceeding from the east are the most important component of the 2008 wave rose with 25% and maximum heights exceeding 3 meters, while the southern component affects the coast more than 25% of the year with significantly lower wave heights. The northeastern component affects the coast during 15% of year with significant sea events. As for 2007, relevant wave heights are not detected for the southeastern component and the percentage of time distressing the coastline is low (7%). In 2008, sea storms are detected in March, May, October, November and December. A "calm period" without events that exceed 1,2 meters in height is detected from the end of May until the end of August. The first event that exceeds 1.5 meters is detected at the end of September.

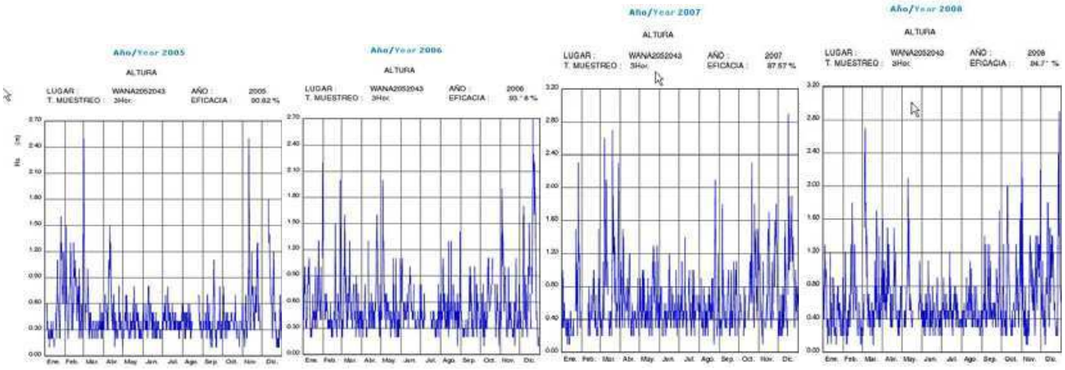


Figure 4: WANA 2052043 - 4 annual evolutions for 2005, 2006, 2007 and 2008; time series of significant wave heights

The annual wave roses from point WANA 2052043 indicate that the coast is influenced by waves from the northeast to the south, but it is apparent that waves with the highest heights come from the northeast and east. Wave components from the south and southeast are important in the percentage of time they affect the coastline, but only exceed 1,5m of significant wave height in 2007 and 2008.
If we compare the 4 year data, it is possible to establish some kind of annual pattern of sea storm events. Significant sea storms are quite possible in the first five months of the year (January – May), but it is not possible to predict the specific month or frequency of important events. After these five months, during the summer, sea storms are absent, but atypical events that usually do not exceed 1 meter in height can take place. This summer "calm period" (with irrelevant events):

- can be interrupted with significant events at the end of August (as during 2007)
- can last until the end of September or October (as during 2006 and 2008)
- can extend until the first part of November (as was detected in 2005)

Therefore, if we analyze sea storm frequency, some kind of seasonal periodicity does exist, but it is not possible to establish a clear monthly pattern.

Precipitation series

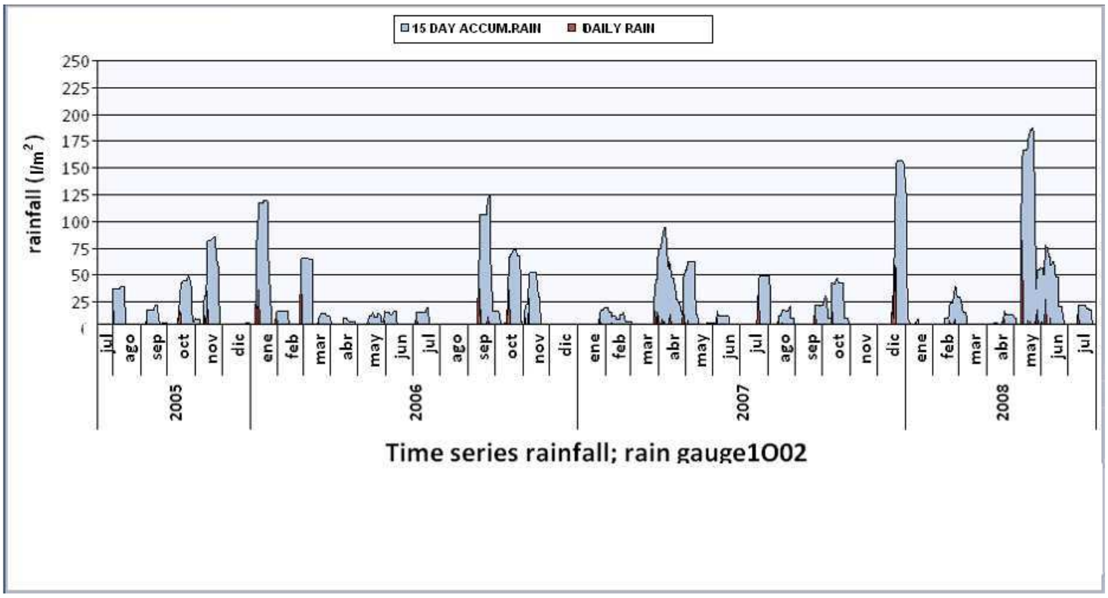


Figure 5: 15-day accumulated rainfall and daily rain; rain gauge 1O02 - corresponding to water body 001

On the precipitation diagram of 15-day accumulated rainfall for the northern coast of Valencia, it is difficult to identify any pattern or regularity. Comparing the three year data, it can be concluded that summer is the season with low rainfall, especially in 2006 which includes a period of very low precipitation (April - August). During the summers of 2005 and 2007, rain events are detected, but the 15 day accumulated values are low and do not exceed 50 l/m^2.

Like summer, winter is also historically considered a low precipitation period. This is confirmed for the winters of 2006/2007, where the 15 day rain accumulation does not exceed 20l/m^2 . The historical pattern can be supported with the fact that the winter data confirms at least a one month period without rain.
During 2005 and 2006, December has a low precipitation index while in 2007 it exceeds 150l/m^2 of 15-day accumulated rainfall. January and the end of February 2006 have significant rain events that accumulate in 15 days more than 100l/m^2 and 60 l/m^2 respectively. During January 2008, no significant rain events are detected, as are not in the end of the December 2007.

During April and May of 2007, important precipitation events are detected (exceeding 60l/m^2 of 15-day accumulated rainfall) which are repeated in May and the beginning of June 2008 (during which period the highest rate of precipitation is detected, with accumulation in May of more than 175l/m^2). The spring of 2006 has significant rain events in March but April and May belong to a five months period of extremely low precipitation.
For all three years, significant rain events are detected during autumn but with variations in monthly frequency and intensity. November 2007 was a completely dry month while rain fell frequently during the end of September and October. During 2006, it rained during the second part of September and with significant frequency and intensity during October and November. For 2005, significant rain events with high accumulation are detected in November.

As for the wave series, regularity on a monthly basis is absent, and it is complicated to establish a pattern using the precipitation diagram for the three year period even on a seasonal scale.

<u>Salinity value series</u>

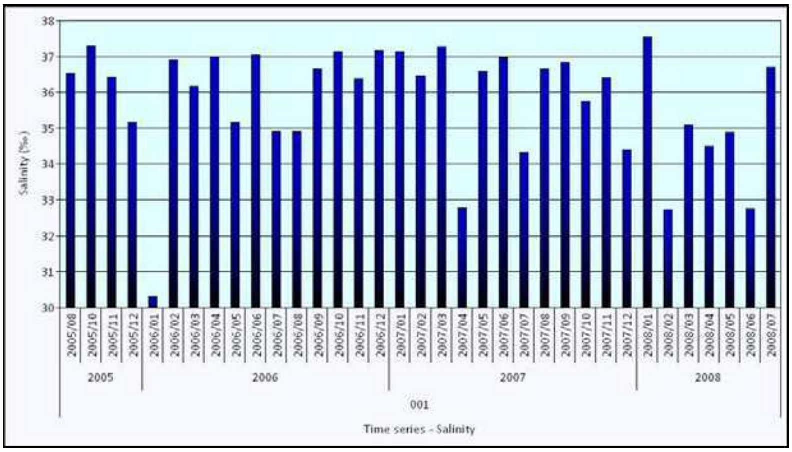


Figure 6: Campaign average salinity values; water body 001

Observing the salinity campaign average values for water body 001, it is obvious that neither monthly nor seasonal regularity exist. The minimum value (30,2‰) and the maximum value (37,5‰) of the water body are detected in same month but in different years; January 2006 and January 2008. The majority of water body average values vary between 36‰ and 37,2‰, which is to be expected for a type II-A water body. Values below 36‰ present significant continental influence and are probably a response to precipitation events in the area. These low salinity values are detected in December 2005; January, May, July and August in 2006; April, July, October and December in 2007; February, March, April, May and July during the 2008.

<u>Chlorophyll a series</u>

Since phytoplankton biomass is influenced by nutrient concentration and hydrodynamic conditions, the temporal series of campaign average values from the chlorophyll a diagram is analyzed in the context of wave conditions, pluviometry series and the salinity diagram for water body 001. The data from the rain gauges is analyzed by:

- Time series of daily and 15 day accumulated rainfall for the entire year (Figure 6 rain series water body 001)
- Time series of 15 day rainfall accumulation that corresponds to the campaign sampling date (if sampling is performed on the 1st of November, significant rainfall data that corresponds to the sample belongs to the second part of October and rain events detected after the 1st of November are irrelevant)

Analyzing the two precipitation diagrams (Figure 5-precipitation diagram and Figure 7,15 day accumulation diagram) the significant rain events detected by rain gauge 1O02 do not always correspond to the sampling date of the related 15 day accumulation value.
During October 2005, accumulation reaching 50l/m^2 is detected while the campaign date value corresponds with less than 5l/m^2 . The same occurs during March 2006 (60l/m^2 reached during the month while sampling date value is less than 10l/m^2), March, April and May 2007 (these months exceeded 50, 100 and 60 l/m^2 while the sampling date values are 0.20 and less than 10l/m^2 respectively). These differences between the two precipitation diagrams (detected rain and corresponding accumulation on the sampling day) are due to:

- *water sampling performed before significant precipitation events (but during the same month)*
- *water sampling performed 16 or more days after significant precipitation events,* because of which it is not expected that salinity values reflect continental influence

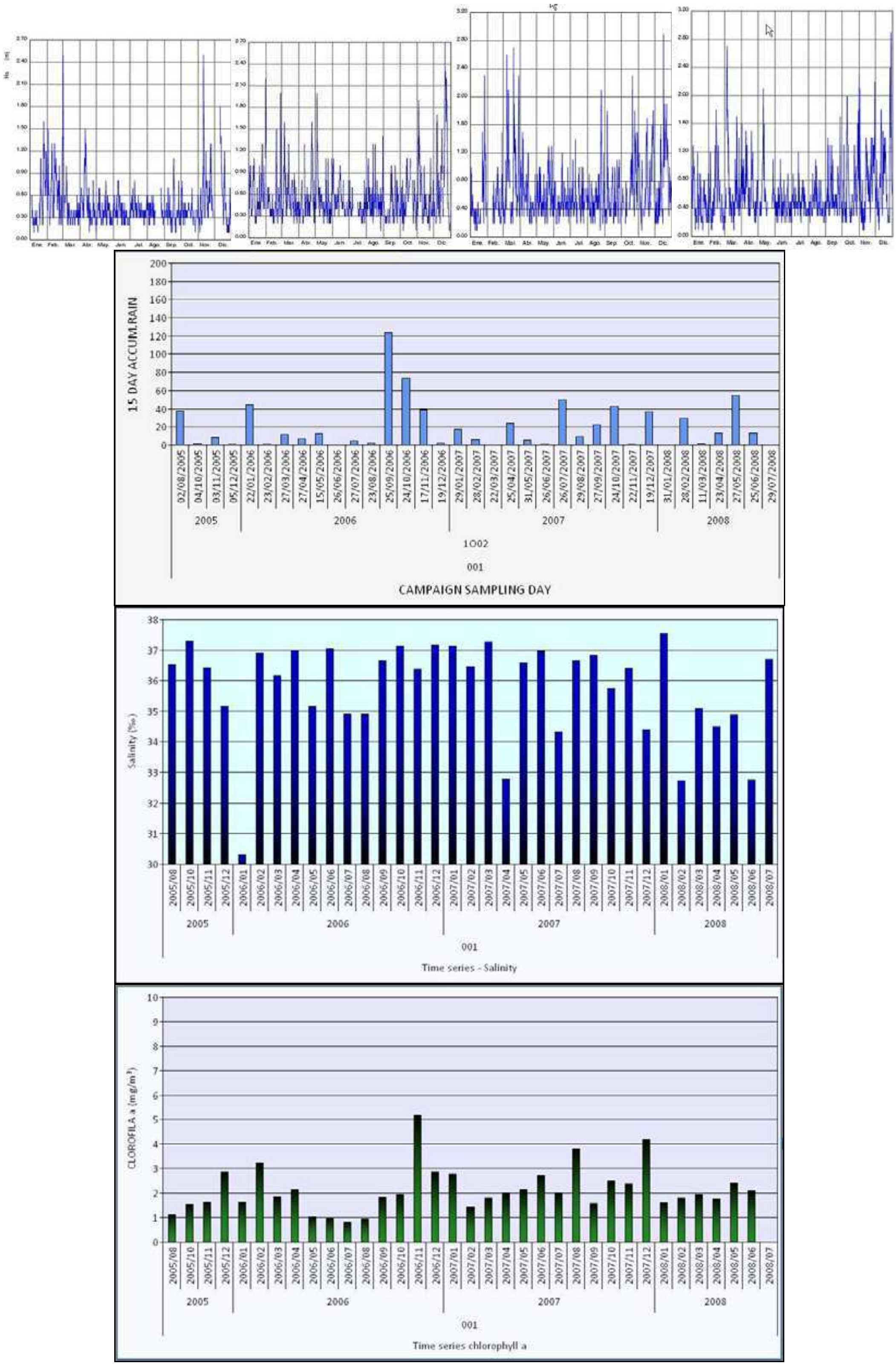


Figure 7: WANA 2052043 time series of significant wave heights; 15-day accumulated rainfall (sampling day); campaign average salinity values; campaign average chlorophyll a values w.b.001

Observing the chlorophyll a three year temporal series, it is apparent that there is no regularity on a monthly basis. This is a result that could be expected due to the lack of established patterns of environmental factors that have a significant influence on biomass dynamics.

Water body 001 has a significant yearly precipitation volume and rainfall events that exceed average values for the Valencian coast (Figure 1.). The time series of 15 day accumulated rainfall indicate that low salinity values correspond to rain events in January 2006, April 2007, July 2007, October 2007, December 2007, February 2008 and May 2008. These rain events are around or exceed 40 l/m^2 of accumulated rainfall during the last 15 days and the process of mixing (freshwater and seawater) results in a salinity concentration of less then 35,0‰. As mentioned above, the high precipitation level in October 2007 results in a low salinity concentration, but with a value around 35.7 ‰. The salinity concentration in April 2007 and February 2008 is less than 33‰ but the accumulated rainfall for these months slightly exceeds 20l/m2 .

During August of 2005 and September, October and November of 2006 (15 day rainfall accumulation on sampling date of at least 40 l/m^2) the expected salinity values were not detected (<35‰). The time series from point WANA 2052043 which records and provides significant wave heights in the offshore area (for water body 001) explains the non-corresponding salinity values for September, October and November of 2006. The buoy detected sea events along with waves exceeding a minimum of 1,0 m of significant height, accelerating the mixing and dispersion processes. During August 2005 sea storm events were not detected (possibly because of buoy data discontinuation) and the salinity value is slightly lower (around 36‰) but still higher than expected.

Finally, the rain/salinity link is tighter due to anthropic influence in the water body. Water samples during some campaigns (in some environmental conditions determined by winds and currents) are directly influenced by anthropic freshwater discharges. This can explain the low salinity values for periods during which precipitation was not detected (July and August 2006 and March 2008)

The December campaign has the highest chlorophyll a values for 2005 and lower salinity values that reflect nutrient enrichment and biomass augment. Although the lowest salinity value from the whole series was detected in January 2006 (30,2‰), the corresponding chlorophyll a value is below 2,00 µg/l. Also, the rest of the detected low salinity concentrations (May, July, August of 2006; April, July of 2007 and almost all values for 2008) apparently do not reflect the expected phytoplankton biomass growth values.
Therefore, the salinity concentration can be used as an indicator that provides information about possible nutrient enrichment or current oligotrophic conditions of coastal water ecosystems, but it should be considered as a qualitative indicator and not a quantitative one. A single salinity value can only provide information about continental influence (which is usually accompanied by nutrient enrichment), but not fresh water volume or nutrient concentration that affects coastal ecosystems (Figure 8).

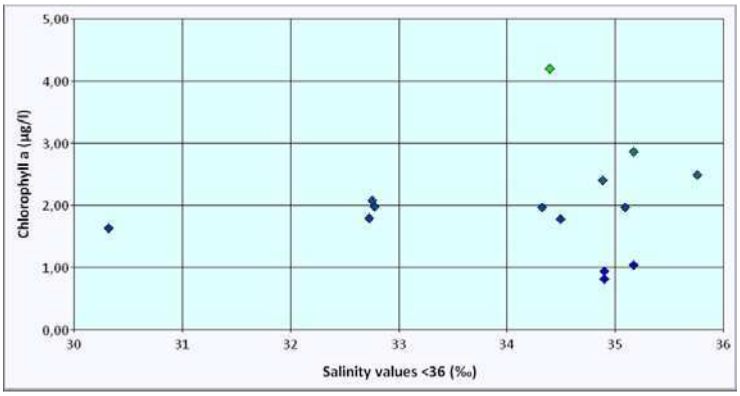


Figure 8: Chlorophyll a VS Salinity <36,0 ‰; scattered diagram points do not indicate a correlation between biomass and continental loads

Biomass reaction to nutrient enrichment is fast but not synchronous, and sampling was sometimes performed during the phytoplankton growth process. This time-lag (nutrient loads/biomass enrichment) could explain some of the above-mentioned non-corresponding chlorophyll a /salinity values.

Finally, most of the chlorophyll a values for water body 001 are below 3,0 µg/l, and for this range of values it is complicated to detect a clear correlation between biomass and continental loads (Figure 8).

Although a pressure-biomass concentration link is lacking, a kind of regularity exists on a seasonal basis with a winter biomass increment (without frequency and without monthly regularity) that certainly cannot be used to establish an annual chlorophyll a pattern.

2nd Maritime climate regime

The second maritime climate regime is determined by the Balearic Islands which close the fetch on the east and reduce wave-related coastal influence. The maritime regime fetch is wide open on the northeast side, from where waves with maximum values are expected. Due the Balearic Islands, but also to coastal morphology, wave influence is only possible from the northeast to the southeast (a mere 90º of open sea).

<u>*Waves series*</u>

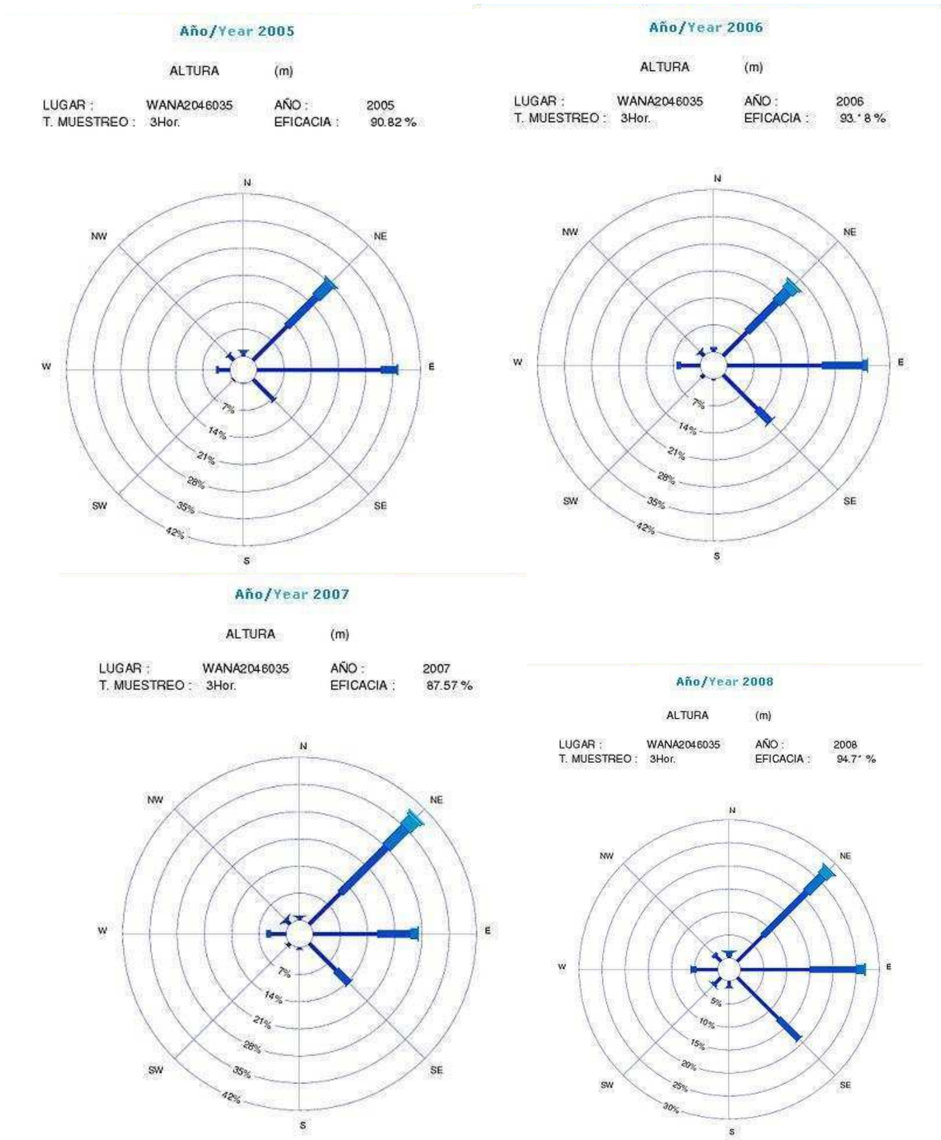


Figure 9: Wave roses; annual evolution for 2005, 2006, 2007 and 2008; data from point WANA 2046035;

On the wave rose diagrams derived from the data provided by point WANA2046035, it is possible to detect three significant wave components (northeastern, eastern and southeastern) and their variation during the four one-year periods. For the northeastern component, maximum significant wave heights were detected as expected during all one-year periods (due the fetch length),.

During 2005, the northeastern (27%) and eastern (36%) components affect the coast line more than 60% of the year. At the end of February, a single event for 2005 was registered where the significant wave height exceeded 2m (proceeding from the northeast). During the first three months (January-March) there is an important number of events that continually exceed the 1m value. After an event in April that exceeded 1,8m of significant height, the buoy register demonstrates a very long “calm period”, lasting until the beginning of November.

The time influence percentage for 2006 for the northeastern (28%) and eastern (36%) components is almost equal to that for 2005, but with a considerable increase in the southeastern (17%) factor. The time series for 2006 presents a higher frequency of events that exceed 1.5m (almost every two months) and the storms and “calm period” are not separated as in 2005. Storms proceeding from the North-East with values that exceed 2.2m were detected in January and December.

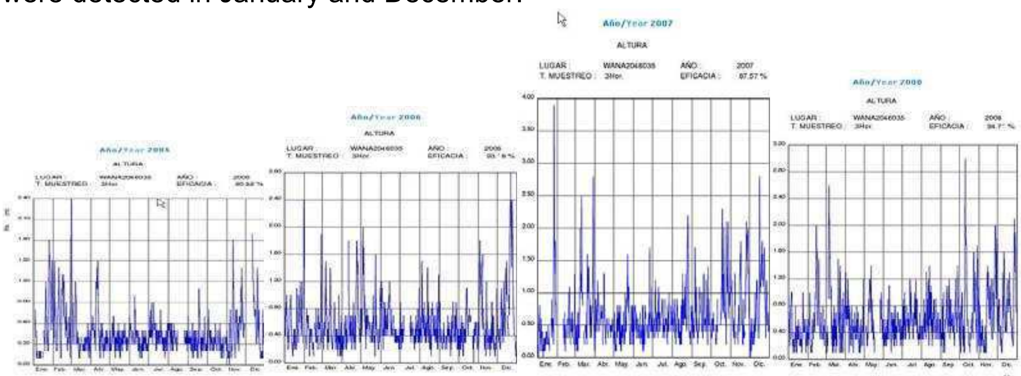


Figure 10: WANA 2046035 - 4 annual evolution 2005,2006, 2007 and 2008; time series significant wave heights

During 2007, the northeastern component affects the coast about 40% of the time and significant wave heights exceed the values from 2005 and 2006. The eastern component is significantly reduced, affecting the coast 25% of the time, compared to the periods from the first two years. Waves proceeding from the southeast affect the coast line about 14% of the year. The maximum significant wave height was detected during January proceeding from the northeast. All of these sea storms mostly proceed from the northeast and the maximum detected significant wave height reached almost 4m. Sea storms that exceed 2m in significant height were detected during the end of January, two times in March, at the end of August, three events in October, two events in November and one in December.

Finally, in 2008, the time during which the northeastern and eastern components affected the coast is balanced but reduced (about 28%), and the southeastern component has a more significant role, affecting the coast nearly 20% of the time. *Sea storms and waves with maximum values (about 3.2m of significant height) proceed from the northeast.* Five events that exceed 2m of significant wave height were detected; February, March, October, November and December. Between March and October we have a period without sea storms but with frequent events with significant heights of around 1m, and it cannot be compared to the calm episode of 2005.

The data obtained from point WANA 2046035 demonstrates that a kind of seasonal pattern exists: The first part of January is calm and the values do not exceed 1m of significant height during all four years. During the end of January (2005, 2006 and 2007) or in the middle of February (2008), a period with frequent sea storms begins which lasts until March (2008), April (2005 and 2007) or May (2006). The storm period is followed by a “calm period” that ends in autumn (early November; 2005, 2006 or in October 2008), though during some years it is possible to detect particular significant events in the end of summer (August 2006, August 2008). The “calm period” of 2007 is divided into shorter calm episodes due to various sea storms detected in May, June and August. During the period from November to December, it is common to detect at least two sea storms.

Despite the above described seasonality, the buoy data presents significant differences between years. During 2005, a long “calm period” was detected (April – early November) that was not been repeated in the following three years. The seasonal pattern is mostly unaligned for the entirety of 2006 (significant heights more frequently exceed the 1m value and the “calm periods” are shorter). The values of significant heights and the time that the coast is affected by waves were significantly higher in 2007 and the “calm period” was not detected as in other years. In 2008 the summer “calm period” has a lower significant wave height series compared to the spring and autumn values, but they frequently exceed 1m. Finally, the “calm period” from 2005 compared to the same period during 2007 (four events exceeding 1.5 m detected) presents two extreme conditions for the same part of the year.
From the presented data, it is clear that monthly regularity cannot be established, even if some kind of seasonality is detected.

Precipitation series

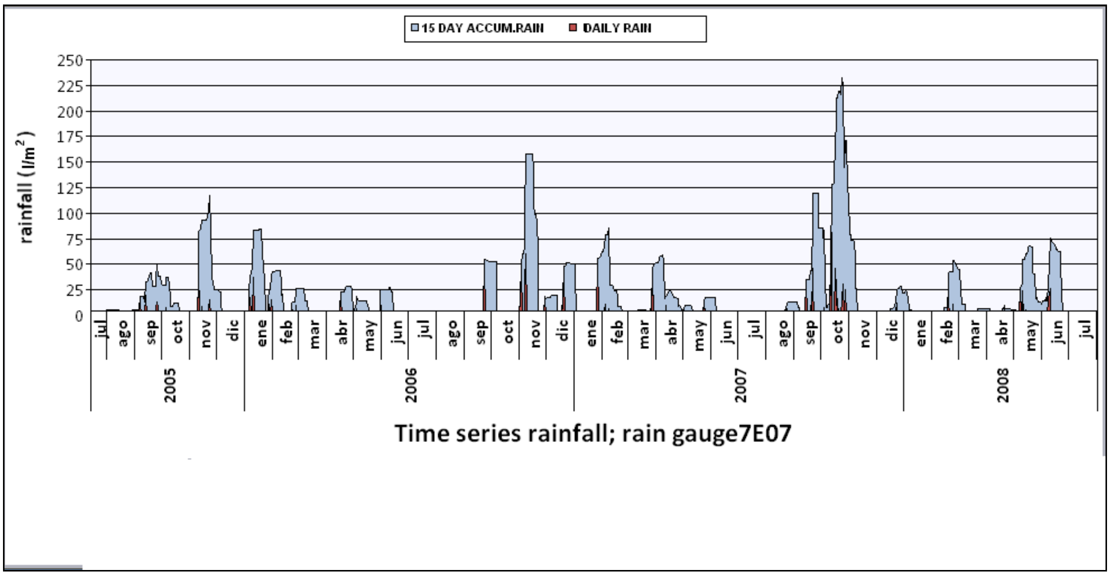


Figure 11: 15-day accumulated rainfall and daily rainfall; corresponding to water body 008

For water body 008 and the second maritime climate, data from rain gauge 7E07 is used. During four years, four events were detected that exceed 100 l/m^2 of accumulated rainfall during a fifteen day period.

The summer periods, especially July and August, are the months with the lowest precipitation index for the whole data set.
The most significant rain events for this area were detected during autumn; the maximum 15 day rainfall accumulation of 225l/m^2 was detected during October 2007. The precipitation index for October 2005 and 2006 is low (almost dry), but events were detected in November that exceed 100 l/m^2.

Winter periods for 2005/2006 and 2006/2007 have significant rain events for the area (rain events exceeding more than 75l/m^2 of accumulation). Intensive and frequent rains in October 2007 were followed by a drought in November and a low precipitation rate for the winter period.

Spring is historically considered the secondary precipitation period, which is not proven by the presented diagram (Figure 11). During 2006 and 2007, rain frequency was high, but the accumulation and intensity of the rain were low. During 2008, March and April have an extremely low precipitation index, with significantly high values for May and June, compared to data for 2006 and 2007.

For 2006 and 2007, the rain diagram has a similar concave shape, but transposed by one month. Due to these irregularities, it is complicated to establish a seasonal pattern for the precipitation series. Seasonal regularity is lacking, which excludes the possibility of establishing a monthly year pattern that can be used for temporal reduction surveys.

Salinity series

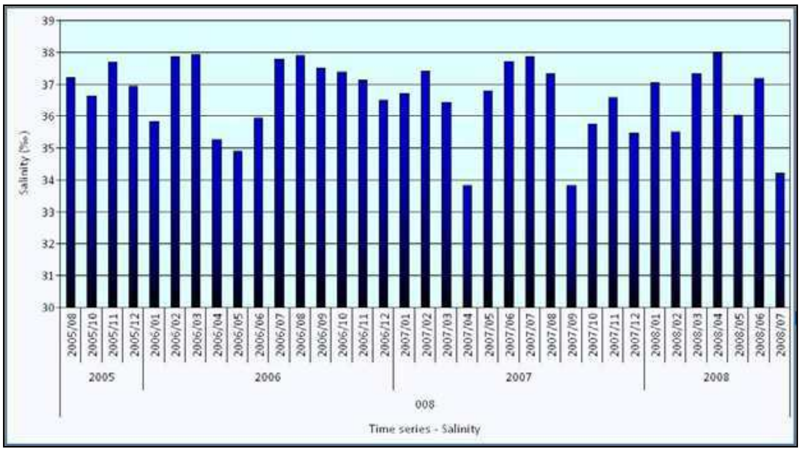


Figure 12: Campaign average salinity values; water body 008

The salinity variation has a dynamic that can be expected for a water body of type II-A. Regularity or a seasonal pattern is hard to identify for parameters strongly influenced by precipitations and sea storms, such as those described above. The lowest salinity values were detected during April and September 2007 (salinity value of about 33,80‰), July 2008 (34,10‰) and during the (35,10‰) and May (34,90‰) of 2006. Values indicating the absence of fresh water influence (equal or slightly lower than 38‰) were detected during February, March, July and August 2006, July 2007 and April 2008.
The values of temporal series show that regularity and seasonality is absent for salinity concentrations and that no pattern can be established.

Chlorophyll a series

Finally, after analyzing the temporal series of sea storm events, precipitation and salinity values, chlorophyll a is presented as the “sum” (as for the 1st maritime clime) of the above parameters which influence biomass and determine the condition of the coastal ecosystem.

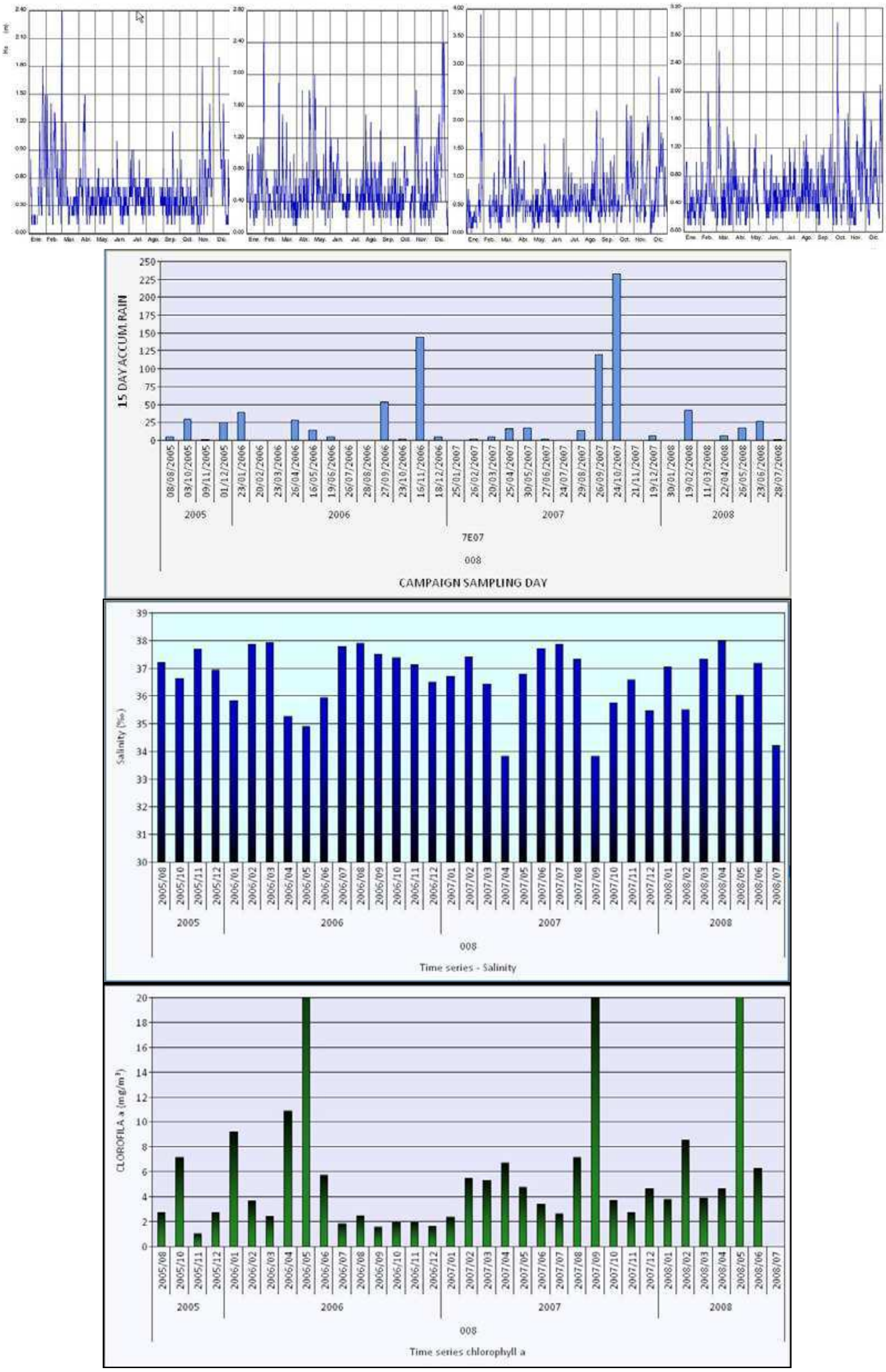


Figure 13: WANA 2046035 time series of significant wave heights; 15-day accumulated rainfall (sampling day); campaign average salinity values; campaign average chlorophyll a values w.b.008

Performing the same analysis for water body 008, the (anti)reciprocity between chlorophyll *a* and salinity is clearly visible from the diagrams, and high chlorophyll *a* values are accompanied by decreased salinity concentrations.
Observing the accumulated rainfall values, it seems that salinity concentration values do not respond as expected. The lowest salinity values are detected in April and May of 2006, April and September of 2007 and July of 2008. The low salinity values of September (>34,0‰) and October(<36‰) 2007 correspond to rain events detected that exceeded $100l/m^2$ and $200l/m^2$ of 15 day accumulation for the sampling date. April and May of 2006 and April of 2007 have low salinity values (near to and greater than 35‰) but the detected precipitation for campaign months are low and accumulated rainfall for the campaign date is less than $10l/m^2$.
The discordance between salinity values and precipitation, as explained for water body 001, is due to the mixing processes of sea and freshwater accelerated by sea storms, as well as anthropogenic influence in the area. Water body 008 is located in an area highly influenced by hypereutrophic Albufera lagoon, and this influence is determined by three water gates that pump fresh water depending on irrigation processes and rice crop agricultural activities in the area. As was mentioned before, Albufera is a hypereutrophic lagoon and the continental loads have high concentrations of chlorophyll a and the water samples occasionally detect direct influence values, as during May 2006, September 2007 and May 2008 (exceeding 20µg/l). The elevated chlorophyll a values are accompanied by low salinity, which indicates a mixing process of continental loads and sea water.
Because of this, low salinity values can be detected during periods of low precipitation such as April and May of 2006 (15 days accumulated rainfall with less than $25l/m^2$) when the corresponding salinity values are low and indicate fresh water mixing with average values 35,1‰ and 34,9‰.
After an analysis of environmental parameters, the links between them and the identified significant anthropogenic influence in the area, it can be deduced that regularity of phytoplankton biomass and corresponding chlorophyll *a* values exists neither on a monthly nor even a seasonal basis.

3rd Maritime climate regime

Water bodies south of Cape Saint Antonio are identified as belonging to the 3rd maritime regime of Valencian coastal waters. This maritime regime is determined by a northeast – southwest coastline. The fetch is open from the northeast to the southwest with the Balearic Islands preventing waves proceeding from a straight angle - about 30º.

Wave series

Figure 14: Wave roses; annual evolution for 2005, 2006, 2007 and 2008; data from point WANA 2049029

During 2005 the southern, southeastern and northeastern components were almost equal (about 20% of the time) in the amount of time they affected the coast line, with the eastern factor being the most dominant (25%). The maximum value of significant wave height (exceeding 1.8m) was detected in early March proceeding from the northeast. The presented significant time series indicate that significant sea storms were absent during 2005. After this, a series of moderate sea storms began in the second part of January that ended by the beginning of March. After March, a “calm period” began lasting until December with a few isolated events (April, July, August and September) that exceeded 1m of significant wave height.

In 2006 the number of sea storms increased, as did the levels of the significant wave height. The eastern component was dominant (as during 2005) with an increased 32% of time affecting the coastline. The northeastern component also increased in value (26%) compared with 2005. The southern component increased in waves maximums, but the percentage of time affecting the coast (about 18%) is more or less the same as in 2005. Wave maximum was detected proceeding from the northeast with a 3.4m significant height. During 2006, various sea storms were detected that exceeded 2m of significant wave height: two in April, one in May, in early November and in late December. During 2006, sea storms were more frequent and a long “calm period” was absent.

In 2007, the northeastern component affected the coastline about 40% of the time and the eastern component 25% of the time. The southern and southwestern component stood at about 12% with waves with lower significant height. The maximum wave significant height was detected proceeding from the same direction as in 2005 and 2006 (northeast) but without a similar value. For this year, it is characteristic that two significant components (northeastern and eastern) affect the coast more than 60% of the time and that the number of sea storms is significantly higher. In 2007, three significant events were detected in March (with a yearly maximum) and events that exceeded 2.0m were detected in late January, in August, October, November and December. As was noted for 2006, a “calm period” does not exist and during all months at least one event exceeding 1.2 m of significant wave height was detected.

For 2008, the North-East is a principal time component (29%), with similar values for waves proceeding from the east (24%) as was detected for 2007. During 2008, the buoy detected significantly increased wave heights for the southwestern component that affected the coast line 15% of the year. The southern component affected the coast the same percentage of time of the year as in 2007. During 2008, a significant wave height maximum was detected for the four year period (4.0m) proceeding from the northeast. Significant events whose wave height exceeded 2m were detected in February, two events in March, two events in April, more then a ten day long series in October, November and two events in December*. The values between April and October do not indicate that a “calm period” existed in 2008.*

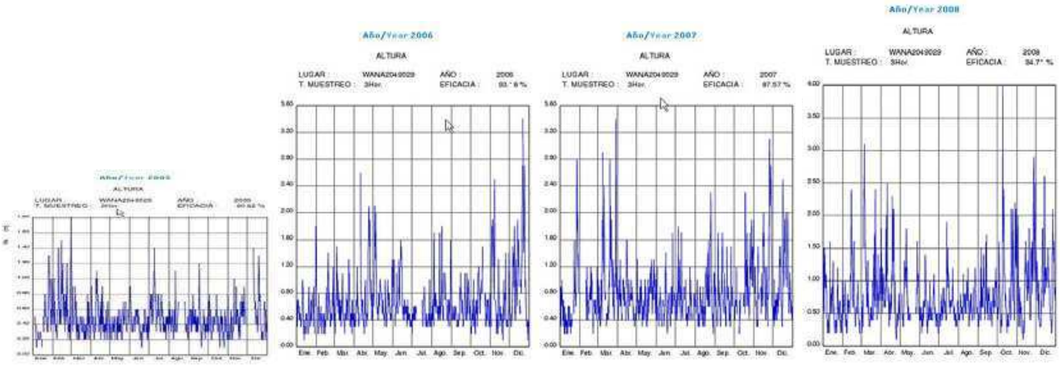


Figure 15:WANA 2049029 - 4 annual evolution 2005,2006, 2007 and 2008; time series significant wave heights

Observing the serial data for the 2005-2008 period, it is complicated to establish a pattern based upon monthly regularity. Some kind of seasonal pattern exists for the year-long periods of 2005, 2007 and 2008. It starts with a calm first half of January that is followed by a period of frequent sea storms that ends during the spring period (beginning in March in 2005, April in 2007 and May in 2008). In 2006, only one event exceeding 1.5m of significant wave height was detected before the second half of April (in late January).

The “calm period” is not as stratified as the stormy winter/spring interval. As was mentioned above, it lasted from March until December in 2005, could not be identified in 2006, lasted from April until the end of August with few sea storms during July 2007, and from May to October during 2008.

A period of frequent sea storm events was detected at the end of the year for all four years. This sea storm period started in December in 2005, in early October in 2006, in late August in 2007 and in October in 2008.

From the presented data, it is clear that monthly regularity does not exist and that a weak seasonal pattern can be established only for 2005, 2007 and 2008. No kind of pattern can be established for the four year data which could be used to reduce survey frequency based on wave seasonality.

Precipitation series

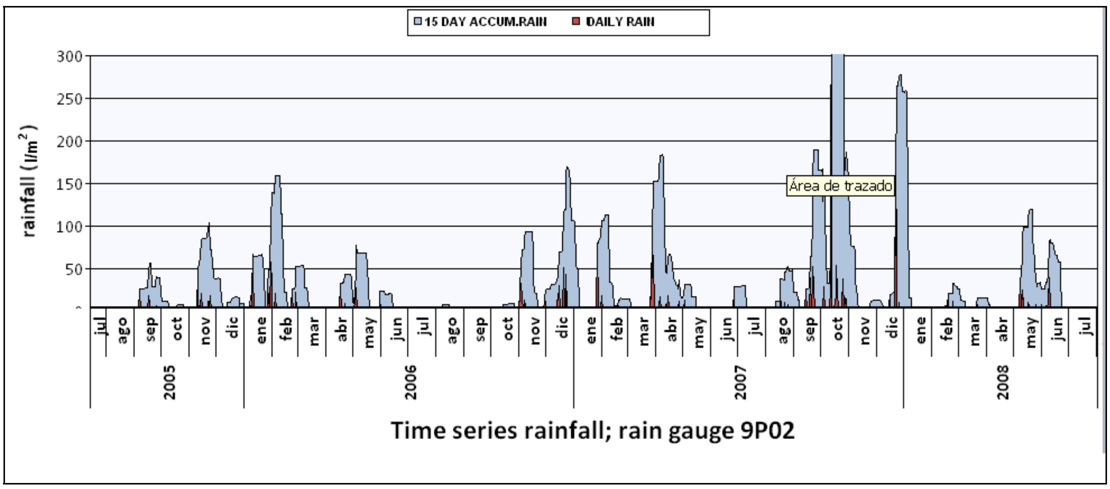


Figure 16: 15-day accumulated rainfall and daily rainfall; rain gauge 9P02 - corresponding to water body 013

The precipitation series that corresponds to water body 013 shows similarity to water body 008. Rain events have the same frequency but with significant differences in the intensity and 15 day accumulation values. October of 2007 was the month that exceeded 1000l/m^2 of 15 day accumulated rain while for the same event the pluviometric station corresponding to water body 008 detected less than 250l/m^2 .
Using the data series, it is possible to establish similar seasonality as was determined for water body 008.

Rain events were almost absent during the summer, especially July and August, throughout all four years. June of 2006 and 2007 had a low precipitation index, but in June 2008 rain events were detected that exceeded 80l/m^2 of 15 day accumulation.
As mentioned previously, rain events with the highest intensity and accumulation values were detected during autumn of 2007 in September and October. During 2005 and 2006, the same months had a low precipitation index (rain events were only detected during September 2005 with accumulation of around 50l/m^2).

No significant rain was detected in December of 2005, but January 2006 saw the beginning of rain events frequently exceeding 50l/m^2 of 15 day accumulation, lasting until the end of February. The winter of 2006/2007 saw frequent rain events during December and February (values exceeding 100l/m^2 of accumulation), but a nearly one-month period without rain was detected in January. In contrast, February and January of 2008 were months with low precipitation values and during December 2007 events were detected that exceeded 250l/m^2 of accumulation.

In spring of 2007, rain was frequent throughout March and April and 15 day accumulation exceeded 150l/m^2 for various events. During 2006 the rain frequency for March and April is the same, however the intensity is significantly lower. During the same months in 2008, precipitation was not detected, but during May and June of 2008 – rainless months in 2006 and 2007 – contrastive rain events were detected for this part of year.

As with water body 008, the concave shape of the yearly precipitation diagram exists for 2006 and 2007 for water body 013, which can demonstrate some kind of seasonality, but which differs in the intensity, frequency and month when rain events are detected. The seasonality pattern that is detected for 2006 and 2007 cannot be applied to 2008. Considering all data and the seasonality identified that cannot be applied to the whole data series, it can be concluded that regularity on a seasonal or monthly basis for rain events cannot be established.

Salinity series

Average temporal salinity values for water body 013 have a low variation rank, as should be expected for a water body of type III (low continental influence). The minimal values are above 36,0‰, and five values around 38,0‰ are detected in the temporal series.
Observing the graph, it is difficult to detect any kind of significant regularity, even on a seasonal basis. Until September 2007, the values are mainly in the rank between 37-38‰ with four months when salinity values below 37‰ were detected (November 2005, May and September 2006 and April 2007). Due to significant precipitation events during the

autumn and winter of 2007/2008 decreased salinity concentrations were detected in the water samples (between 36-37‰) until March 2008.

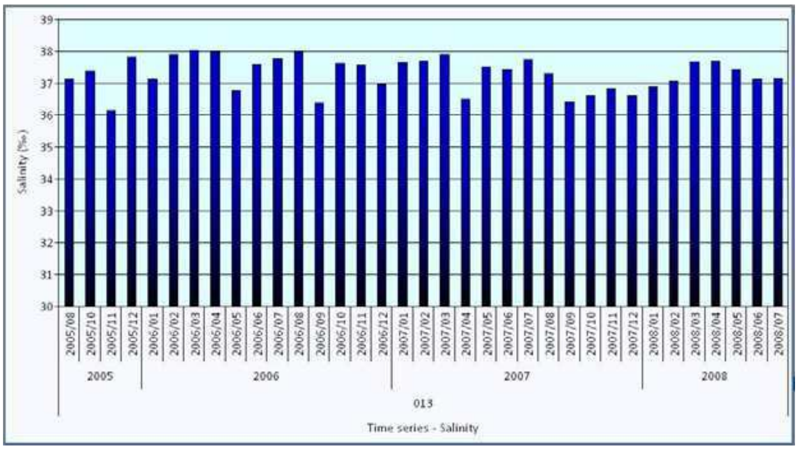


Figure 17: Campaign average salinity values; water body 013

<u>Chlorophyll</u> <u>a</u> *<u>series</u>*

Chlorophyll a series are analyzed as was chlorophyll a for water bodies 001 and 008. Trying to find a pattern or regularity for chlorophyll a includes analysis of: chlorophyll *a* temporal series, salinity values, environmental parameters that affect biomass such as sea waves and rainfall.

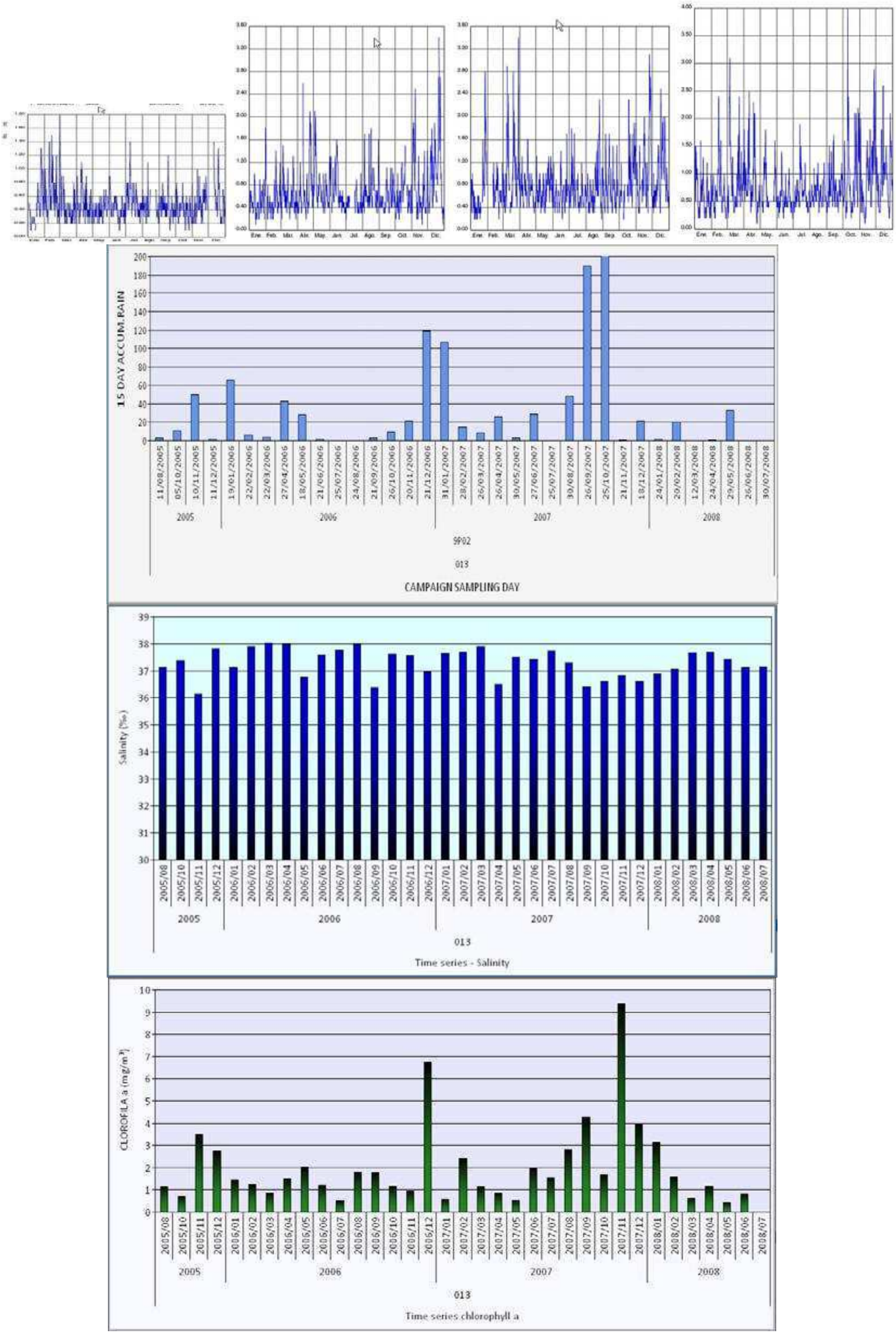


Figure 18: WANA 2049029 time series significant wave height; 15-day accumulated rainfall (sampling day); campaign average salinity values; campaign average chlorophyll a values w.b.013

As mentioned previously, water body 013 is identified as type III - coastal water bodies not affected by fresh water inputs. The definition of type III claims that water bodies which are declared as type III are not affected (significantly) by continental waters whose influence (flow and volume) depends on precipitation. Observing the salinity values for water body 013, it is clear that variance is low (most values between 37,0-38,0) as should be expected, but during a detected period of significant rain **(**September, October and December of 2007) salinity values vary in the range of 36,0-37,0‰ due to increased continental influence. For this period increased chlorophyll a values were detected due to nutrient enrichment that increased the phytoplankton biomass.

The chlorophyll a maximum value for the autumn/winter period of 2007/2008 exceeding a concentration of 9,0 µg/l was detected during November, a month during which no important rain events were detected (without continental influence and nutrient loads). The November phytoplankton "bloom" could be explained by elevated concentrations of a limiting nutrient (significant values of phosphorus, the source of which is difficult to detect) that is likely anthropic in character.
However, during November 2007 various sea storms were detected (one exceeding 3,00 m of significant wave height) that might indicate the unsettling of thermocline stability, which can cause a kind of "upwelling" and nutrient enrichment of superficial layers. This mixing effect due to the vertical flow of nutrient rich deep water can result in a biomass increment on the surface and can be detected through increased chlorophyll a values. This process of vertical mixing was studied for the Mediterranean Sea on the large scale by the Bakun and Agostini at al.2001.

A phytoplankton biomass increment was also detected during December 2006 with values that exceed 6,0 µg/l. Salinity values are around 37,0‰, which might indicate continental influence corresponding to detected rain events (15 day accumulation on the sampling date reaching 120l/m^2). Observing the wave series for this period, an increasing frequency of sea storms was detected during December with increasing values of significant wave height peaking at 3,5m. Like November 2007, this kind of sea storm can break through the thin thermocline layer and unsettle deep water stability, which can result in superficial nutrient enrichment.
Both elevated chlorophyll a values (December 2006 and November 2007) are events that could be caused by increased continental influence or the "upwelling" effect described above. Neither one of these causes exclude the other, and there is a high probability that the increased biomass is a result of a combination of two sources of nutrients; continental loads and vertical mixing of the sea column.

During November 2005, a significant chlorophyll a value increase was also detected (about 3,5 µg/l) that corresponds to the lowest salinity value in the series. Salinity values correspond with rain events (accumulated 50l/m^2) and the sea buoy did not detect significant wave events.
Perhaps these differences in hydrodynamic and continental inputs with lower nutrient concentrations produce the reduced chlorophyll a response.

Increased chlorophyll a values are strongly connected to marine hydrodynamic conditions and precipitation, which possess some kind of qualitative seasonal regularity. The chlorophyll a series diagram for the three-year period indicates three wintertime increases in biomass. This qualitative seasonal regularity does not include frequency or intensity and directly depends on the superposition of hydrodynamic conditions and nutrient loads.

4.2.2.Discussion

Coastal ecosystems are complex (as is presented above) and the level of eutrophic balance depends on anthropogenic influence and environmental factors that are determined by clime seasonality and irregularity. Aside from this, precipitation and wave data sometimes confirms seasonality, but not regularity. The clime has a generally seasonal pattern during the year (rain and sea storms) but not monthly regularity that could make it possible to know when it is going to rain and which months are going to have significant precipitation or sea storm events.

Significant rain events can be expected in the autumn, winter and spring, but it is impossible to establish a pattern using a reduced survey frequency that would cover most important precipitation periods.

Like precipitation, sea storm events are seasonal but not regular (on a monthly basis) and are not always accompanied by significant rain events. Sea storms as analyzed and described above can have different effects on biomass growth:

- Waves can accelerate the mixing process of fresh water and superficial sea layers pulled from the open sea. This process can reduce nutrient concentration during periods of significant continental loads and biomass stability will not be disturbed as it would be during calm sea periods.
- During sea storms with elevated wave heights, vertical mixing of the sea column can occur. Deep nutrient-rich sea water can become mixed with superficial sea layers because of storm hydrodynamic conditions. The superficial layers are enriched with nutrients from the deep by the "upwelling", which can accelerate phytoplankton biomass growth.

Even with the possibility to establish an annual scenario of sea storm events, it would be complicated to predict in which way biomass would be affected: accelerated mixing or "upwelling" (Bakun and Agostini at al. 2001 presented large scale Mediterranean Sea seasonal patterns for the upwelling processes that can not be used on the monthly basis).

Coastal water salinity is strongly affected by the above described environmental processes and its temporal series values sometimes do not present stable seasonality. Salinity concentration can be used as an indicator that provides information about possible nutrient enrichment or the current oligotrophic conditions of coastal water ecosystems. Low salinity concentrations indicate fresh water influence, but it should be considered a qualitative and not quantitative value. The main reason is the dynamic of coastal waters which provides completely unhomogenic conditions (due to the above mentioned hydrodynamic and other environmental components), especially during freshwater plume mixing. A single salinity value can only provide information about continental influence (which is usually accompanied by nutrient enrichment), but not the nutrient concentration or volume of fresh water that affects coastal ecosystems.

Salinity is a good indicator of nutrient enrichment (except in the case of "upwelling"), with a significant reciprocal link to chlorophyll a values. Finally, it is possible to conclude that phytoplankton biomass, replaced with chlorophyll a values, presents a slight tendency towards seasonality due to the influence of environmental factors.

A global coastal survey reduction to any frequency (bimonthly or trimestral) provides unpredicted results, mostly because of climatic irregularity that can lead to erroneous conclusions by omitting periods of accelerated phytoplankton biomass growth or periods of stable biomass. Finally, actual climate change complicates the task of locating regimes of environmental parameters (such as rain and sea storms) that could potentially be used to establish eutrophic temporal scenarios.

4.3.Ecological classification fuzziness

The previous analysis of temporal reduction demonstrates unpredicted consequences to the ecological classification results due to the irregularity of environmental factors (precipitation and hydrodynamic conditions) and anthropogenic influence. Results of the previous analysis have been parsed to see if the reduction in the frequency of campaigns has the same unpredicted consequences on the establishment of ecological status for each water body.
20 ecological statuses are established (en total) for each water body, and the campaign reduction method is performed on 4 different periods with five reductions in the number of annual campaigns using the SPSS algorithms. The results in percentage of unchanged water status are represented in the graph (Figure 1):

Figure 1: Water body result correspondence (%) of reduction analysis performed by SPSS (20 ecological statuses per w.b.)

All water bodies are represented in the graph by a bar constituting 100%, which consists of a red and green part. The green part represents the percentage of corresponding ecological classes (with the original data set) while the red part of the bar represents the percentage of non-corresponding classes. Observing the results, 100% of the ecological classes established using the reduced data sets in water body 003 correspond. Water bodies 001, 002, 004, 008, 014, 015, 016 and 017 have results that approximately meet or exceed 80% unchanged classes. These results clearly demonstrate that decreasing the annual frequency of campaigns is possible in some water bodies, whose results show a considerable number of corresponding ecological statuses.
The same analysis as before is carried out by EXCEL algorithm, and the results separated water bodies 001, 002, 003 and 014 corresponds 100% with original data set. The 90[th] percentile values calculated by EXCEL are lower, because of which all classifications (for three and one year periods) are High ecological statuses that correspond to the original ones.

The water bodies represented in the table below are type II-A (coastal waters influenced by fresh water inputs) and type III (coastal waters not influenced by fresh water inputs). To consider phytoplankton dynamics for each water body, variance and average value of chlorophyll a for a three year period are calculated. Continental influence for each water body is represented with variance and average value of salinity concentration for the same period. The next column of the Table 13 represents chlorophyll *a* distance values (buffer distance) to the next ecological status. These distance values are calculated on the basis of water body chlorophyll *a* average values and the numerical differences from the nearest limit of upgraded or degraded ecological classes. Finally the last column represents the average error 90th percentile bias values compared to the quantitative results calculated using complete data

water body	w.b. type	average value Salinity(g/Kg)	variance Salinity	average value-chl a(mg/m³)	variance-chl a	"buffer" distance average value, (mg/m³)	average error bias value(mg/m³)
001	II-A	36,55	3,5	1,97	1,53	2,83	0,76
002	II-A	36,57	1,58	1,93	1,72	2,87	0,56
003	II-A	37,24	0,71	1,31	0,45	3,49	0,47
004	II-A	36,69	2,62	2,1	2,76	2,70	0,62
005	II-A	36,02	6,71	2,83	3,79	1,97	0,8
006	II-A	37,41	0,8	2,85	5,8	1,98	1,33
007	II-A	36,64	7,26	3,25	5,19	1,55	0,99
008	II-A	36,77	3,16	4,51	36,89	0,29	3,22
009	II-A	36,66	10,52	3,72	10,25	1.08	1,57
010	II-A	36,86	1,52	2,61	5,69	2,19	0,93
011	III	37,56	0,64	1,11	1,03	1,09	0,54
012	III	37,45	0,47	1,3	2,75	0,9	1,37
013	III	37,37	0,44	1,41	3,11	0,79	1,74
014	III	37,58	0,24	1,02	0,28	1,18	0,30
015	III	37,48	0,2	2,16	3,8	0,04	1,54
016	III	37,45	0,32	2,79	4,66	0,59	0,86
017	III	37,56	1,26	2,98	24,92	0,62	3,47
018	III	37,59	0,26	1,79	2,03	0,41	1,01
019	III	37,49	0,21	1,45	2,41	0,75	1,18

Table 13: W. bodies average and variance values of salinity and chlorophyll a; "buffer" distance value of each w.b.to the nearest threshold

Instability in the ecological classification of a water body occurs when the 90th percentile is located near the border that separates the ecological status High / Good or border that separates the ecological status Good / Moderate. These instabilities are not present when chlorophyll *a* levels are sufficiently high to ensure that the water body belongs unequivocally to Moderate ecological status. On this basis, we propose a methodology to locate water bodies where it would be potential to propose a reduction in the frequency of the campaigns, based on the sample mean and standard deviation of the observed chlorophyll a values.

w.b average chl a. value + w.b. standard deviation value > Good/Moderate threshold

Defined (Moderate) ecological classification

In the event that the sum of average and standard deviation of chlorophyll a does not exceed the Good /Moderate threshold and, assuming that observed values come from a normally distributed variable, if the product of the sample standard deviation by the 90th percentile of a standard normal (z_{90} = 1.28) does not exceed the buffer distance average value we consider the classification is defined and it is reasonable to propose a temporal reduction in the campaigns.

(w.b. standard deviation chl a value) X (z_{90} = 1.28) < "buffer" distance average value

Defined (High) ecological classification

In type II-A, water bodies 001, 002, 003 and 004 meet this requirement and the percentages of equivalent classifications meet or exceed 80%. In type III, the water body 014 meets this requirement and the percentage of equivalent classifications also exceeds 80% equivalence

In the event that the sum of the average and standard deviation of chlorophyll a exceed the Good / Moderate threshold, we consider the classification is defined, regardless of the variability of the variable. In type II-A, water body 008 is in this situation and in type III can be found water body 015, 016 and 017. The percentages of equivalent classifications for all of them also meet or exceed 80%.

4.3.2.Water bodies selected by previous results and reduction conditions

At this part will be analyzed water bodies that in the exploratory study have correspondence more than 80% of ecological classifications with the classifications calculated using the original data. Moreover analyzed water bodies fulfill one of the reduction conditions based on the standard deviation, buffer values or on the sum of an average and standard deviation values.

Water body 001

Water body 001, declared as type II-A, has a high variance of salinity as expected (3,50), which confirms high continental influence and the inability of the system to buffer it. Aside from this, water body 001 is a stable ecosystem that can sustain pressure and can maintain a relatively stable phytoplankton biomass. The variance of chlorophyll *a* (1,53) is significantly high but most values are below the High/Good threshold (Figure 2).

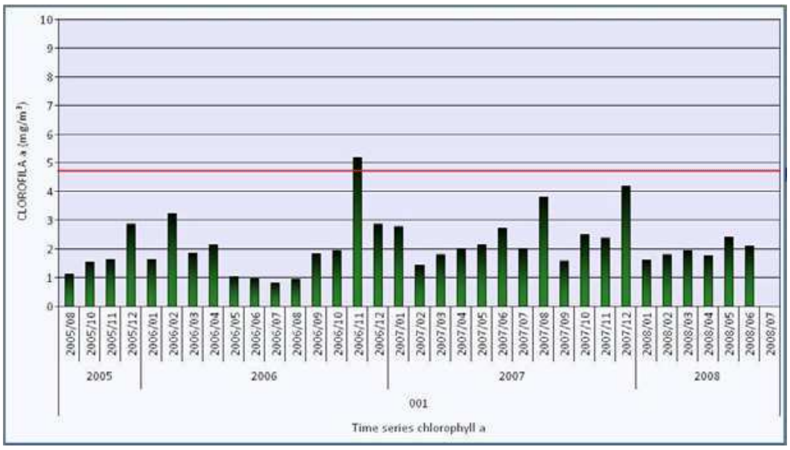


Figure 2: 001 w.b. time series of chlorophyll *a* monthly campaign average values

The indicator value (1.59) is less than the buffer distance (2.83) and the campaign reduction frequency is possible due to the significant distance "buffer" of quantitative results to the High/Good threshold (Table 13). The ecological classification corresponds irrespective of 90th percentile bias (average error 0,76 µg/l; Table 13).
Regarding the quantile plots, it is recommended to substitute monthly campaigns with trimestral survey frequency. Observing the three year quantile plots, it is clear that the trimestral curves have less bias than curves representing bimonthly campaigns, especially the part that affects the 90th percentile, but due to the stability of the ecosystem a bimonthly survey can be considered as a second option.

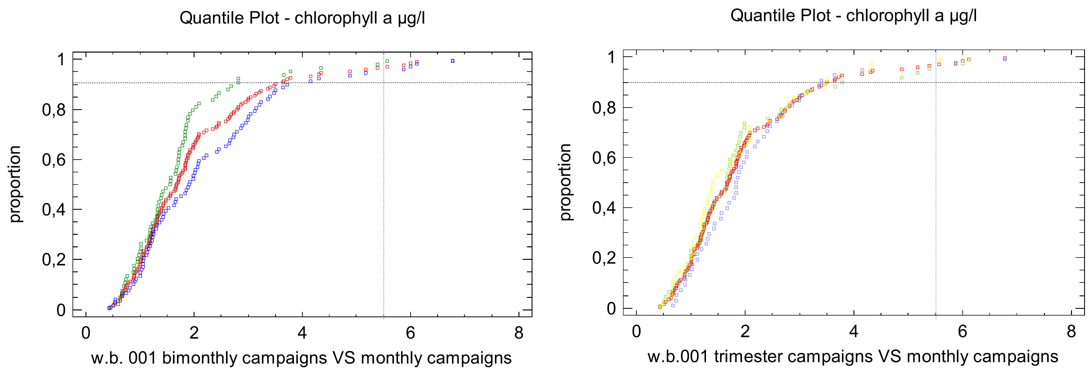


Figure 3: W.b. 001 quantile plots; bimonthly campaigns (green and blue curves) VS monthly campaigns (red curve); trimestral campaigns (green, blue and yellow curves) VS monthly campaigns (red curve); black line-90th percentile and blue line High/Good threshold

Water body 002

Water body (002) represents a stable ecosystem, probably with lowest anthropogenic pressure at the Valencian Coastal waters. As water body type II-A, it has significant continental influence (mostly from submarine ground water discharge) with a salinity variance of 1,58, low chlorophyll *a* values, but also significant variance of chlorophyll a (1,72). The biomass is stable and chlorophyll *a* values vary due to continental influence, but always within the limits of the high ecological class (Figure 4).

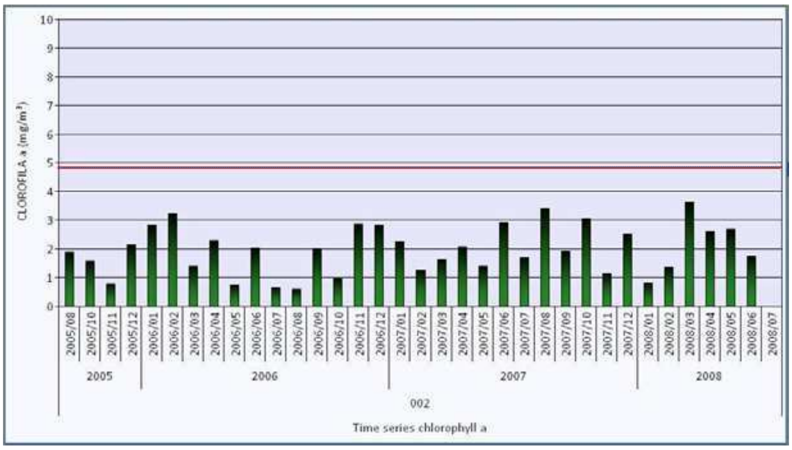


Figure 4: 002 w.b. time series of chlorophyll *a* monthly campaign average values

The indicator value (1.67) is less than the buffer distance (2.87) and as was noted for all water bodies that check this relation, reduced survey frequency is possible in bimonthly or trimestral campaigns without consequences to the ecological classification status (Figure 5).

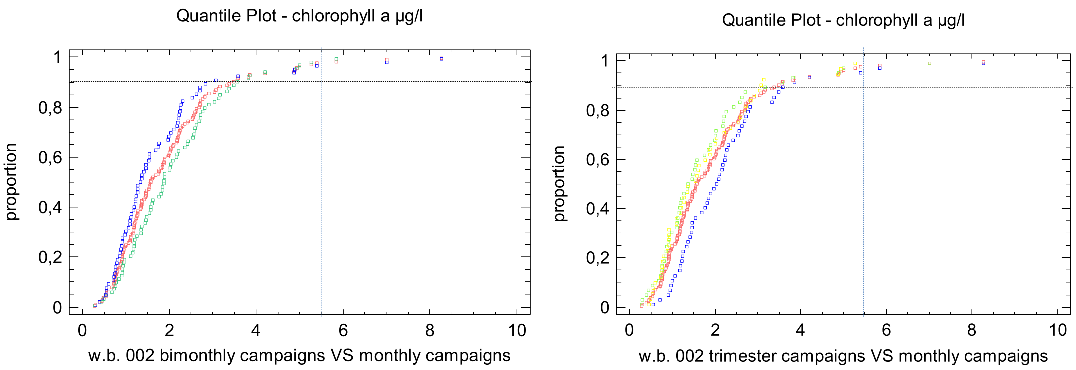


Figure 5: W.b. 002 quantile plots; bimonthly campaigns (green and blue curves) VS monthly campaigns (red curve); trimestral campaigns (green, blue and yellow curves) VS monthly campaigns (red curve); black line-90th percentile and blue line High/Good threshold

Water Body 003

All 20 analyses (Figure 1) for water body 003 confirm that there are no consequences to the ecological classification status due to the bimonthly or trimestral reduction. Continental influence (salinity variance 0,71; Table 13) is low and does not exceed the described hydrodynamic system's capacity to buffer continental pressure. That capacity of absorption of pressure is mainly due to the morphology and hydrodynamic characteristics of the water body. The littoral zone seabed gradient is elevated and the area has the capacity to buffer the continental influence due to the currents and wave mixing processes. The mixing process that occurs throughout the whole water column amplify the dispersion/dilution effect of nutrients and decelerate phytoplankton biomass growth. This results in stable water body phytoplankton biomass (chlorophyll a variance 0,45) and none of the chlorophyll a values exceed the High/Good threshold .

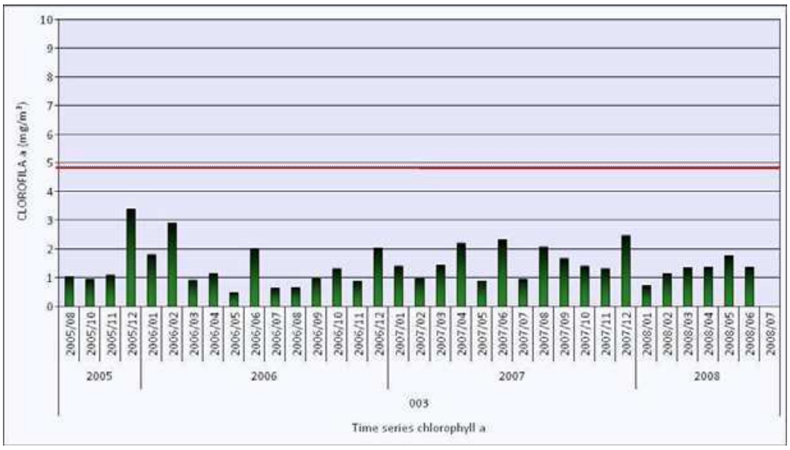
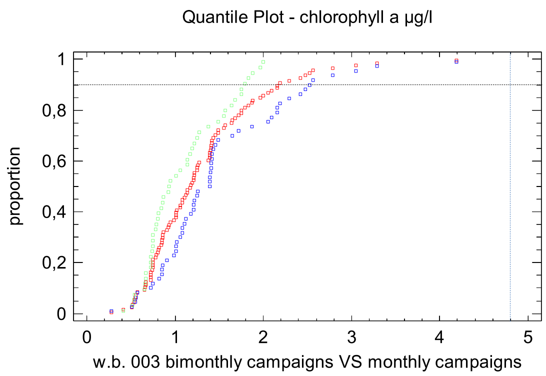


Figure 6: 003 w.b. time series of chlorophyll a monthly campaign average values

If we compare the n-tile curves of reduced data (Figure 7; green, blue and yellow) with the original data set (red), it is clear there is no discrepancy between quality results provided by reduced frequency campaigns and the monthly survey. But, if we compare the quantity results, it appears that trimestral campaign curves have a lower bias in the zone of the 90th percentile and that the calculated results are slightly better than results provided from denser campaign frequency. Beyond the quantitative biases the qualitative results for all reduced data sets are the same due to reduced indicator value (0,85) and the extended "buffer" distances (3,49).



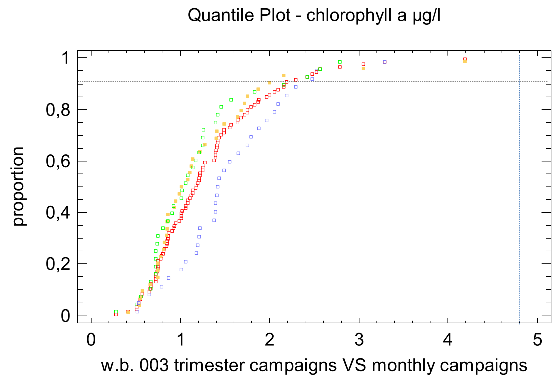


Figure 7: W.b. 003 quantile plots; bimonthly campaigns (green and blue curves) VS monthly campaigns (red curve); trimestral campaigns (green, blue and yellow curves) VS monthly campaigns (red curve);black line-90th percentile and blue line High/Good threshold

Water Body 004

Water body 004, declared as type II-A, has a high variance of salinity as expected (2,62), which confirms high continental influence. Beyond that, water body 004 is a stable ecosystem that can sustain pressure and can maintain a relatively stable phytoplankton biomass. The variance of chlorophyll *a* (2,76) is significantly higher in respect to beyond analyzed water bodies 001,002 and 003, but still the indicator values are mostly below the High/Good threshold (Figure 8).

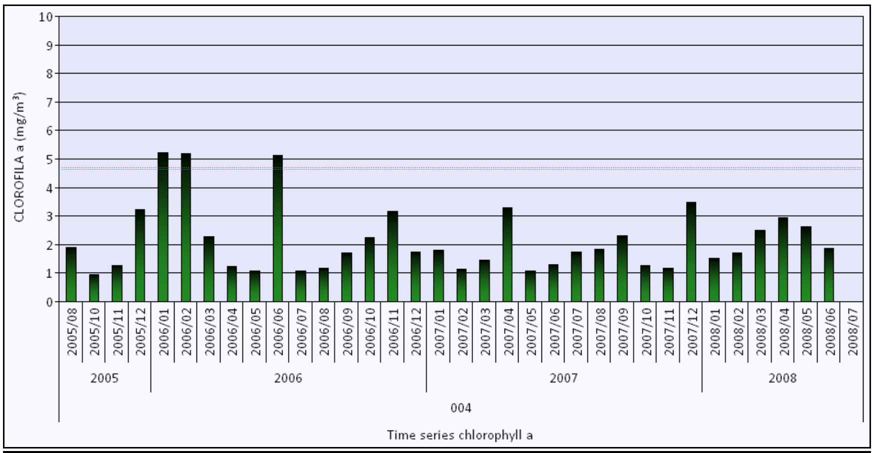


Figure 8: 004 w.b. time series of chlorophyll a monthly campaign average values

The indicator value (2.12) is less than the buffer distance (2.70) and the campaign reduction frequency is possible due to the significant distance "buffer" of quantitative results to the High/Good threshold (Table 13). Reduced survey frequency is possible in bimonthly or trimestral campaigns without consequences to the ecological classification status (Figure 9). The reduction decision for the water body 004 has a more risk (than for the above analyzed water bodies) due the lower "buffer" value to the High/Good threshold (Figure 9).

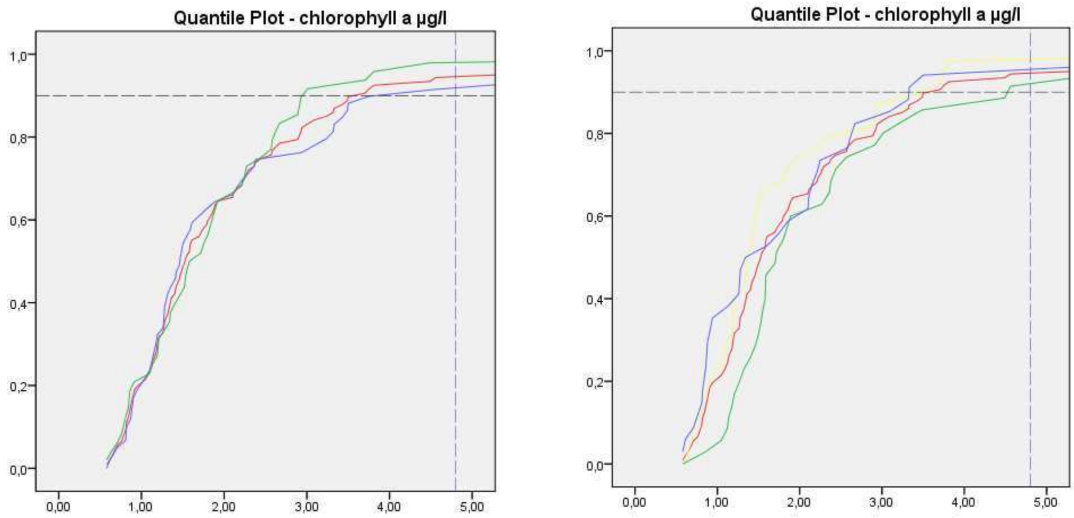


Figure 9: W.b. 004 quantile plots; bimonthly campaigns (green and blue curves) VS monthly campaigns (red curve); trimestral campaigns (green, blue and yellow curves) VS monthly campaigns (red curve);black line-90th percentile and blue line High/Good threshold

Water body 008

Water body 008 represents the most unstable ecosystem for the coastal waters of Valencia with a variance of chlorophyll a 36,89. An expected, but not very high, salinity variance (3,16) confirms continental influence, likely with an important anthropic component that pressures the phytoplankton biomass. On the diagram below (Figure 10) it is possible to observe the variation of chlorophyll *a* campaign average values that are mostly due to the influence of Albufera hypereutrophic lagoon, nutrient-rich, characterized by frequent and severe algal blooms. The Albufera continental loads, as nutrient rich fresh water with an extremely high concentration of chlorophyll a, could be the cause of direct influence sampling. The water samples for this area can detect extremely high chlorophyll a concentrations which do not represent the coastal waters' phytoplankton biomass response to continental pressure. The elevated chlorophyll a concentrations might reflect the water quality of the freshwater inputs and do not indicate the real trophic condition of the water body.

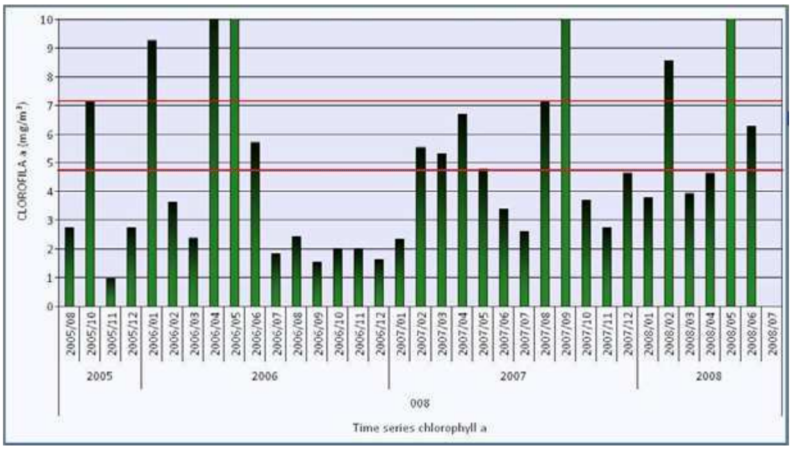


Figure 10: 008 w.b. time series of chlorophyll a monthly campaign average values

Reducing campaigns to bimonthly or trimestral frequency results in 10% (SPSS algorithm) discordant classifications compared to the original (which is unexpectedly low for this kind of unstable eco-system) – this is caused by the dense frequency of chlorophyll *a* values that exceed the Good/Moderate threshold (Figure 10).

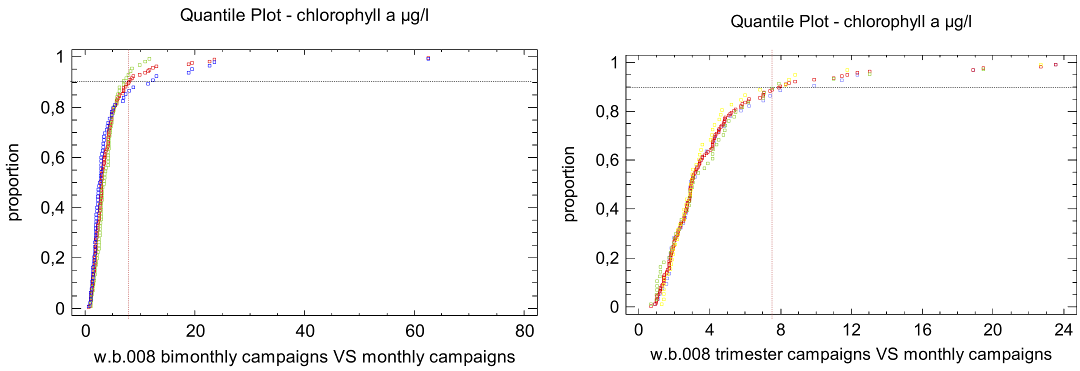


Figure 11: W.b. 008 quantile plots; bimonthly campaigns (green and blue curves) VS monthly campaigns (red curve); trimestral campaigns (green, blue and yellow curves) VS monthly campaigns (red curve); black line-90th percentile and red line Good/Moderate threshold; observing the graphics is important to pay attention on the chlorophyll a scale values

The sum of average and standard deviation of chlorophyll a observed values is 10,58 µg/l which is well above the Good/Moderate threshold (7,20 µg/l), This should indicate that it is expected that approximately 16% of the values are above this value and, consequently, it is difficult to alter the ecological classification results due to a reduction in the frequency of campaigns. It should be noted that has been detected in the data collected in this body of water the presence of an outlier, which is not surprising given the high variability of chlorophyll in this area. However, this value can significantly alter the results. Repeating the calculation of the average and standard deviation, excluding this value, the statistical sum of these is reduced to 7.87, indicating that the value will be exceeded by 16% of the data is the latter. This value is significantly closer to the boundary Good / Moderate, which suggests the possibility of some alteration in the ecological classification of the area when it reduces the frequency of campaigns.

It is possible through Figure 11 to observe that qualitative results are changed for the even-month campaigns simulated for the three year period (in Figure 11 the green n-tile line is below the Good/Moderate threshold due to the significant curve dispersion for the n-tile values >0,8). The trimestral curves better correspond to the original data than bimonthly campaigns and the qualitative results remain unchanged.
Since water body 008 (for the three year period) has percentile values around the threshold, reducing the frequency of sampling could have significant consequences for the ecological classification. Although there were good correspondence results for all 20 reduction analyses (90%), campaigns should have a monthly frequency.

Water body 014

The water body 014 has the lowest value of variance of chlorophyll *a* (0,28, Table 13) of the entire coast of the Valencian Community, which confirms that phytoplankton biomass is stable. The salinity variance is a low 0,24 (which is to be expected of a water body without continental influence -type III), as is the average bias of 90th percentile calculated using the reduced data set (0,30 µg/l). However, analysis provided by SPSS calculated “Good” instead of “High” ecological classification status 3 times for the one year period (the same analysis performed by Excel with 100% correspondence). This is essentially because the water body is classified as type III and the limit of High/Good is drastically lower (2,2 µg/l) than for type II-A . All three changed classifications are calculated for the one year period 2006/2007, where reduced data sets number less than 25 samples for the bimonthly sets and about 15 samples for the trimestral reduction data sets. On the graph below, two peaks of chlorophyll *a* are visible for October and December of 2006 (mainly caused by rain storm events) that produced non-corresponding classification results for the even months and the second and third trimestral data of the one year period from August 2006/July 2007.

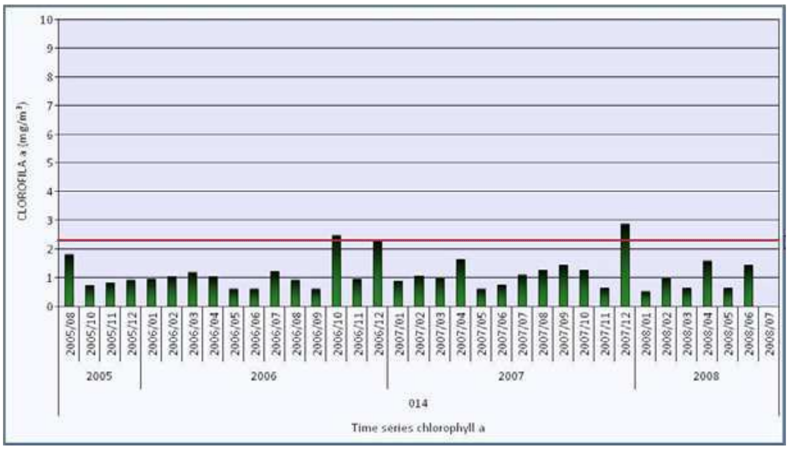


Figure 12: 014 w.b. time series of chlorophyll a monthly campaign average values

However, water body 014 still is a ecosystem with stability in the ecological classification, because the indicator value (0,68) is less than the buffer distance (1,18), , and reducing the monthly campaigns to bimonthly or trimestral frequency will not have any consequences for ecological classification for a period longer than one year.

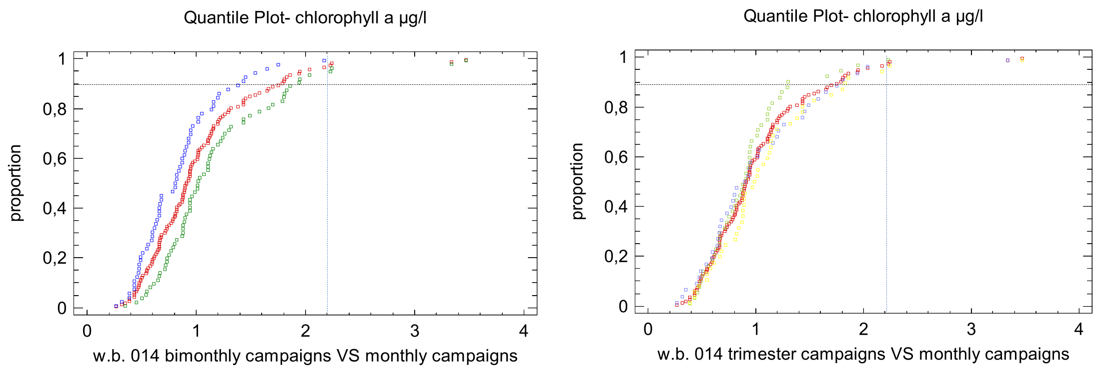


Figure 13: W.b. 014 quantile plots; bimonthly campaigns (green and blue curves) VS monthly campaigns (red curve); trimestral campaigns (green, blue and yellow curves) VS monthly campaigns (red curve); black line-90th percentile and blue line High/Good threshold

Presented n-tile curves confirm that reducing campaign frequency for the three year period results in no significant consequences for ecological status.. Also, the graphics of the quantile plots make it clear that trimestral quantity results correspond better than bimonthly campaigns

Water body 015

The sum of average and standard deviation of chlorophyll a observed values is 4,11 µg/l which is above the Good/Moderate threshold (3,6 µg/l), This should indicate that it is expected that approximately 16% of the values are above this value and, consequently, it is difficult to alter the ecological classification results due to a reduction in the frequency of campaigns.
The problem issue with a water body 015 is that during the studied period was surveyed by only one sampling station. During the three years were analyzed 34 samples, what is an amount of data that is inadequate to complete temporal reduction analysis.

Water body 016

Water body 016 has a 94,12% correspondence of classifications calculated using the reduced data sets. The diagram below presents a time series of chlorophyll *a* campaign average values that frequently exceed the Good/Moderate threshold. It is an unstable ecosystem and the indicator chlorophyll *a* (with variance value 4,66) portrays an unbalanced biomass that is responding to significant anthropic pressure (urban wastewater). Since water body 016 is identified as type III, the value of the salinity variance is expectedly low (0,32) and confirms the insignificant continental load volume.

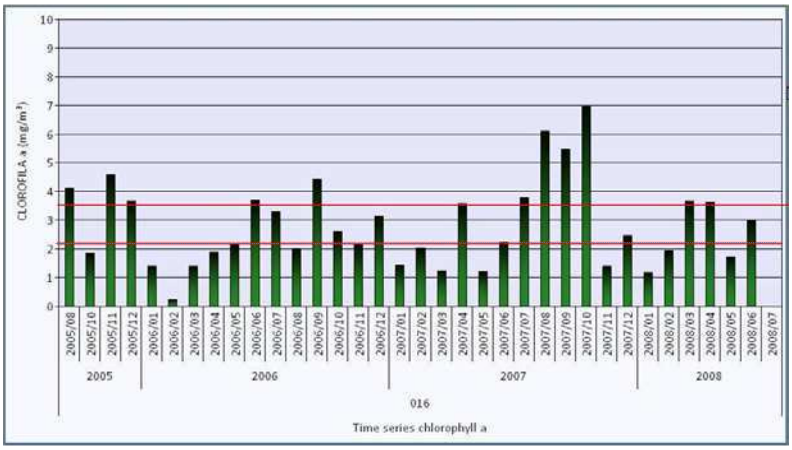


Figure 14: 016 w.b. time series of chlorophyll a monthly campaign average values

As is mentioned, the correspondence of water body 008's quality results is due to the dense frequency of chlorophyll *a* values which exceed the Good/Moderate threshold (Figure 14). The sum of average and standard deviation of chlorophyll a observed values is 4,95 µg/l which is well above the Good/Moderate threshold (3,6 µg/l), This should indicate that it is expected that approximately 16% of the values are above this value and, consequently, it is difficult to alter the ecological classification results due to a reduction in the frequency of campaigns.

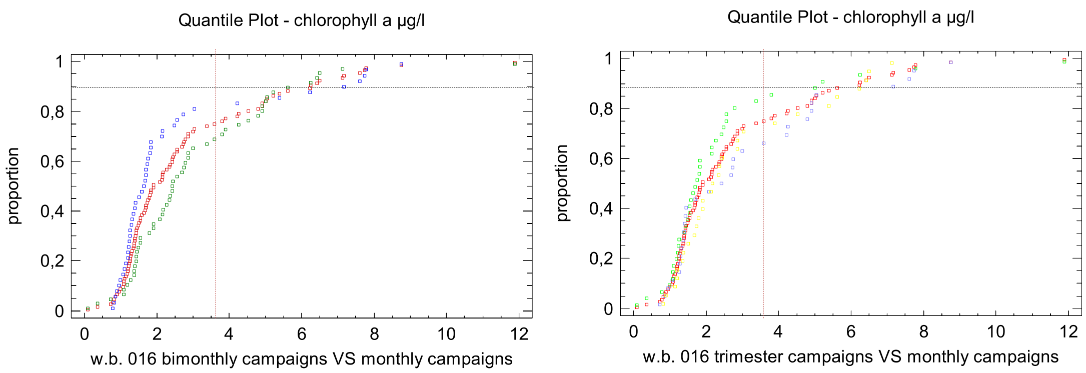


Figure 15: W.b. 016 quantile plots; bimonthly campaigns (green and blue curves) VS monthly campaigns (red curve); trimestral campaigns (green, blue and yellow curves) VS monthly campaigns (red curve); black line-90th percentile and red line Good/Moderate threshold; regarding the graphics, it is important to pay attention to the chlorophyll a scale values

The presented quantile curves (Figure 15) confirm that reducing campaign frequency for a three year period results in no significant consequences for ecological status. All 90th percentile values (calculated from reduced or complete data) have a significant distance "buffer" to the Good/Moderate limit and are settled deep within the moderate range. Despite the bias error of the quantitative results, all qualitative results have the same ecological classification -"Moderate".

Comparing the n-tile diagrams it is clear that bimonthly curves have less dispersion in the area of the 90th percentile, which results in lower bias error in quantitative results.

Water body 017

Water body 017 has the highest variance value (24,92 ; Table 13) of type III as well as a high percentage of corresponding ecological statuses (about 90%; Figure 1). As a type III water body, it has an elevated salinity variance value (1,26; Table 13) which indicates significant continental loads in the area. Due to continental influence during 2007 and 2008, chlorophyll a values (Figure 16) frequently exceeding the Good/Moderate threshold indicate that the system is unable to absorb continental loads (that contain urban wastewater) and to preserve the stability of the phytoplankton biomass.

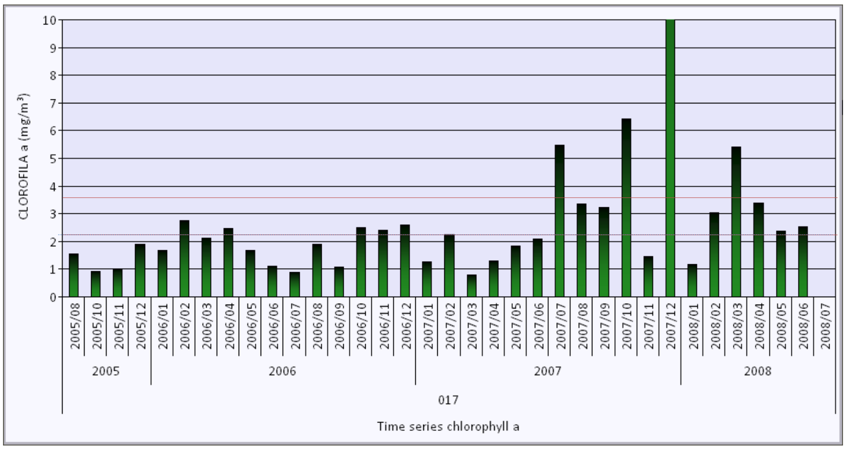


Figure 16: 017 w.b. time series of chlorophyll a monthly campaign average values

The phytoplankton biomass is unstable and the water body is classified as Moderate with a 90th percentile value that significantly exceeds the threshold. The sum of average and standard deviation of chlorophyll a observed values is 7,97 µg/l which is well above the Good/Moderate threshold (3,60 µg/l), This should indicate that it is expected that approximately 16% of the values are above this value and, consequently, it is difficult to alter the ecological classification results due to a reduction in the frequency of campaigns. Again we find outliers that can alter the adjust of a normal distribution to the data by raising the value of the variance. This indicates that, although it is expected that the ecological classification remain unchanged by reducing the frequency of campaigns, there is the possibility that some changes occur

The average error bias of the 90th percentiles calculated using reduced data sets has an elevated value (3,47 µg/l; Table 13) but this does not affect the quality classification.

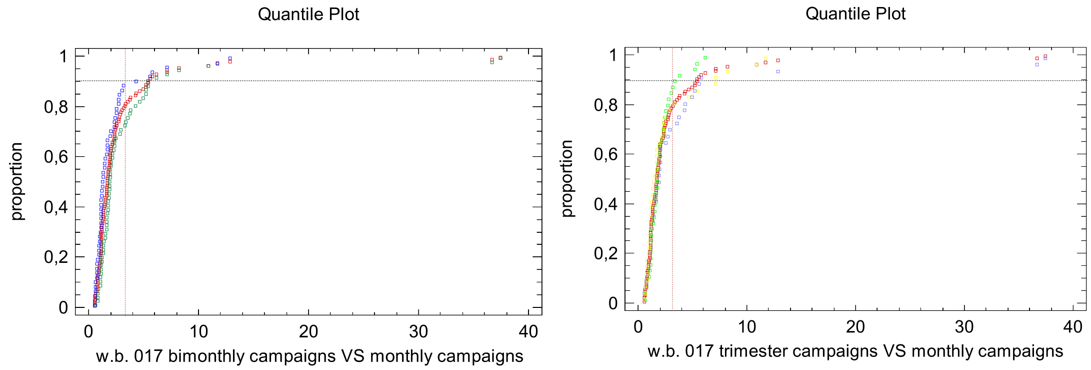


Figure 17: W.b. 017 quantile plots; bimonthly campaigns (green and blue curves) VS monthly campaigns (red curve); trimestral campaigns (green, blue and yellow curves) VS monthly campaigns (red curve); black line-90th percentile and red line Good/Moderate threshold; regarding the graphics, it is important to pay attention to the chlorophyll a scale values

On the quantile plots (Figure 17), it is evident that the 90th percentile values calculated using complete and reduced data sets have a significant “buffer” distance from the Good/Moderate limit. Finally, it can be concluded that sampling frequency reduction will not have any consequences for the ecological classification of water body 017.

4.3.3.Water bodies not selected by previous results

The rest of the water bodies have a less than 80% correspondence of ecological statuses established using reduced data sets. For this type of water body, the variation of chlorophyll *a* is usually around the High/Good or Good/Moderate threshold. Due to these indicator variations, reducing campaigns to bimonthly or trimestral frequency is likely to result in the establishing of erroneous ecological statuses.

Water Body 019

Water body 019 in the chapter ”Provisional ecological classification” is classified as Good ecological status for all four periods. However, with reduced data sets it is classified more than 50% as High or Moderate. Results show higher correspondence for the three year period.

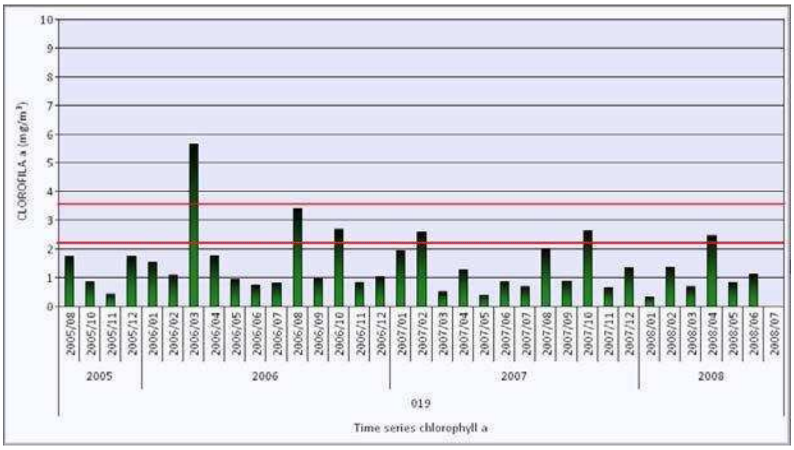


Figure 18: 019 w.b. time series of chlorophyll a monthly campaign average values

The indicator value (1.997) is grater than the buffer distance (0,75). On the diagram (Figure 18) of the time series it is possible to observe the variation (variance value 2,41; Table 13) and the range of chlorophyll a that make campaign frequency reduction impossible.

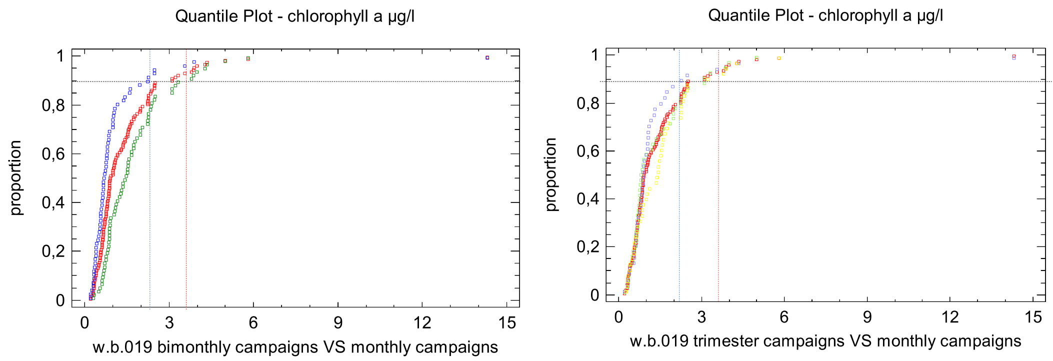


Figure 19: W.b. 019 quantile plots; bimonthly campaigns (green and blue curves) VS monthly campaigns (red curve); trimestral campaigns (green, blue and yellow curves) VS monthly campaigns (red curve); black line-90th percentile, blue line High/Good and red line Good/Moderate threshold; regarding the graphics, it is important to pay attention to the chlorophyll a scale values

The quantile plots (Figure 19) for the three year periods present a discrepancy in the area of the 90th percentile for the bimonthly and trimestral campaign frequencies. The results for the reduced data are settled in the area around the High/Good and Good/Moderate threshold which increases the probability that the ecological classification will differ from that calculated using the complete data. Due to this, reduction analysis results have a 50% discrepancy with the original data and the recommended frequency is one month.

Water body 011

Of all type III water bodies (coastal sites not affected by fresh water inputs) for the coastal waters of Valencia, 011 has almost the highest salinity variance value (0,64). The chlorophyll a variance (1,03) and average (1,11) values are low, but the indicator value (1.30) is greater than the buffer distance (1,09). The percentage of corresponding ecological classifications for water body 011 is less than 60%. For a type III, the thresholds of ecological classes (High – Good and Good – Moderate) are low, and 011 for the three year period is classified as Good.

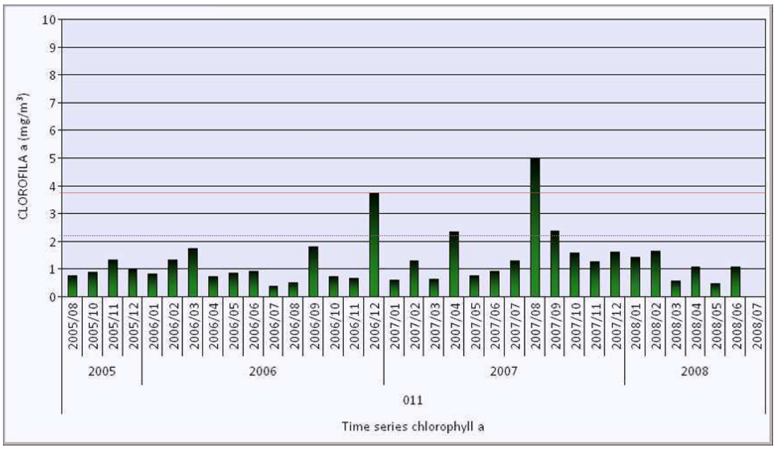


Figure 20: 011 w.b. time series of chlorophyll a monthly campaign average values

If we observe the chlorophyll a temporal series of campaign average values, it can be concluded that the system was reasonably stable during 2005, 2006 and 2008. During the winter of 2006 and the spring and summer of 2007, the chlorophyll a values even exceed the Good /Moderate threshold, and this biomass disturbance causes low correspondences of ecological classes calculated with reduced data.

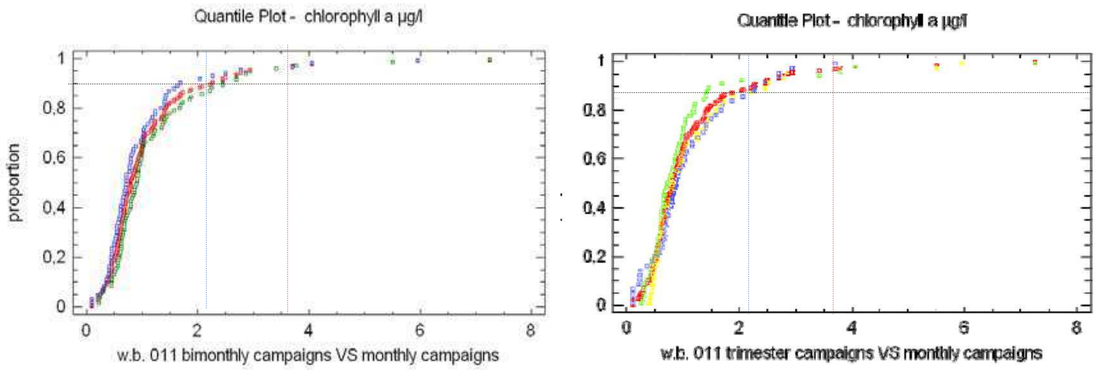


Figure 21: W.b. 011 quantile plots; bimonthly campaigns(green and blue curves) VS monthly campaigns (red curve); trimestral campaigns (green, blue and yellow curves) VS monthly campaigns (red curve); black line-90th percentile, blue line High/Good and red line Good/Moderate threshold; regarding the graphics, it is important to pay attention to the chlorophyll a scale values

The quantile plots for the three year period demonstrate that the 90th percentile value is situated on the High/Good threshold and minimal bias of 90th percentile calculated with reduced data result in all ecological classification being changed. As a system with relatively low chlorophyll a variation, but with 90th percentile values near the High/Good threshold, it is recommended to survey by monthly frequency sampling.

4.3.4.Discussion

- Using average, standard deviation and "buffer" distance average values, it is possible to establish the following coastal ecosystem categorization:

1. Stable systems with ecological classification defined: 001,002,003,004, and 014

-indicator value < "buffer" distance average value (Table 13) and
-sum of average and standard deviation < Good/Moderate threshold

All these analyses (performed using SPSS and EXCEL algorithms) indicate that temporal reduction is possible for systems with ecological classification defined; the quality results correspond to the original ones and the 90th percentile bias values are low because of a narrow chlorophyll *a* variance. The reduction analysis for these water bodies performed using the EXCEL algorithm provided 100% correspondence mainly because of the difference in 90th percentile values (the values are slightly lower). The same analysis provided by the SPSS algorithm results in more than an 80% correspondence. The quality error was mainly detected for one year periods where the reduced data set suffers from a lack of samples (around 20 or less for bimonthly and 10 to 15 for trimester reduced data sets) and due to specific environmental pressure (extended rain events, sea storms that can produce rapid dispersion or nutrient enrichment...), ecological classifications were degraded from the original ones. The results for the three year period have a 100% correspondence with the complete data set classification.
The most important reason for the good correspondence of quality results is the "buffer" distance to the High/Good threshold (see the quantile plots of w.b. 001, 002, 003, 004 and 014 for the three year period) that amortize the bias of the 90th percentile values. Since the stable ecosystems have a low variance of chlorophyll a values, the threshold High/Good is rarely or never exceeded.
Finally, the reduction of campaign frequency for water bodies 001,002,003,004 and 014 is possible and it could be recommended that monthly samplings be replaced with bimonthly or trimester campaigns.
By reducing campaign frequency, ecological classification will not suffer any consequences and this is mainly due the capacity of the systems to sustain present pressure and maintain a stable phytoplankton biomass.
With bimonthly and trimestral campaigns, it is possible to detect significant changes in water body quality (detecting chlorophyll a concentrations that exceed the common values for the series) and if is necessary to reestablish the monthly frequency samplings.
Before reestablishing the campaign frequency, it is necessary to study environmental data series (especially rain and waves), and the decision on sampling frequency has to be made on the basis of integrated data management.

2. Unstable systems with ecological classification defined: 008, 015, 016 and 017

- sum of average and standard deviation > Good/Moderate threshold

Reducing campaign frequency for unstable systems appears not to have any significant consequences on ecological classification status, especially in water bodies with a dense frequency of chlorophyll *a* values which do not fulfill environmental objectives. This is mostly due to the fact that the Moderate/Bad threshold was never established (mainly because the WFD states that European waters have to reach Good quality beyond 2015) and all water bodies that exceed the Good/Moderate limit (irrespective of the quantitative value) are classified as Moderate quality.
Despite corresponding qualitative results, it is recommended to establish an operational or investigative survey for unstable ecological systems which are identified as being at risk of failing their environmental objectives. The Water Frame Directive does not specify the frequency of operational monitoring, and it says that it shall be chosen as to achieve an acceptable level of confidence and precision. However, campaign intervals do not exceed those specified for surveillance monitoring (minimum frequency-three months for chlorophyll *a*).
With the stable systems, the most important reason for the good correspondence of quality results is “buffer” distance, while for the unstable systems it the distance from the Good/Moderate threshold. Due to the lack of the Moderate condition upper limit, the “buffer” value does not have a terminal point. This “buffer” distance excludes water body 008 (the three year period has a 90th percentile value near the threshold) and the campaign frequency reduction can upgrade the ecological classification
The quantitative results of reduced surveys are not as good as qualitative ones. The difference between 90th percentiles calculated using data from reduced campaigns and monthly campaigns is significant and error is not negligible (see the quantile plots of w.b. 008, 015, 016 and 017 for the three year period). Provided temporal reduction will not affect the ecological classification, but the precision of time series will be seriously decreased and temporal data will provide information with a lower level of confidence.
Temporal reduction is possible for unstable water bodies with large “buffers” (016 and 017) and the risk that they will be classified as Good does not exist. For these water bodies (which do not fulfill environmental objectives) WFD indicates the need for establishing a number of measures for reestablishing ecological status. Due to these expected measures and corresponding coastal water quality changes, it is not recommended to reduce the sample frequency.

3. **Systems with ecological classification fuzzy: 005,006,007,009,010,011,012, 013,018 and 019**

   --indicator value >"buffer" distance average value (Table 13) and
   -sum of average and standard deviation < Good/Moderate threshold

Reduction analysis for systems with ecological classification fuzzy, results in less than 80% correspondence. As the 90th percentile values vary around and between the thresholds, the probability for upgrading or degrading ecological classification is elevated. The presented correspondence results vary from 80% to less than 40%, which confirms that unpredicted and uncertain ecological quality results can be obtained by bimonthly and trimestral campaigns.
The consequences of temporal reduction will be:

- Two water bodies (010 and 012) have between 60% and 70% of corresponding classifications.
- Four water bodies have corresponding classifications between 50 and 60% (005,006,011 and 012)
- More than 50% of classifications for water bodies 007, 009 and 018 were changed.
- Water body 019 has the lowest percentage of correlation for the whole reduction analysis (less than 40%) due to chlorophyll *a* concentrations that result in a 90th percentile value on the High/Good threshold (Figure 19). The slight bias of the 90th percentile value calculated using the reduced data set changes the ecological classification which does not correspond to the classification calculated using the complete data set.

The presented results show that, for systems with ecological classification fuzzy, temporal reduction has unpredictable consequences on ecological classification, and for this reason, reducing sampling frequency is not recommended for this type of water body and the campaigns should be maintained monthly.

5.Spatial reduction

5.1.Exploratory study

5.1.1. Introduction

During the three-years of the survey, the monitoring network suffered many changes in design, mostly in the number of stations. After the pilot campaign, the number of sampling stations was almost halved. However, during the following campaigns, due to the results, analysis and modifications of the water bodies this number constantly increased. Specifically, the Valencian coastal waters suffered two changes. In August 2005 14 coastal water bodies were identified, during 2006 that number was increased to 16 and during 2007 was finally increased to 19 coastal water bodies (without port influenced w.b.).

After two reorganizations of the coastal water bodies, in January 2008 the monitoring network was again reorganized. One of the goals of the last reorganization was that each water body be surveyed with a minimum of four sampling stations. The historical survey data (since August 2005) did not accomplish the minimum of four samples per water body and it was modified again. In the chapter “Provisional ecological classifications” the actual survey network, number of samples and number of active/inactive stations for each water body are described, as are the multiples changes that have taken place in the survey networks since August 2005.

Because of the dynamic changes on the number of survey network stations mentioned, the research work could not be performed for all water bodies. In the present study, the water bodies have been divided into three groups:

1. Water bodies that have not been changed during the reorganization process and where the number of sampling stations has always fulfilled the minimum for the survey (water bodies 001, 002, 005, 011).
2. Water bodies that have been changed and where the number of sampling stations has fulfilled the minimum for all campaigns during the three-year period of August 2005 – July 2008 (water bodies 008, 009, 010, 014, 019).
3. Water bodies that have been changed and where the number of sampling stations did not fulfill the minimum of four samples for at least one year (water bodies 003, 004, 006, 007, 012, 013, 015, 016, 017,018).

Only the historical data on water bodies properly surveyed during the three-year period throughout the entire data acquisition have been used for analysis of the monitoring network using spatial reduction:

4.In particular, the following data has been considered in the study:

1.Water body 001
- 156 water samples
- Five active sampling stations, four active since the pilot campaign, last station established in February 2006.

2.Water body 002
- 147 water samples

- Five active sampling stations, one active since the pilot campaign and the other four since the February 2006 campaign.

3. Water body 005
   - 168 water samples
   - Five active sampling stations, all sampled since the pilot campaign.

4. Water body 008
   - 149 water samples
   - Five active sampling stations, four active since the pilot campaign, one station established in April 2007.

5. Water body 009
   - 166 water samples
   - Five sampling stations active since the pilot campaign.

6. Water body 010
   - 187 water samples
   - Six sampling stations active since the pilot campaign.

7. Water body 011
   - 166 water samples
   - Six active sampling stations, three active since the pilot campaign and the rest active since March 2006.

8. Water body 014
   - 128 water samples
   - Four sampling stations active since the pilot campaign.

9. Water body 019
   - 133 water samples
   - Four sampling stations active since the pilot campaign

As in the previous study (temporal reduction), all port influenced water bodies have been excluded. These water bodies have been properly surveyed since July 2008 (four sampling stations per water body) and historical data does not permit spatial reduction analysis.

Significant numbers of water bodies are excluded from the analysis due the lack of historical data. The water bodies 004, 006, 007, 012, 013, 015, 016 and 017 were surveyed by less than four sampling stations for a significant time period (more than one year). Water bodies surveyed by three (or less) sampling stations cannot be analyzed for the spatial reduction when data do not fulfill the minimum, which cannot be reduced. The problem with improperly surveyed water bodies is that they do not have a reference (for a minimum number of sampling stations) and the quality values that are established could be erroneous. Analyzing these (incomplete) campaigns, the final study result could be disrupted and could have a significant (but inaccurate) influence on the final conclusion. These water bodies (due to the lack of data) are not appropriate for this analysis at this moment, but this does not exclude the possibility of implementing the same study in the future (when this lack of data becomes insignificant).

For all the water bodies included into the analysis, the number of possible k-observers is calculated using the following equation:

$$nC\,k = \frac{n!}{k!(n-k)!}$$

n- number of active sampling stations for the water body

k- the number of sampling stations in the observer

w.b. number of active s. stations	number of s. stations in the observer	number of observers	all possible observers
4	3	4	**14**
	2	6	
	1	4	
5	4	5	**30**
	3	10	
	2	10	
	1	5	
6	5	6	**62**
	4	15	
	3	20	
	2	15	
	1	6	

Table 1: Number of possible limited observers for water bodies surveyed by four, five and six sampling stations

For each observer (302 for entire analysis) of sampling stations, the 90th percentile is calculated and related ecological status (quantity and quality values) during the four different periods (three-yearly periods and one three-year period):

- 2005/08-2008/07
- 2005/08-2006/07
- 2006/08-2007/07
- 2007/08-2008/07

The results calculated with the reduced number of stations are compared with the original quality and quantity values obtained using the complete data series.

This analysis attempts to answer one question: is it possible to reduce the number of sampling stations in the Valencian coastal monitoring network without affecting ecological classifications?

As mentioned in paragraph 3.5, from the point of view of space reduction, it has been considered that the different k-observers of sampling stations are different *observers* that provide different monthly measurements of *chlorophyll a* concentration in the area. Therefore, our aim has been to study the concordance between the measurements taken by the reference *observer,* which corresponds to the *reference set of sampling stations* and the measurements taken by the *observers,* which correspond reduced subsamples of the *reference set.*

5.1.2.Water body 001

During these three years, water body 001 was surveyed by five sampling stations (DP001, DP003, DP005, DP007 and DP133, (Annex 2, figure 1). If we consider all possible reductions in the number of stations reviewed, it is possible to define 30 different observers, who have excluded data from at least one of the sampling stations. These observers will be called *limited observers.*

Water body 001 is classified as "High" for the four periods considered (chapter "ecological classifications Provisional), by which, an erroneous ecological classification will always be a degradation.

Based on average, standard deviation and buffer values it has been categorized as a stable system with ecological status defined (chapter : “Ecological classification fuzziness”).

For the first annual period (August 2005-July 2006), the 90th percentile could not calculate at station DP133, due to a lack of data in this period. It was considered, therefore, that the observer who had data available only from the sampling station DP133, could not take measurements and we have worked on 29 *limited observers*. For 97% (28) of these observers the ecological classification done based on their measurements coincided with the ecological classification made from measurements of the reference observer (Figure 1).

For the second period, the 90^{th} percentiles have been calculated for the 30 observers. The ecological status obtained for 80% (24) of them agreed with the original classification. In the third period, again three observers, who taken measurements only in one sampling station (DP003, DP005 and DP133) were discarded due to a lack of data. Of the 27 observers considered, 70% (19) have indicated the same ecological status as the reference observer.

For the three-year period, there has been no observer excluded for lack of data. Of the 30 observers 7% (2) have degraded the ecological classification.

The lowest values of chlorophyll a are detected in the first period (2005/08-2006/07) (90^{th} Pctl=3.36 mg/m³; s^2=1,36) The variance value increases during the second (90^{th} Pctl=3.69 mg/m³ s^2=1.53) and third period (90^{th} Pctl=3.43 mg/m³ s^2=1.64). The high proportion of degraded ecological classifications, in the second and third period, is consequence of the high variability of the data and the small sample size. Because of the breadth of three-year period, and reduced variability of the data in this period (90^{th} Pctl=3.54 mg/m³, s^2=1,53), the percentage of affected ratings is very low in this period.

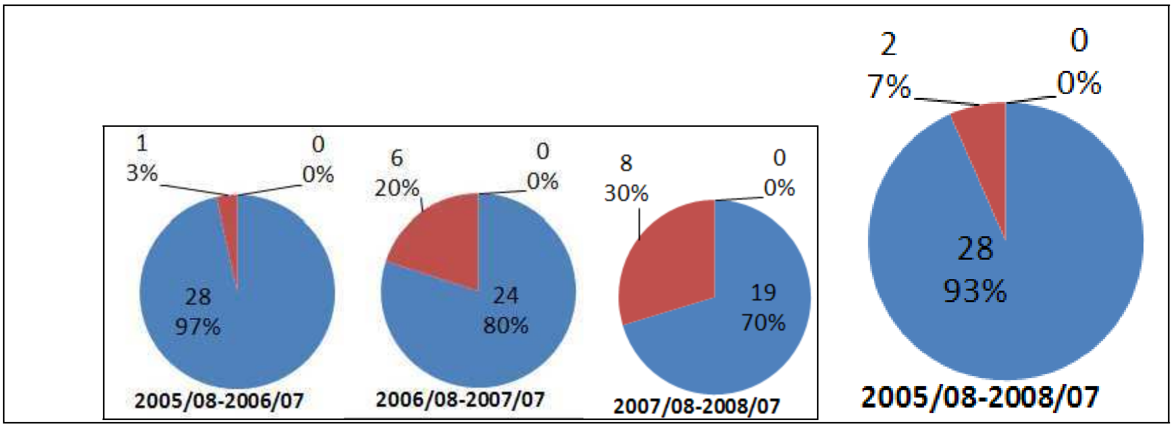


Figure 1: Number (and %) of limited observers where calculated eco. status corresponds to the original one (blue); number (and %) of limited observer that degrade eco. Class (red). water body 001- for different periods

The number and percentage of limited observers (grouped by number of excluded stations) that do not degrade ecological classification for all four periods and for the three-year period are listed In Table 2. It is possible to observe that the five observers, which exclude only one sampling station, do not modify the ecological classification calculated by the *reference observer*. By reducing the number of sampling stations, the percentage of corresponding ecological classifications, drops below 60%. If only the three-year period is analyzed, the results are significantly improved, which is confirmed by the fact that the majority of the non-corresponding ecological classifications are calculated for the one-year periods. Because of this, 4-stations or 3-stations observers have 100% unchanged ecological classifications, 2-stations observers have 90% and 1-station observer have 80% of ecological classifications that correspond to those calculated by the *reference observer*. These results were expected, mostly due to interannual variations, which are analyzed and determined in the chapter “Temporal reduction”.

N. of excluded s. stations	N. of comb. realized for all 4 periods	N. of comb. that do not modify eco. class. (4 periods)	% of comb. that do not modify eco. class. (4 periods)	N. of comb. realized for 3 years period	N. of comb. that do not modify eco. class. (3 years period)	% of comb. that do not modify eco. class. (3 years period)
1	5	5	100,00%	5	5	100%
2	10	5	50,00%	10	10	100%
3	10	6	60,00%	10	9	90%
4	2(from5)	1	50,00%	5	4	80%

Table 2: Number and percentage of limited observers (grouped by number of excluded stations) that do not degrade ecological classification in any of the four periods considered and in the three-year period considered separately.

Detailed results for water body 001 are listed in the annex “Descriptive statistic results”. It includes, for each *limited observer,* and for the four periods considered, the biases between the 90th percentiles, calculated using measurements of the sampling stations that characterize the observer, and the 90th percentile, calculated using the measurements of the 5 sampling stations that characterize the *reference observer*. Also in this table, it is highlighted if the ecological classifications assigned by the *limited observers* correspond to the reference one.

The Table 3 shows that in the majority of observers, the bias of the 90th percentile calculated for the period of three years is generally lower than the bias of the 90th percentile, calculated for annual periods. By comparing the average absolute biases of the 90^{th} percentiles for the annual periods (annual average 0.75 mg/m³) and the average absolute bias for the three- year period (0.47 mg/m³), it is clear that quantity error decreases for the longer sampling period.

average absolute bias 2005-2006	average absolute bias 2006-2007	average absolute bias 2007-2008	annual average absolute bias	average absolute bias 2005-2008
0,64	0,61	1,00	0,75	0,47

Table 3: Water body 001 interannual average absolute biases; annual average absolute bias, average absolute bias for the three years period

5.1.3.Water body 002

Water body 002 was sampled by five sampling stations during last three years: DP134, DP135, DP136, DP137 and DP010. Station DP010 has been sampled since the pilot campaign and the 4 other stations were added in February 2006. For water body 002, 30 possible *limited observers* have been identified.

Like 001, water body 002 is classified as "High" for the four periods considered (chapter "ecological classifications Provisional), by which, an erroneous ecological classification will always be a degradation.

Based on average, standard deviation and buffer values water body 002 has been categorized as a stable system with ecological classification defined (chapter : “Ecological classification fuzziness”)

For the first annual period, some ecological statuses could not be established, due to a lack of samples from DP134, DP135, DP136 and DP137. Four *1-station observers* have been excluded due to the above-mentioned lack of data. For water body 002 and for the first period (August 2005-July 2006) 90^{th} percentile values have been estimated for 26 observers. All 26 ecological classifications established by the different *limited observers* correspond to the ones established by the *reference observer*

For the second annual period, 30 ecological classifications were established. 17% (5) of the ecological classifications degraded the original ecological classification and 83% (25) of them corresponded to the original ecological classification.

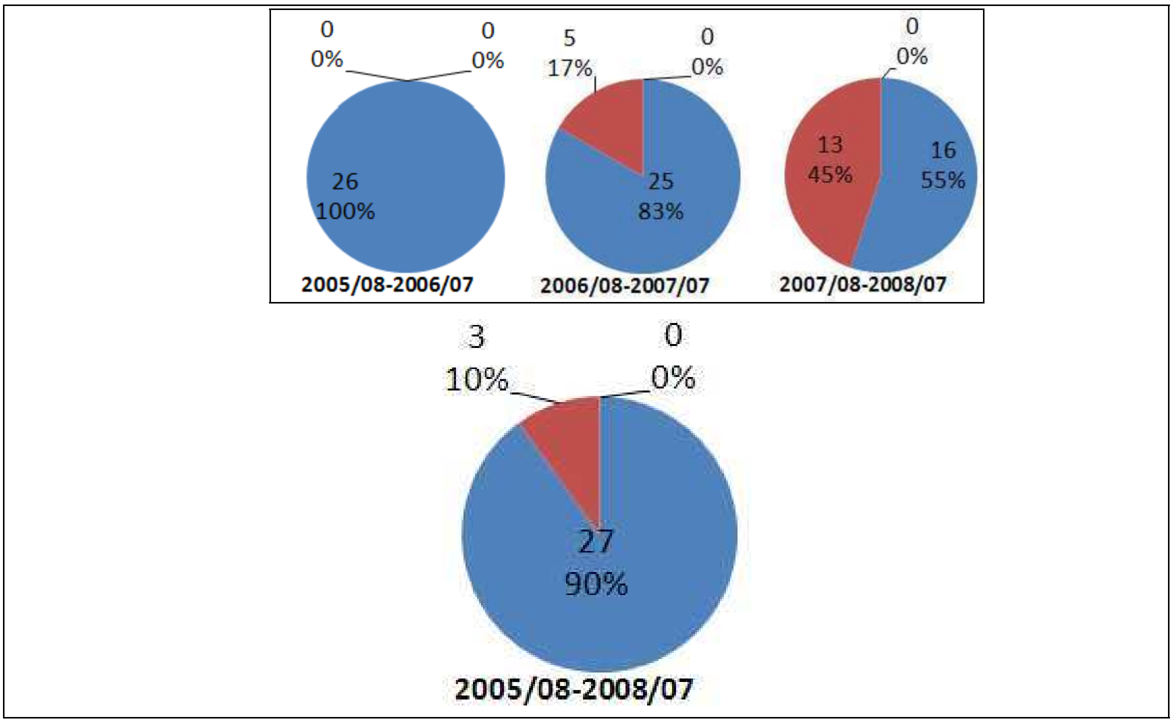


Figure 2: Number (and %) of reduced observers where calculated ecological status corresponds to the original one (blue); number (and %) of reduced observers that degrade ecological class (red). Water body 002 for the four periods.

For the third annual period, ecological classifications were established for 29 observers and *1-station observer* is excluded from analysis (due to a lack of data from DP137). 55% (16) of the ecological classifications from the third annual period correspond to the original quality value and 45% (13) of the classifications degrade the status of the water body

90^{th} percentile values have been calculated for the three-year period for the maximum number of observers (30). Of 30 possible observers, 10% (3) upgrade the ecological classification (do not correspond to the one calculated with all five stations).

Chlorophyll *a* variance value for the first (90^{th} Pctl=3.33 mg/m³) and second (90^{th} Pctl=3.38 mg/m³) annual period has the same low value (1.34). The third annual period has a significantly higher variance (2.40) and 90th percentile (4.49 mg/m³ with lowest buffer distance 0,31 mg/m³), values that are reflected in the percentage of degraded ecological classifications established by the *limited observers*.

The three year period (90^{th} Pctl=3,50 mg/m³) has a variance value of 1.72 , but due to the extended period it appears that the percentage of modified ecological classifications is not comparable with the annual periods.

The number and percentage of *limited observers* (grouped by number of excluded stations) that do not degrade ecological classification for all four periods and for the three-year period are listed in Table 4. In this table, listed results for the four periods show that the percentage of corresponding observers (depending on the number of reduced stations) varies from 60% to 0%. From five *4-station observers*, two of them have a degraded ecological classification for at least one period. *3-station observers* have 60% corresponding ecological classifications while *2-station observers* have 50%. Due to the five campaigns' lack of data for DP134, DP135, DP136 and DP137 for the first one-year period, only one 90th percentile value has been calculated on the basis of one sampling station. Because of this lack of data, 90th percentile values were only calculated for the one *1-station observer* (DP010) for all four periods and the ecological classifications do not correspond to the one established by the *reference observer*.

N. of excluded s. stations	N. of comb. realized for all 4 periods	N. of comb. that do not modify eco. class. (4 periods)	% of comb. that do not modify eco. class. (4 periods)	N. of comb. realized for 3 years period	N. of comb. that do not modify eco. class. (3 years period)	% of comb. that do not modify eco. class. (3 years period)
1	5	3	60,00%	5	5	100%
2	10	6	60,00%	10	10	100%
3	10	5	50,00%	10	9	90%
4	1(from5)	0	0,00%	5	3	60%

Table 4: Number and percentage of *limited observers* (grouped by number of excluded stations) that do not degrade ecological classification in any of the four periods considered and in the three-year period considered separately.

By analyzing the results for the three year period, we observe that the table presents significantly improved results. For the 4-station and 3-station observers, ecological classifications correspond 100% to the original ones. For 2-station and 1-station observers, the percentage is lower at 90% and 60%, respectively. As with water body 001, it can be deduced that a majority of degraded ecological classifications are established for the one year periods (mostly due to interannual variations).

Detailed results for water body 002 are listed in the annex “Descriptive statistic results” (all possible observers with 5 sampling stations).

Observing the annex table, the 90th percentile biases calculated for the three-year period for most observers are lower than 90th percentiles biases calculated for the annual periods (at least one). Comparing the 90th percentile biases for the annual periods (annual average 0.66 mg/m³) with average bias for the three year period (0.57 mg/m³), it is clear that quantity error decreases as sample time increases.

average absolute bias 2005-2006	average absolute bias 2006-2007	average absolute bias 2007-2008	annual average absolute bias	average absolute bias 2005-2008
0,43	0,66	0,90	0,66	0,57

Table 5: Water body 002 interannual average absolute biases; annual average absolute bias, average absolute bias for the three years period

5.1.4.Water body 005

Water body 005 was surveyed by five sampling stations: DP032, DP034, DP036, DP038 and DP040.

As with water bodies 001 and 002, 30 possible *limited observers* were identified. For each observer, 90th percentile values and related ecological status (four periods) were obtained and compared with the results obtained by the reference observer.

Water body 005 is classified as "Good" ecological quality for the two annual (2006-2007 and 2007-2008) and three-year periods. For the first one-year period (2005-2006) it is classified as "High" ecological quality. For the three-year and the other two annual periods, the ecological classification established by the *limited observers* can either degrade or upgrade the original quality values. For the first annual period, the ecological classification established by the *limited observers* can only degrade the original classification.

Based on average, standard deviation and buffer values water body 005 has been categorized as a system with ecological classification fuzzy (chapter : "Ecological classification fuzziness")

Since water body 005 does not have missing data, 30 ecological classifications are established for each of the four periods.

For the annual period August 2005–July 2006, 60% of the ecological classifications (18) were degraded compared to the original ecological status.

For the next annual period (August 2006–July 2007) 77% (23) of observers preserve their original ecological classifications. Since the water body is classified as "Good", One (3%) observer degrades its ecological status to "Moderate" and 6 (20%) of them upgrade it to "High".

For the third annual period (August 2007 – July 2008), 37% (11%) of the reduced observers preserve their ecological classification. 20% (6) of the reduced observers upgrade the ecological classification to "High" and 43% (13) degrade it to "Moderate".

Analyzing the three year period, 20% (6) observers upgrade the ecological classification to "High" and 7% (2) degrade it to "Moderate" status. 73% (22) of the observers preserve the original classification.

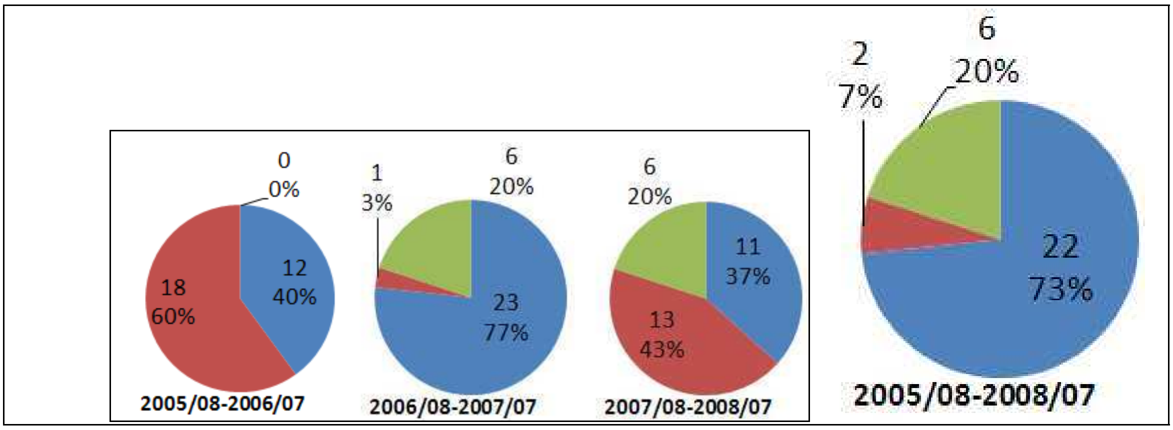


Figure 3: Number (and %) of reduced observers where calculated ecological status corresponds to the original one (blue); number (and %) of reduced observers that degrade ecological status (red); number (and%) of reduced observers that upgrade ecological status (green); water body 005-four periods

For the third annual period (90th Pctl=-6.98 mg/m³) the highest variance value is calculated (4.98). For the two other annual periods (2005-2006; 90th Pctl=4.80 mg/m³ and 2006-2007; 90th Pctl=5.74 mg/m³) variance value is high compared to water bodies 001 and 002, with values of 3.36 for the first and 2.95 for the second. Variance value is reflected in the percentage of modified ecological classifications for the second and third period. Therefore, during the first annual period, the 90th percentile is more significant because of the value near to the High-Good threshold (short buffer value, less than 0,00 mg/m³).

The three-year period has a 90th percentile (5.69 mg/m³) between two thresholds. This establishes two buffer value (0.89 mg/m³ for the High/Good and 1.51 mg/m³ for the Good/Moderate) that are exceeded by 27% of observers due to the considerable variance value of 3.79.

The number and percentage of limited observers (grouped by number of excluded stations) that do not degrade ecological classification for all four periods and for the three-year period are listed in Table 6.

It is possible to observe that the 4-station observer have the highest percentage (40%) of observers that preserve the ecological classifications established by the *reference observer* (4 periods). The observer that excludes more than one sampling station has poor correspondence (10% for the 3- and 2-station observers and 0% for the 1-station observer) to the classifications established by the reference observer.

Analyzing only the three years period, results are improved significantly for the 4-station (100%) and 3-station (90%) observers. This confirms that all modified ecological classifications for the 4-station observers are calculated for the annual periods and the majority for the 3-station observers. The 2-stations (60%) and 1-station (40%) observers results are higher, but a significant number of modified ecological classifications is calculated for the three-year period.

N. of excluded s. stations	N. of comb. realized for all 4 periods	N. of comb. that do not modify eco. class. (4 periods)	% of comb. that do not modify eco. class. (4 periods)	N. of comb. realized for 3 years period	N. of comb. that do not modify eco. class. (3 years period)	% of comb. that do not modify eco. class. (3 years period)
1	5	2	40,00%	5	5	100%
2	10	1	10,00%	10	9	90%
3	10	1	10,00%	10	6	60%
4	5	0	0,00%	5	2	40%

Table 6: Number and percentage of *limited observers* (grouped by number of excluded stations) that do not degrade ecological classification in any of the four periods considered and in the three-year period considered separately.

Detailed results for water body 001 are listed in the annex "Descriptive statistic results".

Observing this annex table, the 90th percentile biases calculated for the three-year period for most observers are lower than the 90th percentiles biases calculated for the annual periods (at least one). Comparing the 90th percentile biases for the annual periods (annual average 0.91 mg/m³) with the average bias for the three-year period (0.71 mg/m³), it is clear that quantity error decreases as sample time increases.

average absolute bias 2005-2006	average absolute bias 2006-2007	average absolute bias 2007-2008	annual average absolute bias	average absolute bias 2005-2008
0,90	0,56	1,26	0,91	0,71

Table 7: Water body 005 interannual average absolute biases; annual average absolute bias, average absolute bias for the three-year period

5.1.5.Water body 008

Water body 008 was surveyed by five sampling stations: DPU001, DP051, DP052, DP054 and DP055. The station DPU001 is active from the April 2007 and has a 15 samples analyzed.

For this water body, 30 possible *limited observers* have been identified for spatial reduction analysis. For each observer, 90th percentile values and related ecological status (quantity and quality values) are calculated for all four different periods. The quality and 90th percentile values are compared with the results obtained by the *reference observer* (5 stations).

Water body 008 is classified as "Moderate" ecological quality for the two annual (2005-2006 and 2007-2008) and three-year periods. For the second one year period (2006-2007) it is classified as "Good" ecological quality. For the three-year and for two annual periods, the ecological classification established by the *limited observers* can only upgrade the original quality values. For the second annual period, the ecological classification established by the *limited observers* can either degrade or upgrade original classification.

Based on average, standard deviation and buffer values, water body 008 has been categorized as a unstable system with ecological classifications defined (chapter : “Ecological classification fuzziness”). These systems must be revised because the high variance and the presence of outliers may alter the ecological classification.

For the first period 29 ecological classifications are established. The 90th percentile is not calculated only for the 1-station observer, due to the lack of data for DPU001 which was inactive during the first year. 59% (17) ecological classifications are unchanged and upgraded ecological classifications are established for the 41 (12) of the limited observers.

The second period is the same as the first, with a lack of data from sampling station DPU001. Because of this, 29 ecological classifications are established. Just one (3%) observer degraded the original classification, while the other 28 ecological classifications remained unchanged.

For the last annual period, 90th percentile values were calculated 30 times. All 30 ecological classifications were unchanged compared to the classifications established *by the reference observer*.

Analyzing the three-year data, 20 (67%) of 30 ecological classifications were unchanged. 10 (33%) observers upgraded the ecological classification for water body 008.

The variance values for the first (90th Pctl=9.19 mg/m³) and third annual period (90th Pctl=11.81 mg/m³) are extremely high (91,97 and 21,12), likely due to the fact that the sampling area was in the near field of continental influence (Albufera hypertrophic lagoon). The second annual period has a low variance of 2.85 (for water body 008) and in observer with considerable buffer values (established by 90th Pctl=5.35 mg/m³, 0.55 on the High/Good and 1,85 mg/m³ on the Good/Moderate threshold) has a low percentage of modified ecological classifications. The elevated 90th percentile value calculated for the third period establishes a considerable buffer value (4.61 mg/m³) on the Good-Moderate threshold that results in 100% unchanged ecological classifications, beyond the elevated variance value. The first annual period has a considerable buffer (1.99 mg/m³) but an extremely high variance value, which results in an elevated percentage of observers that do not preserve the “original” ecological classification.

The variance for the three-year period has a value of 36.89 and due to the 90thPCTL.-7.96 mg/m³ value (the low buffer on the Good-Moderate threshold – 0.76 mg/m³) the percentage of unchanged ecological classifications is lower than for the second and third annual periods.

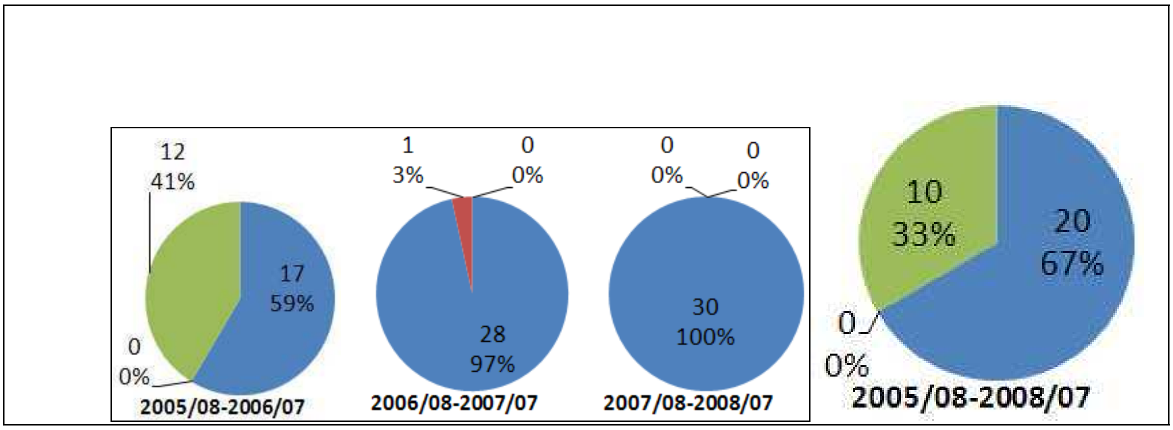


Figure 4: Number (and %) of reduced observers where calculated ecological status corresponds to the original one (blue); number(and %) of reduced observers that degrade ecological status (red); number (and %) of reduced observers that upgrade ecological status (green); water body 008-four periods

Table 8 lists the number and percentage of *limited observers* (grouped by number of excluded stations) that do not degrade ecological classification for all four periods and for the three-year period.

It is possible to observe that four of five observers that exclude one sampling station do not modify the ecological classification established by the *reference observer* for all four periods. Reducing the number of sampling stations in the observers (3-station, 2-station and 1-station), the percentage of corresponding ecological classifications reduces to 60%, 50% and 35 % correspondently.

N. of excluded s. stations	N. of comb. realized for all 4 periods	N. of comb. that do not modify eco. class. (4 periods)	% of comb. that do not modify eco. class. (4 periods)	N. of comb. realized for 3 years period	N. of comb. that do not modify eco. class. (3 years period)	% of comb. that do not modify eco. class. (3 years period)
1	5	4	80,00%	5	4	80%
2	10	6	60,00%	10	8	80%
3	10	5	50,00%	10	5	50%
4	4(from5)	1	25,00%	5	3	60%

Table 8: Number and percentage of *limited observers* (grouped by number of excluded stations) that do not degrade ecological classification in any of the four periods considered and in the three-year period considered separately.

If only the three-year period is analyzed, the results are slightly improved for the 3- (80%) and 1-station (60%) observers. For the four- and two- station observers, the results are same (80%). From these results, it can be concluded that a significant number of upgraded classifications is calculated using the three-year data.

In the annex “Descriptive statistic results”, detailed results are listed for water body 008

Observing the annex table, the 90th percentile biases calculated for the three-year period for most observers are lower than the 90th percentile biases calculated for the annual periods (at least one). Comparing the 90th percentile biases for the annual periods (annual average 3.44 mg/m³) with the average bias for the three-year period (1.61 mg/m³), it is clear that the quantity of error decreases as sample time increases.

average absolute bias 2005-2006	average absolute bias 2006-2007	average absolute bias 2007-2008	annual average absolute bias	average absolute bias 2005-2008
7,35	0,55	2,40	3,44	1,61

Table 9: Water body 008 interannual average absolute biases; annual average absolute bias, average absolute bias for the three-year period

5.1.6.Water body 009

Water body 009 was sampled by five sampling stations during the last three years: DP056, DP057, DP059, DP061 and DP062. For water body 009, 30 possible *limited observers* of sampling stations are identified.

For each observer, 90th percentile values and related ecological status are calculated for all four periods (three annual periods and the three-year period).

Ecological classifications established by the *limited observers* are compared to the quality results obtained by the *reference observer*.

The water body is classified as "Moderate" for the three-year and third annual periods (classification can only be upgraded). For each period, only one buffer value is established on the Good/Moderate threshold; three-year period – 1.02 mg/m³ and third annual period 3.18 mg/m³. The first and second annual periods are classified as "Good" ecological quality (classification can be degraded as well as upgraded).

Based on average, standard deviation and buffer values water body 009 has been categorized as a system with ecological classification fuzzy (chapter : "Ecological classification fuzziness"). Additionally, the sum of average and standard deviation (6.92) is very close to the Good / Moderate threshold (7.2) and the presence of outliers can also alter the ecological classification in this water body.

For the first annual period 30 ecological statuses are established. Of the 30 established ecological classifications, 11 (37%) were unchanged. Nine (30%) observers degraded and 10 (33%) upgraded the original ecological classification.

For the second annual period, 30 ecological classifications were established. For 16 (53%) observers, ecological classification remained unchanged. Eight (27%) observers upgraded and six (20%) degraded the original ecological classification.

For the second annual period , all 30 *limited observers,* did not upgrade or degrade ecological classification compared to the original one.

Of the 30 observers, 22 (73%) have unchanged ecological classifications for the three-year period. Eight (27%) of the observers have upgraded the original ecological classification.

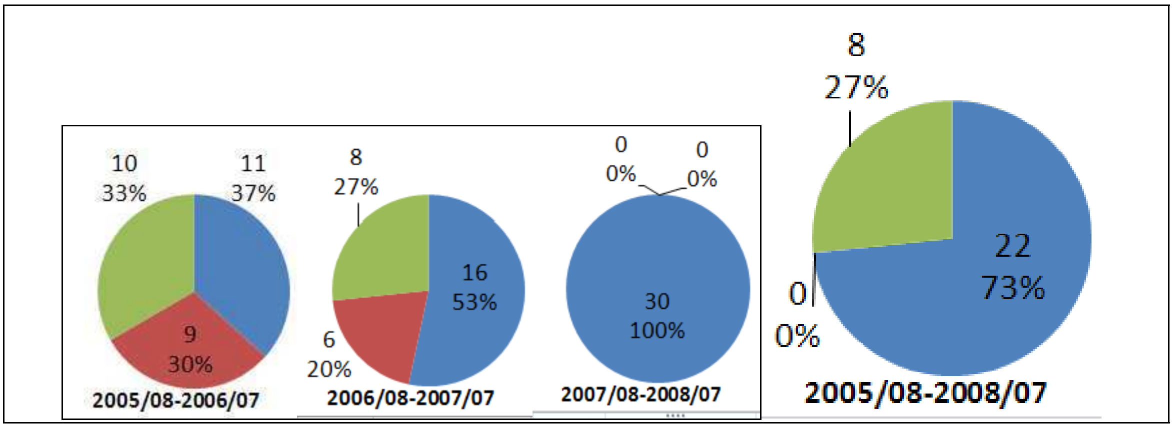


Figure 5: Number (and %) limited observers where calculated ecological status corresponds to the original one (blue); number (and %) of limited observers that degrade ecological status (red); number (and %) of limited observers that upgrade ecological status (green); water body 009-four periods

The variance values for water body 009 are high as is expected for a water body of type II-A. For the first annual period (90th Pctl=5.05 mg/m³) the variance value is 6.10 , for the second (90th Pctl=5.51 mg/m³) 4.86 and for the third (90th Pctl=10.38 mg/m³) 16.62 . As with water body 008, the percentage of unchanged ecological classifications depends on variance but also on 90th percentile value to determine the distance(s) to the threshold(s) – buffer(s). For the first and second annual periods, the water body is classified as good, because of which two thresholds exist which can be exceeded. With a variance value that decreases from the first to the second annual period, the percentage of unchanged ecological classifications increased. Water body 009 was classified as "Moderate" for the third annual period which determines only one threshold to exceed and a related (considerable) buffer for the 90thPCTL.-10.38 mg/m³. Due to this, the water body with elevated variance has 30 observers that establish unchanged ecological classifications (all elevated 90th percentile biases extend the difference to the Good-Moderate threshold)

Variance for the three-year period (90th Pctl=8.22 mg/m³) has a high value (10.25), but this is expected for a water body that is significantly influenced by continental and hypereutrophic waters.

The number and percentage of limited observers (grouped by number of excluded stations) that do not degrade ecological classification for all four periods and for the three-year period are listed in table 010,

It is possible to observe that two of five (40%) observers that exclude one sampling station do not modify the original ecological classification established by the *reference observer* for all four periods. Reducing two sampling stations, the percentage of corresponding ecological classifications reduces to 30%. 0% of the 2 and 1-station observers preserve the original ecological classifications for all four periods.
Analyzing only the three years period, the results are significantly higher. These results confirm that the majority of non-corresponding eco classifications are calculated for one year periods, due to annual temporal variations.

N. of excluded s. stations	N. of comb. realized for all 4 periods	N. of comb. that do not modify eco. class. (4 periods)	% of comb. that do not modify eco. class. (4 periods)	N. of comb. realized for 3 years period	N. of comb. that do not modify eco. class. (3 years period)	% of comb. that do not modify eco. class. (3 years period)
1	5	2	40,00%	5	4	80%
2	10	3	30,00%	10	9	90%
3	10	0	0,00%	10	7	70%
4	5	0	0,00%	5	2	40%

Table 10: Number and percentage of *limited observers* (grouped by number of excluded stations) that do not degrade ecological classification in any of the four periods considered and in the three-year period considered separately.

In the annex "Descriptive statistic results", detailed results for water body 009 are listed .

Observing the annex table, the 90th percentile biases calculated for the three-year period for most observers are lower than the 90th percentile biases calculated for the annual periods (at least one). Comparing the 90th percentile biases for the annual periods (annual average 1.38 mg/m³) with the average bias for the three-year period (1.24 mg/m³), it is clear that quantity error decreases as sample time increases.

average absolute bias 2005-2006	average absolute bias 2006-2007	average absolute bias 2007-2008	annual average absolute bias	average absolute bias 2005-2008
2,03	1,16	0,96	1,38	1,24

Table 11: Water body 009 interannual average absolute biases; annual average absolute bias, average absolute bias for the three-year period

5.1.7.Water body 010

Water body 010 was surveyed by six sampling stations: DP066, DP068, DP070, DP072, DP073 and DP074.

For water body 010, 62 possible *limited observers* were identified. For each observer, 90th percentile values and related ecological status are calculated (four periods) and compared to the results obtained by the *reference observer*

Water body 010 for the three-year and first annual periods is classified as "Good" ecological quality. For the second one-year period (2006-2007) it is classified as "High" ecological quality and for the third annual period (2007-2008) it is classified as "Moderate". For the three-year period and the first annual periods, ecological classification established by *limited observers* can either degrade or upgrade the original quality values. For the second annual period, ecological classifications established by *limited observers* can only degrade the original classification, while for the third period it can only upgrade it.

Based on average, standard deviation and buffer values water body 010 has been categorized as a system with ecological classification fuzzy (chapter: "Ecological classification fuzziness")

For the first annual period (August 2005–July 2006), due to a lack of samples from DP066 and DP074, all 62 ecological statuses were not established. Two 1-station observers have been excluded from the analysis. 33 (55%) observers calculated unchanged ecological classifications. 14 (23%) of the observers degraded and 13 (22%) upgraded the ecological classification calculated using the complete data.

During the next annual period (August 2006 – July 2007) ecological classifications were established for all 62 observers for the water body. 30 (48%) observers preserve the original status and the rest of 32 reduced observers degrade the ecological classification established by the reference observer.

For the third annual period (August 2007 – July 2008) due to a lack of data from DP073, one 1-station observer is excluded from the analysis. 40 (66%) of the observers preserved ecological status and 21 (34%) of the observers upgraded ecological classification.

Analyzing the three year period, of 62 90^{th} percentiles, 38 of them preserved the original ecological classification (61%). 8 (13%) observers degrade and 16 (26%) upgrade the original ecological status for water body 010.

Variance values for water body 010 are expectedly high (type II-A). For the first period (90^{th} Pctl=6.02 mg/m³) variance has a value of 5.88 for the second (90^{th} Pctl=4,51 mg/m³) 2.53 and for the third (90^{th} Pctl=8.02 mg/m³) 8.04 The lowest variance value was calculated during the second period, however, so was the lowest percentage of observers that preserved ecological classification. The small buffer value (at the second period – 0.22 mg/m³) causes the low percentage of observers with unchanged ecological classifications. For the first and third annual periods, the percentage of observers that preserve their ecological classifications grows with the variance value. Due to this, it can be concluded that the percentage of unmodified ecological classifications depends (for water body 010) more on the buffer value(s) and less on the variance.

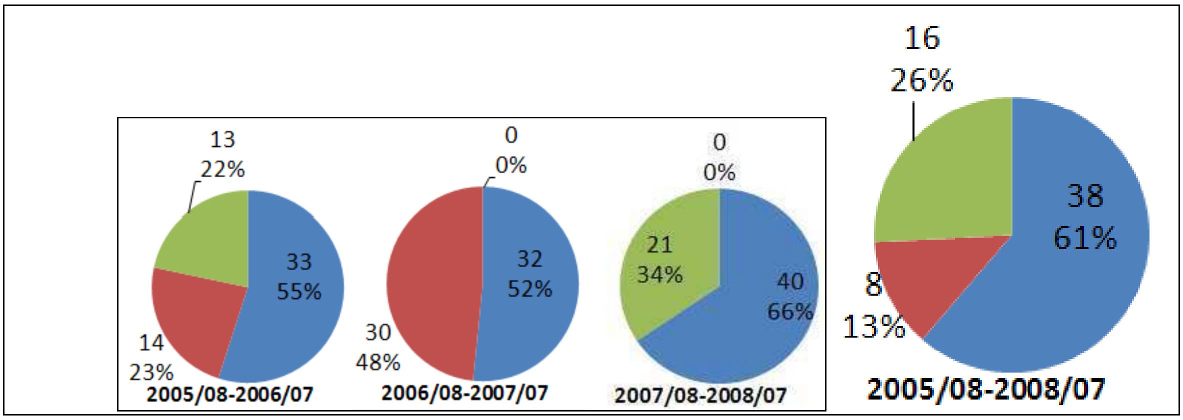


Figure 6: Number (and %) of reduced observers where calculated ecological status corresponds to the original one (blue); number (and %) of reduced observers that degrade ecological status (red); number (and%) of reduced observers that upgrade ecological status (green); water body 010-four periods

The three-year period has considerable High-Good and Good-Moderate buffer values (established by 90thPCTL.-5.91 mg/m³) and a significant variance value (5.69). Due to these values, 24 observers do not preserve the “original“ ecological classification.

In Table 12 the number and percentage of observers (grouped by number of excluded stations) that do not degrade ecological classification for the all four periods and for the three-year period are listed. It is possible to observe that three of six (50%) observers that exclude one sampling station do not modify the original ecological classification established by the reference observer for all four periods. Reducing two sampling stations, the percentage of corresponding ecological classifications lowers by 20,0%. Reducing the number of sampling stations, the percentages of classifications that correspond to the original quality value decrease to 0% (for the 1-station observer). 3- (30%) and 2- (13.33%) sampling station observers have a low percentage of ecological classifications that correspond to the one calculated using all sampling stations. Generally it could be concluded that water body 010 has low percentages of unmodified classifications for all limited observers.

Analyzing only the three-year period, the results are significantly higher, especially for the 5- (83.33%), 4- (93.33%) 3- (60%) and 2-station (46,67%) observers. The 1-station observer did not establish unchanged ecological classifications for either the three-year period or for the four periods.

These results confirm that a significant part of the non-corresponding eco classifications is calculated for the one-year periods (due to the interannual variations).

N. of excluded s. stations	N. of comb. realized for all 4 periods	N. of comb. that do not modify eco. class. (4 periods)	% of comb. that do not modify eco. class. (4 periods)	N. of comb. realized for 3 years period	N. of comb. that do not modify eco. class. (3 years period)	% of comb. that do not modify eco. class. (3 years period)
1	6	3	50,0%	6	5	83,33%
2	15	3	20,0%	15	14	93,33%
3	20	6	30,00%	20	12	60,00%
4	15	2	13,33%	15	7	46,67%
5	3(from 6)	0	0,00%	6	0	0,00%

Table 12: Number and percentage of *limited observers* (grouped by number of excluded stations) that do not degrade ecological classification in any of the four periods considered and in the three-year period considered separately.

The annex "Descriptive statistic results" lists detailed results for water body 010 (all possible observers with 6 sampling stations).

Observing the annex table, the 90th percentile biases calculated for the three-year period for most observers are lower than the 90th percentile biases calculated for the annual periods (at least one). Comparing the 90th percentile biases for the annual periods (annual average 1.16 mg/m^3) with the average bias for the three-year period (0.92 mg/m^3), it is clear that quantity error decreases as sample time increases.

average absolute bias 2005-2006	average absolute bias 2006-2007	average absolute bias 2007-2008	annual average absolute bias	average absolute bias 2005-2008
1,25	0,64	1,59	1,16	0,92

Table 13: Water body 019 interannual average absolute biases; annual average absolute bias, average absolute bias for the three-year period

5.1.8.Water body 011

Water body 011 was surveyed by six sampling stations: DP002a, DP076, DP078, DP079, DP080 and DP081.

For water body 011, 62 possible limited observers were identified . For each observer, a 90^{th} percentile and related ecological status were established (four periods) and compared with the results obtained by the reference observer.

Water body 011 for the second annual period (2006-2007) and three-year period is classified as “Good” ecological quality. For the first and third one-year periods (2005-2006 and 2007-2008) it is classified as “High” ecological quality. For the three-year and second annual period, the ecological classification established by the limited observers can either degrade or upgrade the “original” quality values. For the first and third annual periods, the ecological classification established by the limited observers can only degrade the original classification.

Based on average, standard deviation and buffer values water body 011 has been categorized as a system with ecological classification fuzzy (chapter : “Ecological classification fuzziness”)
For the annual period (August 2005–July 2006), ecological classifications are established for 59 of 62 possible observers, due to a lack of data from DP002a, DP079 and DP081 (three 1-station observers are excluded). 52 (88%) observers calculated unchanged ecological classifications. Seven (12%) observers degraded the original ecological classification.

During the next annual period (August 2006 – July 2007) ecological classifications were established for 59 observer due to the same lack of data as for the 2005-2006 period. 52 (88%) observers preserved the original ecological classification. Five (9%) observers degraded and two (3%) upgraded the original ecological classification.

For the third annual period (August 2007 – July 2008) ecological classifications were established for 60 limited observers. Two 1-station observers are excluded from analysis due to a lack of data from DP078 and DP079. 30 (50%) observers preserve the ecological classification while the other 30 (50%) degrade it.

Analyzing the three year period, 90^{th} percentiles are calculated for 62 observers. The original ecological classification (calculated using six stations) was preserved by 38 (61%) of the limited observers One (2%) observer degraded and 23 (37%) upgraded the original ecological status for water body 011.

Variance values for water body 011 are expectedly low (type III). For the first period (90^{th} Pctl=1.56 mg/m³) variance has a value of 0.46 , for the second (90^{th} Pctl=2.69 mg/m³) 1.07 and for the third period (90^{th} Pctl=2.00 mg/m³) 1.48. For the first and third period the water body is classified as “High” and the 90^{th} percentile determines the buffer value (0.64 mg/m³ for the first and 0.20 mg/m³ for the third period) which is more often exceeded during the third period. The variance is lower than the buffer for the first period, which results in a high percentage of observers that preserve their ecological classifications. The third period has the opposite variance/buffer value situation, which results in 50% of the observers changing their “original” ecological classification. The second period has two considerable buffer values (0.49 at the High/Good and 0.91 mg/m³ at the Good/Moderate threshold), due to which 88% of the observers have unchanged ecological classifications.

The variance value for the three-year period (90th Pctl=2.26 mg/m³) is not as significant (1.03) as the buffer value (0.06 mg/m³) at the High-Good threshold. Because of this, most of the observers that do not preserve their ecological classification upgrade the original status for the three-year period.

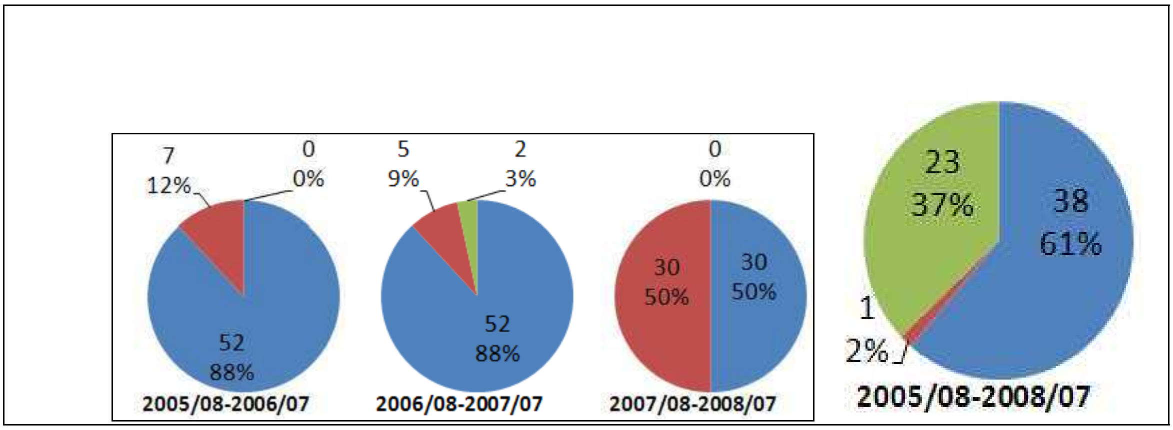


Figure 7: Number (and %) of limited observers where calculated ecological status corresponds to the original one (blue); number (and %) of limited observers that degrade ecological status (red); number (and %) of limited observers that upgrade ecological status (green); water body 011-four periods

Table 14 lists the number and percentage of limited observers (grouped by the number of excluded stations) that do not degrade ecological classification for all four periods and for the three-year period.

It is possible to observe that two of six (33,33%) observers that exclude one sampling station do not modify the original ecological classification calculated using the complete data for all four periods. Reducing two sampling stations, the percentage of corresponding ecological classifications lowers to 20%. Reducing three sampling stations, the percentage of corresponding ecological classifications reduces to 15%. For 2- and 1-station observers, the percentage of classifications that corresponds to the original quality value decreases to 0%. As with water body 010, it can be concluded that 011 has a low percentage of unmodified classifications for all reduced observers.

Analyzing only the three year period, the results are higher for all limited observers. 5-stations observers have 66.67% of ecological classifications that correspond to the quality values calculated using all six sampling stations. 4- stations observers have a higher percentage result (73,33%) than 5-stations observers. 3- and 2-stations observers have 60% of corresponding ecological classifications, while 1-station observers have only 33,33% .

These results confirm that a significant part of the non-corresponding eco classifications is calculated for the one-year periods (due to interannual variations).

N. of excluded s. stations	N. of comb. realized for all 4 periods	N. of comb. that do not modify eco. class. (4 periods)	% of comb. that do not modify eco. class. (4 periods)	N. of comb. realized for 3 years period	N. of comb. that do not modify eco. class. (3 years period)	% of comb. that do not modify eco. class. (3 years period)
1	6	2	33,33%	6	4	66,67%
2	15	3	20,00%	15	11	73,33%
3	20	3	15,00%	20	12	60,00%
4	15	0	0,00%	15	9	60,00%
5	2(from 6)	0	0,00%	6	2	33,33%

Table 14: Number and percentage of *limited observers* (grouped by number of excluded stations) that do not degrade ecological classification in any of the four periods considered and in the three-year period considered separately.

The annex “Descriptive statistic results” lists detailed results for water body 011 (all possible observers with 6 sampling stations).

Observing this annex table, the 90^{th} percentile biases calculated for the three-year period for most observers are lower than the 90th percentile biases calculated for the annual periods (at least one). Comparing the 90^{th} percentile biases for the annual periods (annual average 0.46 mg/m^3) with average bias for the three-year period (0.31 mg/m^3), it is clear that the quantity of error decreases as sample time increases.

average absolute bias 2005-2006	average absolute bias 2006-2007	average absolute bias 2007-2008	annual average absolute bias	average absolute bias 2005-2008
0,26	0,35	0,77	0,46	0,31

Table 15: Water body 011 interannual average absolute biases; annual average absolute bias, average absolute bias for the three years period

5.1.9.Water body 014

Water body 014 was sampled by four sampling stations during the last three years : DP096, DP098, DP100 and DP102. For water body 014, 14 possible limited observers of sampling stations are identified.

For each observer, 90^{th} percentile values and a related ecological status are established for the four periods (three annual and the three-year period).

Water body 014 (chapter "Provisional ecological classifications") is classified as "High" for all four periods, for which reason erroneous ecological classifications (calculated using a reduced number of stations) can only be degraded.

Based on average, standard deviation and buffer values water body 011 has been categorized as a system with stable ecological classification defined (chapter : "Ecological classification fuzziness")

For the first period 13 ecological classifications are established. One 1-station observer is excluded from the analysis due to a lack of data from DP096. 12 analyzed observers (92%) are unchanged, and one (8%) limited observer establishes a degraded ecological status.

During the next annual period (August 2006 – July 2007) 11 (79%) observers of 14 preserve the original ecological classification. Three (21%) of the limited observers degrade the original ecological classification.

For the last annual period, 90^{th} percentile values were calculated 13 times. A 1-station observer was excluded from analysis due to a lack of data from DP102. All 13 ecological classifications were unchanged compared to the classification established using all four stations.

Analyzing the three-year period, all 14 (100%) observers established unchanged ecological classifications.

The variance values for 014 are the lowest compared to the rest of the water bodies. For the first annual period (90^{th} Pctl=1.63 mg/m³) is 0.20 , for the second (90^{th} Pctl=1.82 mg/m³) 0.38 and for the third (90^{th} Pctl=1.83 mg/m³) 0.27 classifications. These values are expected for a stable type III ecosystem. The second annual period has the highest variance value and percentage of degraded ecological classifications.

The variance for the three-year period (90^{th} Pctl=1.79 mg/m³) has a low value (0.28) and a considerable buffer (0.41 mg/m³) for water body 014, which results in all observers preserving the original classification.

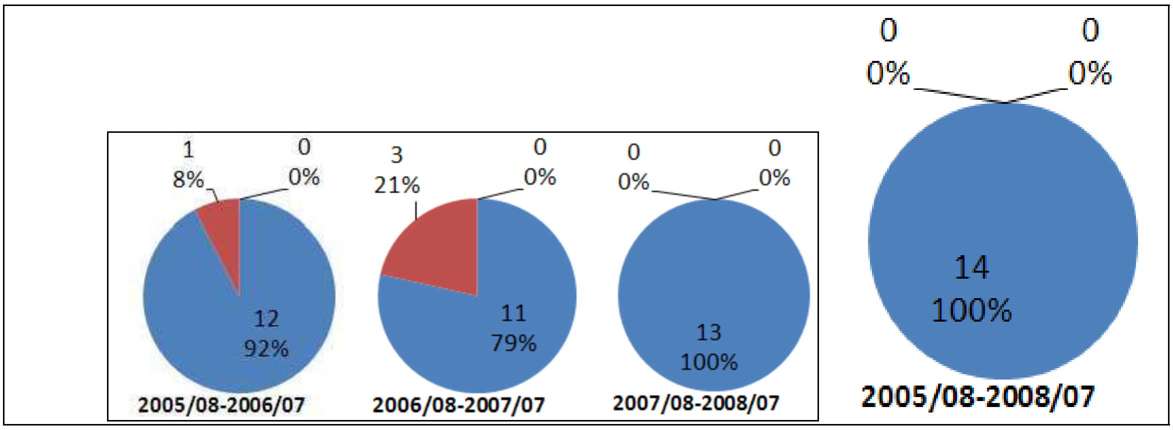


Figure 8: Number (and %) of limited observers where calculated ecological status corresponds to the original one (blue); number (and %) of limited observers that degrade ecological status (red); number (and %) of limited observers that upgrade ecological status (green); water body 014-four periods

Table 16 lists the number and percentage of limited observers (grouped by number of excluded stations) that do not degrade ecological classification for all four periods and for the three-year period.

It is possible to observe that all four observers that exclude one sampling station do not modify the ecological classification calculated using the complete data for all four periods. Reducing the number of sampling stations , the percentage of corresponding ecological classifications lowers to 83.33% (for 2-stations observers) and 50% (for 1-station observers).
Analyzing only the three-year period and all limited observers, it can be observed that all types of limited observers classify the water body as "High" and do not degrade the ecological classification calculated using four stations.

N. of excluded s. stations	N. of comb. realized for all 4 periods	N. of comb. that do not modify eco. class. (4 periods)	% of comb. that do not modify eco. class. (4 periods)	N. of comb. realized for 3 years period	N. of comb. that do not modify eco. class. (3 years period)	% of comb. that do not modify eco. class. (3 years period)
1	4	4	100,00%	4	4	100%
2	6	5	83,33%	6	6	100%
3	2	1	50,00%	4	4	100%

Table 16: Number and percentage of limited observers (grouped by number of excluded stations) that do not degrade ecological classification in any of the four periods considered and in the three-year period considered separately.

The annex "Descriptive statistic results" lists detailed results for water body 014 (all possible observers with 4 sampling stations).

Observing the annex table, the 90th percentile biases calculated for the three-year period for most observers are lower than the 90th percentiles biases calculated for the annual periods (at least one). Comparing the 90th percentile biases for the annual periods (annual average 0.23 mg/m^3) with the average bias for the three year period (0.05 mg/m^3), it is clear that quantity error decreases as sample time increases.

average absolute bias 2005-2006	average absolute bias 2006-2007	average absolute bias 2007-2008	annual average absolute bias	average absolute bias 2005-2008
0,27	0,32	0,11	0,23	0,05

Table 17: Water body 014 interannual average absolute biases; annual average absolute bias, average absolute bias for the three-year period

5.1.10.Water body 019

Water body 019 was sampled by four sampling stations during the last three-years: DP125, DP127, DP129 and DP131. For water body 019, 14 possible limited observers of sampling stations are identified.

For each observer, a 90th percentile value and related ecological status are calculated for the four periods (three annual and the three-year period). Water body 019 (chapter “Provisional ecological classifications”) is classified as “Good” for all four periods, for which reason erroneous ecological classifications can be either degraded or upgraded.

Based on average, standard deviation and buffer values water body 011 has been categorized as a system with ecological classification fuzzy (chapter: “Ecological classification fuzziness”)
For the first period, all 14 possible ecological classifications are established. Eight (57%) of them are unchanged and six (43%) establish a degraded ecological status.

During the next annual period, seven (50%) observers of 14 preserve the original ecological classification. Seven (50%) of the limited observers degrade the original ecological classification.

For the last annual period, 90th percentile values were calculated 14 times, and the ecological classification was unchanged for 11 (79%) observers. Three (21%) ecological classifications degraded the ecological classification established by all four stations.

Analyzing the three year period, 11 of 14 observers (79%) established unchanged ecological classifications. Three (21%) of the limited observers degraded the original ecological classification.

The variance values for 019 are significantly higher compared to water bodies 014 and 011. For the first annual period (90th Pctl=3.39 mg/m³) the value is is 4.83 for the second (90th Pctl=3.25 mg/m³) 1.54 and for the third (90th Pctl=2.46 mg/m³) 0.92 . For the first two annual periods, the buffer for the Good/Moderate threshold is significantly lower than the buffer for the High/Good threshold. Because of this, all modified classifications degrade the ecological classification established by the reference observer. The third period has a significantly lower buffer for the High/Good threshold, however all three observers that did not preserve the original status degrade the ecological classification. Analyzing the 90th percentile biases for the third period, it can be observed that all elevated errors are negative (degraded quantity and quality value) while positive errors do not exceed 0.1 mg/m³.

The variance for the three-year period (90thPCTL.-3.10 mg/m³) has a value of 2.41 and the buffer value for the Good/Moderate threshold is significantly lower. Due to this, all limited observers that modify the "original" ecological classification degrade the water body status.

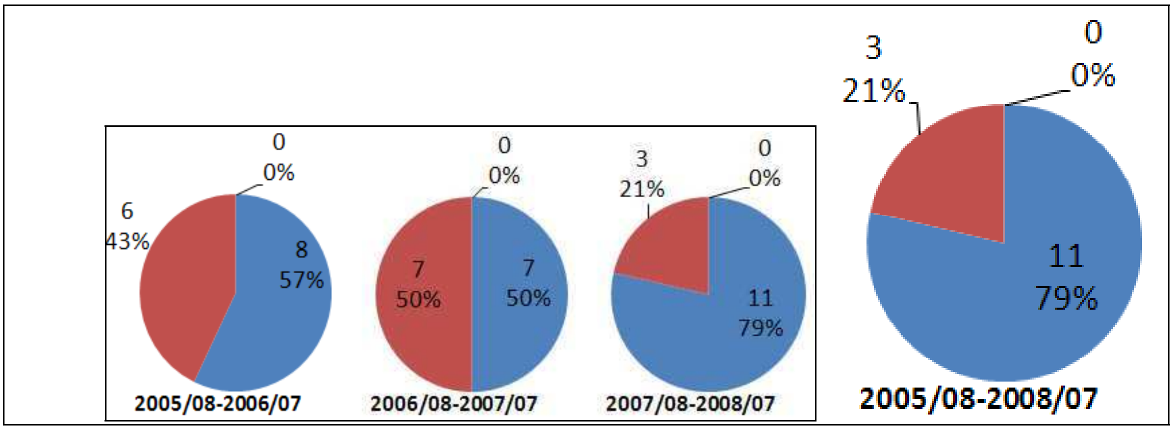


Figure 9: Number (and %) of limited observers that calculated eco.status corresponds to the original one (blue); number (and %) of limited observers that degrade eco. status; water body 019-four periods

Table 18 lists the number and percentage of limited observers (grouped by number of excluded stations) that do not degrade the ecological classification for all four periods and for the three-year period. It is possible to observe that two of the four observers that exclude one sampling station do not modify the ecological classification calculated using the complete data for all four periods. Reducing the number of sampling stations in the observers, the percentage of corresponding ecological classifications reduces to 33,33% (for 2-stations observers) and to 25% (for 1-station observers).

Analyzing only the three-year period, the results are significantly higher. 3-stations observers correspond 100% to the ecological classifications calculated using four sampling stations. 2-station and 1-station observers have percentages (83% and 50%) that are higher from the ones calculated using four periods. These results (as with the majority of water bodies in this analysis) confirm that the majority of non-corresponding eco classifications are calculated for the one-year periods.

N. of excluded s. stations	N. of comb. realized for all 4 periods	N. of comb. that do not modify eco. class. (4 periods)	% of comb. that do not modify eco. class. (4 periods)	N. of comb. realized for 3 years period	N. of comb. that do not modify eco. class. (3 years period)	% of comb. that do not modify eco. class. (3 years period)
1	4	2	50,00%	4	4	100%
2	6	2	33,33%	6	5	83,33%
3	4	1	25,00%	4	2	50%

Table 18: Number and percentage of *limited observers* (grouped by number of excluded stations) that do not degrade ecological classification in any of the four periods considered and in the three-year period considered separately.

The annex "Descriptive statistic results" lists detailed results for water body 019.

Observing this annex table, the 90th percentile biases calculated for the three-year period for most observers are lower than the 90th percentile biases calculated for the annual periods (at least one). Comparing the 90th percentile biases for the annual periods (annual average 0.72 mg/m³) with the average bias for the three-year period (0.44 mg/m³), it is clear that quantity error decreases as sample time increases.

average absolute bias 2005-2006	average absolute bias 2006-2007	average absolute bias 2007-2008	annual average absolute bias	average absolute bias 2005-2008
1,09	0,56	0,51	0,72	0,44

Table 19: Water body 019 interannual average absolute biases; annual average absolute bias, average absolute bias for the three-year period

5.1.11.Discussion

- Analyzing annual periods, the ecological classification established using limited observers often modify the ecological classification established by the reference observer Due to interannual variations (analyzed in the chapter “Seasonal pattern regularity”) the same observer do not preserve the “original” quality status for all three annual periods. Analyzing the longer period (three years) the interannual variations do not have the same influence on the 90th percentiles and on the ecological classification of the water body (quality value). The quality results for the three-year period obtained through limited observers preserve the “original” ecological classification more frequently (extreme values do not influence the quality results for the longer time the mime as they do for the annual periods). These results indicate that sampling station reduction is more adequate for longer periods.
- The percentages of limited observers (that do not upgrade or degrade ecological classification for the four periods) are different for each water body and in general depend on the stability of the system ecological status (water body). Results can be compared with temporal reduction analysis (chapter “Ecological classification fuzziness”), where the results of a reduction in the frequency of the campaigns depend on the stability of the system ecological status
- “Stable systems” have the highest number of limited observers that preserve ecological status, those being analyzed water bodies 001, 002 and 014. For all four analyzed periods, these water bodies are classified as “High”. Due to ecological classification, the “High/Good” threshold is the only thing that can be exceeded and change the quality value. All three water bodies have a low variance and significant buffer distance (at the High/Good threshold), which reduces the percentage of limited observers that do not preserve the ecological classification established using the complete data. In these systems, for the three-year period, the percentage of overlapping classifications is above 90%. It is possible to study in them, the reduction of two to three sampling stations.
- In the only “Unstable system”, in which we studied the possibility of spatial reduction (w.b. 008), the percentage of classifications that have been altered to reduce one or two sampling stations is 80%.
- The rest of the analyzed water bodies were classified as unstable systems. The results of this analysis demonstrate that there are a significantly lower percentage of observers that can provide the same ecological classifications for all four periods. In these systems, for the three-year period, the percentage of overlapping classifications is under 80%. It is possible to study, for the three-year period, the reduction of one or at most two stations in water body 005, 009 y 019. The water bodies 010 and 011 require a more detailed study of the sampling stations, as the percentage of matching classifications is 61% and the changes that occur, depending on number of stations that are removed, is very high.

5.2.Selecting the most appropriate k-observers

5.2.1.Introduction

The comparison between the 90th percentile of the limited observers, and the 90th percentile of the reference observer, has allowed us preselecting the observers with less bias and to discard the observers with more bias. Since the results of the preliminary study indicated that better results were obtained when working with a three-year period, we have focused the study in this period.

The order established between the pre-selected observers, based on the 90th percentile of the empirical CDF, may be easily modified with small changes in the data. The reason is that this order is based only on the value of the 90th percentile of the empirical CDF in a relatively small sample. That is why, we have tried to find a more stable ordination, based on the proximity between the monthly averages of chlorophyll *a* measurements collected at the sampling stations checked every month.

In order to locate limited observers with a complete and unbiased known of chlorophyll *a* level in each water body and with a high concordance with the reference observer, it has been analyzed, in each water body, the distribution of the measurements taken in the different sampling stations, it has been calculated the ICC and the Euclidean distance between the monthly average of the measurements taken by the reference observer and the preselected observers, and it has been represented the Euclidean distances in a two-dimensions space using MDS.

The 90th percentiles of the measurements taken by the selected observers, and the 90th percentile of the reference observe has been compared with a Quantile Plot.

Approach to establishing the optimal k-observer of sampling stations based on monthly average

The order established between the pre-selected combinations, based on the 90th percentile of the empirical Cumulative Distribution Function (CDF), can be easily modified by small changes in data. The reason for is that the order is based only on the value of the 90th percentile of the empirical CDF in a relatively small sample. For this reason, a more stable ordination has been proposed, based on the proximity between the monthly averages of chlorophyll *a* measurements collected at the sampling stations.

As mentioned in paragraph 4.4, the problem has been approached from the perspective of reproducibility of measurements or of inter-observer variability. From the standpoint of space reduction, it has been considered that the different k-observers of sampling stations are different *observers* that provide different monthly measurements of chlorophyll *a* concentrations in a given water body. Therefore, the aim has been to quantify the concordance between the measures taken by the *observer,* which corresponds to the *reference set of sampling stations* and the measures taken by the *observers,* which correspond to the preselected *k-observers* of the *reference set*.

The information about the status of a water body in a given month, obtained by a particular *observer,* is synthesized by averaging the measurements of *chlorophyll a* collected in the subset of sampling stations that characterize this *observer.*

The average is a statistic sensitive to outliers and it will allow us to know if the sampling stations that characterize an observer contain the extreme values of the original dataset. If, in months when abnormally high concentrations of chlorophyll *a* have been measured in the water body, there is proximity between the monthly averages for two different observers, this indicates that both observers have detected these alterations, or rather that the campaigns in which these values have been obtained belong to entities that characterize the two observers.

Moreover, as mentioned in paragraph 4.4, if, month by month, the averages of measurements made in the sampling stations that characterize an observer are similar to the averages of measurements collected at the sampling stations that characterize the reference observer, this will be an indicator of proximity between the empirical CDF of measurements collected in the group of stations that characterize the observer and the empirical CDF of the original dataset.

The following have been considered as observers of the water body: the reference observer, characterized by the all stations of the water body, and preselected observers, characterized by those combinations that have obtained the lowest bias for the 90th percentile. For each campaign, the monthly average has been calculated for the measurements taken by the reference observer and they have been compared with the monthly average for the measurements taken by the preselected observers.

Two measures have been considered to compare the proximity or distance between the monthly averages of two different observers: Euclidean distance and the interclass correlation coefficient.

Interclass correlation coefficient

The correlation coefficient is considered as a proximity measure, and it is calculated for the monthly average of the *preselected observers* and the monthly average of the *reference observer*. As mentioned in paragraph 4.4, to obtain more accurate results, this analysis uses the interclass correlation coefficient (ICC). ICC considers not only if the individuals change their values in parallel (which can give us an inaccurate picture); but also the absolute measured values and their variance. All these factors have to be similar to obtain a high correlation coefficient.

Euclidean distance and Multidimensional Scaling

Calculating and comparing the Euclidean distance between chlorophyll *a* average value series of analyzed k-observers, the output should confirm the results of the ICC analysis. Results will be presented using the Multidimensional Scaling technique. In the two-dimensional space in which the distances are presented, it is easy to identify which observers verify that the Euclidean distance between the monthly averages of chlorophyll *a* concentration in their stations and the monthly averages of chlorophyll *a* concentration in the sampling stations of the reference observer is less relevant.

5.2.2.Analysis and results

Analysis for Water Body 001

Water Body 001 is surveyed by five different sampling stations (Annex 2). It is classified as a type II-A water body (significant continental influence and salinity variance). The previous exploratory study detected that spatial reduction is possible without consequences to ecological classification. The following analysis will attempt to identify the most adequate k-observers for the spatial reduction.

<u>Comparative study</u>

Multiple box-and-whisker plots have been used to compare the range, median, quartiles and outliers of the measurements of *chlorophyll a* obtained from the five sampling stations located in water body 001. Graphics have been made for the three annual periods and the three-year period.

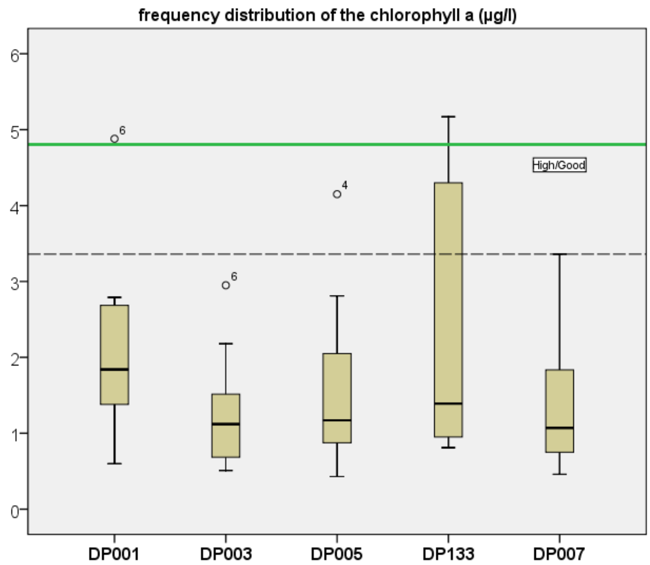


Figure 1: Box-and-whisker plot of chlorophyll a concentration for the annual period of August 2005 - July 2006 in water body 001; The green line represents the High/Good threshold, and the broken line represents the 90th percentile

During the annual period of August 2005-July 2006, it is possible to observe that the highest concentrations of chlorophyll *a* in water body 001 have been observed mainly at station DP133 (Figure 1). Some samples analyzed for DP001 and DP133 have concentrations that exceed the High/Good threshold.

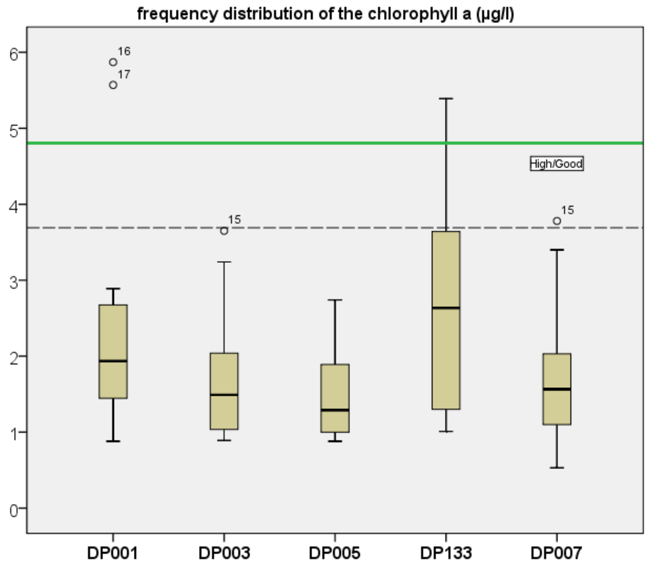


Figure 2:Box-and-whisker plot of chlorophyll a concentration for the annual period of August 2006 - July 2007 in water body 001; The green line represents the High/Good threshold, and the broken line represents the 90th percentile

For the second annual period, it is clear that the highest concentrations of chlorophyll *a* have been taken by sampling stations DP001 and DP133 (Figure 2). Only stations DP001 and DP133 took samples that exceeded the High/Good threshold.

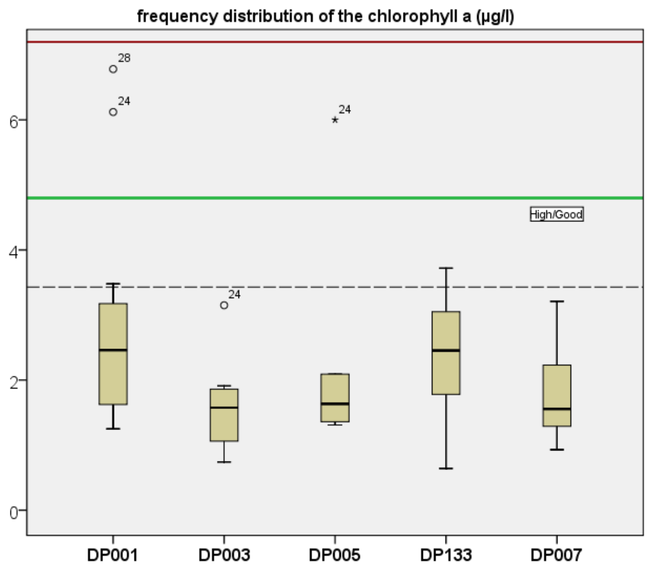


Figure 3: Box-and-whisker plot of chlorophyll a concentration for the annual period of August 2007 - July 2008; water body 001; The green line represents the High/Good threshold, the red line represents the Good/Moderate threshold, and the broken line represents the 90th percentile

For the third annual period, the box-and-whisker plot shows that the highest concentrations of chlorophyll *a* have been recorded by sampling stations DP133, DP005 and DP001 (Figure 3). During the August 2007 campaign, samples analyzed for DP001 and DP005 have chlorophyll *a* concentrations that exceed the High/Good threshold.

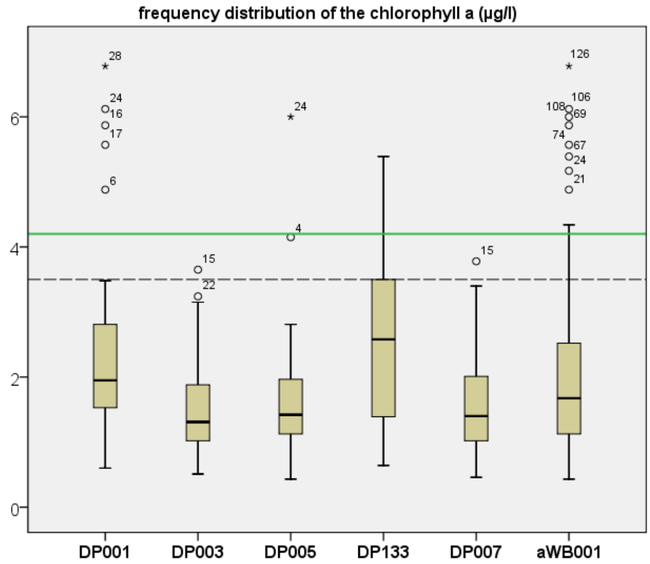


Figure 4: Box-and-whisker plot of chlorophyll a concentration for the three year period of August 2005 - July 2008 in water body 001; The green line represents the High/Good threshold, and the broken line represents the 90th percentile

The box-and-whisker plot for the three year period clearly shows that the highest concentrations of chlorophyll *a* for water body 001 were recorded mostly by sampling stations DP001 and DP133 (Figure 4). During just one campaign (August 2007), sampling station DP005 took a sample which had a concentration of chlorophyll *a* that exceed 4.8 µg/l (High/Good threshold). It is clear that sampling stations can be divided into two groups. The first group consists of stations DP003, DP005 and DP007, which have lower values for the water body. The second group consists of DP001 and DP133, which have significantly higher values than the first group.

On the last box-and-whisker plot, aWB001 is included (box-and-whisker plot for all of water body 001). As commented on *in chapter 4.4,* in order to attain the greatest similarity between the 90^{th} percentile of the complete dataset and the 90^{th} percentile of a subset of the complete dataset, it is desirable that the empirical CDF of chlorophyll *a* concentration in the complete dataset does not differ substantially from the empirical CDF of this subset, especially in the higher percentiles. Additionally, it would also be advantageous if the subsample included the greatest values of the original sample. Therefore, the measurements taken in the optimal k-observer of sampling stations should have an empirical CDF similar to the empirical CDF of aWB001.

The overlapping graphs corresponding to the empirical CDF of each sampling station allow us to compare the 90th percentiles of these distributions (Figure 5). The datasets taken by the sampling stations that belong to the first group have a 90^{th} percentile lower than the 90^{th} percentile of the reference dataset, while the datasets taken by the sampling stations that belong to the second group have a 90^{th} percentile higher than this value. This tells us that if we want to select a k-observer of sampling stations where the set of measurements taken have a 90^{th} percentile similar to the 90^{th} percentile of the full sample, it will be necessary to select sampling stations of both groups.

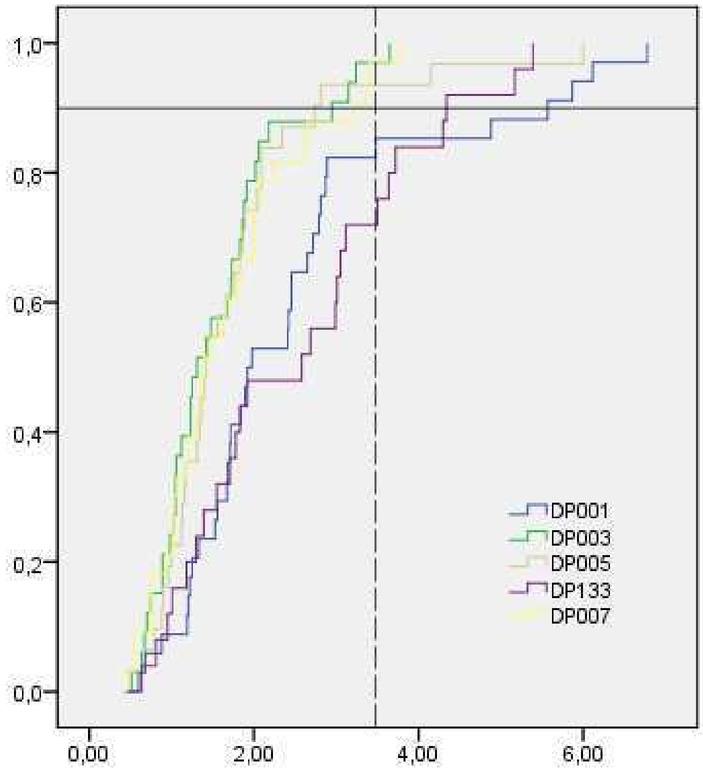


Figure 5: Quantile plot of chlorophyll a concentration for all sampling stations for the three year period August 2005 - July 2008, water body 001

<u>*Pre-selection of k-observers of the sampling stations based on the deviation of the 90th percentile*</u>

Observing 5.2.2.Table 1 (taken from the Exploratory study and aligned by the three year bias), the results confirm what was easily observable on the previous quantile plot – the k-observers (where k<4) have a lower bias on the 90th percentile if they include sampling stations from both groups. Also, it can be observed that these combinations include only one sampling station that belongs to the second group, DP001 or DP133.

k-observer	Absolute bias PCTL (3 years)	eco. class. (3 years)	- n stations	average bias (4 periods)	Corresponding eco. class. (4 periods)
DP001 DP007	**0,002**	**corresp.**	**3**	**0,898**	**no**
DP001 DP003	**0,03**	**corresp.**	**3**	**1,025**	**no**
DP003 DP133 DP007	**0,06**	**corresp.**	**2**	**0,140**	**yes**
DP005 DP133 DP007	**0,10**	**corresp.**	**2**	**0,285**	**yes**
DP003 DP005 DP133	**0,10**	**corresp.**	**2**	**0,283**	**yes**
DP001 DP003 DP133 DP007	**0,10**	**corresp.**	**1**	**0,130**	**yes**
DP003 DP133	**0,12**	**corresp.**	**3**	**0,495**	**yes**
DP001 DP005 DP007	**0,13**	**corresp.**	**2**	**0,755**	**no**
DP003 DP005 DP133 DP007	**0,15**	**corresp.**	**1**	**0,135**	**yes**
DP001 DP003 DP005 DP133	**0,15**	**corresp.**	**1**	**0,505**	**yes**
DP133 DP007	0,18	corresp.	3	0,520	yes
DP001 DP003 DP007	0,19	corresp.	2	0,178	yes
DP001 DP005 DP133 DP007	0,21	corresp.	1	0,478	yes
DP007	0,24	corresp.	4	0,170	yes
DP001 DP003 DP005	0,28	corresp.	2	0,893	no
DP001 DP003 DP005 DP007	0,31	corresp.	1	0,325	yes
DP003 DP007	0,37	corresp.	3	0,328	yes
DP003	0,47	corresp.	4	0,397	yes(3periods)
DP005 DP007	0,53	corresp.	3	0,353	yes
DP003 DP005 DP007	0,55	corresp.	2	0,418	yes
DP001 DP003 DP133	0,58	corresp.	2	0,838	no
DP001 DP133 DP007	0,60	corresp.	2	0,783	no
DP005 DP133	0,65	corresp.	3	0,738	yes
DP003 DP005	0,66	corresp.	3	0,428	yes
DP005	0,75	corresp.	4	0,763	yes(3periods)
DP001 DP005 DP133	0,79	corresp.	2	1,455	no
DP001 DP005	0,90	corresp.	3	1,265	no
DP133	1,13	corresp.	4	1,365	no
DP001 DP133	1,63	degrade	3	1,940	no
DP001	2,18	degrade	4	2,163	no

Table 1: **90th percentile biases calculated by k- combinations and ecological classification correspondence for water body 001;**

The highest biases of the 90th percentile have been obtained for the 1-observers that include DP001 or DP133 and the 2- combination DP001, DP133, because they only include sampling stations from the second group. The 3- combination DP001, DP005, DP133, which includes 2 stations from second group and one from the first, has the highest bias. However, the lowest bias of the 90th percentile belongs to the 4- combination DP001, DP003, DP007, DP133, excluding DP005. These results might indicate that DP005 has the lowest influence on the 90th percentile for w.b.001 due to the fact that the station includes outliers.

The low bias of the 90th percentile lets us suppose that the possibility of spatial reduction exists, and the first 10 k-observers with the lowest bias are pre-selected for a more in-depth analysis. Nevertheless, the same analysis will be performed for the k-observers with the highest 90th percentile bias in order to highlight the difference.

__Interclass correlation coefficient__

Results obtained by the ICC have been interpreted using a *qualitative scale* developed by *Landis and Koch*, typical for these types of studies. According to this scale, correlation between items is considered very good (0.90 to 1), good (0.80 to 0.90), moderate (0.60 to 0.80) or poor (0 to 0.60). k-observers that have an ICC > 0,950 are marked as Very Good*.

k-observer	I.Correlation coeffic.	Confidence interval 95%		
		Lower	Upper	Concordance
DP001 DP007	**,957c**	**,914**	**,979**	**Very Good***
DP001 DP003	,914c	,829	,957	Very Good
DP003 DP133 DP007	,936c	,861	,969	Very Good
DP001 DP003 DP133 DP007	**,986c**	**,972**	**,993**	**Very Good***
DP003 DP005 DP133	**,963c**	**,926**	**,981**	**Very Good***
DP005 DP133 DP007	**,950c**	**,900**	**,975**	**Very Good***
DP003 DP133	,936c	,873	,968	Very Good
DP001 DP005 DP007	**,975c**	**,951**	**,988**	**Very Good***
DP133 DP007	,918c	,837	,959	Very Good
DP001 DP003 DP007	**,979c**	**,958**	**,989**	**Very Good***
DP133	*,813c*	*,553*	*,919*	*Good*
DP001 DP133	*,861c*	*,546*	*,944*	*Good*
DP001	*,784c*	*,520*	*,897*	*Moderate*

Table 2: ICCs of the preselected k-observers, confidence intervals and interpretation of ICCs according to the Landis and Koch qualitative scale.

As expected, the 4-observer that excludes DP005 has the highest ICC (0.986). The two 3-observers with high ICC are DP001,DP005,DP007 (0.975) and DP001,DP003,DP007 (0.979). The first one includes one sampling station from the first group, two from the second group, but one station (DP005) includes outliers. The second one has two from the first group and one from the second group. The 2-station combination that has the highest ICC coefficient is DP001, DP007 (0.957). All 10 *preselected observers* have an ICC value higher than 0.9 and have “Very Good” concordance with the *reference observer*. Six k-observers have been selected according to ICC that represent a good alternative for the reference observer.

For the preselected observers that produce the highest bias with respect to the reference 90th percentile for the three year period (DP133; DP001,DP133 and DP001), “Good” or even “Moderate” concordance to the complete data has been detected. The ICC coefficients calculated for the k-observers are 0.813; 0.861 and 0.784 respectively.

Euclidean distance and MDS

Calculating and comparing the Euclidean distance between chlorophyll *a* average series of analyzed k-observers, the output confirms the results of the previous analysis. Results are presented using the Multidimensional Scaling technique.

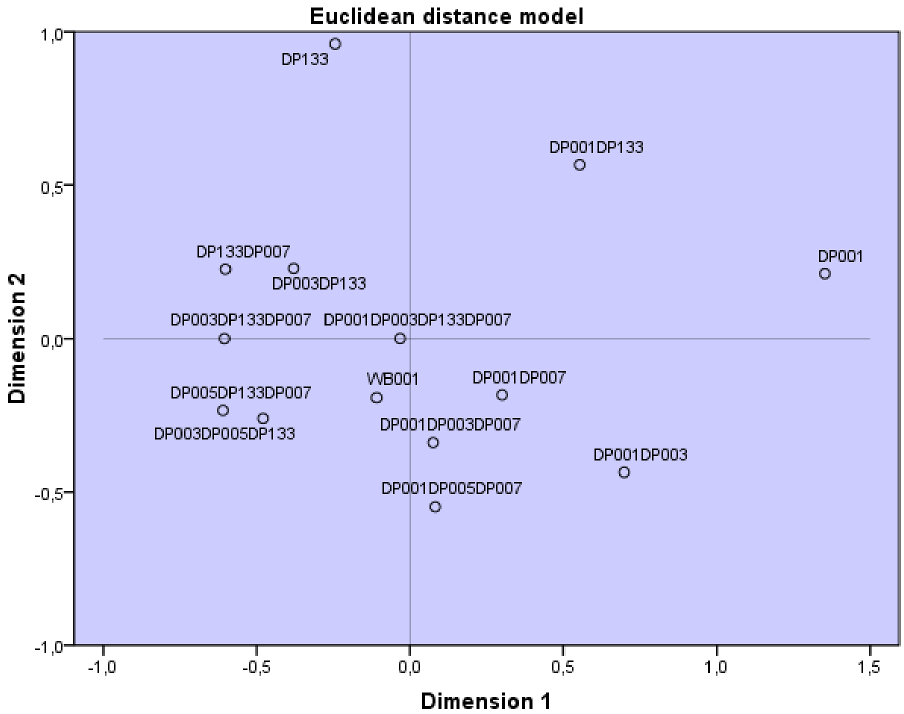


Figure 6: MDS for the Euclidian distance model for the pre-selected k-observers, w.b.001

The 4-station combination that excludes DP005 has the lowest Euclidean distance compared to reference observer WB001. Combination DP001, DP003, DP007 has the lowest Euclidian distance of the 3-observers, and DP001, DP007 has the lowest Euclidian distance of the 2-observers.
The two axes are associated to sampling stations DP001 and DP133 respectively.

Optimal k-observer of sampling stations for water body 001

The previous analyses confirmed that spatial reduction is possible for water body 001 and there are various alternatives. The best results for the spatial reduction have been obtained by the 4-observer that excludes DP005.

Finally the 6 k-observers are selected as most appropriate for the spatial reduction :

Sampling stations	N. of reduced s.s.	90th Percentile
DP001 DP003 DP005 DP133 DP007	**0**	**3,54**
DP001 DP003 DP133 DP007	1	3,64
DP001 DP003 DP007	2	3,35
DP001 DP005 DP007	2	3,41
DP003 DP005 DP133	2	3,64
DP005 DP133 DP007	2	3,64
DP001 DP007	3	3,54

Table 3: Selected k-observers and the 90th percentiles obtained by their data subsets

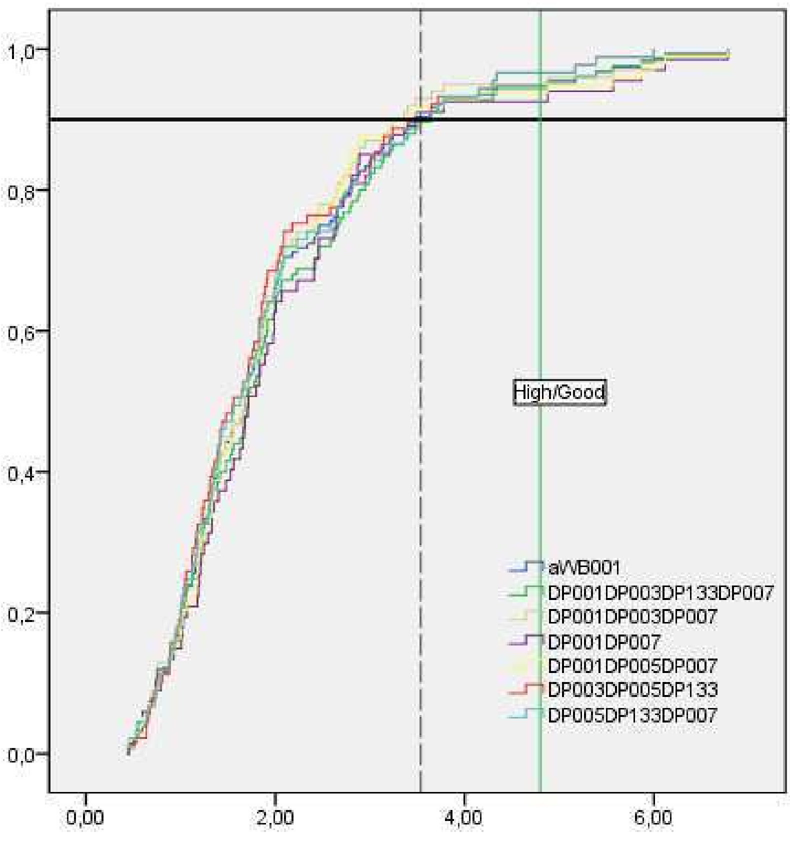


Figure 7- Quantile plot for the selected k-observers of chlorophyll a concentration for the three year period of August 2005 - July 2008, water body 001

Finally, the quantile plot (Figure 7) shows that the empirical CDF of the data subset taken by the selected k-observers is similar to the empirical CDF of the “original” dataset taken by the reference observer. Also, the 90th percentiles of these subsets are close to the 90th percentile of the reference dataset, and the biases are irrelevant.

Analysis for Water Body 002

Water Body 002 is surveyed by five different sampling stations (Annex 2). It is classified as a type II-A water body (significant continental influence and salinity variance). The previous exploratory study determined that a spatial reduction is possible without consequences to the ecological classification. The next analysis will attempt to identify the most adequate k-observers for the spatial reduction.

Comparative study

Multiple box-and-whisker plots have been used to compare the range, median, quartiles and outliers of the measurements of *chlorophyll a* obtained from the five sampling stations located in water body 002. Graphics have been made for the three annual periods and the three-year period.

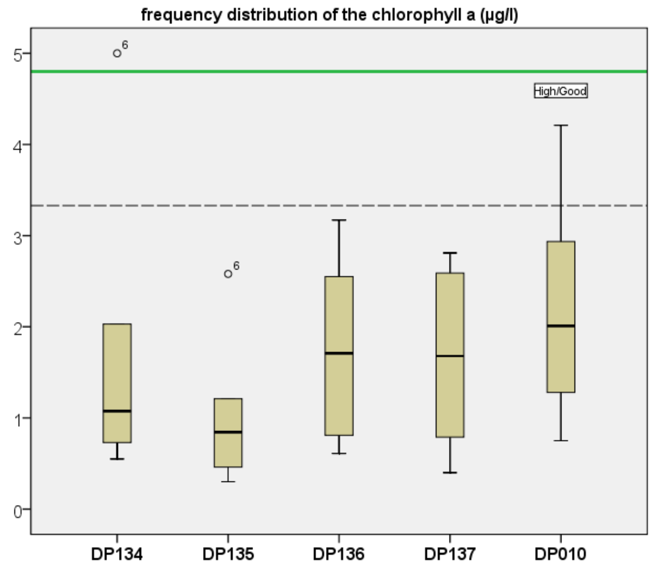


Figure 8: Box-and-whisker plot of chlorophyll a concentration for the annual period of August 2005 - July 2006 in water body 002; The green line represents the High/Good threshold, and the broken line represents the 90th percentile

During the annual period of August 2005-July 2006, it is possible to observe that the highest concentrations of chlorophyll *a* in water body 002 have been observed mainly at station DP010 (Figure 8). Of all stations, only DP134 analyzed one sample that exceeded the High/Good threshold, therefore the other samples represented low chlorophyll *a* concentrations.

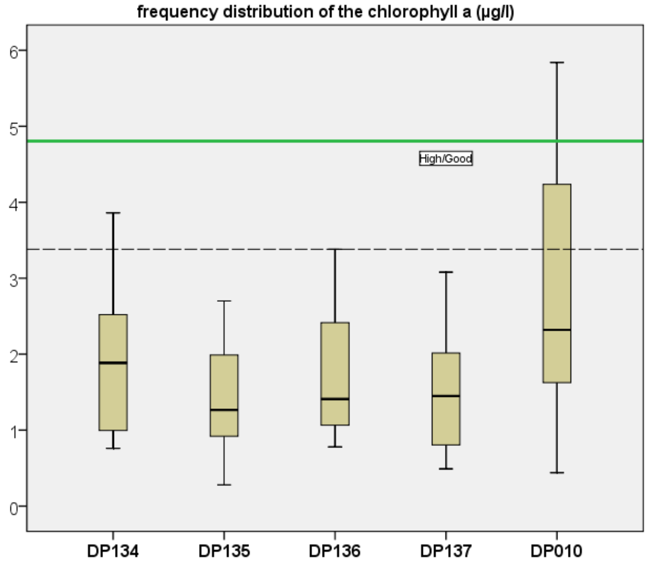


Figure 9:Box-and-whisker plot of chlorophyll a concentration for the annual period of August 2006 - July 2007 in water body 002; The green line represents the High/Good threshold, and the broken line represents the 90th percentile

For the second annual period, it is clear that the highest concentrations of chlorophyll *a* have been taken by sampling stations DP010 and DP134 (Figure 9). Only station DP010 took samples that exceeded the High/Good threshold.

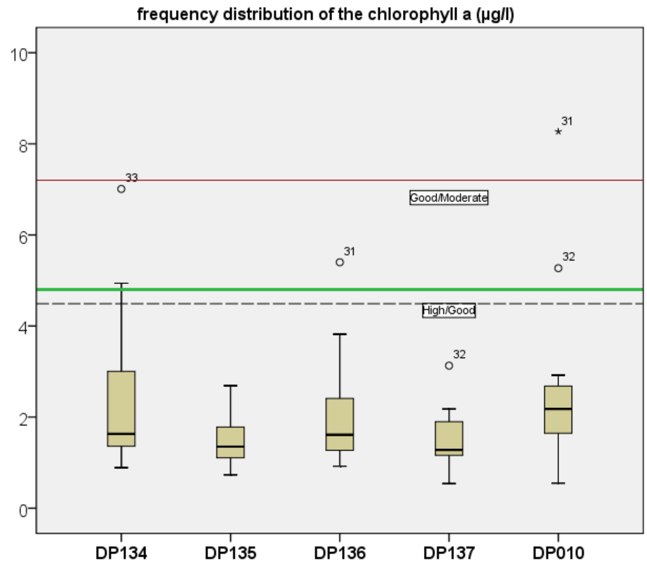


Figure 10: Box-and-whisker plot of chlorophyll a concentration for the annual period of August 2007 - July 2008; water body 002; The green line represents the High/Good threshold, the red line represents the Good/Moderate threshold, and the broken line represents the 90th percentile

For the third annual period, the box-and-whisker plot shows that the highest concentrations of chlorophyll *a* have been recorded by sampling stations DP010, DP134 and DP136 (Figure 10). All three sampling stations detected concentrations that exceed the High/Good threshold.

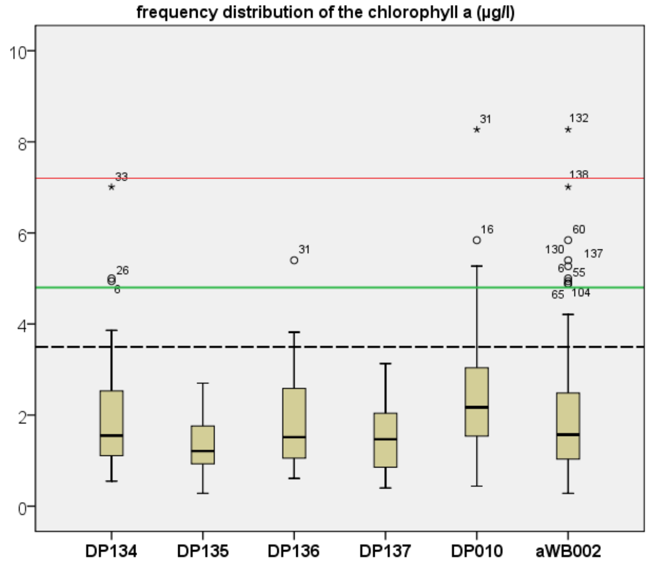


Figure 11: Box-and-whisker plot of chlorophyll a concentration for the three year period of August 2005 - July 2008 in water body 001; The green line represents the High/Good threshold, the red line represents the Good/Moderate threshold, and the broken line represents the 90th percentile

The box-and-whisker plot for the three year period clearly shows that the highest concentrations of chlorophyll *a* for water body 001 were recorded mostly by sampling stations DP010, DP134 and DP136 (Figure 11). It is also clear that DP135 and DP137 belong to the group with the lowest chlorophyll a concentrations.

The optimal k-observer of sampling stations should have an empirical CDF similar to empirical CDF of aWB002, which is included in the last box-and-whisker plot.

The overlapping graphs corresponding to the empirical CDF of each sampling station, allow us to compare the 90th percentiles of these distributions (Figure 12). The datasets taken by the sampling stations that belong to the first group have a 90th percentile lower than the 90th percentile of the reference dataset, while the datasets taken in the sampling stations that belong to the second group have a 90th percentile higher than this value. Observing the quantile diagram, it is not quite clear to which group DP136 belongs, mainly due to the 90th percentile, which is close to the 90th percentile of the w.b.

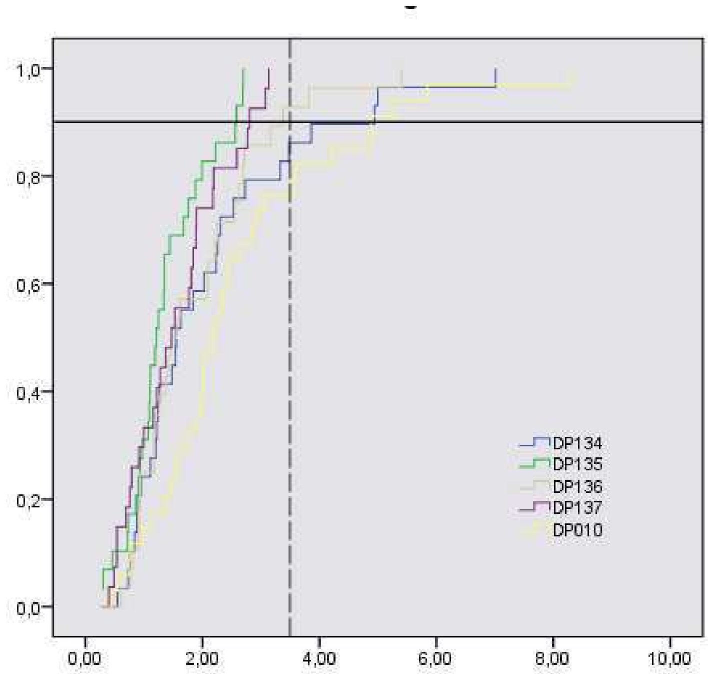


Figure 12: Quantile plot of chlorophyll a concentration for all sampling stations for the three year period August 2005 - July 2008, water body 002

Pre-selection of k-observers of the sampling stations based on the deviation of the 90th percentile

Observing Table 5(taken from the Exploratory study and aligned by the three year bias), the results confirm what was easily observable on the previous quantile plot – the k-observers (where k<4) have a lower bias on the 90th percentile if they include sampling stations from both groups. Also, it is visible that these combinations include only one sampling station that belongs to the second group, DP010 or DP134.
Therefore, the quantile diagram shows that 1-observer DP136 will have a low 90th percentile bias. It should be noted that the 4-observer that excludes DP136 has the lowest 90th percentile bias for all k-observers.

As with water body 001, the highest 90th percentile biases have been obtained by the 1-combinations that include DP010 or DP134 and the 2-observer DP010, DP134 because they only include sampling stations from the second group.
DP010,DP134,DP136 has the highest bias for the 3-observers, which is to be expected due to the hierarchy of empirical CDFs of the sampling stations (Figure 12).

The low bias of the 90th percentile lets us suppose that the possibility of spatial reduction exists, and the first 10 k-observers with the lowest bias are pre-selected for a more in-depth analysis. Nevertheless, the same analysis will be performed for the k-observers with the highest 90th percentile bias in order to highlight the difference.

k-observer	bias PCTL (3 years)	eco. class. (3 years)	- n stations	average absolute error (4 periods)	Corresponding eco. class. (4 periods)
DP134 DP135 DP137 DP010	**0,072**	**corresp.**	**1**	**0,140**	**yes**
DP136	**0,074**	**corresp.**	**4**	**0,267**	**no**
DP135 DP136 DP010	**0,088**	**corresp.**	**2**	**0,275**	**yes**
DP136 DP137 DP010	**0,092**	**corresp.**	**2**	**0,348**	**no**
DP134 DP137	**0,123**	**corresp.**	**3**	**0,403**	**yes**
DP134 DP136 DP137	**0,143**	**corresp.**	**2**	**0,145**	**yes**
DP134 DP135 DP136	**0,153**	**corresp.**	**2**	**0,145**	**yes**
DP134 DP135	**0,153**	**corresp.**	**3**	**0,363**	**yes**
DP134 DP135 DP136 DP010	**0,299**	**corresp.**	**1**	**0,353**	**no**
DP135 DP136 DP137 DP010	**0,307**	**corresp.**	**1**	**0,348**	**yes**
DP134 DP136 DP137 DP010	0,326	corresp.	1	0,378	no
DP134 DP136	0,330	corresp.	3	0,560	no
DP135 DP137 DP010	0,373	corresp.	2	0,600	yes
DP134 DP135 DP136 DP137	0,388	corresp.	1	0,513	yes
DP136 DP137	0,398	corresp.	3	0,458	yes
DP135 DP010	0,464	corresp.	3	0,603	no
DP134 DP135 DP137	0,580	corresp.	2	0,620	yes
DP137 DP010	0,588	corresp.	3	0,758	no
DP134 DP135 DP010	0,607	corresp.	2	0,668	no
DP137	0,634	corresp.	4	0,510	yes(2 periods)
DP134 DP137 DP010	0,677	corresp.	2	0,700	no
DP135 DP136 DP137	0,753	corresp.	2	0,830	yes
DP135 DP136	0,794	corresp.	3	0,703	yes
DP135 DP137	0,805	corresp.	3	0,975	yes
DP135	0,918	corresp.	4	1,193	Yes(3 periods)
DP136 DP010	1,181	corresp.	3	0,983	no
DP134 DP136 DP010	1,248	corresp.	2	0,905	no
DP134 DP010	1,422	degrade	3	1,493	no
DP134	1,442	degrade	4	1,287	no
DP010	1,582	degrade	4	1,923	no

Table 4: 90th percentile biases calculated by k- combinations and ecological classification correspondence for water body 002;

Interclass correlation coefficient

Results obtained by the ICC have been interpreted using a *qualitative scale* developed by *Landis and Koch*, typical for these types of studies. According to this scale, correlation between items is considered very good (0.90 to 1), good (0.80 to 0.90), moderate (0.60 to 0.80) or poor (0 to 0.60). k-observers that have an ICC > 0,950 are marked as Very Good*.

k-observer	Intreclass correlation coef.	Confidence interval 95%		
		Lower	Upper	
DP134DP135DP137DP010	**,984^{c}**	**,967**	**,992**	**Very Good***
DP136	,767c	,492	,892	Moderate
DP135DP136DP010	**,954c**	**,908**	**,977**	**Very Good***
DP135DP137DP010	**,961c**	**,923**	**,981**	**Very Good***
DP134DP137	,847c	,672	,928	Good
DP134DP136DP137	,949c	,893	,976	Very Good
DP134DP135DP137	,881c	,718	,947	Good
DP134DP135	,810^{c}	,600	,910	Good
DP134DP135DP136DP010	**,989^{c}**	**,976**	**,995**	**Very Good***
DP135DP136DP137DP010	**,967^{c}**	**,934**	**,983**	**Very Good***
DP134DP136	,912c	,813	,959	Very Good
DP134DP010	,881c	,617	,952	Good
DP134	,689c	,338	,854	Poor
DP010	,697c	,341	,855	Poor

Table 5: ICCs of the preselected k-observers, confidence intervals and interpretation of ICCs according to the Landis and Koch qualitative scale.

The 4-observer that excludes DP137 has the highest ICC (0.989). The 3-observer with the highest ICC is DP135,DP137,DP010 (0.961), which excludes DP136 from the first group and from the second group DP134. The 2-observer that has the highest ICC coefficient is DP134,DP136 (0.912).
Four pre-selected k-observers have an ICC < 0,9 and Good or Moderate concordance. The other six k-observers have Very Good concordance (ICC>0.9) according to the Landis scale.

For the preselected observers that produce the highest bias with respect to the reference 90th percentile for the three year period (DP010; DP134 and DP134,DP010), “Good” for the 2-observer and “Poor” for the 1-observers have been detected. The ICC coefficients calculated for the k-observers are 0.697; 0.689 and 0.881 respectively.

Euclidean distance and MDS

Calculating and comparing the Euclidean distance between chlorophyll *a* average series of analyzed k-observers, the output confirms the results of the previous analysis. Results are presented using the Multidimensional Scaling technique. In the two dimensional space in which the distances are presented, it is easy to identify which observers verify that the Euclidean distance between the monthly averages of chlorophyll *a* concentration in their stations and the monthly averages of chlorophyll *a* concentration in the sampling stations of the reference observer (WB002) is less relevant.

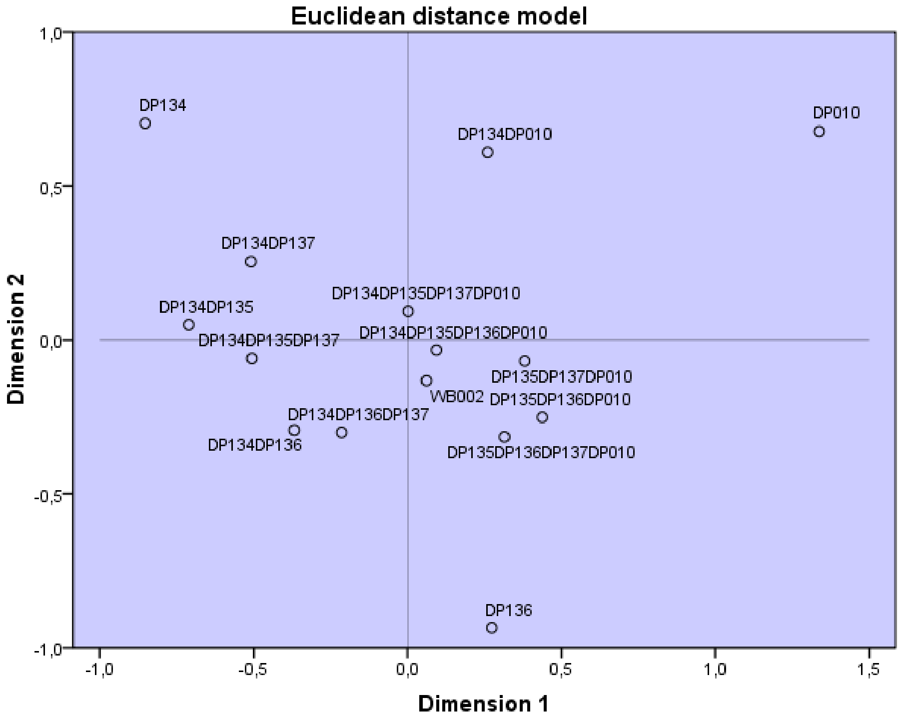


Figure 13: MDS for the Euclidian distance model for the pre-selected k-observers, w.b.002

The 4-observer that excludes DP137 has the lowest Euclidean distance compared to reference observer WB002. DP135,DP137,DP010 has the lowest Euclidian distance for the 3-station combinations, and DP134,DP136 has the lowest Euclidian distance for the 2-observers.

The distance for the 1-observers DP134 and DP010 is significant and comparable with the distance of 1-observer DP136, which has one of the lowest 90^{th} percentile biases for the three-year period. DP136 has been pre-selected due to its low 90^{th} percentile bias, with an ICC and MDS of Euclidian distance detected as an inappropriate combination for the spatial reduction.

Optimal k-observer of sampling stations for water body 002

The previous analyses confirmed that spatial reduction is possible for water body 002 and there are a few alternatives.

Finally, 5 k-observers are selected as most appropriate for the spatial reduction:

Sampling stations	N. of reduced s.s.	90th Percentile
DP134 DP135 DP136 DP137 DP010	**0**	**3,498**
DP134 DP135 DP137 DP010	1	3,540
DP135 DP136 DP137 DP010	1	3,191
DP134 DP135 DP136 DP010	1	3,797
DP135 DP136 DP010	2	3,586
DP135 DP137 DP010	2	3,125

Table 6: Selected k-observers and the 90th percentiles obtained by their data subsets

Finally, the quantile plot (Figure 14) shows that the empirical CDF of the data subset taken by the selected k-observers are similar to the empirical CDF of the "original" dataset taken by the reference observer. It can also be observed that the CDFs have a highest dispersion in the area from the 80th to the 95th percentile.

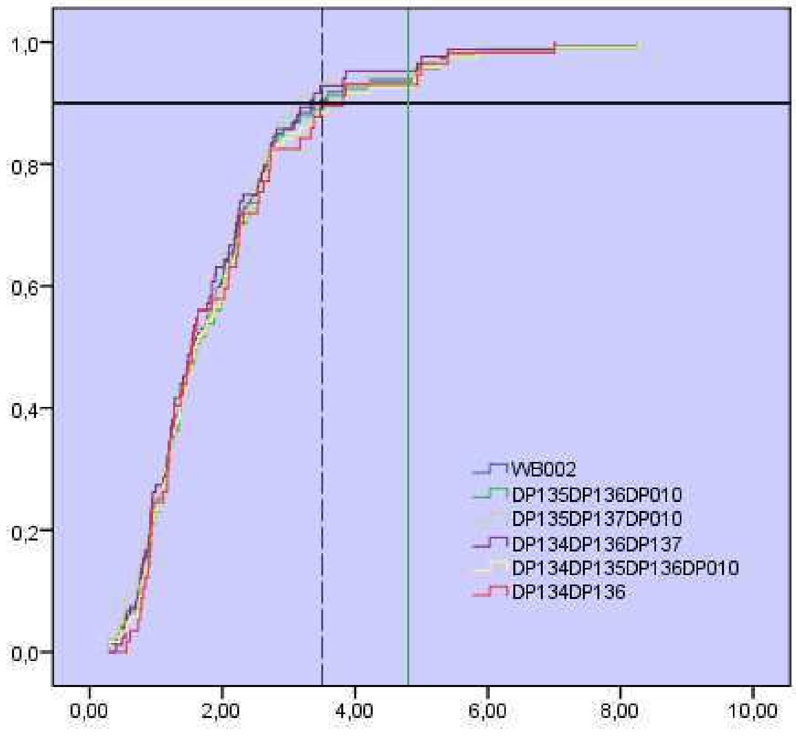


Figure 14- Quantile plot for the selected k-observers of chlorophyll a concentration for the three year period of August 2005 - July 2008, water body 002

Analysis for Water Body 005

Water Body 005 is surveyed by five different sampling stations (Annex 2). It is classified as a type II-A water body (significant continental influence and salinity variance). The previous exploratory study detected that spatial reduction is possible without consequences to ecological classification. The following analysis will attempt to identify the most adequate k-observers for the spatial reduction.

Comparative study

Multiple box-and-whisker plots have been used to compare the range, median, quartiles and outliers of the measurements of *chlorophyll a* obtained from the five sampling stations located in water body 005. Graphics have been made for the three annual periods and the three-year period.

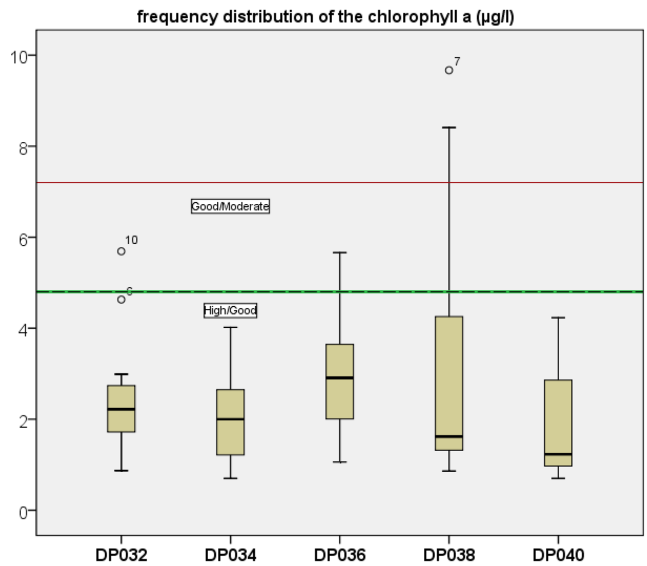


Figure 15:Box-and-whisker plot of chlorophyll a concentration for the annual period of August 2005 - July 2006 in water body 005; The green line represents the High/Good threshold, the red line represents the Good/Moderate threshold, and the broken line represents the 90th percentile

During the annual period of August 2005-July 2006, it is possible to observe that the highest concentrations of chlorophyll *a* in water body 005 have been observed mainly at station DP038 (Figure 15). Samples were analysed from stations DP032, DP036 and DP038 with concentrations of the chlorophyll a exceeding the High/Good threshold. A significant number of samples from station DP038 exceed the Good/Moderate threshold.

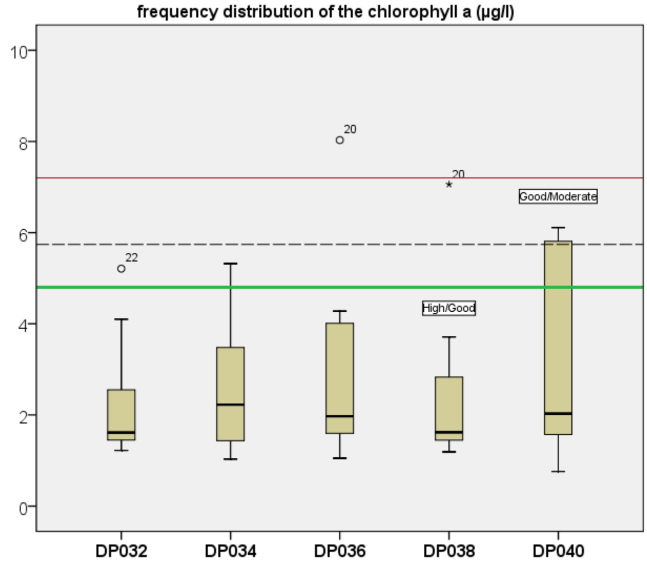


Figure 16: Box-and-whisker plot of chlorophyll a concentration for the annual period of August 2006 - July 2007; water body 005; The green line represents the High/Good threshold, the red line represents the Good/Moderate threshold, and the broken line represents the 90th percentile

For the second annual period, station DP040 analyzed the highest number of samples (in the water body) that exceed the High/Good threshold. The chlorophyll *a* concentrations at station DP038 are significantly lower than for the first annual period.

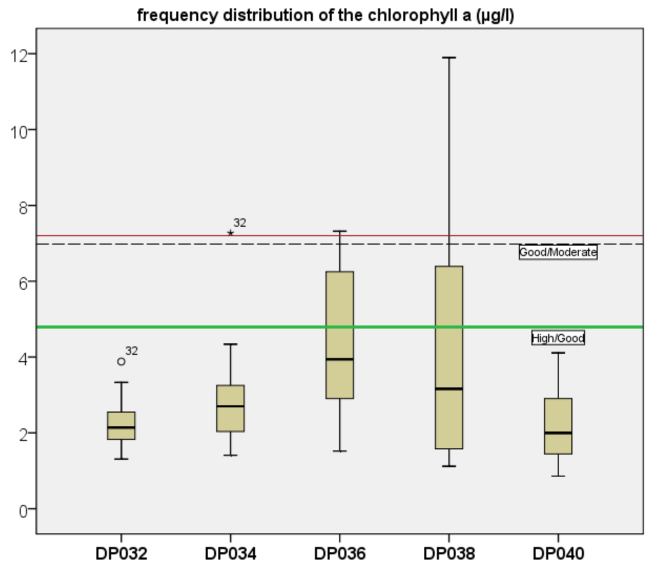


Figure 17: Box-and-whisker plot of chlorophyll a concentration for the annual period of August 2007 - July 2008; water body 005; The green line represents the High/Good threshold, the red line represents the Good/Moderate threshold, and the broken line represents the 90th percentile

For the third annual period, the box-and-whisker plot shows that the highest concentrations of chlorophyll *a* have been recorded by sampling stations DP036 and

DP038. During the campaign carried out in August 2007, samples analyzed for DP036 and DP038 have chlorophyll *a* concentrations that exceed the Good/Moderate threshold. DP032, DP034, DP040 took samples throughout the year that did not usually exceed the High/Good threshold.

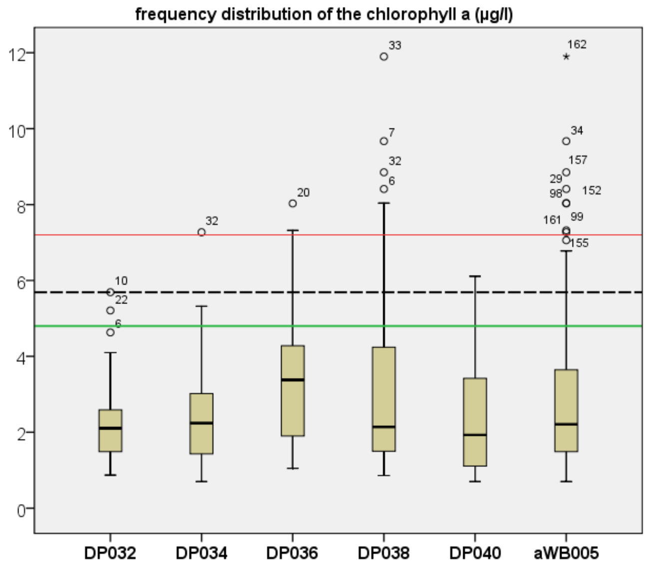


Figure 18: Box-and-whisker plot of chlorophyll a concentration for the three year period of August 2005 - July 2008 in water body 005; The green line represents the High/Good threshold, the red line represents the Good/Moderate threshold, and the broken line represents the 90th percentile

The box-and-whisker plot for the three year period clearly shows that the highest concentrations of chlorophyll *a* for water body 005 were recorded mostly by sampling stations DP036 and DP038 (Figure 18). It is clear that the sampling stations can be divided into two groups. The first group is composed of stations DP032 and DP034, which have lower values for the water body. The second group consists of DP036 and DP038, which have significantly higher values than the first group. Station DP040 detected high chlorophyll *a* concentrations during the second annual period, but not for the other two annual periods. Due to this, DP040 clearly belongs to neither the first nor the second group.

The overlapping graphs corresponding to the empirical CDF of each sampling station allow us to compare the 90th percentiles of these distributions (Figure 19). The datasets taken by the sampling stations that belong to the first group have a 90th percentile lower than the 90th percentile of the reference dataset, and the datasets taken by the sampling stations that belong to the second group have a 90th percentile higher than this value. Observing the quantile diagram, it is clear why it is complicated to assign a group for sampling station DP040 (90th percentile similar to the 90th percentile of the w.b.).

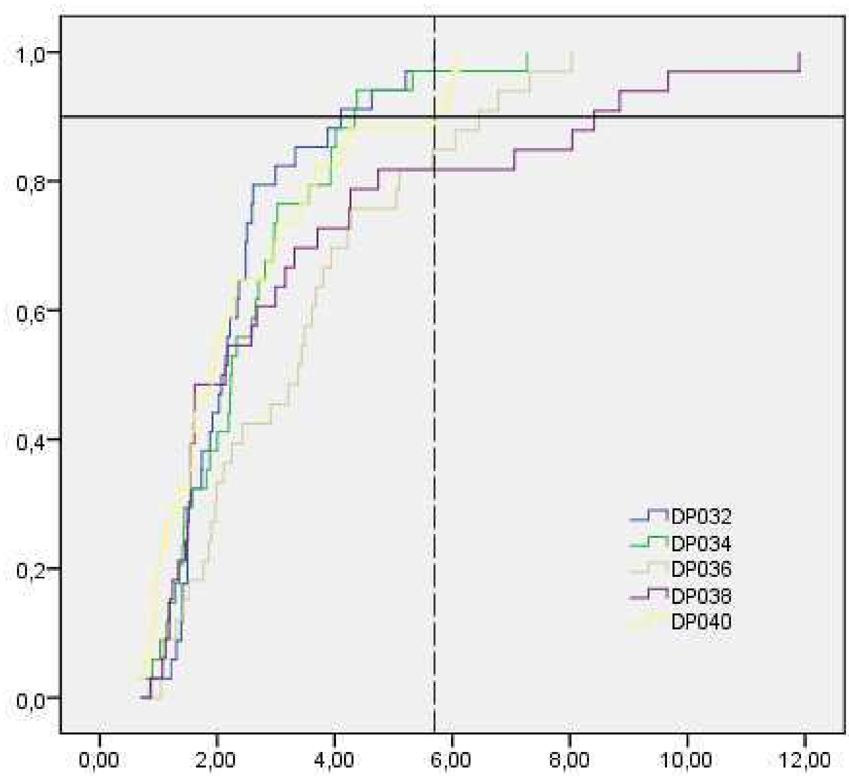


Figure 19:Quantile plot of chlorophyll a concentration for all sampling stations for the three year period August 2005 - July 2008, water body 005

Pre-selection of k-observers of the sampling stations based on the deviation of the 90th percentile

Observing Table 7(taken from the Exploratory study and aligned by the three year bias), the results confirm what was easily observable on the previous quantile plot – the k-observers (where k<4) have a lower bias on the 90th percentile if they include sampling stations from both groups. Also, it is visible that these combinations include only one sampling station that belongs to the second group (DP036 or DP038).

Also, the quantile diagram anticipated that 1-observer DP040 would have a low 90th percentile bias. The 4-observer that excludes DP040 has a lower 90th percentile bias for all 4-observers.

Similar as for the previous water bodies, the highest 90th percentile biases are obtained for 1-observer DP038 and 2-observer DP036,DP038 because they only include sampling stations from the second group.
DP032,DP034,DP040, which includes stations that belong to the first group and DP040 (which belongs to neither group), has the highest bias for the 3-observers.

The low bias of the 90th percentile lets us suppose that the possibility of spatial reduction exists, and the first 10 k-observers with the lowest bias are pre-selected for a more in-depth analysis. Nevertheless, the same analysis will be performed for the k-observers with the highest 90th percentile bias in order to highlight the difference.

w.b.	k-observer	bias PCTL (3 years)	eco. class. (3 years)	- n stations	average absolute error (4 periods)	Corresponding eco. class. (4 periods)
005	**DP032 DP036 DP040**	**0,009**	**corresp.**	**2**	**0,245**	**no**
005	**DP034 DP036 DP040**	**0,015**	**corresp.**	**2**	**0,300**	**yes**
005	**DP032 DP034 DP036 DP038**	**0,018**	**corresp.**	**1**	**0,503**	**no**
005	**DP032 DP038 DP040**	**0,021**	**corresp.**	**2**	**0,238**	**no**
005	**DP032 DP036**	**0,027**	**corresp.**	**3**	**0,478**	**no**
005	**DP034 DP036**	**0,045**	**corresp.**	**3**	**0,263**	**yes**
005	**DP040**	**0,117**	**corresp.**	**4**	**1,063**	**no**
005	**DP034 DP038 DP040**	**0,171**	**corresp.**	**2**	**0,418**	**no**
005	**DP032 DP036 DP038 DP040**	**0,257**	**corresp.**	**1**	**0,270**	**no**
005	**DP032 DP038**	**0,271**	**corresp.**	**3**	**1,418**	**no**
005	DP036 DP040	0,317	corresp.	3	0,240	no
005	DP034 DP036 DP038 DP040	0,332	corresp.	1	0,245	no
005	DP032 DP034 DP036 DP040	0,439	corresp.	1	0,390	yes
005	DP032 DP034 DP038 DP040	0,439	corresp.	1	0,445	yes
005	DP032 DP034 DP036	0,505	corresp.	2	0,488	no
005	DP032 DP034 DP038	0,577	corresp.	2	0,710	no
005	DP038 DP040	0,607	corresp.	3	1,235	no
005	DP032 DP036 DP038	0,717	corresp.	2	0,843	no
005	DP036 DP038 DP040	0,723	corresp.	2	0,610	no
005	DP036	0,955	corresp.	4	0,870	no
005	DP032 DP040	1,005	upgrade	3	1,193	no
005	DP034 DP036 DP038	1,054	corresp.	2	0,835	no
005	DP034 DP040	1,219	upgrade	3	1,238	no
005	DP032 DP034 DP040	1,325	upgrade	2	1,245	no
005	DP032	1,328	upgrade	4	1,520	no
005	DP034	1,333	upgrade	4	0,780	no
005	DP032 DP034	1,349	upgrade	3	1,353	no
005	DP034 DP038	1,409	corresp.	3	1,550	no
005	DP036 DP038	1,840	degrade	3	1,688	no
005	DP038	2,981	degrade	4	3,140	no

Table 7: 90th percentile biases calculated by k- combinations and ecological classification correspondence for water body 005

Interclass correlation coefficient

Results obtained by the ICC have been interpreted using a *qualitative scale* developed by *Landis and Koch*, typical for these types of studies. According to this scale, correlation between items is considered very good (0.90 to 1), good (0.80 to 0.90), moderate (0.60 to 0.80) or poor (0 to 0.60). k-observers that have an ICC > 0,950 are marked as Very Good*.

		Confidence interval 95%		
	Intraclass correlation[a]	Lower limit	Upper limit	Value
DP032DP036DP040	**,952**	**,905**	**,976**	**Very Good***
DP034DP036DP040	**,959**	**,918**	**,980**	**Very Good***
DP032DP034DP036DP038	**,983**	**,966**	**,992**	**Very Good***
DP032DP038DP040	**,983**	**,965**	**,992**	**Very Good***
DP032DP036	**,953**	**,907**	**,977**	**Very Good***
DP034DP036	**,962**	**,923**	**,981**	**Very Good***
DP040	,726	,459	,862	Moderate
DP034DP038DP040	**,985**	**,969**	**,992**	**Very Good***
DP032DP036DP038DP040	**,993**	**,987**	**,997**	**Very Good***
DP032DP038	,929	,858	,965	Very Good
DP036	,882	,643	,951	Good
DP036DP038	,902	,724	,958	Very Good
DP038	,806	,608	,904	Good

Table 8: ICCs of the preselected k-observers, confidence intervals and interpretation of ICCs according to the Landis and Koch qualitative scale.

The 4-observer that excludes DP034 has the highest ICC (0.993). The two 3-observers with high ICC are DP032DP038DP040 (0.983) and DP034DP038DP040 (0.985). Both 3-observers include one station from each group and DP040. Nine of ten preselected observers have an ICC value higher than 0.9 and “Very Good” concordance with the reference observer.

For the preselected observers that produce the highest bias with respect to the reference 90^{th} percentile for the three year period (DP036; DP038 and DP036,DP038), “Very Good” is assigned for the 2-observer and “Good” for the 1-station combinations. The ICC coefficients calculated for the k-observers are 0.882; 0.806 and 0.902 respectively.

Euclidean distance and MDS

Calculating and comparing the Euclidean distance between chlorophyll *a* average series of analyzed k-observers, the output confirms the results of the previous analysis. Results are presented using the Multidimensional Scaling technique. In the two dimensional space in which the distances are presented, it is easy to identify which observers verify that the Euclidean distance between the monthly averages of chlorophyll *a* concentration in their stations and the monthly averages of chlorophyll *a* concentration in the sampling stations of the reference observer (WB005) is less relevant.

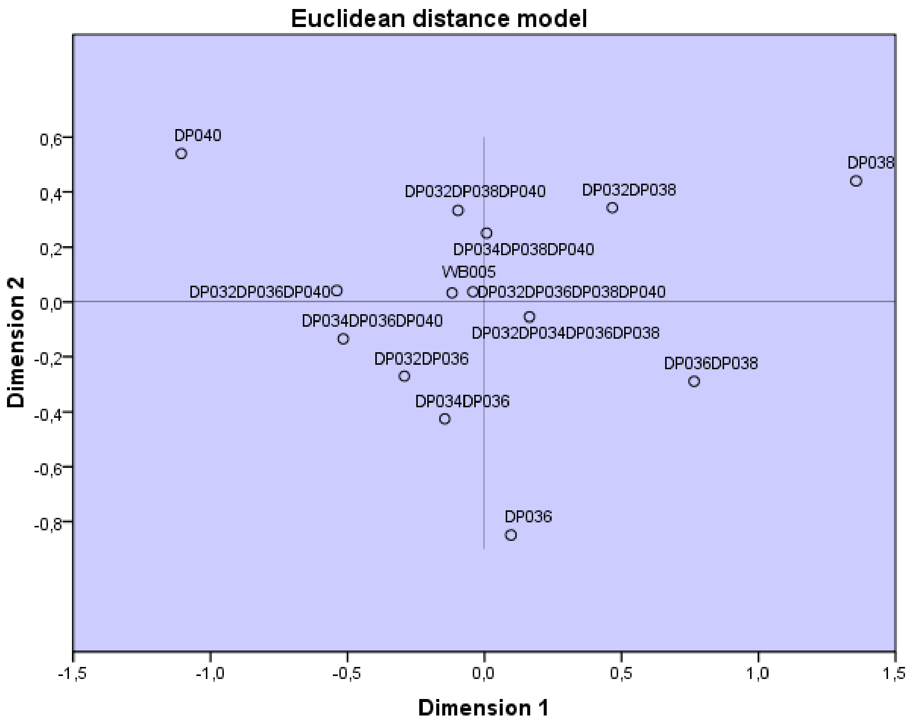


Figure 20: MDS for the Euclidian distance model for the pre-selected k-observers, w.b.005

The 4-station k-observer that excludes DP034 has the lowest Euclidean distance compared to reference observer WB005. DP034,DP038,DP040 has the lowest Euclidian distance for the 3-observers, and DP032,DP036 has the lowest Euclidian distance for the 2-observers.
The highest distance is presented by the 1-observers DP036, DP038 and DP040.

Optimal k-observer of sampling stations for water body 005

The previous analyses confirmed that spatial reduction is possible for water body 005 and there are various alternatives.

Finally, 8 k-observers are selected as most appropriate for the spatial reduction:

Sampling stations	N. of reduced s.s.	90^{th} Percentile
DP032 DP034 DP036 DP038 DP040	0	5,693
DP032 DP034 DP036 DP038	1	5,675
DP032 DP036 DP038 DP040	1	5,950
DP032 DP036 DP040	2	5,684
DP032 DP038 DP040	2	5,714
DP034 DP036 DP040	2	5,7080
DP034 DP038 DP040	2	5,8640
DP032 DP036	3	5,666
DP034 DP036	3	5,7380

Table 9: Selected k-observers and the 90th percentiles obtained by their data subsets

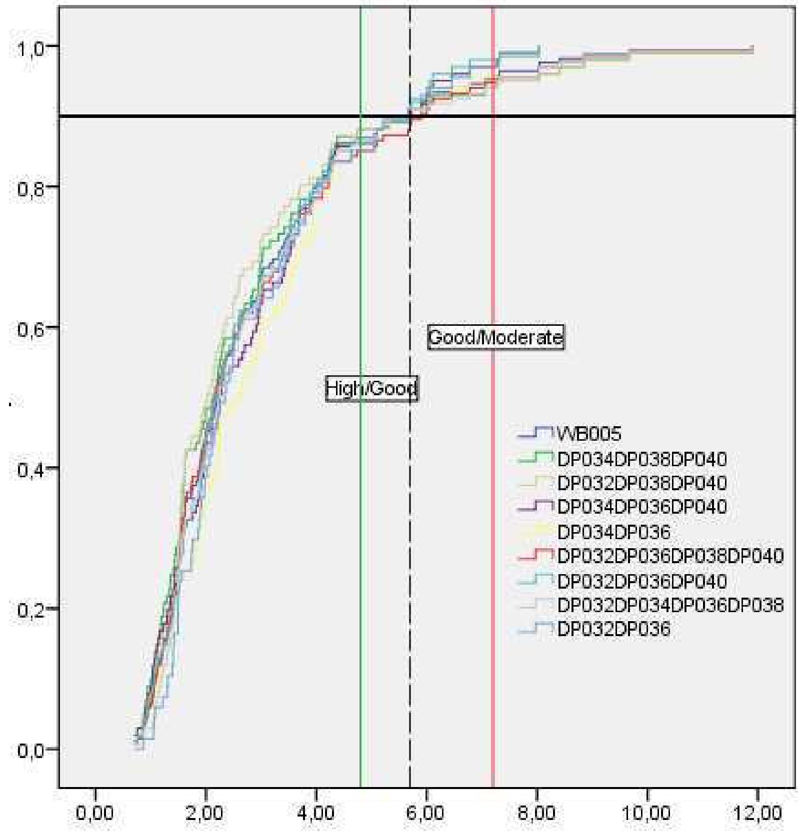


Figure 21: Quantile plot for the selected k-observers of chlorophyll a concentration for the three year period of August 2005 - July 2008, water body 005

Finally, the quantile plot (Figure 21) shows that the empirical CDF of the data subset taken by the selected k-observers is similar to the empirical CDF of the "original" dataset taken by the reference observer. Also, the 90^{th} percentiles of these subsets are close to the 90^{th} percentile of the reference dataset, and the biases are irrelevant.

Analysis for Water Body 008

Water Body 008 is surveyed by five different sampling stations (Annex 2). It is classified as a type II-A water body (significant continental influence and salinity variance). Sampling station DPU001 was active from the second year of monitoring and analyzed only 15 samples during the three years of the study. The previous exploratory study detected that spatial reduction is possible without consequences to ecological classification. The following analysis will attempt to identify the most adequate k-observers for the spatial reduction.

<u>*Comparative study*</u>

Multiple box-and-whisker plots have been used to compare the range, median, quartiles and outliers of the measurements of *chlorophyll a* obtained from the five sampling stations located in water body 008. Graphs have been made for the three annual periods and the three-year period.

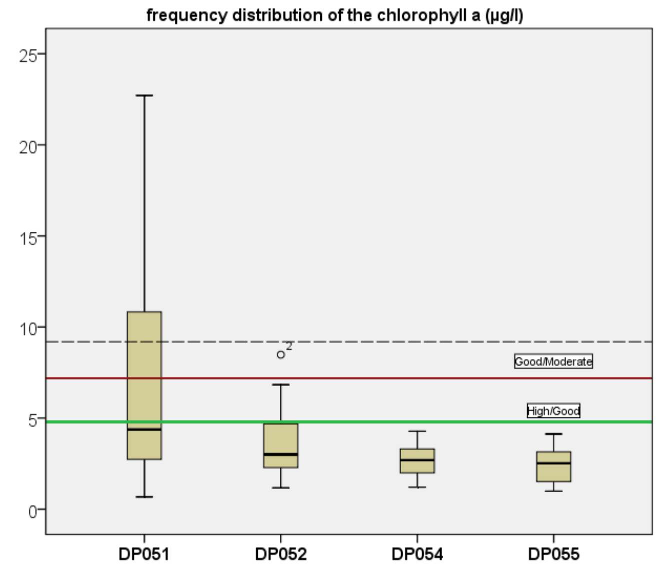


Figure 22:Box-and-whisker plot of chlorophyll *a* concentration for the annual period of August 2005 - July 2006 in water body 008; The green line represents the High/Good threshold, the red line represents the Good/Moderate threshold, and the broken line represents the 90th percentile

During the annual period of August 2005-July 2006, it is clear that the highest concentrations of chlorophyll *a* in water body 005 have been observed mainly at station DP051 (Figure 22). A significant number of samples from DP051 exceeds the Good/Moderate threshold. The concentrations sampled by DP054 and DP056 do not exceed the High/Good threshold.

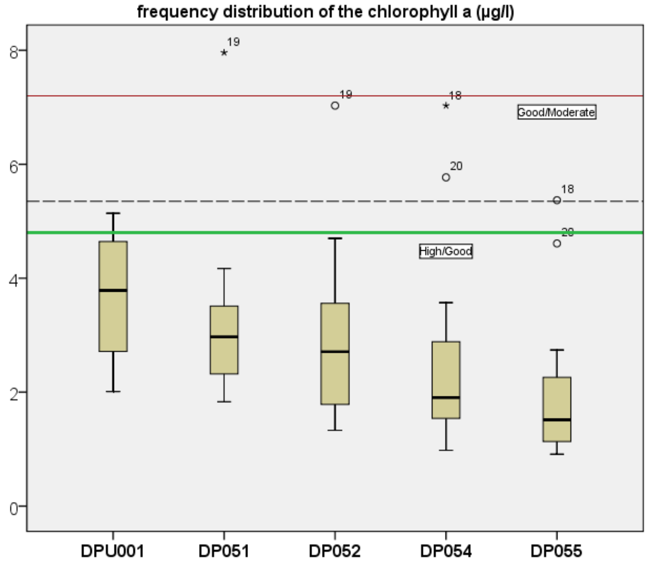


Figure 23:Box-and-whisker plot of chlorophyll a concentration for the annual period of August 2006 - July 2007; water body 008; The green line represents the High/Good threshold, the red line represents the Good/Moderate threshold, and the broken line represents the 90th percentile

During the second annual period, elevated concentrations were not detected for water body 008. In this year, sampling station DPU001 was activated, which analyzed the samples with the highest chlorophyll *a* concentrations. The other stations in the water body analyzed samples that have chlorophyll *a* concentrations which are for the most part below the High/Good threshold. Nevertheless, during campaigns in February, March and April 2007, elevated values for all water body sampling stations were detected except DPU001.

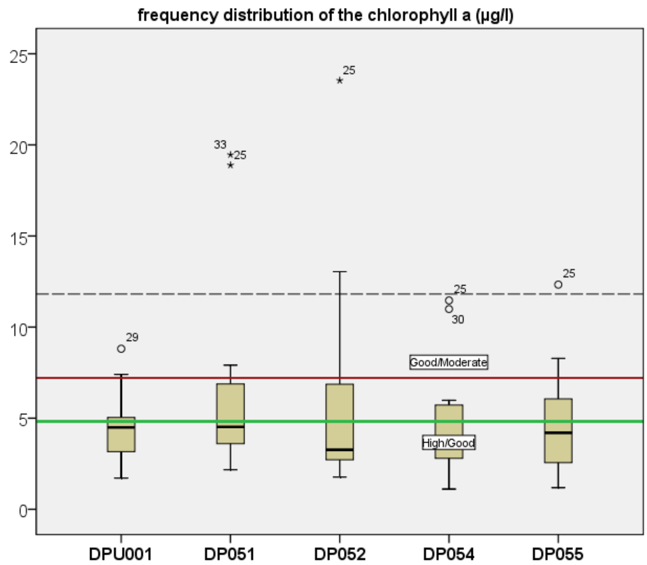


Figure 24:Box-and-whisker plot of chlorophyll a concentration for the annual period of August 2007 - July 2008; water body 008; The green line represents the High/Good threshold, the red line represents the Good/Moderate threshold, and the broken line represents the 90th percentile

For the third annual period, the box-and-whisker plot shows that the highest concentrations of chlorophyll *a* have been recorded by sampling stations DP051 and DP052 (Figure 24). For the three year period, all stations analyzed samples with chlorophyll *a* concentrations that exceed the Good/Moderate threshold.

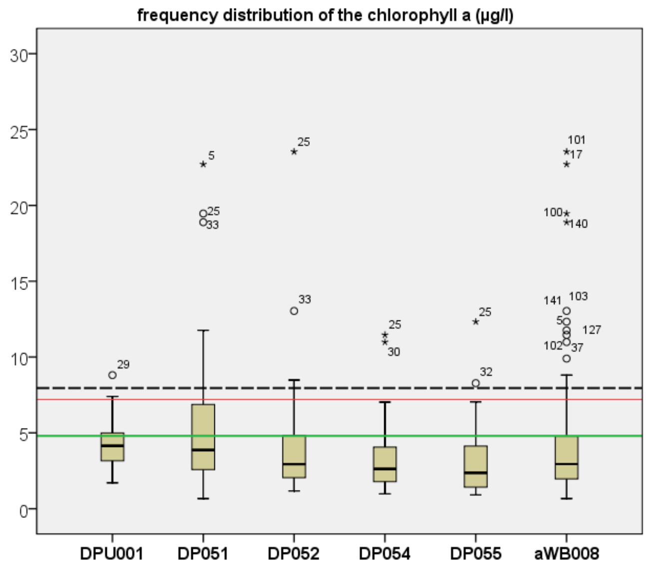


Figure 25: Box-and-whisker plot of chlorophyll a concentration for the three year period of August 2005 - July 2008 in water body 008; The green line represents the High/Good threshold, the red line represents the Good/Moderate threshold, and the broken line represents the 90th percentile

The box-and-whisker plot for the three year period clearly shows that the highest concentrations of chlorophyll *a* for water body 010 were recorded mostly by sampling stations DP051 (Figure 25). Therefore, all sampling stations analyzed samples that have chlorophyll *a* concretions beyond the Good/Moderate threshold.
A division of the stations into two groups is not as clear as for the other analyzed water bodies. The first group is made up of stations DP054 and DP055, which represent lower values for the water body. The second group should be made up of DP051, which has significantly higher values than the first group. Stations DPU001 and DP052 can be assigned to the third intermediate group.

The optimal k-observer of sampling stations should have an empirical CDF similar to the empirical CDF of WB008, which is included on the last box-and-whisker plot.

The overlapping graphs corresponding to the empirical CDF of each sampling station allow us to compare the 90th percentiles of these distributions (Figure 26). The datasets taken by the sampling stations that belong to the first group have a 90th percentile lower than the 90th percentile of the reference dataset, and the datasets taken by the sampling stations that belong to the second group have a 90th percentile higher than this value. Observing the quantile diagram, it is clear that sampling station DPU001 and DP052 make up a third intermediate group with a 90th percentile close to the 90th percentile of the w.b.

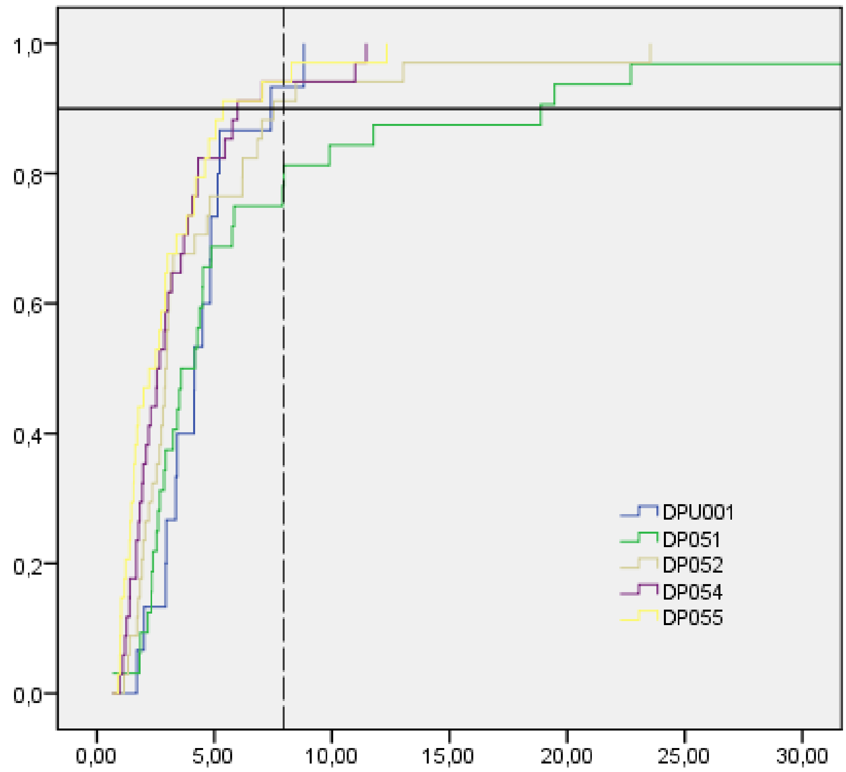


Figure 26: Quantile plot for the all sampling stations of chlorophyll *a* concentration for the three year period August 2005 - July 2008, water body 008

Pre-selection of k-observers of the sampling stations based on the deviation of the 90th percentile

The quantile diagram anticipated that 1-observers DPU001 and DP052 would have a low 90th percentile bias. The problem with the first 1-observer is that for DPU001 data was only analyzed for the 15 campaigns during the second and third year of the studied period. Due to this significant lack of data, this result should be taken with caution, as should results obtained using k-combinations that include DPU001.
The two 4-observers that exclude DPU001 or DP052 have the lowest 90th percentile bias of all k-observers (where k>1). The 3-observer that excludes the same stations has the lowest bias of the 3-observers.
Finally, the 2-observer with the lowest bias includes DPU001 and DP051.
Since water body 008 has just one station (DP051) that belongs to the second group (elevated chlorophyll *a* concentrations), the highest biases of the 90th percentile are obtained for 1-observer DP051 and for all 2-observers that include DP051.

The low bias of the 90th percentile allows us to suppose that the possibility of spatial reduction exists, and the first 10 k-observers with the lowest bias are pre-selected for a more in-depth analysis. This analysis should pay special attention to the k-observers that include DPU001 because of this station’s lack of data for the first and second year of the analyzed period.
Nevertheless, the same analysis will be performed for 1-observer DP051 which has the highest 90th percentile bias, in order to highlight the difference.

k-observer	bias PCTL (3 years)	eco. class. (3 years)	- n stations	average absolute error (4 periods)	Corresponding eco. class. (4 periods)
DPU001	**0,004**	**corresp.**	**4**	**1,640**	**Yes(2periods)**
DP052	**0,05**	**corresp.**	**4**	**2,925**	**yes**
DPU001 DP051 DP054 DP055	**0,13**	**corresp.**	**1**	**0,650**	**yes**
DP051 DP052 DP054 DP055	**0,16**	**corresp.**	**1**	**0,295**	**yes**
DP051 DP054 DP055	**0,29**	**corresp.**	**2**	**0,628**	**yes**
DPU001 DP051 DP052 DP055	**0,40**	**corresp.**	**1**	**0,823**	**yes**
DPU001 DP052	**0,43**	**corresp.**	**3**	**0,468**	**yes**
DP051 DP052 DP055	**0,50**	**corresp.**	**2**	**1,845**	**yes**
DPU001 DP051 DP052 DP054	**0,65**	**corresp.**	**1**	**0,868**	**yes**
DPU001 DP052 DP055	**0,70**	**corresp.**	**2**	**1,080**	**no**
DPU001 DP052 DP054	0,71	corresp.	2	1,088	no
DPU001 DP051 DP055	0,74	corresp.	2	3,003	yes
DP052 DP055	0,86	upgrade	3	1,185	no
DP052 DP054	0,88	upgrade	3	1,310	no
DPU001 DP055	0,91	upgrade	3	2,333	no
DP052 DP054 DP055	0,93	upgrade	2	1,248	no
DPU001 DP052 DP054 DP055	0,93	upgrade	1	1,315	no
DPU001 DP054	0,93	upgrade	3	2,038	no
DPU001 DP054 DP055	1,35	upgrade	2	2,045	no
DP054	1,46	upgrade	4	2,040	no
DPU001 DP051 DP054	1,72	corresp.	2	3,325	yes
DPU001 DP051 DP052	1,72	corresp.	2	4,265	yes
DP055	1,75	upgrade	4	1,843	no
DP051 DP052 DP054	1,80	corresp.	2	2,355	yes
DP054 DP055	1,88	upgrade	3	1,918	no
DP051 DP055	2,50	corresp.	3	4,515	yes
DP051 DP054	3,17	corresp.	3	4,893	yes
DP051 DP052	4,18	corresp.	3	5,720	yes
DPU001 DP051	5,23	corresp.	3	13,965	yes
DP051	11,33	corresp.	4	16,618	no

Table 10: 90th percentile biases calculated by k- combinations and ecological classification correspondence for water body 008;

Interclass correlation coefficient

Results obtained by the ICC have been interpreted using a *qualitative scale* developed by *Landis and Koch*, typical for these types of studies. According to this scale, correlation between items is considered very good (0.90 to 1), good (0.80 to 0.90), moderate (0.60 to 0.80) or poor (0 to 0.60). k-observers that have an ICC > 0,950 are marked as Very Good*.

		Confidence interval 95%		
	Intraclass correlation[a]	Lower	Upper	Lower
DPU001	,537[c]	-,424	,846	Poor
DP052	,809[c]	,617	,905	Good
DPU001DP051DP054DP055	**,984[c]**	**,967**	**,992**	**Very Good***
DP051DP052DP054DP055	**,996[c]**	**,991**	**,998**	**Very Good***
DP051DP054DP055	**,986[c]**	**,972**	**,993**	**Very Good***
DPU001DP051DP052DP055	**,984[c]**	**,967**	**,992**	**Very Good***
DPU001DP052	,794[c]	,588	,897	Moderate
DP051 DP052 DP055	**,975[c]**	**,949**	**988**	**Very Good***
DPU001DP051DP052DP054	**,983[c]**	**,960**	**,992**	**Very Good***
DPU001DP052DP055	,764[c]	,531	,881	Moderate
DP051	,667[c]	,331	,836	Moderate

Table 11: ICCs of the preselected k-observers, confidence intervals and interpretation of ICCs according to the Landis and Koch qualitative scale.

The 4-observer that excludes DPU001 has the highest ICC (0.996). The 3-observer with the highest ICC is DP051,DP054,DP055 (0.986). The 2-observers and 1-stations combinations have a low ICC. Six of the ten preselected observers have a “Very Good” concordance with the reference observer.

1-observer DPU001 has a “Poor” concordance with the reference observer and this ICC is calculated on the basis of 15 possible campaigns. Observing the station DPU001, the chlorophyll *a* concentrations vary completely independently comparing to the rest of the water body sampling stations. This independent behaviour can be explained by the sampling station’s position, which is usually (North-South sea currents) not influenced by the Albufera (hypereutrophic lagoon).

For the preselected 1-observer DP051 which produced the highest bias with respect to the reference 90th percentile for the three year period, "Moderate" concordance to the complete data was detected. The ICC coefficient calculated for 1-observer DP051 is 0.667.

Euclidean distance and MDS

Calculating and comparing the Euclidean distance between chlorophyll *a* average series of the analyzed k-observers, the output confirms the results of the previous analysis. The results are presented using the Multidimensional Scaling technique. In the two dimensional space in which the distances are presented, it is easy to identify which observers verify that the Euclidean distance between the monthly averages of chlorophyll *a* concentration in their stations and the monthly averages of chlorophyll *a* concentration in the sampling stations of the reference observer (WB008) is less relevant. 1-observer DPU001 (which has an extremely low ICC) is excluded from the analysis due to the significant lack of data, which would reduce analysis to 15 campaigns for all k-combinations in the three year period.

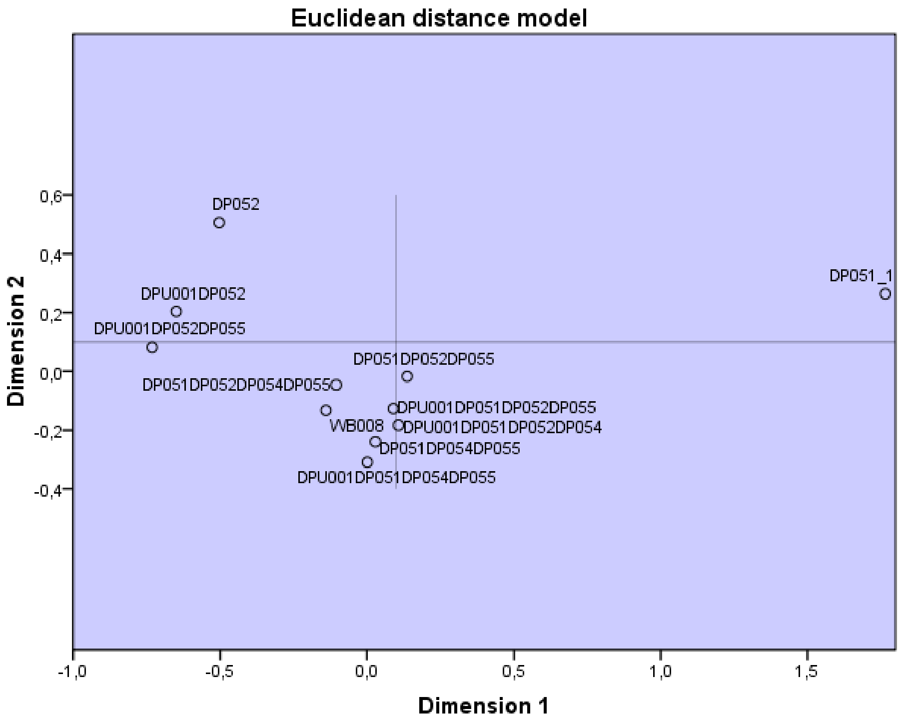


Figure 27: MDS for the Euclidian distance model for the pre-selected k-observers, w.b.001

The 4-observer that excludes DPU001 has the lowest Euclidean distance compared to reference observer WB008. This result coincides with the highest ICC for the k-observers, as station DPU001 was activated in April 2007. The three 4-observers for which a high ICC was calculated ("Very Good*" concordance with a reference observer) have a similar Euclidian distance.

DP051,DP054,DP055 has the lowest Euclidian distance of the 3-observers.
1-staton combination DP051 has the most significant Euclidian distance on a monthly average basis.

<u>*Optimal k-observer of sampling stations for water body 008*</u>

The previous analyses confirmed that spatial reduction is possible for water body 008, and there are various alternatives. The preselected k-combinations that exclude DPU001, ICC analysis and Euclidian distance, are selected as being the most appropriate for the spatial reduction. This result is determined by the fact that station DPU001 was activated during the second year of the survey and analyzed only 15 samples. The significantly low ICC value (for 1-observer DPU001) is explained by independent (and in most cases opposite) variations of chlorophyll *a* concentrations in respect to the other water body sampling stations. Reduction of DPU001 will not change the ecological classification and significantly modify the 90th percentile, but eliminating this station could represent a loss of water body reference condition information (since the station is not usually influenced by the Albufera hypereutrophic lagoon) and other relevant data. Finally, all k-observers that can be selected for the spatial reduction (due to the ICC and Euclidian distance results) of water body 008 which include station DPU001 do not reflect the real situation. This is mainly because the average value of the k-observer for the first and part of the second annual period does not include the values of DPU001. For this reason, all of these results should be taken with caution.

Finally, 6 k-observers are selected as most appropriate for the spatial reduction:

Sampling stations	N. of reduced s.s.	90th Percentile
DPU001 DP051 DP052 DP054 DP055	0	7,960
DPU001DP051DP054DP055	1	8,088
DP051 DP052 DP054 DP055	1	8,120
DP051 DP054 DP055	2	8,248
DPU001DP051DP052DP055	1	8,360
DP051 DP052 DP055	2	8,460
DPU001DP051DP052DP054	1	8,610

Table 12: Selected k-observers and the 90th percentiles obtained by their data subsets

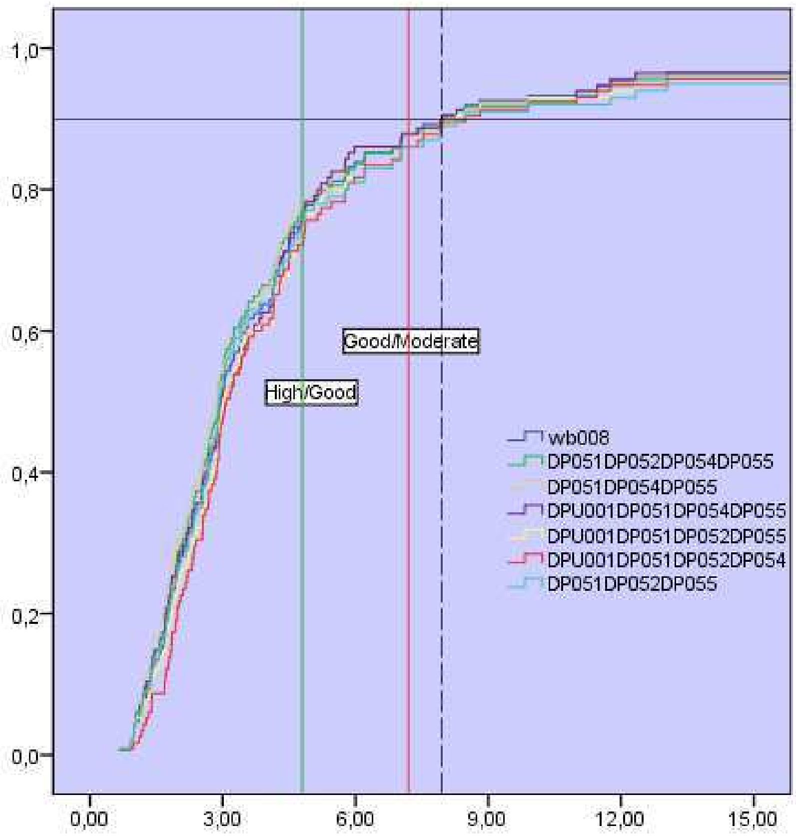


Figure 28: Quantile plot for the selected k-observers of chlorophyll *a* concentration for the three year period of August 2005 - July 2008, water body 001

Finally, the quantile plot (Figure 28) shows that the empirical CDF of the data subset taken by the selected k-observers is similar to the empirical CDF of the "original" dataset taken by the reference observer. Also, the 90^{th} percentiles of these subsets are close to the 90^{th} percentile of the reference dataset, and the biases are irrelevant.

Analysis for the Water Body 009

Water Body 009 is surveyed by five different sampling stations (Annex 2). It is classified as a type II-A water body (significant continental influence and salinity variance). The previous exploratory study detected that spatial reduction is possible without consequences to ecological classification. The following analysis will attempt to identify the most adequate k-observers for the spatial reduction.

<u>*Comparative study*</u>

Multiple box-and-whisker plots have been used to compare the range, median, quartiles and outliers of the measurements of *chlorophyll a* obtained from the five sampling stations located in water body 009. Graphics have been made for the three annual periods and the three-year period.

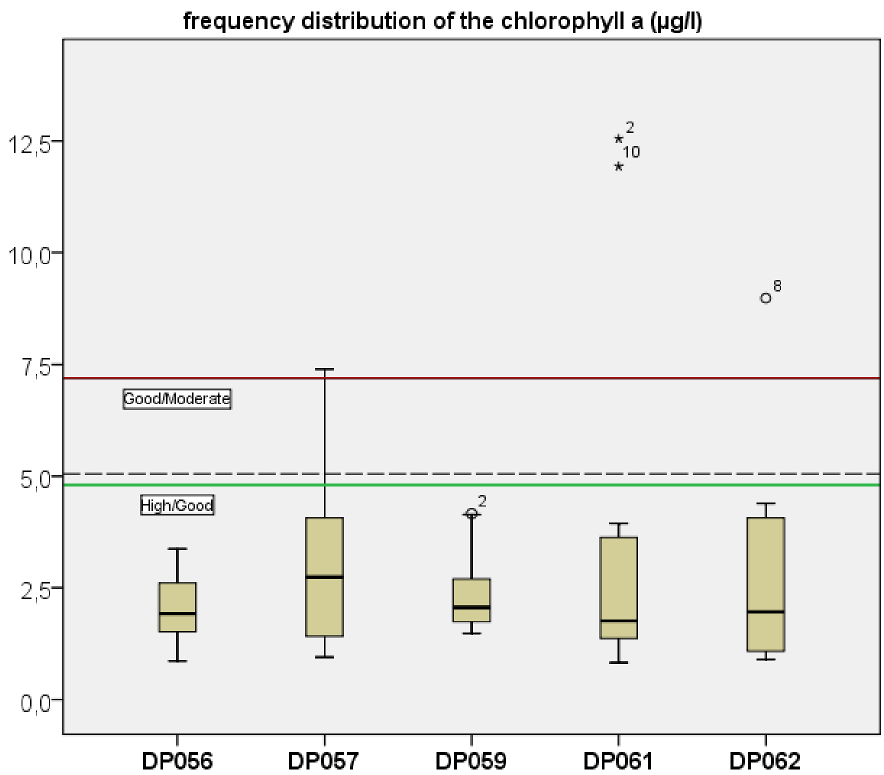


Figure 29: Box-and-whisker plot of chlorophyll a concentration for the annual period of August 2005 - July 2006 in water body 009; The green line represents the High/Good threshold, the red line represents the Good/Moderate threshold, and the broken line represents the 90th percentile

During the annual period of August 2005-July 2006, it is possible to observe that the highest concentrations of chlorophyll *a* in water body 005 have been observed mainly at station DP057 (Figure 29). Some samples analyzed from DP061 and DP062 have concentrations that exceed the Good/Moderate threshold.

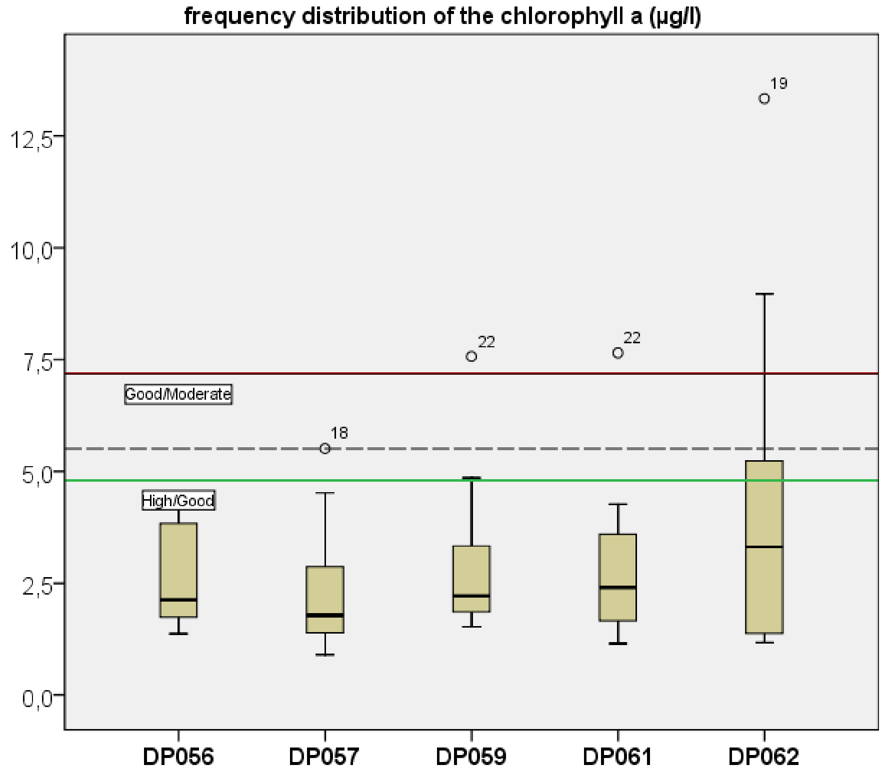


Figure 30:Box-and-whisker plot of chlorophyll a concentration for the annual period of August 2006 - July 2007; water body 009; The green line represents the High/Good threshold, the red line represents the Good/Moderate threshold, and the broken line represents the 90th percentile

For the second annual period, it is clear that the highest concentrations of chlorophyll *a* were taken by sampling station DP062 (Figure 30). Most of the samples taken by the other sampling stations in the water body have chlorophyll *a* concentrations below the High/Good threshold.

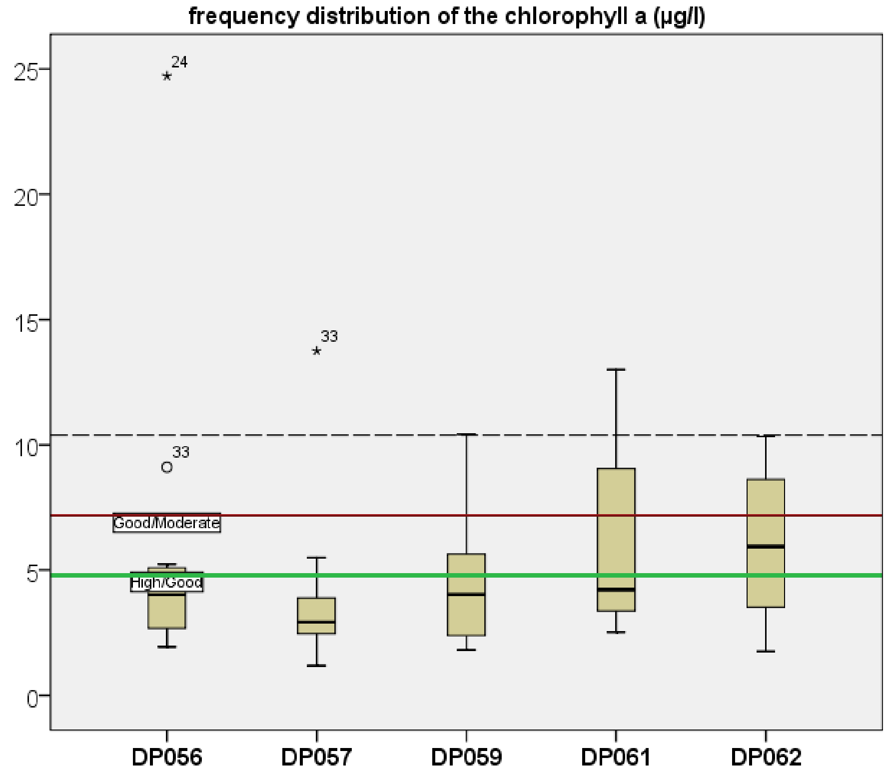


Figure 31: Box-and-whisker plot of chlorophyll a concentration for the annual period of August 2007 - July 2008; water body 009; The green line represents the High/Good threshold, the red line represents the Good/Moderate threshold, and the broken line represents the 90th percentile

For the third annual period, the box-and-whisker plot shows that the highest concentrations of chlorophyll *a* have been detected by sampling stations DP062, DP061 and DP059. Extremely high chlorophyll *a* concentrations were detected for stations DP056 and DP057, but only in two campaigns (outliers).

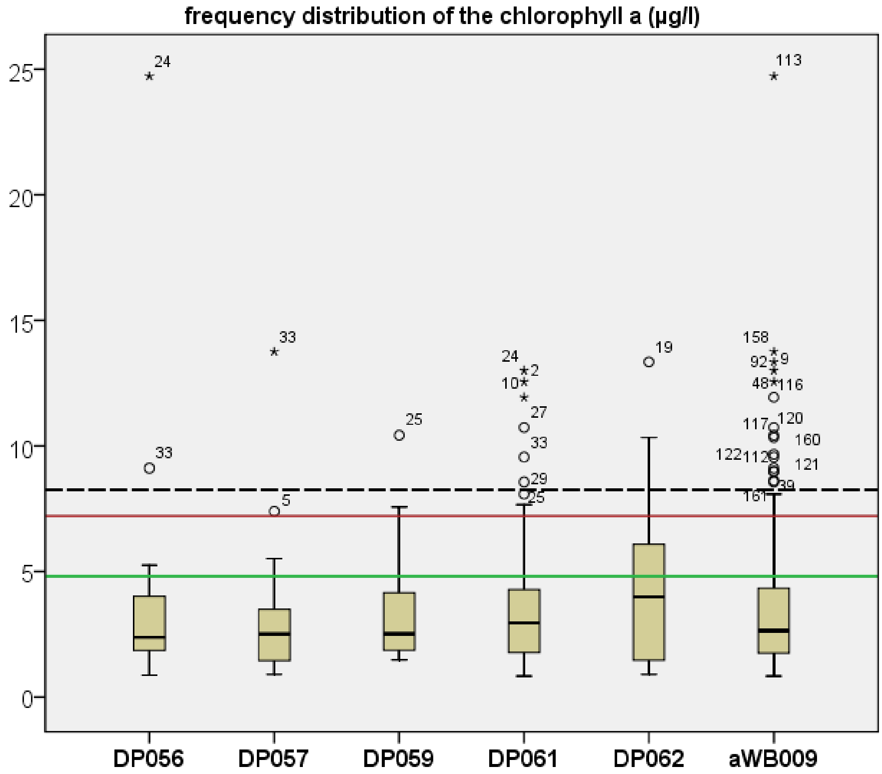


Figure 32: Box-and-whisker plot of chlorophyll a concentration for the three year period of August 2005 - July 2008 in water body 009; The green line represents the High/Good threshold, the red line represents the Good/Moderate threshold, and the broken line represents the 90th percentile

The box-and-whisker plot for the three year period clearly shows that the highest concentrations of chlorophyll *a* for water body 005 were recorded mostly by stations DP061 and DP062 (Figure 32). It is clear that the sampling stations can be divided into two groups. The first group is composed of stations DP056, DP057 and DP059, which have lower values for the water body. The second group consists of DP061 and DP062, which have significantly higher values than the first group.

The overlapping graphs corresponding to the empirical CDF of each sampling station allow us to compare the 90th percentiles of these distributions (Figure 33). The datasets taken by the sampling stations that belong to the first group have a 90^{th} percentile lower than the 90^{th} percentile of the reference dataset, and the datasets taken by the sampling stations that belong to the second group have a 90^{th} percentile higher than this value. This tells us that if we want to select a k-observer of sampling stations where the set of measurements taken by them have a 90^{th} percentile similar to the 90^{th} percentile of the full sample, it will be necessary to select sampling stations from both groups.

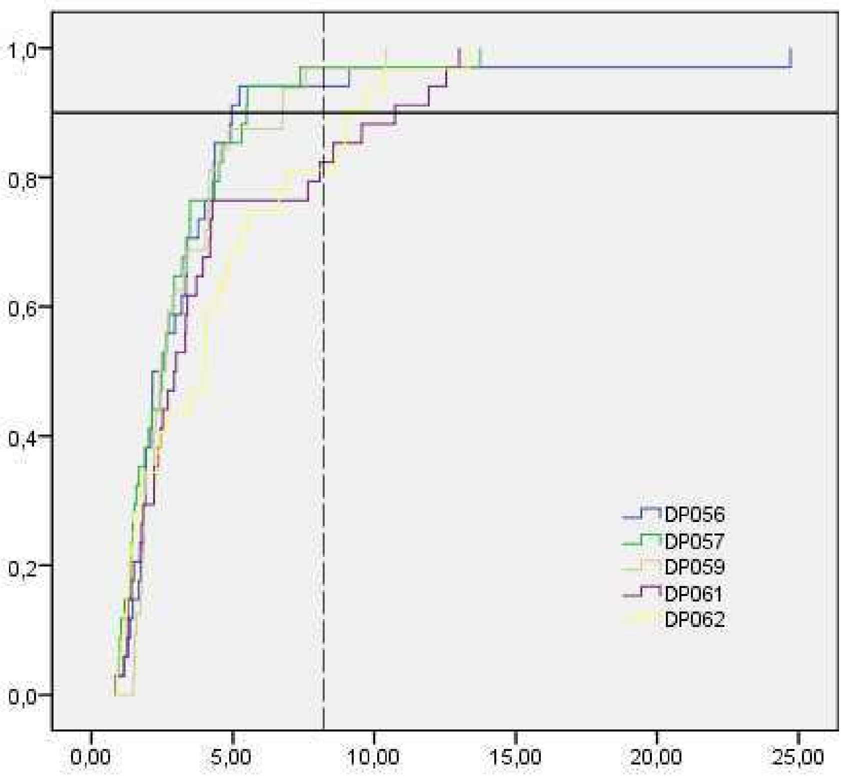


Figure 33: Quantile plot of chlorophyll a concentration for all sampling stations for the three year period August 2005 - July 2008, water body 009

Pre-selection of k-observers of the sampling stations based on the deviation of the 90th percentile

Observing Table 13(taken from the Exploratory study and aligned by the three year bias), the results confirm what was easily observable on the previous quantile plot – the k-observers (where k<4) have a lower bias on the 90th percentile if they include sampling stations from both groups. Also, it is visible that these combinations include only one sampling station that belongs to the second group (DP061 or DP062).

The highest biases of the 90th percentile have been obtained for the 1-observers DP056 (first group) and DP061 (second group). The 3-station and 2-observers that only include the first group of stations have the highest bias.

However, the combination that excludes DP056 has the lowest 90th percentile bias for the 4-observers. These results can indicate that DP056 has the lowest influence on the 90th percentile for w.b. 009 due to the fact that the station includes outliers.

The low 90th percentile bias allows us to suppose that the possibility of spatial reduction exists, and the first 10 k-observers with the lowest bias are pre-selected for a more in-depth analysis. Nevertheless, the same analysis will be performed for the k-observers with the highest 90th percentile bias in order to highlight the difference.

w.b.	k-observer	bias PCTL (3 years)	eco. class. (3 years)	- n stations	average absolute error (4 periods)	Corresponding eco. class. (4 periods)
009	**DP056 DP059 DP061**	**0,187**	**corresp.**	**2**	**0,565**	**no**
009	**DP056 DP057 DP061**	**0,273**	**corresp.**	**2**	**1,150**	**no**
009	**DP057 DP059 DP061 DP062**	**0,378**	**corresp.**	**1**	**0,638**	**yes**
009	**DP057 DP061**	**0,435**	**corresp.**	**3**	**2,100**	**no**
009	**DP057 DP062**	**0,501**	**corresp.**	**3**	**1,090**	**no**
009	**DP056 DP059 DP062**	**0,549**	**corresp.**	**2**	**0,630**	**no**
009	**DP056 DP057 DP061 DP062**	**0,571**	**corresp.**	**1**	**0,570**	**yes**
009	**DP059 DP062**	**0,571**	**corresp.**	**3**	**1,083**	**no**
009	**DP057 DP059 DP061**	**0,582**	**corresp.**	**2**	**0,730**	**yes**
009	**DP059 DP061**	**0,633**	**corresp.**	**3**	**1,800**	**no**
009	DP056 DP059 DP061 DP062	0,641	corresp.	1	0,455	no
009	DP056 DP057 DP059 DP061	0,744	corresp.	1	0,465	no
009	DP056 DP062	0,749	corresp.	3	0,850	no
009	DP057 DP061 DP062	0,755	corresp.	2	1,275	no
009	DP059 DP061 DP062	0,813	corresp.	2	1,485	no
009	DP057 DP059 DP062	0,816	corresp.	2	0,508	yes
009	DP056 DP061 DP062	0,873	corresp.	2	1,080	no
009	DP056 DP057 DP062	0,883	corresp.	2	0,300	yes
009	DP056 DP061	0,930	corresp.	3	2,133	no
009	DP062	1,239	corresp.	4	2,838	no
009	DP056 DP057 DP059 DP062	1,360	upgrade	1	0,538	no
009	DP059	1,453	upgrade	4	1,135	no
009	DP061 DP062	1,644	corresp.	3	2,753	no
009	DP057 DP059	2,342	upgrade	3	0,945	no
009	DP056 DP059	2,531	upgrade	3	1,198	no
009	DP057	2,724	upgrade	4	1,668	no
009	DP056 DP057 DP059	2,752	upgrade	2	1,155	no
009	DP056 DP057	2,896	upgrade	3	1,505	no
009	DP061	3,106	corresp.	4	3,448	no
009	DP056	3,129	upgrade	4	4,298	no

Table 13: 90th percentile biases calculated by k- combinations and ecological classification correspondence for water body 009;

Interclass correlation coefficient

Results obtained by the ICC have been interpreted using a *qualitative scale* developed by *Landis and Koch*, typical for these types of studies. According to this scale, correlation between items is considered very good (0.90 to 1), good (0.80 to 0.90), moderate (0.60 to 0.80) or poor (0 to 0.60). k-observers that have an ICC > 0,950 are marked as Very Good*.

	Intraclass correlation[a]	Confidence interval 95%		
		Lower limit	Upper limit	Value
DP056DP059DP061	**,969[c]**	**,937**	**,984**	**Very Good***
DP056DP057DP061	**,971c**	**,942**	**,986**	**Very Good***
DP057DP059DP061DP062	**,975c**	**,950**	**,988**	**Very Good***
DP057DP061	,947c	,894	,974	Very Good
DP057DP062	,922c	,843	,961	Very Good
DP056DP059DP062	**,969c**	**,938**	**,984**	**Very Good***
DP056DP057DP061DP062	**,991c**	**,982**	**,996**	**Very Good***
DP059DP062	,913c	,826	,957	Very Good
DP057DP059DP061	**,954c**	**,908**	**,977**	**Very Good***
DP059DP061	,929c	,858	,965	Very Good
DP061DP062	,907c	,784	,957	Very Good
DP056DP057	,917c	,833	,959	Very Good
DP056	,785c	,567	,893	Moderate

Table 14: ICCs of the preselected k-observers, confidence intervals and interpretation of ICCs according to the Landis and Koch qualitative scale.

The 4-observer that excludes DP059 has the highest ICC (0.991). The three 3-observers with high ICC are DP056,DP059,DP061 (0.969) DP056,DP057,DP061 (0.971) and DP056,DP059,DP062(0.969). All three 3-observers include one sampling station from the second group and two from the first group. The 2-observer that has the highest ICC coefficient is DP057,DP061 (0.947). All the 10 *preselected observers* have an ICC value higher than 0.9 and have a “Very Good” concordance with the *reference observer*.

For the preselected observers that produce the highest bias with respect to the reference 90^{th} percentile for the three year period (DP056; DP056,DP057 and DP061DP062) “Very Good” and “Moderate” concordance to the complete data are assigned. The ICC coefficients calculated for the k-observers are 0.785; 0.917 and 0.907 respectively.

Euclidean distance and MDS

Calculating and comparing the Euclidean distance between chlorophyll *a* average series of analyzed k-observers, the output confirms the results of the previous analysis. Results are presented using the Multidimensional Scaling technique. In the two dimensional space in which the distances are presented, it is easy to identify which observers verify that the Euclidean distance between the monthly averages of chlorophyll *a* concentration in their stations and the monthly averages of chlorophyll *a* concentration in the sampling stations of the reference observer (WB009) is less relevant.

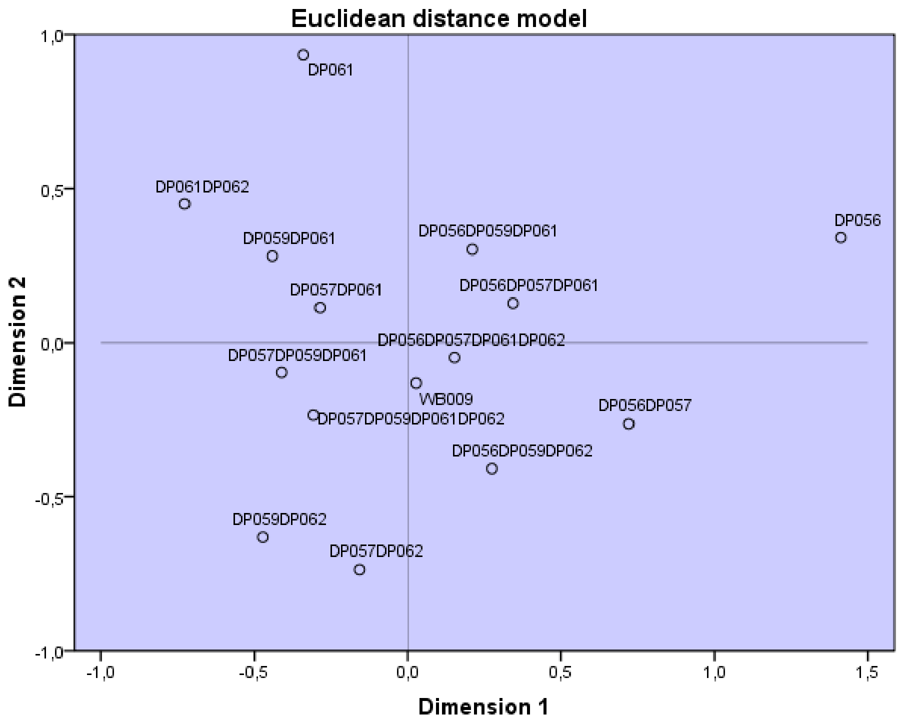


Figure 34: MDS for the Euclidian distance model for the pre-selected k-observers, w.b.009

The 4-station k-observer that excludes DP059 has the lowest Euclidean distance compared to reference observer WB009. DP057,DP059,DP061 has the lowest Euclidian distance of the 3-observers and DP057,DP061 has the lowest Euclidian distance for the 2-observers.

<u>Optimal k-observer of sampling stations for water body 009</u>

The previous analyses confirmed that spatial reduction is possible for water body 009 and there are a few alternatives

Finally, 6 k-observers are selected as most appropriate for the spatial reduction:

Sampling stations	N. of reduced s.s.	90th Percentile
DP056 DP057 DP059 DP061 DP062	0	8,224
DP056 DP057 DP061 DP062	1	8,795
DP057 DP059 DP061 DP062	1	8,602
DP056 DP057 DP061	2	7,951
DP056 DP059 DP061	2	8,037
DP056 DP059 DP062	2	7,675
DP057 DP059 DP061	2	7,642

Table 15: Selected k-observers and the 90th percentiles obtained by their data subsets

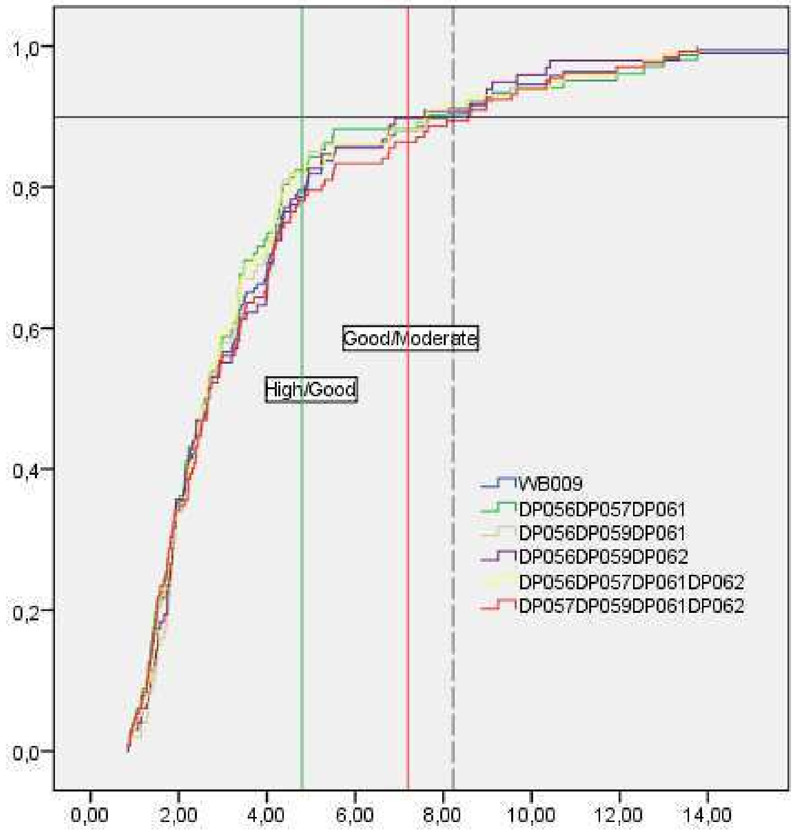


Figure 35: Quantile plot for the selected k-observers of chlorophyll a concentration for the three year period of August 2005 - July 2008, water body 009

Finally, the quantile plot (Figure 35) shows that the empirical CDF of the data subset taken by the selected k-observers is similar to the empirical CDF of the "original" dataset taken by the reference observer. Also, it is apparent that the CDFs have the highest dispersion in the area from the 80th to the 87th percentile.

Analysis for Water Body 010

Water Body 010 is surveyed by six different sampling stations (Annex 2). It is classified as a type II-A water body (significant continental influence and salinity variance). The previous exploratory study determined that spatial reduction is possible without consequences to ecological classification. The following analysis will attempt to identify the most adequate k-observers for the spatial reduction.

<u>*Comparative study*</u>

Multiple box-and-whisker plots have been used to compare the range, median, quartiles and outliers of the measurements of chlorophyll *a* obtained by the six sampling stations located in water body 010. Graphs have been made for the three annual periods and the three-year period.

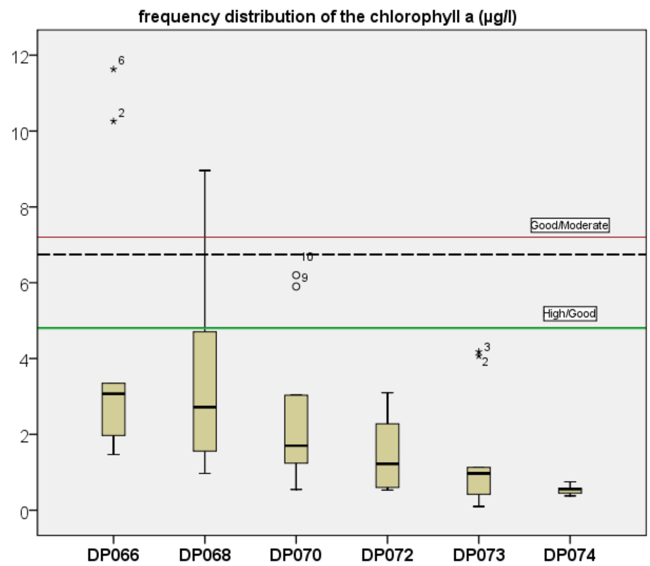


Figure 36 :Box-and-whisker plot of chlorophyll a concentration for the annual period of August 2005 - July 2006 in the water body 010; the red line represents the Good/Moderate threshold, the Green line represents the High/Good threshold, and the broken line represents the 90th percentile

During the annual period of August 2005 - July 2006, it is clear that the highest concentrations of chlorophyll *a* in water body 010 were taken mainly by station DP068 (Figure 36). Some samples analyzed for DP066 have concentrations that exceed the Good/Moderate threshold.

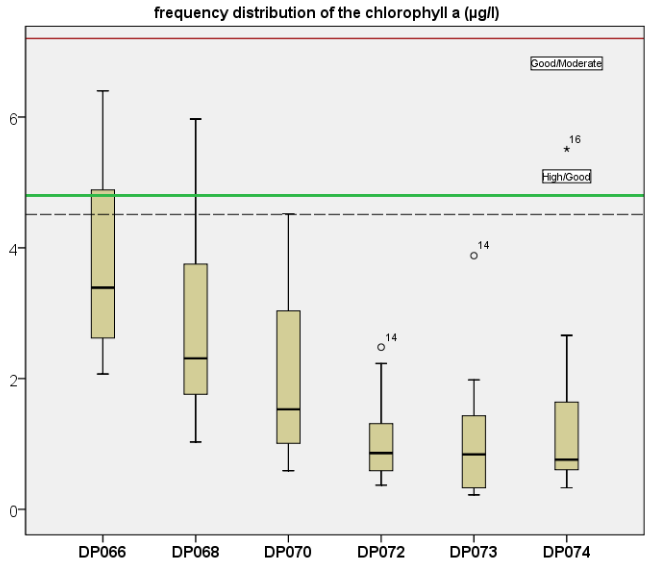


Figure 37: Box-and-whisker plot of chlorophyll a concentration for the annual period of August 2006 - July 2007 in the water body 010; the red line represents the Good/Moderate threshold, the Green line represents the High/Good threshold, and the broken line represents the 90th percentile

For the second annual period, it is clear that the highest concentrations of chlorophyll *a* were taken by sampling stations DP066 and DP068 (Figure 37). Only stations DP066, DP068 and DP075 took samples that exceed the High/Good threshold.

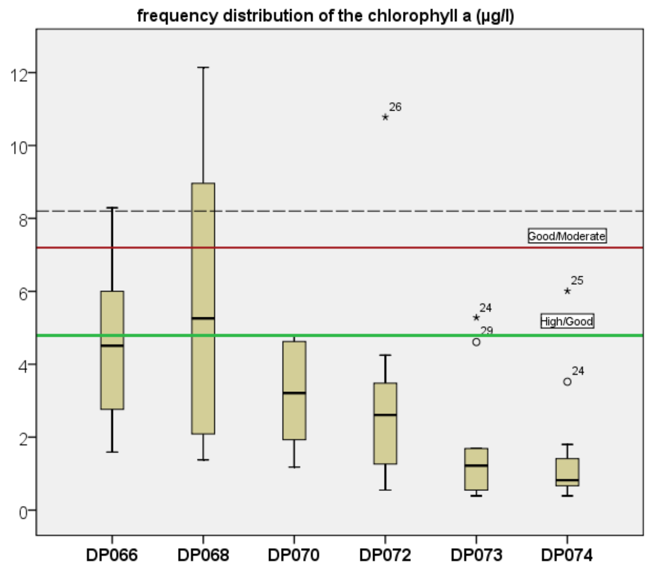


Figure 38: Box-and-whisker plot of chlorophyll a concentration for the annual period of August 2007 - July 2008 in the water body 010; the red line represents the Good/Moderate threshold, the Green line represents the High/Good threshold, and the broken line represents the 90th percentile

For the third annual period, the box-and-whisker plot shows that the highest concentrations of chlorophyll *a* have been recorded by sampling stations DP066 and

DP068 (Figure 38). Station DP072 analyzed one sample with a concentration that exceeded the Good/Moderate threshold.

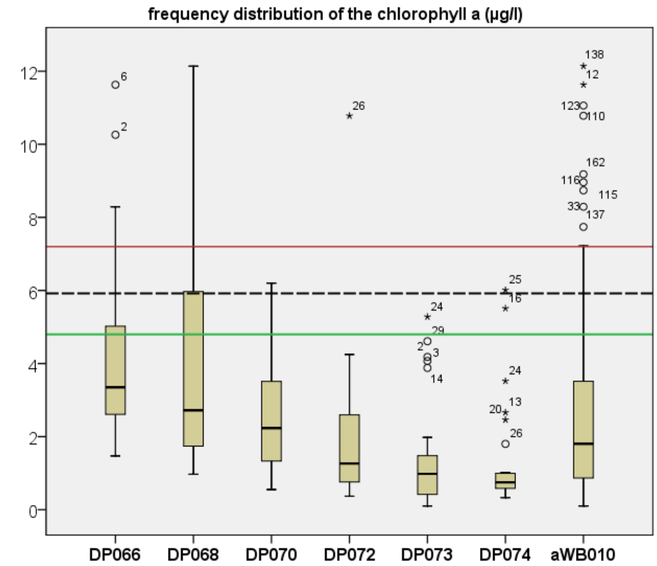


Figure 39: Box-and-whisker plot of chlorophyll a concentration for the three year period of August 2005 - July 2008 in water body 010; The green line represents the High/Good threshold, the red line represents the Good/Moderate threshold, and the broken line represents the 90th percentile

The box-and-whisker plot for the three year period clearly shows that the highest concentrations of chlorophyll *a* for water body 010 were recorded mostly by stations DP066 and DP068 (Figure 39). Analyzing this diagram, it is clear that water body 010 has sampling stations that can be divided into three groups. The first group consists of stations DP073 and DP074, which have lower values of chlorophyll *a* concentrations for the water body. The second group consists of DP066 and DP068, which have significantly higher values, and the third group consists of DP070 and DP072, which have intermediate chlorophyll *a* values for the water body.

The division into three groups is not as clear on the quantile graphs where the empirical CDF of all sampling stations can be compared as it is on the box-and-whisker diagram. On the quantile diagram, only two groups are observable. The datasets taken in the sampling stations that belong to the first and third group have a 90^{th} percentile lower than the 90^{th} percentile of the reference dataset. The datasets taken in the sampling stations that belong to the second group have a 90^{th} percentile higher than the 90^{th} percentile of the reference dataset. This tells us that if we want to select a k-observer of sampling stations where the set of measurements taken have a 90^{th} percentile similar to the 90^{th} percentile of the full sample, it will be necessary to select sampling stations from the second and third or the first group.

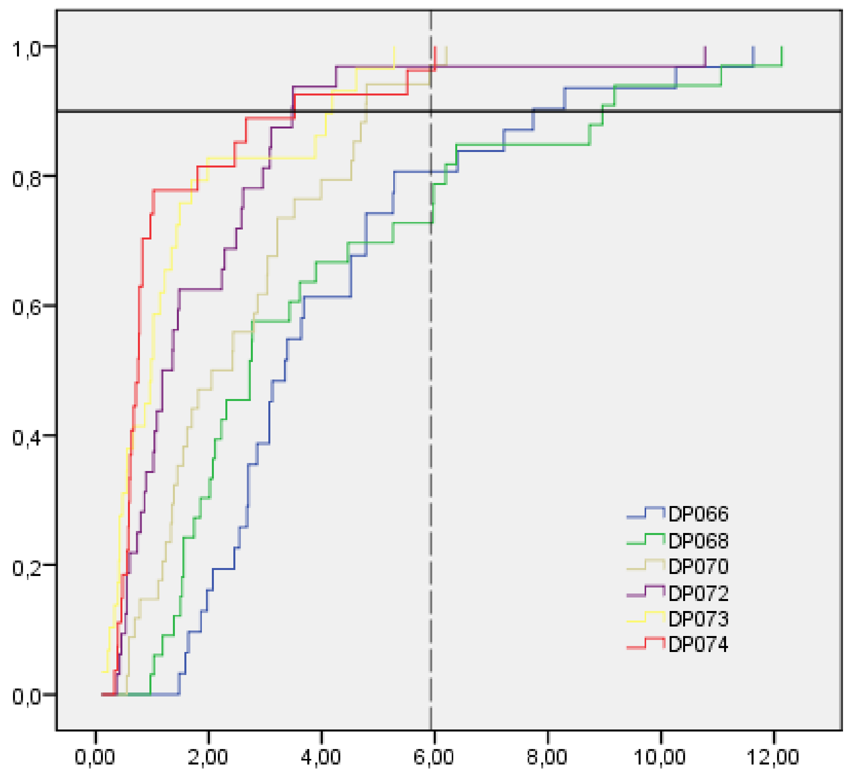


Figure 40: Quantile plot of chlorophyll a concentration for all sampling stations for the three year period August 2005 - July 2008, water body 010

<u>Pre-selection of k-observers of the sampling stations based on the deviation of the 90th percentile</u>

Observing table Table 16 (taken from the Exploratory study and aligned by the three year bias), the results confirm what was easily observable on the previous quantile plot – the k-observers (where k<5) have a lower bias on the 90th percentile if they include sampling stations from the first or third and second group. Also, it is visible that these combinations include only one sampling station that belongs to the second group (DP066 or DP068).

The highest biases of the 90th percentile have been obtained for the 1-observer DP068 and the 2-observer DP066,DP068, because they only include sampling stations from the second group. The 3-station and 4-observers that include only stations from the first and third groups have the highest bias.

k-observer	bias 90 PCTL (3 years)	eco. class. (3 years)	- n stations	average absolute error (4 periods)	Corresponding eco. class. (4 periods)
DP068 DP070 DP072 DP074	**0,00**	**corresp.**	**2**	**0,318**	**yes**
DP068 DP070 DP073	**0,05**	**corresp.**	**3**	**0,365**	**yes**
DP068 DP072 DP073	**0,05**	**corresp.**	**3**	**0,878**	**yes**
DP066 DP068 DP070 DP072 DP073	**0,06**	**corresp.**	**1**	**0,180**	**yes**
DP068 DP073 DP074	**0,06**	**corresp.**	**3**	**0,343**	**yes**
DP068 DP070 DP072	**0,06**	**corresp.**	**3**	**0,643**	**yes**
DP066 DP068 DP070 DP073 DP074	**0,07**	**corresp.**	**1**	**0,150**	**no**
DP068 DP070 DP074	**0,07**	**corresp.**	**3**	**0,310**	**yes**
DP066 DP068 DP072 DP073 DP074	**0,08**	**corresp.**	**1**	**0,260**	**no**
DP068 DP072 DP074	**0,08**	**corresp.**	**3**	**0,625**	**yes**
DP066 DP068 DP070 DP072 DP074	**0,12**	**corresp.**	**1**	**0,243**	**yes**
DP066 DP070 DP074	**0,14**	**corresp.**	**3**	**0,995**	**no**
DP068 DP070 DP073 DP074	**0,18**	**corresp.**	**2**	**0,120**	**yes**
DP066 DP068 DP070 DP073	0,32	corresp.	2	0,365	no
DP068 DP070	0,32	corresp.	4	0,745	yes
DP066 DP068 DP070 DP074	0,35	corresp.	2	0,653	no
DP066 DP070	0,36	corresp.	4	1,348	no
DP066 DP073	0,37	corresp.	4	1,428	no
DP068 DP073	0,40	corresp.	4	0,883	yes
DP066 DP072 DP074	0,43	corresp.	3	0,615	no
DP066 DP068 DP070 DP072	0,44	corresp.	2	0,465	no
DP066 DP068 DP073 DP074	0,44	corresp.	2	0,765	no
DP068 DP074	0,44	corresp.	4	1,153	no
DP068 DP072 DP073 DP074	0,46	corresp.	2	0,853	no
DP066 DP068 DP072 DP073	0,47	corresp.	2	0,525	no
DP066 DP068 DP072 DP074	0,47	corresp.	2	0,860	no
DP066 DP070 DP072	0,52	corresp.	3	0,385	no
DP066 DP074	0,56	corresp.	4	1,713	no
DP068 DP070 DP072 DP073	0,58	corresp.	2	0,563	yes
DP066 DP073 DP074	0,60	corresp.	3	0,900	no
DP066 DP070 DP073	0,64	corresp.	3	0,608	no
DP068 DP070 DP072 DP073 DP074	0,65	corresp.	1	0,390	yes
DP066 DP070 DP072 DP074	0,65	corresp.	2	0,593	no
DP066 DP070 DP073 DP074	0,65	corresp.	2	0,815	no
DP066 DP072 DP073	0,65	corresp.	3	0,848	no
DP066 DP072 DP073 DP074	0,66	corresp.	2	1,043	no
DP066 DP072	0,97	corresp.	4	1,518	no
DP066 DP070 DP072 DP073	0,99	corresp.	2	0,668	no
DP066 DP070 DP072 DP073 DP074	1,14	upgrade	1	1,150	no
DP070	1,14	upgrade	5	1,160	no
DP070 DP074	1,16	upgrade	4	1,155	no

DP070 DP073	1,26	upgrade	4	1,378	no
DP070 DP072	1,32	upgrade	4	1,515	no
DP070 DP073 DP074	1,32	upgrade	3	1,610	no
DP066 DP068 DP070	1,35	degrade	3	1,368	no
DP070 DP072 DP074	1,38	upgrade	3	1,900	no
DP070 DP072 DP073	1,39	upgrade	3	1,848	no
DP068 DP072	1,40	degrade	4	1,175	no
DP070 DP072 DP073 DP074	1,48	upgrade	2	1,900	no
DP066 DP068 DP073	1,61	degrade	3	1,525	no
DP066 DP068 DP074	1,72	degrade	3	1,768	no
DP073	1,74	upgrade	5	1,473	no
DP073 DP074	1,82	upgrade	4	1,853	no
DP072 DP073	1,89	upgrade	4	2,208	no
DP066 DP068 DP072	1,98	degrade	3	1,880	no
DP074	2,00	upgrade	5	1,647	no
DP072 DP073 DP074	2,02	upgrade	3	2,333	no
DP066	2,26	degrade	5	1,363	no
DP072 DP074	2,43	upgrade	4	2,470	no
DP072	2,44	upgrade	5	2,228	no
DP066 DP068	2,93	degrade	4	2,693	no
DP068	3,17	degrade	5	2,653	no

Table 16: 90th percentile biases calculated by k- combinations and ecological classification correspondence for water body 010

The low 90th percentile bias allows us to suppose that the possibility of spatial reduction exists, and the first 13 k-observers with the lowest bias are pre-selected for a more in-depth analysis. For water body 010 (six sampling stations - 62 k-observers) 13 k-observers have been preselected due to the significant difference between the 13th (0.18) and 14th (0.32) k-observer bias value.
Nevertheless, the same analysis will be performed for the k-observers with the highest 90th percentile bias in order to highlight the difference.

Interclass correlation coefficient

Results obtained by the ICC have been interpreted using a *qualitative scale* developed by *Landis and Koch*, typical for these types of studies. According to this scale, correlation between items is considered very good (0.90 to 1), good (0.80 to 0.90), moderate (0.60 to 0.80) or poor (0 to 0.60). k-observers that have an ICC > 0,950 are marked as Very Good*.

	Intraclass correlation[a]	Confidence interval 95%		
		Lower limit	Upper limit	Value
DP068DP070DP072DP074	**,964[c]**	**,926**	**,982**	**Very Good***
DP068DP070DP073	**,963c**	**,927**	**,982**	**Very Good***
DP068DP072DP073	**,952c**	**,904**	**,976**	**Very Good***
DP066DP068DP070DP072DP073	**,990c**	**,957**	**,996**	**Very Good***
DP068DP073DP074	**,964c**	**,927**	**,982**	**Very Good***
DP068DP070DP072	,943c	,886	,971	Very Good
DP066DP068DP070DP073DP074	**,988c**	**,972**	**,994**	**Very Good***
DP068DP070DP074	,936c	,872	,968	Very Good
DP066DP068DP072DP073DP074	**,990c**	**,980**	**,995**	**Very Good***
DP068DP072DP074	**,972c**	**,941**	**,986**	**Very Good***
DP066DP068DP070DP072DP074	**,991c**	**,968**	**,997**	**Very Good***
DP066DP0702DP074	,946c	,893	,973	Very Good
DP068DP070DP073DP074	**,971c**	**,936**	**,986**	**Very Good***
DP066DP068	,789c	,085	,927	Moderate
DP066	,745c	,097	,905	Moderate
DP068	,738c	,387	,879	Moderate

Table 17: ICCs of the preselected k-observers, confidence intervals and interpretation of ICCs according to the Landis and Koch qualitative scale.

The 5-observer that excludes DP073 has the highest ICC (0.991). The 4-observer with the highest ICC is DP068,DP070,DP073,DP073 (0.971). It includes one sampling station from the second group, two from the first group, and one from the third group. The 3-station combination with the highest ICC coefficient is DP068,DP72,DP074 (0.972) and includes one station from each group. All 13 *preselected observers* have an ICC value higher than 0.9 and have “Very Good” concordance with the *reference observer*.

For the preselected k-observers that produce the highest bias with respect to the reference 90^{th} percentile for the three year period (DP068; DP066 and DP066,DP068), “Moderate” concordance to the complete data is assigned. The ICC coefficients calculated for the k-observers are 0.738, 0.745 and 0.789 respectively

Euclidean distance and MDS

Calculating and comparing the Euclidean distance between chlorophyll *a* average series of analyzed k-observers, output confirms the results of previous analysis. The output confirms the results of the previous analysis. Results are presented using the Multidimensional Scaling technique. In the two dimensional space in which the distances are presented, it is easy to identify which observers verify that the Euclidean distance between the monthly averages of chlorophyll *a* concentration in their stations and the monthly averages of chlorophyll *a* concentration in the sampling stations of the reference observer (WB010) is less relevant.

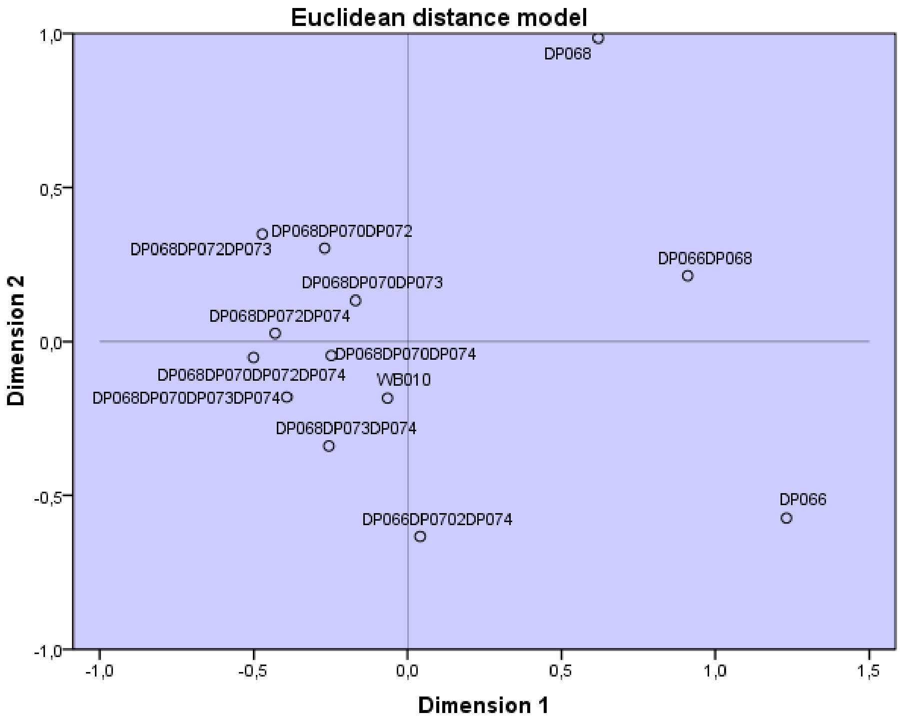


Figure 41: MDS for the Euclidian distance model for the pre-selected k-observers, w.b.010

The 5-observers have been excluded from the MDS because the ICCs are elevated and vary between 0.988 and 0.991. It is clear that these 5-observers have the lowest Euclidean distance value.

The two 3-observers DP068,DP070,DP074 and DP068,DP073,DP074 have the lowest Euclidean distance compared to reference observer WB010. DP068,DP070,DP073,DP074 has the lowest Euclidian distance of the 4-observers.

The two axes are associated to sampling stations DP066 and DP068 respectively.

Optimal k-observer of sampling stations for water body 010

The previous analyses confirmed that spatial reduction is possible for water body 010 and there are various alternatives.

Finally, 10 k-observers are selected as most appropriate for the spatial reduction:

Sampling stations	N. of reduced s.s.	90th Percentile
DP066 DP068 DP070 DP072 DP073 DP074	**0**	**5,921**
DP066 DP068 DP072 DP073 DP074	1	6,001
DP066 DP068 DP070 DP072 DP073	1	5,980
DP066 DP068 DP070 DP073 DP074	1	5,995
DP068 DP070 DP073 DP074	2	5,744
DP068 DP070 DP072 DP074	2	5,921
DP068 DP072 DP073	3	5,975
DP068 DP073 DP074	3	5,980
DP068 DP070 DP074	3	5,995
DP068 DP070 DP073	3	5,973
DP068DP072DP074	3	6,001

Table 18: Selected k-observers and the 90th percentiles obtained by their data subsets

Finally, the quantile plot (Figure 42) shows that the empirical CDF of the data subset taken by the selected k-observers is similar to the empirical CDF of the "original" dataset taken by the reference observer.

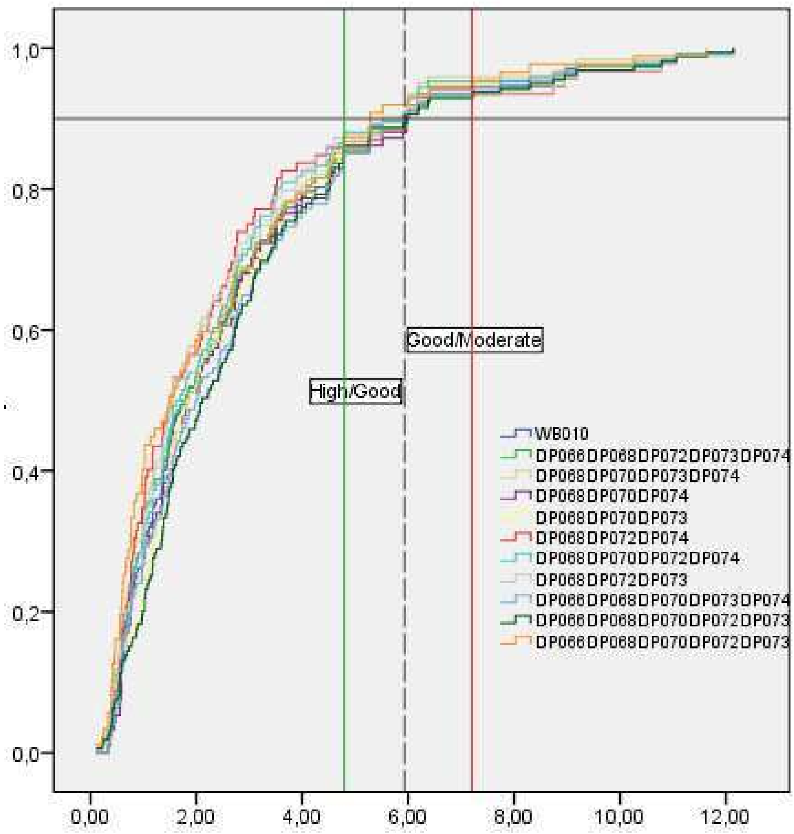


Figure 42: Quantile plot for the selected k-observers of chlorophyll a concentration for the three year period of August 2005 - July 2008, water body 010

Analysis for Water Body 011

Water Body 011 is surveyed by six different sampling stations (Annex 2). It is classified as a type III water body (without continental influence). The previous exploratory study determined that spatial reduction is possible without consequences to ecological classification. The following analysis will attempt to identify the most adequate k-observers for the spatial reduction.

<u>*Comparative study*</u>

Multiple box-and-whisker plots have been used to compare the range, median, quartiles and outliers of the measurements of chlorophyll *a* obtained by the six sampling stations located in water body 011. Graphs have been made for the three annual periods and the three-year period.

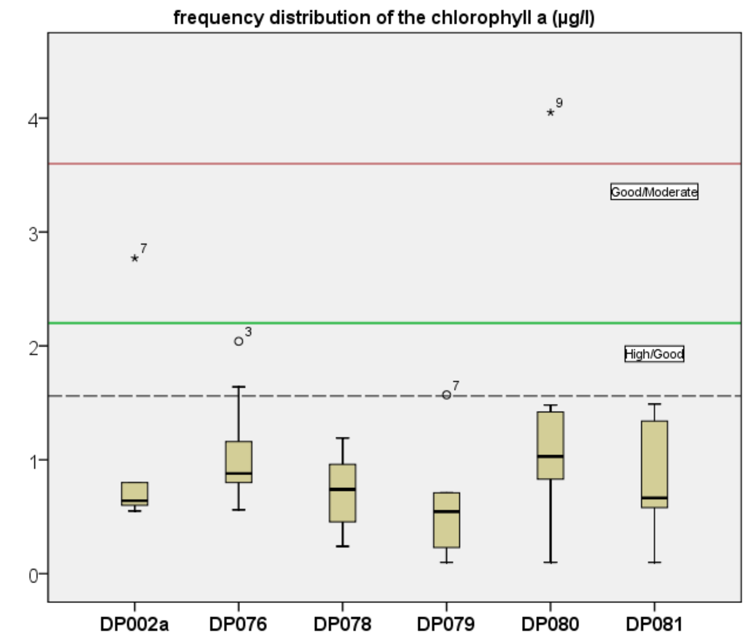


Figure 43: Box-and-whisker plot of chlorophyll a concentration for the annual period of August 2005 - July 2006 in the water body 011; the red line represents the Good/Moderate threshold, the Green line represents the High/Good threshold, and the broken line represents the 90th percentile

During the annual period of August 2005 - July 2006, it is clear that the highest concentrations of chlorophyll *a* in water body 011 were taken by stations DP080, DP076 and DP081 (Figure 43). One sample analyzed from was analyzed one sample that exceeds the Good/Moderate threshold, and one for DP002a that exceeds the High/Good threshold, therefore the other samples present low chlorophyll *a* concentrations.

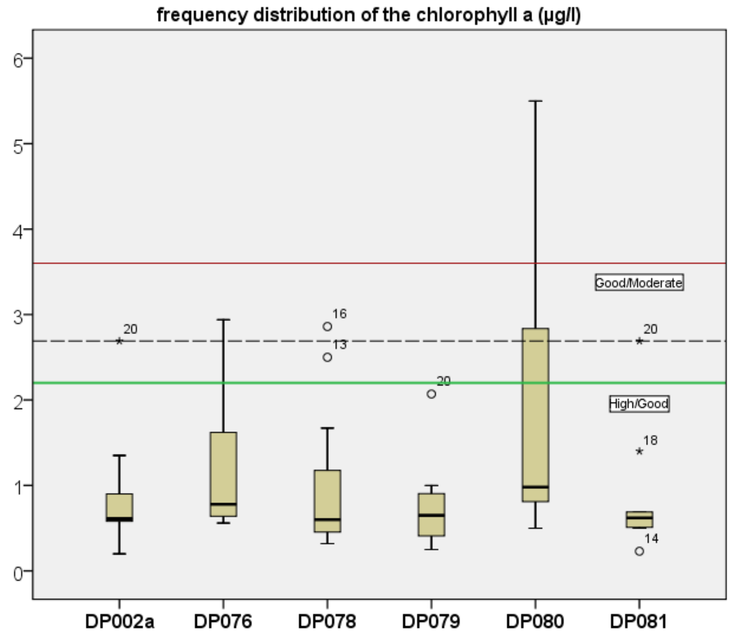


Figure 44: Box-and-whisker plot of chlorophyll a concentration for the annual period of August 2006 - July 2007 in the water body 011; the red line represents the Good/Moderate threshold, the Green line represents the High/Good threshold, and the broken line represents the 90th percentile

For the second annual period, it is clear that the highest concentrations of chlorophyll *a* were taken by sampling stations DP080 and DP076 (Figure 44). Only station DP080 took samples that exceed the Good/Moderate threshold.

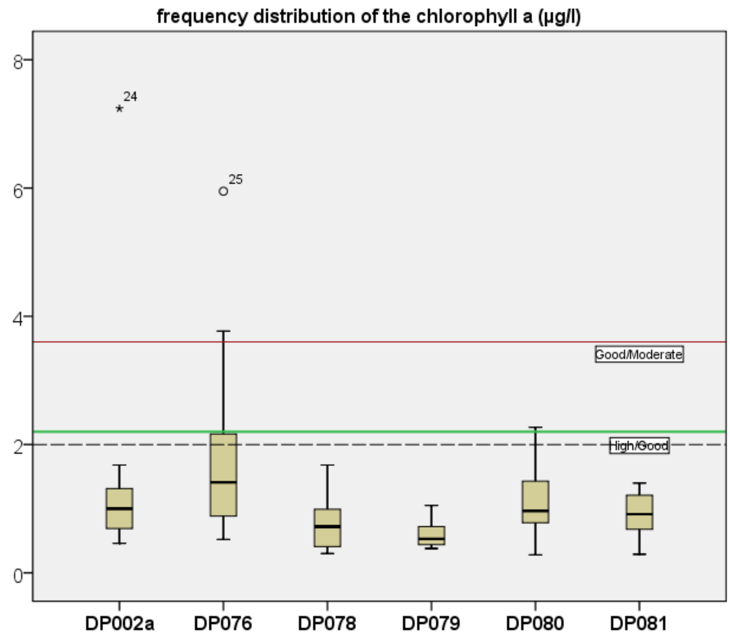


Figure 45: Box-and-whisker plot of chlorophyll a concentration for the annual period of August 2007 - July 2008 in the water body 011; the red line represents the Good/Moderate threshold, the Green line represents the High/Good threshold, and the broken line represents the 90th percentile

For the third annual period, the box-and-whisker plot shows that the highest concentrations of chlorophyll *a* have been mainly recorded by sampling stations DP080 and DP076 (Figure 45). Sampling station DP002a analyzed one sample that has a chlorophyll a concentration that exceeds the Good/Moderate threshold.

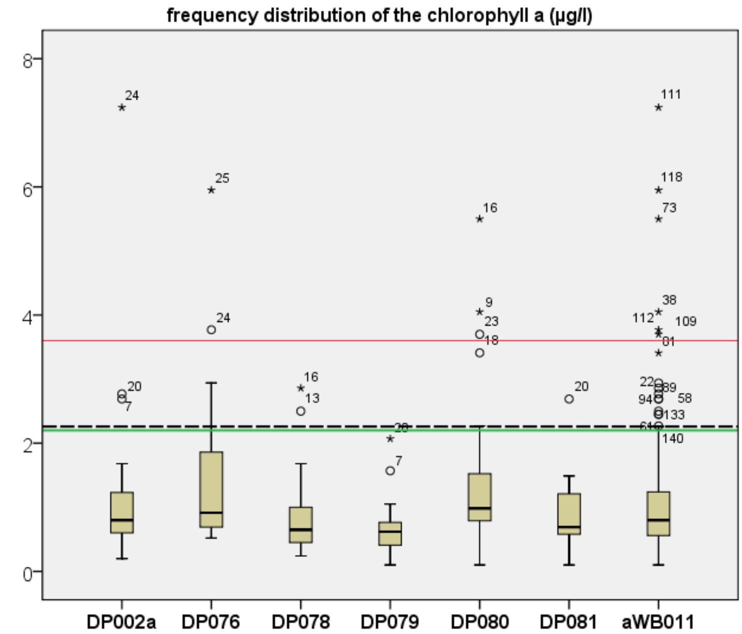


Figure 46: Box-and-whisker plot of chlorophyll a concentration for the three year period of August 2005 - July 2008 in water body 011; The green line represents the High/Good threshold, the red line represents the Good/Moderate threshold, and the broken line represents the 90th percentile

The box-and-whisker plot for the three year period clearly shows that the highest concentrations of chlorophyll *a* for water body 011 were recorded mostly by stations DP080 and DP076 (Figure 46). It is clear that DP078 and DP079 belong to the group with the lowest chlorophyll a concentrations. The second group with elevated chlorophyll *a* concentrations consists of DP076 and DP080, but it is unclear which group DP002a and DP081 belong to.

The optimal k-observer of sampling stations should have an empirical CDF similar to the empirical CDF of aWB011, which is included on the last box-and-whisker plot.

The overlapping graphs corresponding to the empirical CDF of each sampling station allow us to compare the 90th percentiles of these distributions (Figure 47). The datasets taken by the sampling stations that belong to the first group have a 90^{th} percentile lower than the 90^{th} percentile of the reference dataset, and the datasets taken by the sampling stations that belong to the second group have a 90^{th} percentile higher than this value. Observing the quantile diagram, it is clear that DP081 belongs to the first group. Classifying station DP002a is more complicated. The empirical CDF of DP002a follows the empirical CDF of the first group stations, but just before the 90^{th} percentile value jumps into the area of the second group.

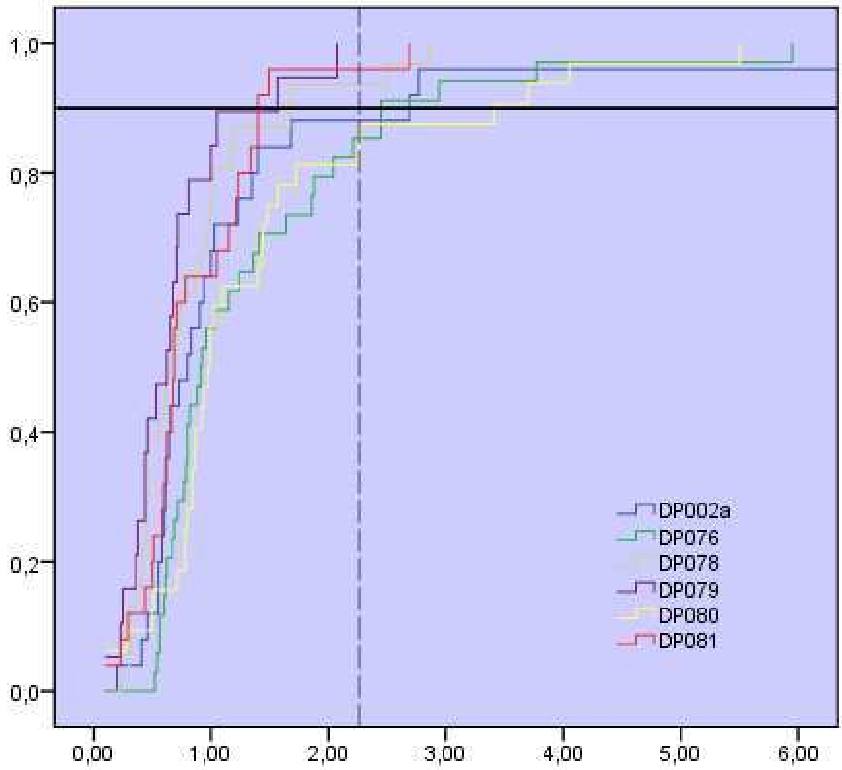


Figure 47:Quantile plot of chlorophyll a concentration for all sampling stations for the three year period August 2005 - July 2008, water body 011

Pre-selection of k-observers of the sampling stations based on the deviation of the 90th percentile

Observing Table 19 (taken from the Exploratory study and aligned by the three year bias), the results confirm what was easily observable on the previous quantile plot – the k-observers (where k<5) have a lower bias on the 90th percentile if they include sampling stations from both groups. Also, it is visible that these combinations include only one sampling station that belongs to the second group (DP076 or DP080).

As for the previous water bodies, the highest biases of the 90th percentile have been obtained for 1-observer DP080 and 2-observer DP080,DP076 because they only include sampling stations from the second group.
DP078,DP079,DP081, which includes only stations that belong to the first group, has the highest bias of the 3-observers.
The low 90th percentile bias allows us to suppose that the possibility of spatial reduction exists, and the first 15 k-observers with the lowest bias are pre-selected for a more in-depth analysis. For water body 011 15 k-observers have been preselected due to the jump in the bias value (15th - 0.09 and 16th -0.11) of the k-observers.
Nevertheless, the same analysis will be performed for the k-observers with the highest 90th percentile bias in order to highlight the difference.

k-observer	bias 90 PCTL (3 years)	eco. class. (3 years)	- n stations	average absolute error (4 periods)	Corresponding eco. class. (4 periods)
DP078 DP080 DP081	**0,00**	**corresp.**	**3**	**0,280**	**yes**
DP079 DP080	**0,00**	**corresp.**	**4**	**0,418**	**no**
DP002a DP078 DP079 DP080	**0,001**	**corresp.**	**2**	**0,130**	**yes**
DP002a DP078 DP080 DP081	**0,002**	**corresp.**	**2**	**0,138**	**yes**
DP076 DP078 DP079 DP080 DP081	**0,01**	**corresp.**	**1**	**0,058**	**yes**
DP076 DP079 DP080 DP081	**0,01**	**corresp.**	**2**	**0,150**	**no**
DP002a DP076 DP078 DP081	**0,04**	**corresp.**	**2**	**0,110**	**no**
DP002a DP079 DP080 DP081	**0,04**	**corresp.**	**2**	**0,178**	**yes**
DP002a DP076 DP078 DP079	**0,05**	**corresp.**	**2**	**0,288**	**no**
DP076 DP078 DP079 DP080	**0,06**	**corresp.**	**2**	**0,160**	**no**
DP078 DP079 DP080	**0,06**	**corresp.**	**3**	**0,295**	**yes**
DP076 DP078 DP081	**0,07**	**upgrade**	**3**	**0,143**	**no**
DP002a DP076 DP079 DP080 DP081	**0,08**	**corresp.**	**1**	**0,118**	**no**
DP080 DP081	**0,09**	**corresp.**	**4**	**0,488**	**no**
DP076 DP079	**0,09**	**corresp.**	**4**	**0,733**	**no**
DP002a DP076 DP079 DP081	0,11	upgrade	2	0,290	no
DP076 DP078 DP079	0,12	upgrade	3	0,330	no
DP076 DP078 DP080 DP081	0,13	corresp.	2	0,163	no
DP002a DP079 DP080	0,13	corresp.	3	0,468	no
DP079 DP080 DP081	0,14	upgrade	3	0,355	no
DP002a DP076 DP078 DP079 DP080	0,15	corresp.	1	0,163	no
DP078 DP080	0,15	corresp.	4	0,370	no
DP076 DP079 DP081	0,18	upgrade	3	0,330	no
DP002a DP076 DP078 DP080 DP081	0,19	corresp.	1	0,125	yes
DP076 DP080 DP081	0,19	corresp.	3	0,325	no
DP002a DP076 DP079 DP080	0,19	corresp.	2	0,363	no
DP076 DP079 DP080	0,19	corresp.	3	0,448	no
DP076 DP078	0,19	corresp.	4	0,465	no
DP076 DP081	0,19	corresp.	4	0,478	no
DP002a DP076 DP078	0,19	corresp.	3	0,480	no
DP002a DP076 DP081	0,19	corresp.	3	0,485	no
DP002a DP076 DP079	0,19	corresp.	3	0,750	no
DP076 DP078 DP080	0,20	corresp.	3	0,285	no
DP002a DP076 DP078 DP079 DP081	0,21	upgrade	1	0,115	no
DP076 DP078 DP079 DP081	0,22	upgrade	2	0,140	no
DP002a DP076 DP078 DP080	0,22	corresp.	2	0,243	no
DP002a DP076 DP080 DP081	0,26	corresp.	2	0,295	no
DP002a DP078 DP080	0,26	corresp.	3	0,343	yes
DP002a DP078 DP079 DP080 DP081	0,30	upgrade	1	0,195	no
DP002a DP080 DP081	0,30	corresp.	3	0,548	no

DP002a DP078	0,34	upgrade	4	0,195	no
DP002a DP079	0,39	upgrade	4	0,803	no
DP076	0,43	corresp.	5	1,110	no
DP002a DP076	0,43	corresp.	4	1,120	no
DP002a	0,46	corresp.	5	2,295	no
DP078 DP079 DP080 DP081	0,47	upgrade	2	0,305	no
DP002a DP076 DP080	0,49	corresp.	3	0,768	no
DP002a DP078 DP079	0,58	upgrade	3	0,293	no
DP002a DP078 DP081	0,59	upgrade	3	0,265	no
DP078	0,59	upgrade	5	0,353	no
DP002a DP078 DP079 DP081	0,60	upgrade	2	0,325	no
DP078 DP079	0,60	upgrade	4	0,430	no
DP002a DP081	0,60	upgrade	4	0,483	no
DP002a DP080	0,64	corresp.	4	0,840	no
DP002a DP079 DP081	0,69	upgrade	3	0,473	no
DP079	0,69	upgrade	5	0,690	no
DP078 DP081	0,72	upgrade	4	0,383	no
DP078 DP079 DP081	0,74	upgrade	3	0,433	no
DP079 DP081	0,82	upgrade	4	0,508	no
DP076 DP080	0,82	corresp.	4	0,895	no
DP081	0,83	upgrade	5	0,725	no
DP080	1,35	degrade	5	1,495	no

Table 19: 90th percentile biases calculated by k- combinations and ecological classification correspondence for water body 011

Interclass correlation coefficient

Results obtained by the ICC have been interpreted using a *qualitative scale* developed by *Landis and Koch*, typical for these types of studies. According to this scale, correlation between items is considered very good (0.90 to 1), good (0.80 to 0.90), moderate (0.60 to 0.80) or poor (0 to 0.60). k-observers that have an ICC > 0,950 are marked as Very Good*.

The 5-observer that excludes DP078 has the highest ICC (0.987). The 4-observer with the highest ICC is DP076,DP079,DP080,DP081 (0.977). It includes two sampling stations from the second and two from the first group. The 3-observer with the highest ICC coefficient is DP078,DP080,DP081 (0.913). The 2-observers, like the 3-observers, have a low concordance with the reference observer.

Four of the 15 *preselected observers* have an ICC value lower than 0.9 and “Good” concordance with the *reference observer*. The rest of the pre-selected k-observers have an ICC value higher than 0.9 and “Very Good” concordance.

For the preselected observers that produced the highest bias with respect to the reference 90^{th} percentile for the three year period (DP080; DP076 and DP076,DP080) “Moderate” and “Very Good” concordance to the complete data is assigned. The ICC coefficients calculated for the k-observers are 0.727, 0.730 and 0.916 respectively

	Intraclass correlation[a]	Confidence interval95%		
		Lower limit	Upper limit	Value
DP078DP080DP081	,913c	,826	,956	Very Good
DP079DP080	,819c	,633	,911	Good
DP002aDP078DP079DP080	,940c	,880	,970	Very Good
DP002aDP078DP080DP081	**,968c**	**,937**	**,984**	**Very Good***
DP076DP078DP079DP080DP081	**,979c**	**,958**	**,990**	**Very Good***
DP076DP079DP080DP081	**,977c**	**,954**	**,988**	**Very Good***
DP002aDP076DP078DP081	,942c	,885	,971	Very Good
DP002aDP079DP080DP081	,920c	,839	,960	Very Good
DP002aDP076DP078DP079	,935c	,870	,967	Very Good
DP076DP078DP079DP080	**,966c**	**,932**	**,983**	**Very Good***
DP078DP079DP080	,836c	,669	,918	Good
DP076DP078DP081	,910c	,821	,955	Very Good
DP002aDP076DP79DP080DP081	**,987c**	**,973**	**,994**	**Very Good***
DP080DP081	,842c	,685	,921	Good
DP076DP079	,856c	,711	,928	Good
DP076DP080	,916c	,742	,965	Very Good
DP076	,730c	,467	,864	Moderate
DP080	,727[c]	,450	,866	Moderate

Table 20: ICCs of the preselected k-observers, confidence intervals and interpretation of ICCs according to the Landis and Koch qualitative scale.

Euclidean distance and MDS

Calculating and comparing the Euclidean distance between chlorophyll *a* average series of analyzed k-observers, the output confirms the results of the previous analysis. Results are presented using the Multidimensional Scaling technique. In the two dimensional space in which the distances are presented, it is easy to identify which observers verify that the Euclidean distance between the monthly averages of chlorophyll *a* concentration in their stations and the monthly averages of chlorophyll *a* concentration in the sampling stations of the reference observer (WB011) is less relevant

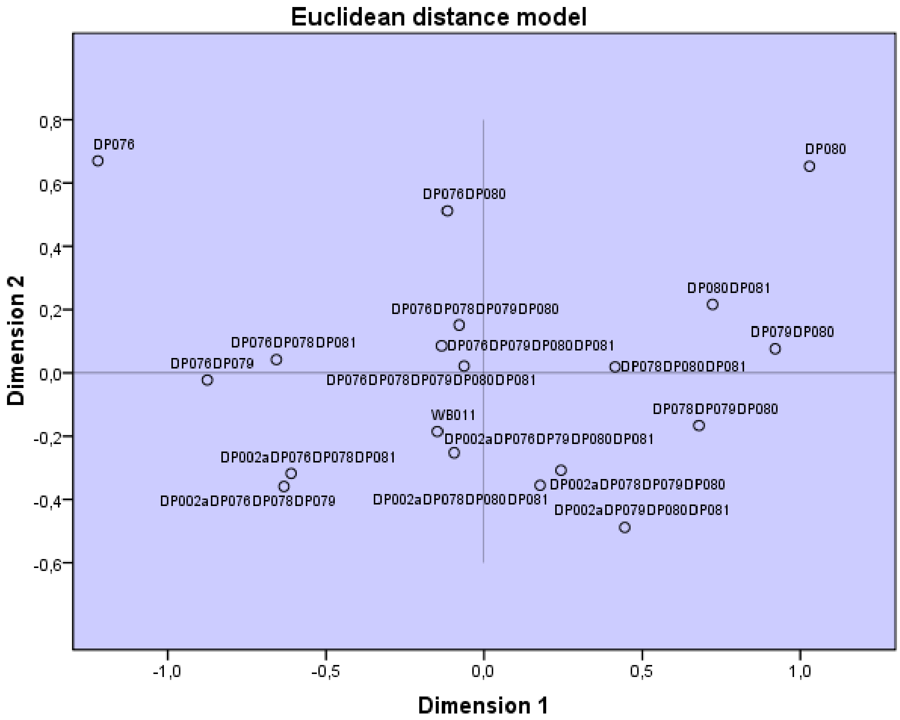


Figure 48: MDS for the Euclidian distance model for the pre-selected k-observers, w.b.011

The 5-observer that excludes DP078 has the lowest Euclidean distance compared to reference observer WB011. DP076,DP079,DP080,DP081 has the lowest Euclidian distance of the 4-observers, and DP078,DP080,DP081 has the lowest Euclidian distance of the 3-observers. All analyzed 2-observers represent a significant Euclidian distance to the reference observer.
1-observers DP076 and DP080 represent the highest Euclidean distance on the model.

Optimal k-observer of sampling stations for water body 011

The previous analyses confirmed that spatial reduction is possible for water body 011 and there are a few alternatives.

Finally, 5 k-observers are selected as most appropriate for the spatial reduction:

Sampling stations	N. of reduced s.s	90th Percentile
DP002a DP076 DP078 DP079 DP080 DP081	0	2,2630
DP076 DP078 DP079 DP080 DP081	1	2,2500
DP002a DP076 DP079 DP080 DP081	1	2,3420
DP076 DP078 DP079 DP080	2	2,3240
DP076 DP079 DP080 DP081	2	2,2690
DP002a DP078 DP080 DP081	2	2,2660

Table 21: Selected k-observers and the 90th percentiles obtained by their data subsets

Finally, the quantile plot (Figure 49) shows that the empirical CDF of the data subset taken by the selected k-observers is similar to the empirical CDF of the "original" dataset taken by the reference observer.

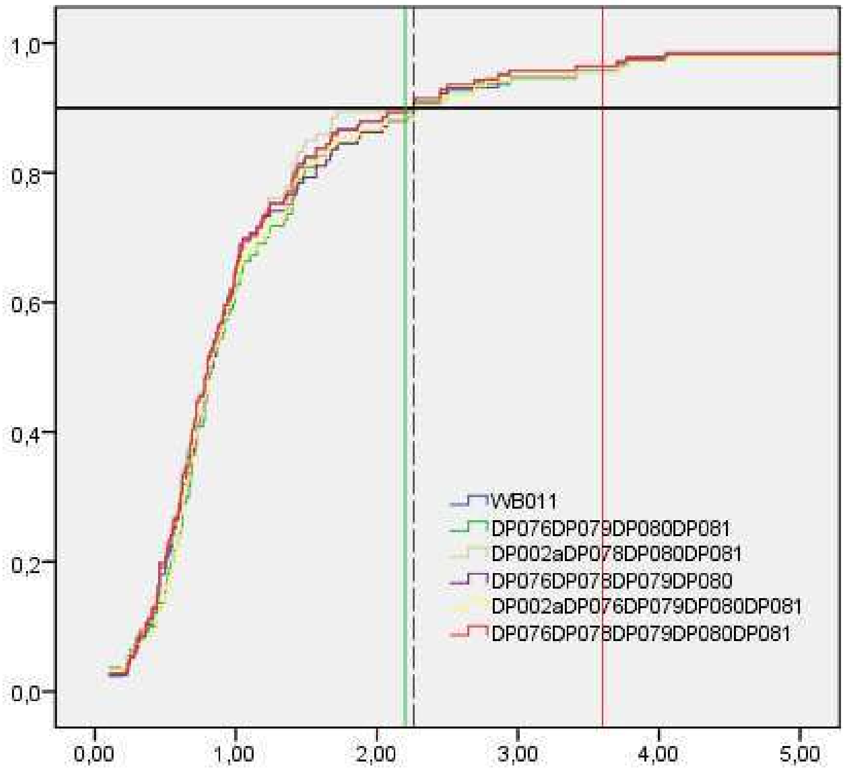


Figure 49: Quantile plot for the selected k-observers of chlorophyll a concentration for the three year period of August 2005 - July 2008, water body 011

The water body has a 90th percentile near the High/Good threshold for the three year period. Even if the quantitative results obtained by the selected k-observers have a low 90th percentile bias, the ecological classification for water body 011 could be different than would be obtained by the reference data.

Analysis for the Water Body 014

Water Body 014 is surveyed by four different sampling stations (Annex 2). It is classified as a type III water body (without continental influence). The previous exploratory study determined that spatial reduction is possible without consequences to ecological classification. The following analysis will attempt to identify the most adequate k-observers for the spatial reduction.

Comparative study

Multiple box-and-whisker plots have been used to compare the range, median, quartiles and outliers of the measurements of chlorophyll *a* obtained by the four sampling stations located in water body 014. Graphs have been made for the three annual periods and the three-year period.

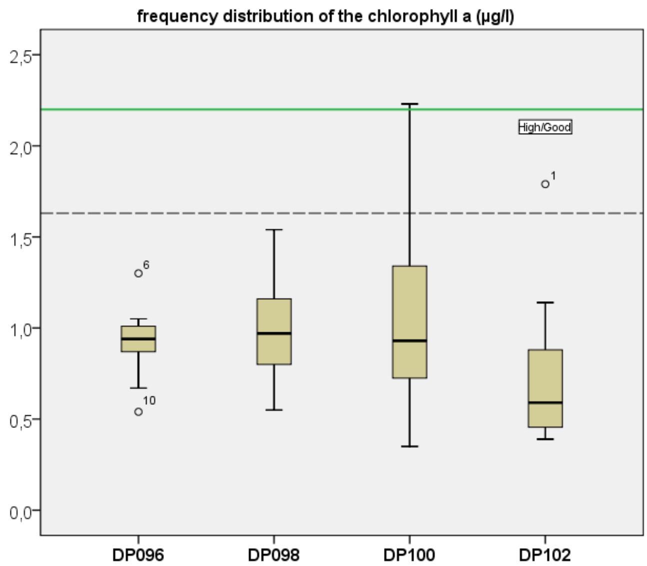


Figure 50: Box-and-whisker plot of chlorophyll a concentration for the annual period of August 2005 - July 2006 in the water body 014; the Green line represents the High/Good threshold, and the broken line represents the 90th percentile

During the annual period of August 2005 - July 2006, it is clear that the highest concentrations of chlorophyll *a* in water body 014 were taken mainly by station DP100 (Figure 50). Of all the stations, only DP100 analyzed samples that exceeded the High/Good threshold.

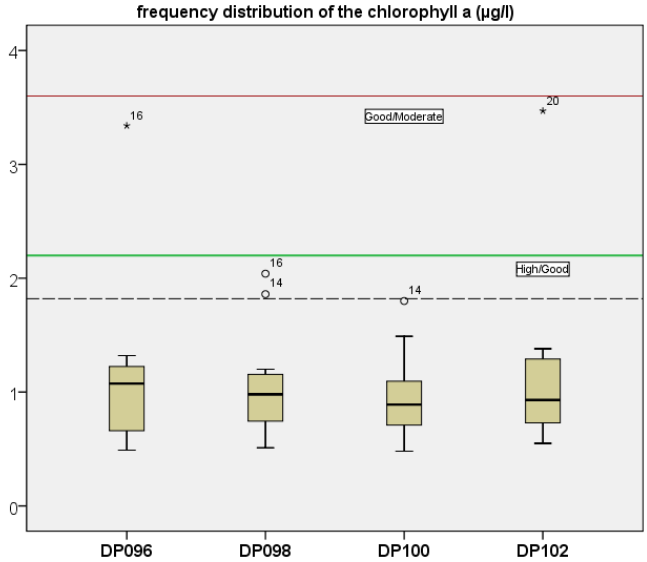


Figure 51: Box-and-whisker plot of chlorophyll a concentration for the annual period of August 2006 - July 2007 in the water body 014; the Green line represents the High/Good threshold, and the broken line represents the 90th percentile

For the second annual period, the analyzed samples mainly have chlorophyll *a* concentrations that not exceed the High/Good threshold. Only station DP096a and DP102 analyzed one sample that significantly exceeds the High/Good threshold.

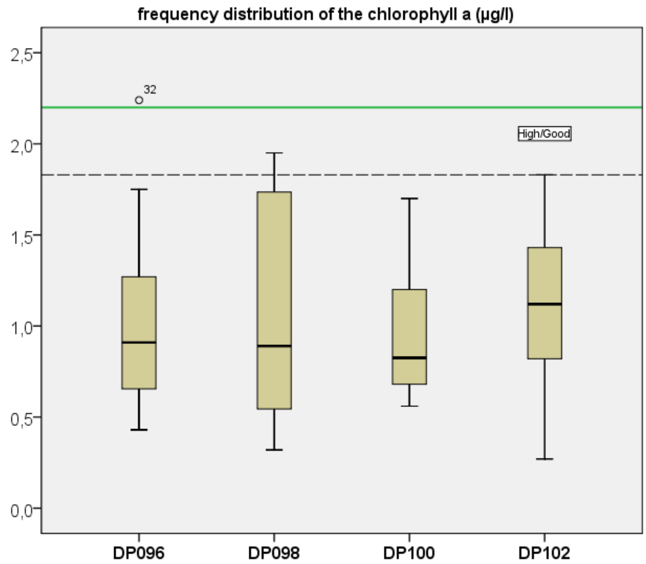


Figure 52: Box-and-whisker plot of chlorophyll a concentration for the annual period of August 2007 - July 2008 in the water body 014; the red line represents the Good/Moderate threshold, the Green line represents the High/Good threshold, and the broken line represents the 90th percentile

For the third annual period, the analyzed samples mainly have chlorophyll *a* concentrations that do not exceed the High/Good threshold. Only station DP096 analyzed one sample that exceeds the High/Good threshold.

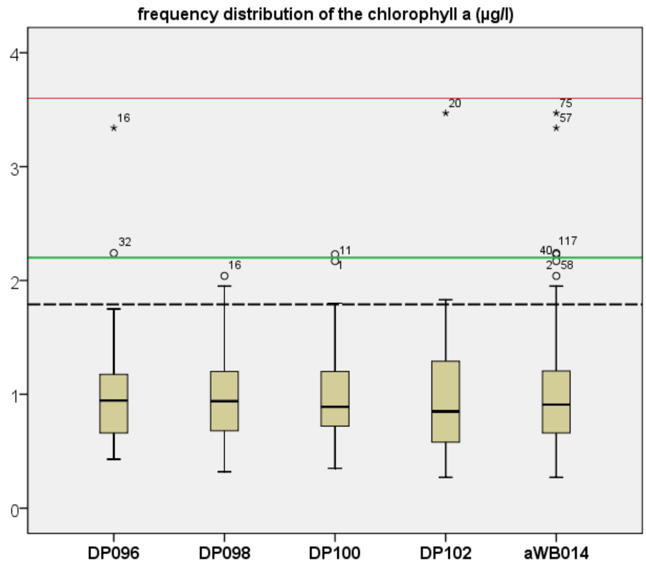


Figure 53: Box-and-whisker plot of chlorophyll a concentration for the three year period of August 2005 - July 2008 in water body 014; The green line represents the High/Good threshold, the red line represents the Good/Moderate threshold, and the broken line represents the 90th percentile

The box-and-whisker plot for the three year period shows that all four sampling stations have a similar empirical CDF. Therefore, the empirical CDF of aWB014 can be compared with the empirical CDF of all four stations.

The overlapping graphs corresponding to the empirical CDF of each sampling station allow us to compare the 90th percentiles of these distributions (Figure 54). It is not possible to establish two groups of sampling stations for water body 014 due to the minimal differences of the empirical CDFs and 90^{th} percentile values. Observing the quantile diagram, it is clear that the CDFs of all sampling stations are aligned and balanced.

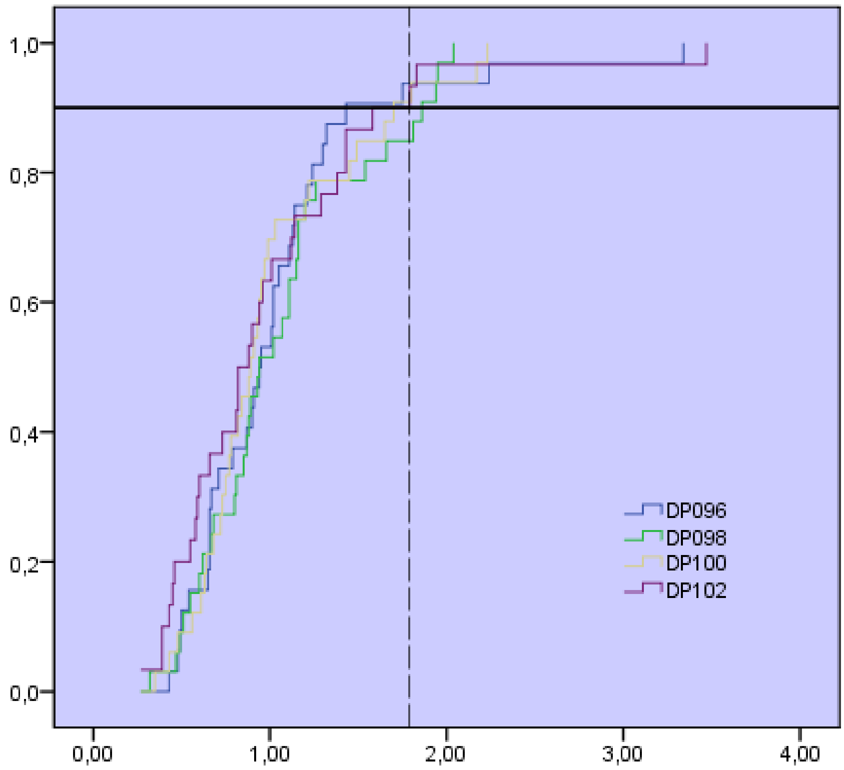


Figure 54: Quantile plot of chlorophyll a concentration for all sampling stations for the three year period August 2005 - July 2008, water body 014

<u>Pre-selection of k-observers of the sampling stations based on the deviation of the 90th percentile</u>

Observing table 00x (taken from the Exploratory study and aligned by the three year bias), the results confirm what was easily observable on the previous quantile plot. All of the k-observers have a low 90th percentile bias (0.007-0.137) due to the extremely low variance and mostly low chlorophyll *a* concentrations that do not exceed the High/Good threshold. The low bias of the k-observers can be expected after analyzing the quantile diagram, which shows a small difference between the 90th percentile of each sampling station.
DP096,DP098,DP102 has the lowest bias for all k-observers (0.007), DP098,DP102 has the lowest bias for the 2-observers (0.031), and DP102 has the lowest bias for the 1-observers (0.022).

The low bias of the 90th percentile for water body 014 allows us to suppose that the possibility of spatial reduction exists for all 14 k-observers

w.b.	k-observer	bias 90 PCTL (3 years)	eco. class. (3 years)	- n stations	average absolute error (4 periods)	Corresponding eco. class. (4 periods)
014	**DP096 DP098 DP102**	**0,007**	**corresp.**	**1**	**0,140**	**yes**
014	**DP096 DP098 DP100**	**0,010**	**corresp.**	**1**	**0,025**	**yes**
014	**DP100 DP098 DP102**	**0,012**	**corresp.**	**1**	**0,035**	**yes**
014	**DP102**	**0,022**	**corresp.**	**3**	**0,497**	**no**
014	**DP102 DP098**	**0,031**	**corresp.**	**2**	**0,110**	**yes**
014	**DP100**	**0,031**	**corresp.**	**3**	**0,220**	**no**
014	**DP100 DP098**	**0,034**	**corresp.**	**2**	**0,143**	**yes**
014	**DP100 DP102**	**0,037**	**corresp.**	**2**	**0,178**	**yes**
014	**DP096 DP098**	**0,039**	**corresp.**	**2**	**0,155**	**yes**
014	**DP096 DP100 DP102**	**0,071**	**corresp.**	**1**	**0,113**	**yes**
014	**DP096 DP100**	**0,071**	**corresp.**	**2**	**0,205**	**yes**
014	**DP096 DP102**	**0,092**	**corresp.**	**2**	**0,340**	**no**
014	**DP098**	**0,117**	**corresp.**	**3**	**0,133**	**yes**
014	**DP096**	**0,137**	**corresp.**	**3**	**0,460**	**no**

Table 22: 90th percentile biases calculated by k- combinations and ecological classification correspondence for water body 014;

Interclass correlation coefficient

Results obtained by the ICC have been interpreted using a *qualitative scale* developed by *Landis and Koch*, typical for these types of studies. According to this scale, correlation between items is considered very good (0.90 to 1), good (0.80 to 0.90), moderate (0.60 to 0.80) or poor (0 to 0.60). k-observers that have an ICC > 0,950 are marked as Very Good*.

	Intraclass correlation[a]	Confidence interval95%		
		Lower limit	Upper limit	Value
DP096 DP098 DP102	**,960[c]**	**,920**	**,980**	**Very Good***
DP096 DP098 DP100	**,955c**	**,910**	**,978**	**Very Good***
DP100 DP098 DP102	**,968c**	**,937**	**,984**	**Very Good***
DP102	,759c	,490	,886	Moderate
DP102 DP098	,905c	,809	,953	Very Good
DP100	,750c	,491	,877	Moderate
DP100 DP098	,873c	,744	,936	Good
DP100 DP102	,926c	,853	,963	Very Good
DP096 DP098	,935c	,869	,968	Very Good
DP096 DP100 DP102	**,973c**	**,946**	**,986**	**Very Good***
DP096 DP100	,910c	,820	,955	Very Good
DP096 DP102	,890c	,778	,945	Good
DP098	,808c	,609	,905	Good
DP096	,870c	,732	,936	Good

Table 23: ICCs of the preselected k-observers, confidence intervals and interpretation of ICCs according to the Landis and Koch qualitative scale.

The 3-observer that excludes DP098 has the highest ICC (0.973). The 2-observer with the highest ICC is DP096,DP098 (0.935). The 1-observer that has the highest ICC coefficient is DP096 (0.870).

Six of the 14 analyzed k-observers have an ICC lower than 0.9; all of them are 1- and 2-observers.

Euclidean distance and MDS

Calculating and comparing the Euclidean distance between chlorophyll *a* average series of analyzed k-observers, the output confirms the results of the previous analysis. Results are presented using the Multidimensional Scaling technique. In the two dimensional space in which the distances are presented, it is easy to identify which observers verify that the Euclidean distance between the monthly averages of chlorophyll *a* concentration in their stations and the monthly averages of chlorophyll *a* concentration in the sampling stations of the reference observer (WB014) is less relevant.

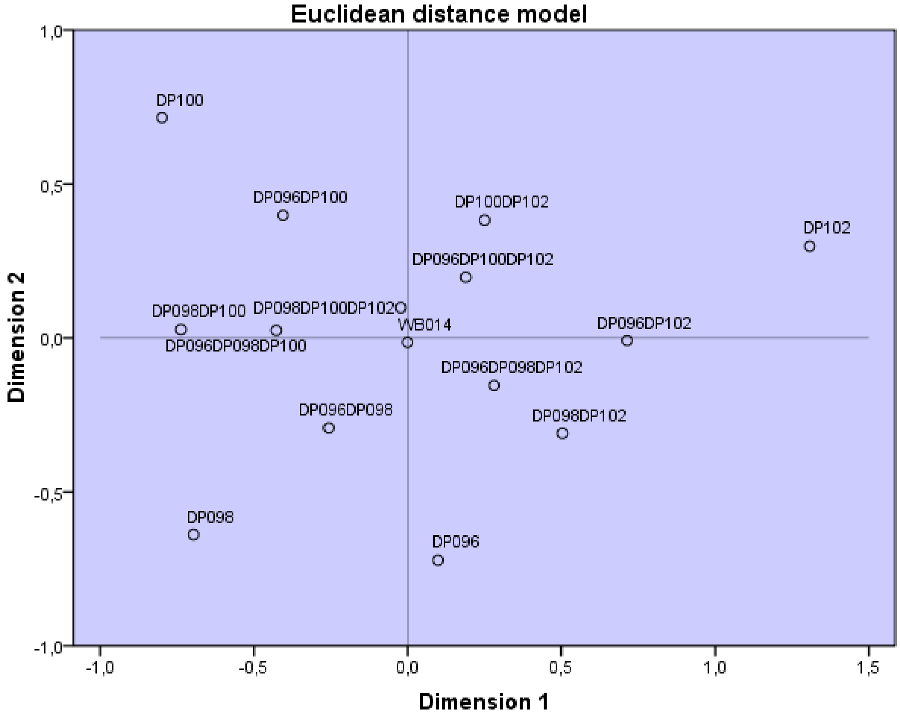


Figure 55: MDS for the Euclidian distance model for the pre-selected k-observers, w.b.014

The MDS diagram shows that all 3-observers possess the lowest Euclidean distance. Of all the 3-observers, DP098,DP100,DP102 has the lowest distance to WB014 . Euclidean distance increases as k is reduced.

<u>*Optimal k-observer of sampling stations for water body 014*</u>

The previous analyses confirmed that spatial reduction is possible for water body 014.

Finally, 4 k-observers are selected as most appropriate for the spatial reduction:

Sampling stations	N. of reduced s.s.	90^{th} Percentile
DP096 DP098 DP100 DP102	**0**	**1,791**
DP096 DP100 DP102	1	1,720
DP098 DP100 DP102	1	1,803
DP096 DP098 DP102	1	1,798
DP096 DP098 DP100	1	1,801

Table 24: Selected k-observers and the 90th percentiles obtained by their data subsets

Finally, the quantile plot (Figure 56) shows that the empirical CDF of the data subset taken by the selected k-observers is similar to the empirical CDF of the "original" dataset taken by the reference observer.

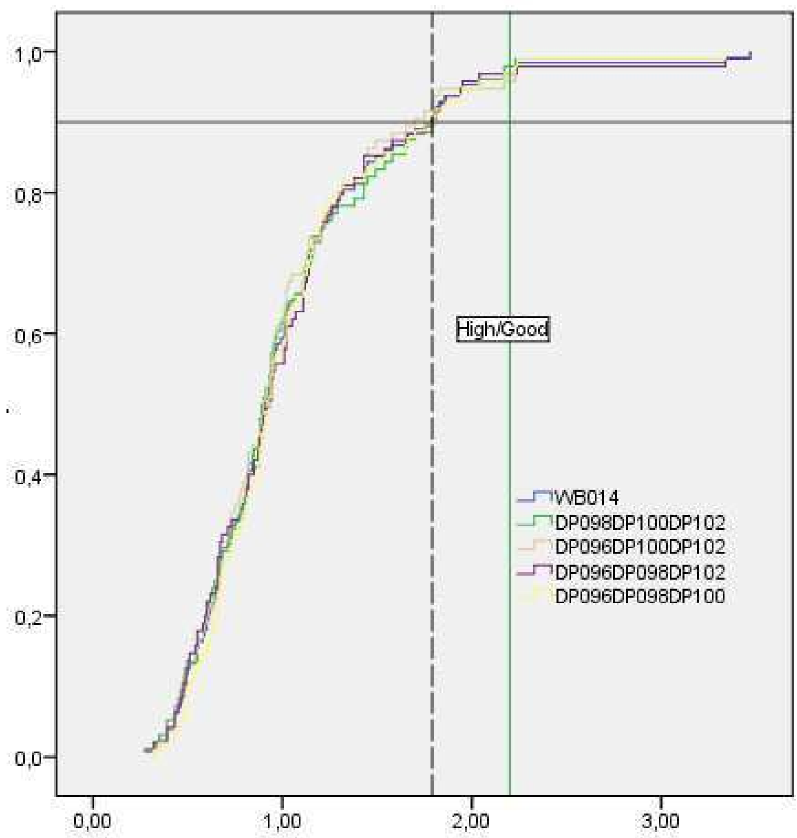


Figure 56: Quantile plot for the selected k-observers of chlorophyll a concentration for the three year period of August 2005 - July 2008, water body 014

Analysis for Water Body 019

Water Body 019 is surveyed by four different sampling stations (Annex 2). It is classified as a type III water body (without continental influence). The previous exploratory study determined that spatial reduction is possible without consequences to ecological classification. The following analysis will attempt to identify the most adequate k-observers for the spatial reduction.

Comparative study

Multiple box-and-whisker plots have been used to compare the range, median, quartiles and outliers of the measurements of chlorophyll *a* obtained by the four sampling stations located in water body 019. Graphs have been made for the three annual periods and the three-year period.

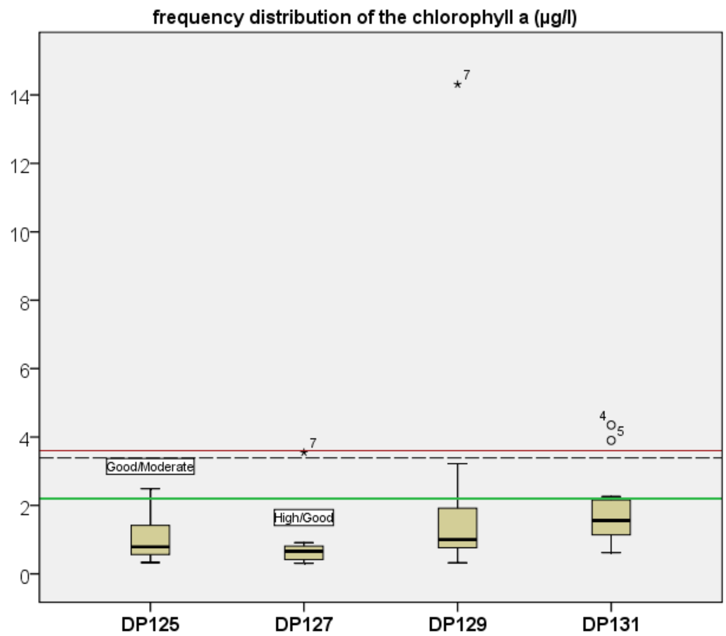


Figure 57: Box-and-whisker plot of chlorophyll a concentration for the annual period of August 2005 - July 2006 in the water body 019; the Green line represents the High/Good threshold, and the broken line represents the 90th percentile

During the annual period of August 2005 - July 2006, it is clear that the highest concentrations of chlorophyll *a* in water body 014 were taken mainly by stations DP129 and DP131 (Figure 57).

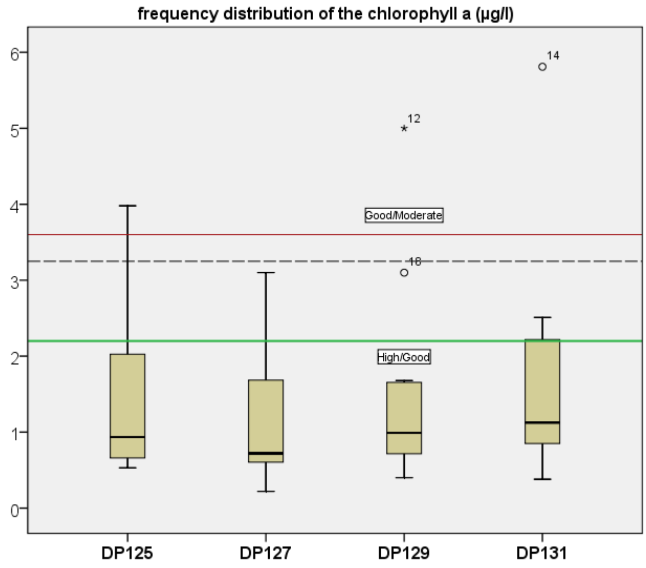


Figure 58: Box-and-whisker plot of chlorophyll a concentration for the annual period of August 2006 - July 2007 in the water body 019; the Green line represents the High/Good threshold, and the broken line represents the 90th percentile

For the second annual period, the analyzed samples mainly have chlorophyll *a* concentrations that do not exceed the High/Good threshold, however there is a significant number of samples for all stations that exceed this threshold.

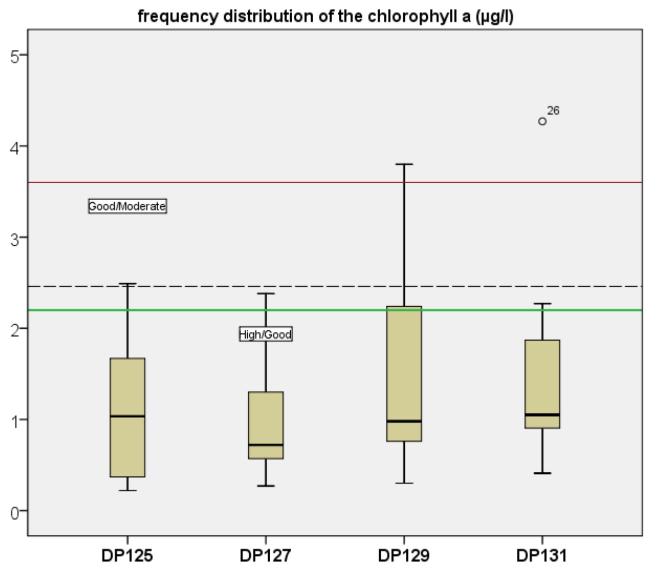


Figure 59: Box-and-whisker plot of chlorophyll a concentration for the annual period of August 2007 - July 2008 in the water body 019; the red line represents the Good/Moderate threshold, the Green line represents the High/Good threshold, and the broken line represents the 90th percentile

For the third annual period, the samples analyzed have chlorophyll *a* concentrations that mostly do not exceed the High/Good threshold. Only station DP129 analyzed a significant number of samples that exceed the High/Good and Good/Moderate thresholds.

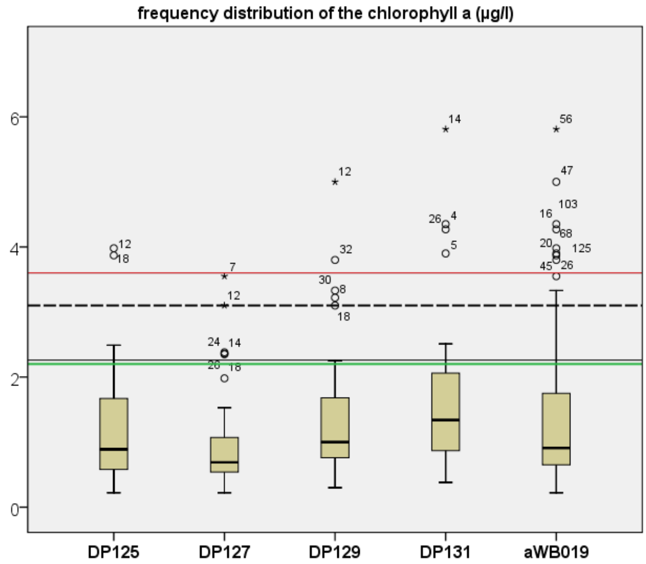


Figure 60: Box-and-whisker plot of chlorophyll a concentration for the three year period of August 2005 - July 2008 in water body 019; The green line represents the High/Good threshold, the red line represents the Good/Moderate threshold, and the broken line represents the 90th percentile

The box-and-whisker plot for the three year period clearly shows that the highest concentrations of chlorophyll *a* for water body 019 have been taken mostly by sampling stations DP125, DP129 and DP131 (Figure 60). For DP127, most of the analyzed samples do not exceed the High/Good threshold, and it can be concluded that this station represents the lowest chlorophyll *a* concentrations for water body 019.

The optimal k-observer of sampling stations should have an empirical CDF similar to the empirical CDF of aWB019, which is included on the last box-and-whisker plot.

The overlapping graphs corresponding to the empirical CDF of each sampling station allow us to compare the 90th percentiles of these distributions (Figure 61). For water body 019, it is not possible to establish two groups of sampling stations. Observing the quantile diagram, it seems that the CDFs divide the stations into two groups due to the dispersion of the empirical CDFs in the area of the 90th percentile.

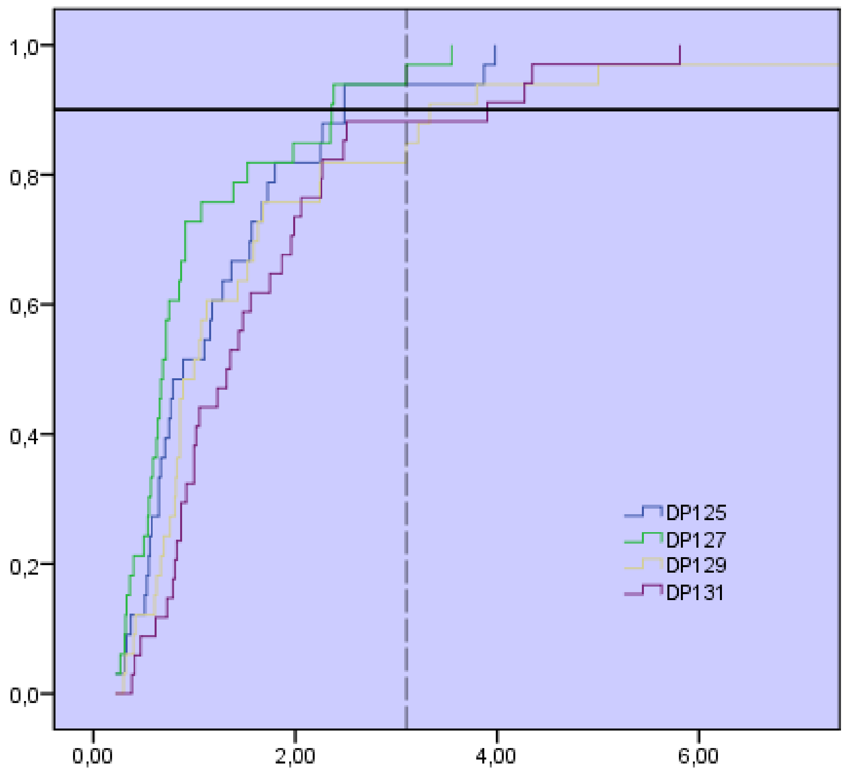


Figure 61: Quantile plot of chlorophyll a concentration for all sampling stations for the three year period August 2005 - July 2008, water body 019

<u>Pre-selection of k-observers of the sampling stations based on the deviation of the 90th percentile</u>

Observing Table 25 (taken from the Exploratory study and aligned by the three year bias), the results confirm what was easily observable on the previous quantile plot. A significant number of k-observers have a low 90th percentile bias, unlike water body 014.

DP125,DP127,DP129 has the lowest bias for all k-observers (0.000), DP129,DP127 for the 2-observers (0.036), and DP129 for the 1-observers (0.512).

The low 90th percentile bias allows us to suppose that the possibility of spatial reduction exists, and the first 10 k-observers with the lowest bias have been pre-selected for a more in-depth analysis.

w.b.	k-observer	bias 90 PCTL (3 years)	eco. class. (3 years)	- n stations	average absolute error (4 periods)	Corresponding eco. class. (4 periods)
019	**DP125 DP127 DP129**	**0,000**	**corresp.**	**1**	**0,155**	**yes**
019	**DP129 DP127**	**0,036**	**corresp.**	**2**	**0,235**	**yes**
019	**DP129 DP127 DP131**	**0,108**	**corresp.**	**1**	**0,308**	**no**
019	**DP125 DP129**	**0,153**	**corresp.**	**2**	**0,498**	**no**
019	**DP125 DP129 DP131**	**0,219**	**corresp.**	**1**	**0,455**	**no**
019	**DP125 DP131**	**0,318**	**corresp.**	**2**	**0,273**	**no**
019	**DP131 DP127**	**0,472**	**corresp.**	**2**	**0,338**	**no**
019	**DP129**	**0,512**	**degrade**	**3**	**2,925**	**no**
019	**DP125 DP127 DP131**	**0,610**	**corresp.**	**1**	**0,275**	**yes**
019	**DP125**	**0,610**	**corresp.**	**3**	**0,590**	**no**
019	**DP125 DP127**	**0,687**	**corresp.**	**2**	**0,550**	**yes**
019	**DP129 DP131**	**0,720**	**degrade**	**2**	**0,900**	**no**
019	**DP127**	**0,728**	**corresp.**	**3**	**0,368**	**yes**
019	**DP131**	**0,985**	**degrade**	**3**	**1,213**	**no**

Table 25: 90th percentile biases calculated by k-observers and ecological classification correspondence for water body 019;

Interclass correlation coefficient

Results obtained by the ICC have been interpreted using a *qualitative scale* developed by *Landis and Koch*, typical for these types of studies. According to this scale, correlation between items is considered very good (0.90 to 1), good (0.80 to 0.90), moderate (0.60 to 0.80) or poor (0 to 0.60). k-observers that have an ICC > 0,950 are marked as Very Good*.

	Intraclass correlation[a]	Confidence interval 95%		
		Lower limit	Upper limit	Value
DP125 DP127 DP129	**,974[c]**	**,948**	**,987**	**Very Good***
DP129 DP127	,931c	,861	,965	Very Good
DP129 DP127 DP131	**,989c**	**,978**	**,995**	**Very Good***
DP125 DP129	,938c	,877	,969	Very Good
DP125 DP129 DP131	**,992c**	**,937**	**,997**	**Very Good***
DP125 DP131	,869c	,736	,935	Good
DP131 DP127	,887c	,774	,943	Good
DP129	,776c	,552	,889	Moderate
DP125 DP127 DP131	,914c	,828	,957	Very Good
DP125	,858c	,713	,930	Good

Table 26: ICCs of the preselected k-observers, confidence intervals and interpretation of ICCs according to the Landis and Koch qualitative scale.

The 3-observer that excludes DP127 has the highest ICC (0.992). The 2-observer with the highest ICC is DP129,DP127 (0.931). The 1-observer with the highest ICC coefficient is DP125 (0.858).

Six of the 10 k-observers analyzed have an ICC higher than 0.9; all 3-observers and two 2-observers.

Euclidean distance and MDS

Calculating and comparing the Euclidean distance between chlorophyll *a* average series of analyzed k-observers, the output confirms the results of the previous analysis. Results are presented using the Multidimensional Scaling technique. In the two dimensional space in which the distances are presented, it is easy to identify which observers verify that the Euclidean distance between the monthly averages of chlorophyll *a* concentration in their stations and the monthly averages of chlorophyll *a* concentration in the sampling stations of the reference observer (WB019) is less relevant.

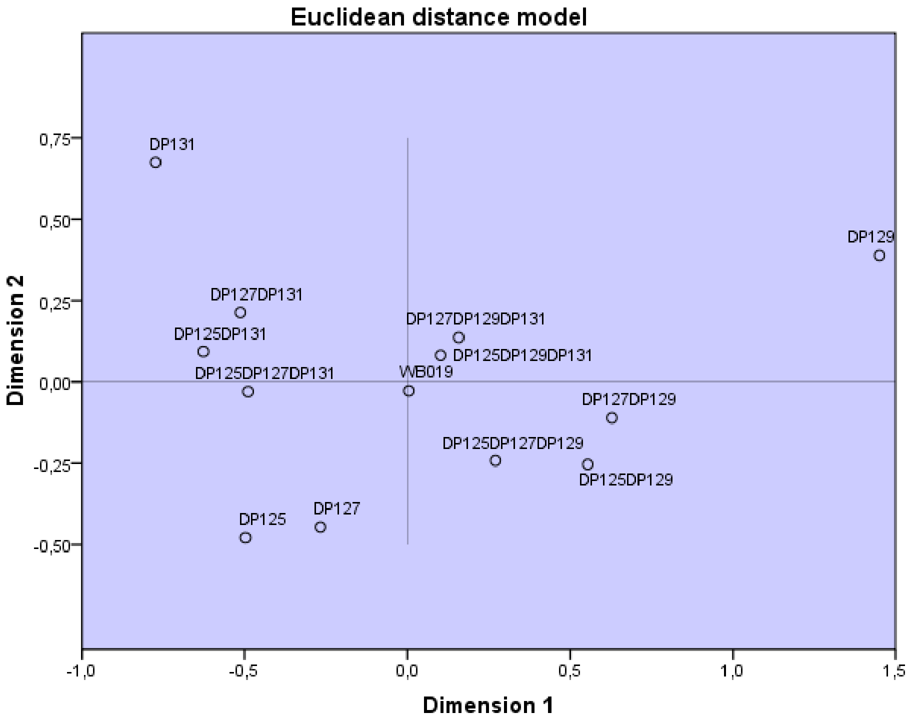


Figure 62: MDS for the Euclidian distance model for the pre-selected k-observers, w.b.019

The MDS diagram shows that all 3-observers possess the lowest Euclidean distance. Of all the 3-observers, DP125,DP129,DP131 has the lowest distance to WB019. Euclidean distance increases as k is reduced.

Optimal k-observer of sampling stations for water body 019

The previous analyses confirmed that spatial reduction is possible for water body 019, and there a few alternatives.

Finally, 3 k-observers are selected as most appropriate for the spatial reduction:

Sampling stations	N. of reduced s.s.	90th Percentile
DP125 DP127 DP129 DP131	**0**	**3,1000**
DP125 DP129 DP131	1	3,3190
DP125 DP127 DP129	1	3,1000
DP129 DP127 DP131	1	3,2080

Table 27: Selected k-observers and the 90th percentiles obtained by their data subsets

Finally, the quantile plot (Figure 63) shows that the empirical CDF of the data subset taken by the selected k-observers is similar to the empirical CDF of the "original" dataset taken by the reference observer.

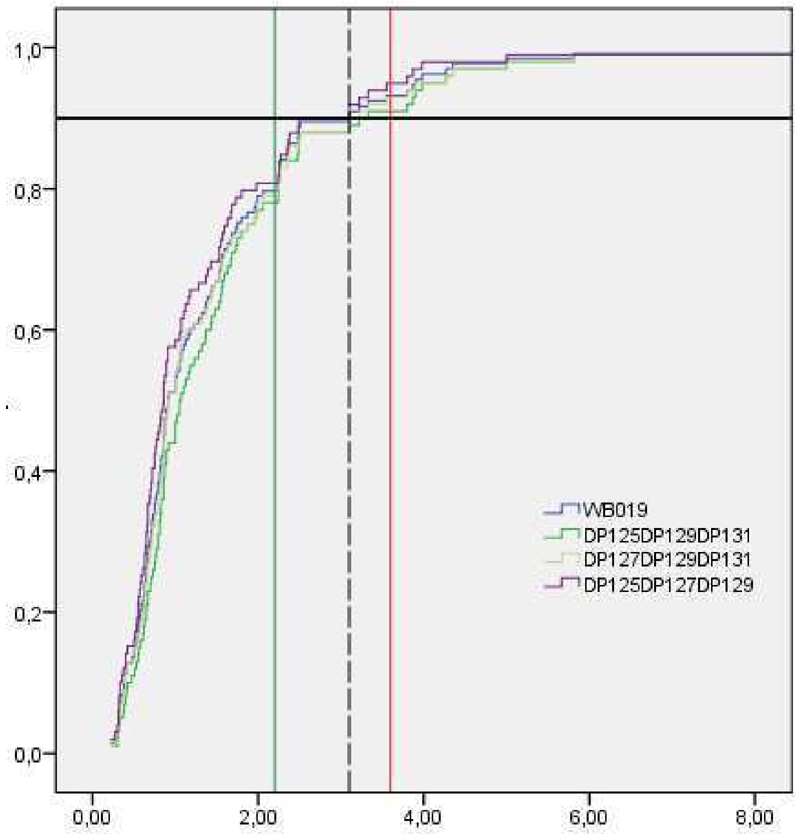


Figure 63: Quantile plot for the selected k-observers of chlorophyll a concentration for the three year period of August 2005 - July 2008, water body 019

5.2.3.Discussion

- The selection of the sampling stations to be reviewed may not be based on the comparison between the 90th percentiles of the empirical cumulative distribution functions of the samples. We need a methodology expandable to other environments.

- Observing the results of the screening process, it was concluded that, in order to get the minimal difference between the 90th percentile of the data analyzed by the reference observer and the 90th percentile of the data analyzed by an observer, it is desirable that the empirical cumulative distribution function *(empirical CDF)* of chlorophyll *a* concentration in the full sample does not differ substantially from the empirical CDF of the subsample, especially in the higher percentiles. Additionally, it would also be advantageous if the subsample includesthe greatest values of the original sample.

- If, month to month, the averages of measurements made in the sampling stations that characterize an observer are similar to the averages of measurements collected at the sampling stations that characterize the reference observer, this will be an indicator of proximity between the empirical CDF of measurements collected in the group of stations that characterize the observer and the empirical CDF of the original dataset. Moreover, the sensitivity of the average to outliers will allow us to know if the sampling stations that characterize an observer contain the extreme values of the original dataset.

- The Interclass correlation coefficient (ICC) and the Euclidean distance are adequate measures to know the concordance between the monthly averages of two different observers. The MDS will be useful in visualizing the proximity between observers.

- This methodology can be applied to spatial reduction for other environmental monitoring networks and it is not limited by the indicator (chlorophyll *a*), geographically or by type of water body. It is recommended to apply the same technique to other environmental indicators such as phytoplankton recounts, nutrients, priority substances etc. The technique can also be applied to other types of water bodies, such as estuaries, transitional, lakes, rivers and groundwater that are monitored in the scope of the WFD.

- A detailed revision of the sampling station of the better observers, using box-whisker and quantile plots, proves that in most water bodies, sampling stations can be divided into two or three groups.

    1. Sampling stations with mainly low chlorophyll *a* concentrations (for a particular water body) belong to the first group. These sampling stations indirectly affect the water body's 90th percentile value (by the number of samples that are included in the quantile calculation and by the water body minimal values). Empirical CDFs for these sampling stations are narrower than for the entire water body.

2. Sampling stations with high chlorophyll *a* concentrations belong to the second group. They have a significant influence on the water body's 90th percentile value (by the number and value of the samples with elevated chlorophyll *a* concentrations). Empirical CDFs for these sampling stations are wider than for the entire water body.

3. Sampling stations with intermediate chlorophyll *a* contractions compared to the first and second groups of sampling stations belong to the third group. This type of sampling station is detected only in few water bodies.

- Good observers include a balanced number of sampling stations from both groups.

- The number of sampling stations reduced for each water body is directly related to the indistinctness of ecological classification and the stability of the system

5.3.Virtual simulation of chlorophyll *a* concentration and post reduction management

5.3.1.Introduction

As seen in the section "Selecting the most appropriate k-observers”, we have recommended several potential observers that could be used to analyze the ecological state of different water bodies. For these observers, the ICC is above 0.95, the agreement between the "reference observer" and every one of them is very good, and the Euclidean distance between the monthly averages of measurements taken by each of them and the monthly averages of measurements taken by the "reference observer' is small. This guarantees that any of them can provide a complete and unbiased view of the changes experienced by the chlorophyll in this system. The final selection of these alternatives should be made taking other factors such as effectiveness, cost, etc. into account.

One of the criteria that can help make the final decision is to evaluate the possibility of obtaining, virtual values of the measurements of chlorophyll a in the unrevised stations from measurements taken at the stations visited, assuming that changes significant to the cdf of the values of chlorophyll a in the unrevised stations do not occur.

We believe that, since there are factors that can significantly alter the ecological stability of the areas where the sampling stations are located, these possible changes should be controlled with occasional revisions. Therefore, the selection of the sampling stations to be reviewed should not be a permanent selection and should be able to be modified if circumstances indicate that it is advisable.

If, in the post-management of the spatial reduction, we take occasional measurements using the unrevised stations and repeatedly find significant discrepancies between the virtual and observed values, this will indicate to us that there have been significant anthropogenic changes in the specific area where this station is located. This may have important implications for the spatial reduction, perhaps because the “selected observer” is no longer the most suitable to provide a complete and stable view of the area, and we should consider performing a new selection.

5.3.2.Water body 001

In water body 001, five alternatives for the spatial reduction were selected.

Sampling stations	N. of reduced s.s.	90th Percentile
DP001 DP003 DP005 DP133 DP007	**0**	**3,54**
DP001 DP003 DP133 DP007	1	3,64
DP001 DP003 DP007	2	3,35
DP001 DP005 DP007	2	3,41
DP003 DP005 DP133	2	3,64
DP005 DP133 DP007	2	3,64
DP001 DP007	3	3,54

Table 1: **Selected k-observers and the 90th percentiles obtained by their data subsets**

To study the possibility of estimating virtual values for each of the sampling stations in w.b. 001 using the other sampling stations of this system, we calculated the Pearson correlation coefficients between sampling stations (Table 2) and obtained stepwise regression models to predict the values of each of the sampling stations in w.b. 001 from the other sampling stations. For each station, stepwise regression has made a selection of the best predictors, and coefficient R^2 indicates the percentage of the variability of the response variable explained by the model (Table 3).

In water body 001, there was missing data in 13 of the 34 revised campaigns. Therefore, allocations have been made using the k-nearest neighbours method in these reviews, obtaining the estimated values for missing data from the average of the values of the 3 closest neighbors. Using the Hot-Deck technique, the data from the 12 campaigns was recuperated for the use of the model.

From the information gathered in these tables, the k-observer was chosen from among the selected observers, since it would make the best predictions.

Correlations

		DP001	DP003	DP005_1	DP133_1	DP007_1
DP001	Pearson Correlation	1,000	,471**	,395*	,607**	,265
	Sig. (2-tailed)		,006	,023	,000	,130
	N	34	33	33	34	34
DP003	Pearson Correlation	,471**	1,000	,417*	,602**	,691**
	Sig. (2-tailed)	,006		,016	,000	,000
	N	33	33	33	33	33
DP005_1	Pearson Correlation	,395*	,417*	1,000	,138	,281
	Sig. (2-tailed)	,023	,016		,444	,113
	N	33	33	33	33	33
DP133_1	Pearson Correlation	,607**	,602**	,138	1,000	,445**
	Sig. (2-tailed)	,000	,000	,444		,008
	N	34	33	33	34	34
DP007_1	Pearson Correlation	,265	,691**	,281	,445**	1,000
	Sig. (2-tailed)	,130	,000	,113	,008	
	N	34	33	33	34	34

**. Correlation is significant at the 0.01 level (2-tailed).

*. Correlation is significant at the 0.05 level (2-tailed).

Table 2: Pearson correlation coefficents for the water body

regression model		
modeled sampling station	station(s) used in the model	R^2
DP001	DP133, DP005	0,48
DP003	DP007, DP001	0,548
DP005	DP003	0,147
DP007	DP003	0,46
DP133	DP001, DP007	0,47

Table 3: Predicting variables and the adjusted R squared

The optimal combination of sampling stations to be reviewed regularly consists of DP001, DP005 and DP007. Using the regression model, virtual values for DP003 and DP133 can be obtained.

Virtual values model for DP003 and DP133

In the model to estimate the values for DP003, DP001 and DP007are incorporated into the model as predictor variables. The model explains 55% of the variability of DP003. In the model to estimate the values for DP133, DP001 and DP007 are incorporated into the model as predictor variables. The model explains 47% of the variability of DP133.

The graph below shows the relationship between predicted and real values for sampling station DP003.

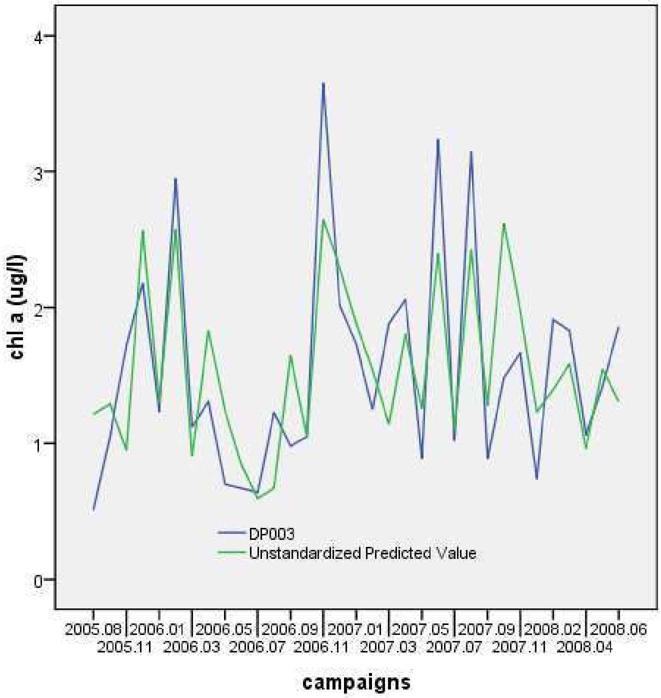


Figure 1:The chart above compares DP003's chlorophyll a series and the virtual series obtained using the model

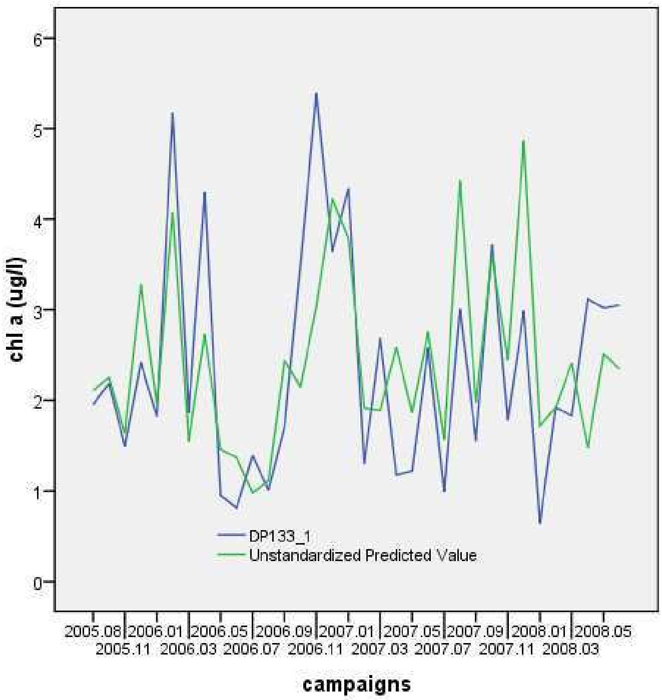


Figure 2: The chart above compares DP133 chlorophyll a series and the virtual series obtained using the model

Data used to obtain 90th percentile	90th percentile	BIAS (error)	Eco.class.
90th PCTL all sampling stations	3,542	-	High
90th PCTL excluding DP003 and DP133	3,408	0,134	High
90th PCTL virtual values of DP003 and DP133	3,309	0,233	High

Table 4: 90th percentiles and ecological classification calculated by using virtual data or excluding sampling stations

Table 4 shows the 90th percentiles and ecological classifications obtained using reference data, chosen observer data and the joining of the chosen observer data and the virtual values of the unrevised sampling stations. The 90th percentiles obtained present negligible biases in respect to the 90th percentile obtained using reference data. All 90th percentiles reflect the same quality value (ecological class). These results confirm the accuracy of the virtual values obtained in this system for the unrevised sampling stations.

Therefore, since it is possible to obtain virtual values quite accurately for unrevised stations belonging to wb001, it is possible in this system to control for potential stable alterations in the anthropogenic characteristics of specific areas where unrevised stations are located by annually reviewing these stations and the comparison between virtual and actual values.

As mentioned above, the continuous presence of significant discrepancies may have implications for the spatial reduction, because the actual observer is potentially no longer the most suitable to provide a complete and stable view of the area, and performing a revision should be considered.

5.3.3.Water body 002

In water body 002, five alternatives for spatial reduction were selected.

Sampling stations	N. of reduced s.s.	90^{th} Percentile
DP134 DP135 DP136 DP137 DP010	**0**	**3,498**
DP134 DP135 DP137 DP010	1	3,540
DP135 DP136 DP137 DP010	1	3,191
DP134 DP135 DP136 DP010	1	3,797
DP135 DP136 DP010	2	3,586
DP135 DP137 DP010	2	3,125

Table 5: Selected k-observers and the 90th percentiles obtained by their data subsets

To study the possibility of estimating virtual values for each of the sampling stations in w.b. 002 using the other sampling stations of this system, we calculated the Pearson correlation coefficients between sampling stations (Table 6) and obtained stepwise regression models to predict the values of each of the sampling stations in w.b. 002 from the other sampling stations. For each station, stepwise regression has made a selection of the best predictors, and coefficient R^2 indicates the percentage of the variability of the response variable explained by the model (Table 7).

In the water body 002, there was missing data in 8 of the 34 revised campaigns. Therefore, allocations have been made using the k-nearest neighbours method in these reviews, obtaining the estimated values for missing data, from the average of the values of the 3 donors closest neighbours. Using the Hot-Deck technique, the data from the three campaigns was recuperated for the use of the model.

From the information gathered in these tables, the k-observer was chosen from among the selected observers, since it would make the best predictions.

Correlations

		DP134	DP135	DP136_1	DP137_1	DP010
DP134	Pearson Correlation	1,000	,652**	,055	,189	,161
	Sig. (2-tailed)		,000	,776	,327	,403
	N	29	29	29	29	29
DP135	Pearson Correlation	,652**	1,000	,240	,319	,356
	Sig. (2-tailed)	,000		,210	,092	,058
	N	29	29	29	29	29
DP136_1	Pearson Correlation	,055	,240	1,000	,581**	,424*
	Sig. (2-tailed)	,776	,210		,001	,022
	N	29	29	29	29	29
DP137_1	Pearson Correlation	,189	,319	,581**	1,000	,488**
	Sig. (2-tailed)	,327	,092	,001		,007
	N	29	29	29	29	29
DP010	Pearson Correlation	,161	,356	,424*	,488**	1,000
	Sig. (2-tailed)	,403	,058	,022	,007	
	N	29	29	29	29	34

**. Correlation is significant at the 0.01 level (2-tailed).

*. Correlation is significant at the 0.05 level (2-tailed).

Table 6: Pearson correlation coefficents for the water body

regression model:		
modled sampling satation	station(s) used in the model	R^2
DP134	DP135	0,40
DP135	DP134	0,40
DP136	DP137	0,31
DP010	DP137	0,21
DP137	DP136	0,31

Table 7: Predicting variables and the adjusted R squared

The optimal combination of sampling stations to be reviewed regularly consists of DP135, DP136 and DP010. Using the regression model, virtual values for DP134 and DP137 can be obtained.

Virtual values model for DP134 and DP137

In the model to estimate the values of DP134, DP135 is incorporated into the model as a predictor variable. The model explains 40% of the variability of DP134. In the model to estimate the values of DP137, DP136 is incorporated into the model as predictor variable. The model explains 31% of the variability of DP137.

The graph below shows the relationship between the predicted and real values for the sampling stations DP134 and DP137.

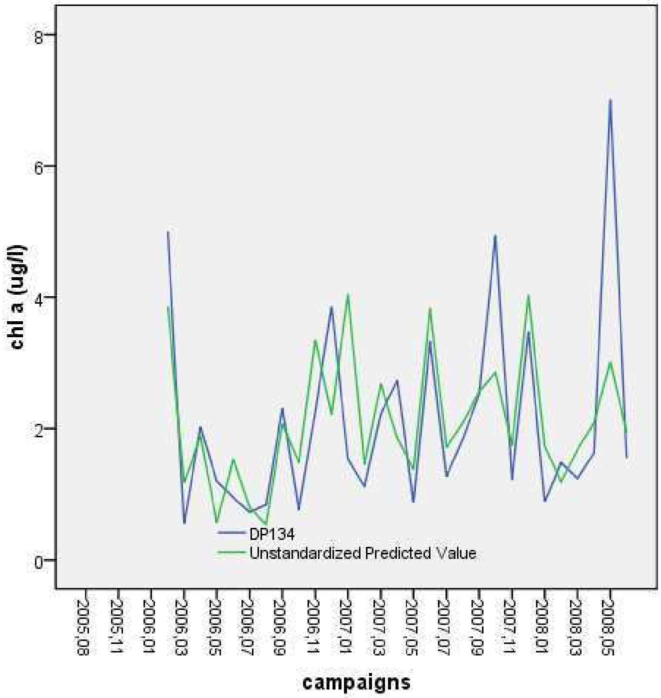


Figure 3: The chart above compares DP134's chlorophyll a series and the virtual series obtained by the model

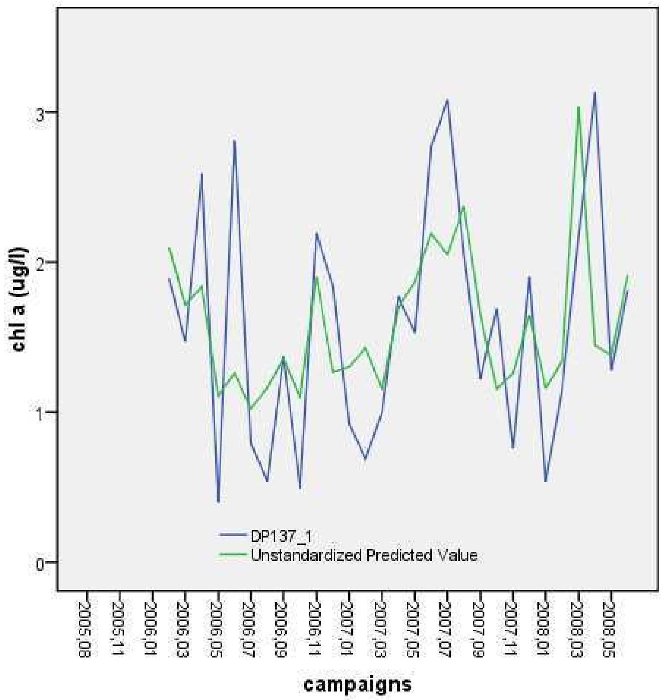


Figure 4: The chart above compares DP137's chlorophyll a series and the virtual series obtained using the model

Data used to obtain 90th percentile	90th percentile	BIAS (error)	Eco.class.
90th PCTL all sampling stations	3,498	-	High
90th PCTL excluding DP134 and DP137	3,586	0,088	High
90th PCTL virtual values of DP134 and DP137	3,296	0,202	High

Table 8: 90th percentiles and ecological classification calculated by using virtual data or excluding the sampling stations

Table 8 shows the 90th percentiles and the ecological classifications obtained using the reference data, chosen observer data and the joining of the chosen observer data and the virtual values of the unrevised sampling stations. The 90th percentiles obtained present negligible biases in respect to the 90th percentile obtained using reference data. All 90th percentiles reflect the same quality value (ecological class). These results confirm the accuracy of the virtual values obtained in this system for the unrevised sampling stations.

Therefore, since it is possible to obtain virtual values quite accurately for unrevised stations belonging to w.b.002, it is possible in this system to control for potential stable alterations in the anthropogenic characteristics of specific areas where unrevised stations are located by annually reviewing these stations and the comparison between virtual and actual values.

As mentioned above, the continuous presence of significant discrepancies may have implications for the spatial reduction, because the actual observer is potentially no longer the most suitable to provide a complete and stable view of the area, and performing a revision should be considered.

5.3.4. Water body 005

In water body 005, eight alternatives for spatial reduction were selected.

Sampling stations	N. of reduced s.s.	90^{th} Percentile
DP032 DP034 DP036 DP038 DP040	**0**	**5,693**
DP032 DP034 DP036 DP038	1	5,675
DP032 DP036 DP038 DP040	1	5,950
DP032 DP036 DP040	2	5,684
DP032 DP038 DP040	2	5,714
DP034 DP036 DP040	2	5,7080
DP034 DP038 DP040	2	5,8640
DP032 DP036	3	5,666
DP034 DP036	3	5,7380

Figure 5: Selected k-observers and the 90^{th} percentiles obtained by their data subsets

To study the possibility of estimating virtual values for each of the sampling stations in w.b. 005 using the other sampling stations of this system, we calculated the Pearson correlation coefficients between sampling stations (Table 9) and obtained stepwise regression models to predict the values of each of the sampling stations in w.b. 005 from the other sampling stations. For each station, stepwise regression has made a selection of the best predictors, and coefficient R^2 indicates the percentage of the variability of the response variable explained by the model (Table 10).

From the information gathered in these tables, the k-observer was chosen from among the selected observers, since it would make the best predictions.

Correlations

		DP032	DP034	DP036	DP038	DP040
DP032	Pearson Correlation	1,000	,537**	,450**	,565**	,154
	Sig. (2-tailed)		,001	,009	,001	,386
	N	34	34	33	33	34
DP034	Pearson Correlation	,537**	1,000	,635**	,639**	,333
	Sig. (2-tailed)	,001		,000	,000	,055
	N	34	34	33	33	34
DP036	Pearson Correlation	,450**	,635**	1,000	,693**	,558**
	Sig. (2-tailed)	,009	,000		,000	,001
	N	33	33	33	33	33
DP038	Pearson Correlation	,565**	,639**	,693**	1,000	,280
	Sig. (2-tailed)	,001	,000	,000		,114
	N	33	33	33	33	33
DP040	Pearson Correlation	,154	,333	,558**	,280	1,000
	Sig. (2-tailed)	,386	,055	,001	,114	
	N	34	34	33	33	34

**. Correlation is significant at the 0.01 level (2-tailed).

Table 9: Pearson correlation coefficents for the water body

regression model:		
modled sampling satation	station(s) used in the model	R^2
DP036	DP038, DP040	0,60
DP038	DP036, DP032	0,53
DP032	DP038	0,30
DP034	DP038	0,39
DP040	DP036	0,29

Table 10: Predicting variables and the adjusted R squared

The optimal combination of sampling stations to be reviewed regularly consists of DP032, DP038 and DP040. Using the regression model, virtual values for DP036 and DP034 can be obtained.

Virtual values model for DP036 and DP034

In the model to estimate the values of DP036, DP038 and DP040 are incorporated into the model as predictor variables. The model explains 60% of the variability of DP036. In the model to estimate the values of DP034, DP038 is incorporated into the model as a predictor variable. The model explains 39% of the variability of DP034.

The graph below shows the relationship between the predicted and real values for sampling stations DP036 and DP034.

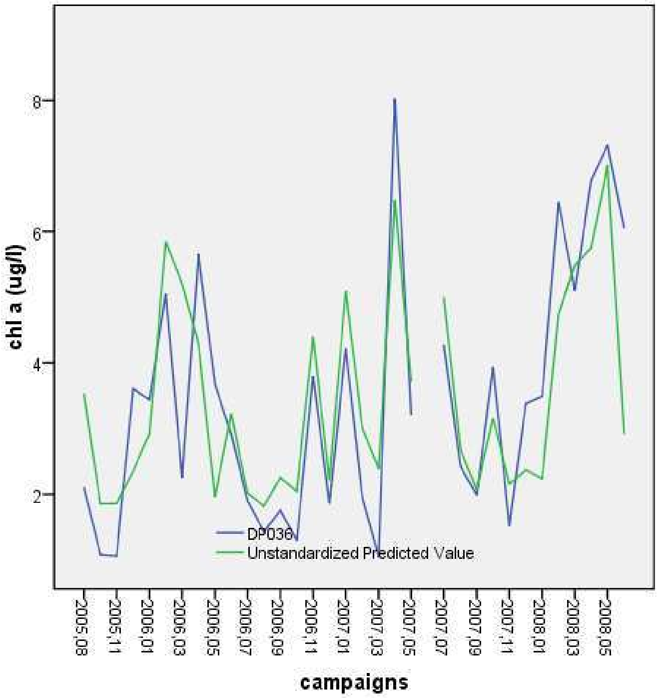


Figure 6: The chart above compares DP036's chlorophyll a series and the virtual series obtained by the model

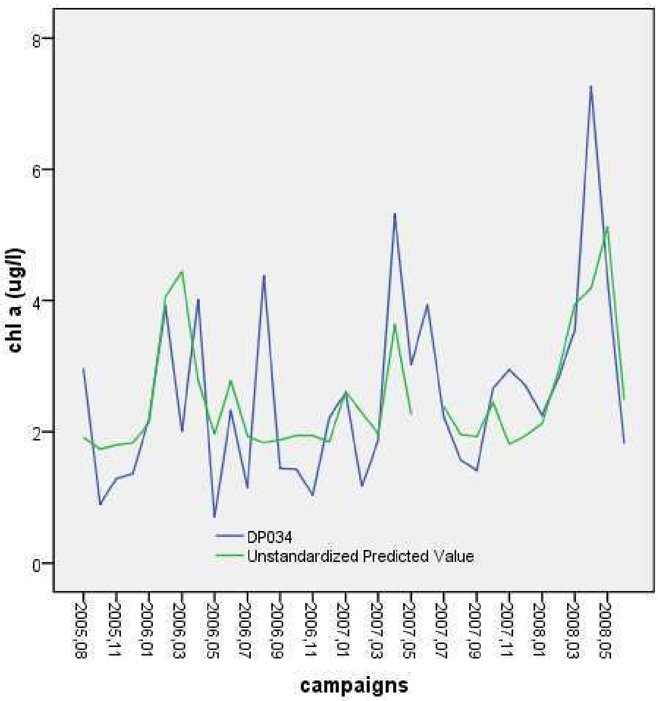


Figure 7: The chart above compares DP034's chlorophyll a series and the virtual series obtained using the model

Data used to obtain 90th percentile	90th percentile	BIAS (error)	Eco.class.
90th PCTL all sampling stations	5,693	-	Good
90th PCTL excluding DP034 and DP036	5,714	0,021	Good
90th PCTL virtual values of DP034 and DP036	5,263	0,430	Good

Table 11: 90th percentiles and ecological classification calculated by using virtual data or excluding the sampling stations

Table 11 shows the 90th percentiles and the ecological classifications obtained using the reference data, chosen observer data and the joining of the chosen observer data and the virtual values of the unrevised sampling stations. The 90th percentiles obtained present negligible biases in respect to the 90th percentile obtained using reference data. All 90th percentiles reflect the same quality value (ecological class). These results confirm the accuracy of the virtual values obtained in this system for the unrevised sampling stations.

Therefore, since it is possible to obtain virtual values quite accurately for unrevised stations belonging to w.b.005, it is possible in this system to control for potential stable alterations in the anthropogenic characteristics of specific areas where unrevised stations are located by annually reviewing these stations and the comparison between virtual and actual values.

As mentioned above, the continuous presence of significant discrepancies may have implications for the spatial reduction, because the actual observer is potentially no longer the most suitable to provide a complete and stable view of the area, and performing a revision should be considered.

5.3.5.Water body 008

In water body 008, six alternatives for spatial reduction were selected.

Sampling stations	N. of reduced s.s.	90^{th} Percentile
DPU001 DP051 DP052 DP054 DP055	**0**	**7,960**
DPU001DP051DP054DP055	1	8,088
DP051 DP052 DP054 DP055	1	8,120
DP051 DP054 DP055	2	8,248
DPU001DP051DP052DP055	1	8,360
DP051 DP052 DP055	2	8,460
DPU001DP051DP052DP054	1	8,610

Table 12: Selected k-observers and the 90th percentiles obtained by their data subsets

To study the possibility of estimating virtual values for each of the sampling stations in w.b. 008 using the other sampling stations of this system, we calculated the Pearson correlation coefficients between sampling stations (Table 13) and obtained stepwise regression models to predict the values of each of the sampling stations in w.b. 008 from the other sampling stations. For each station, stepwise regression has made a selection of the best predictors, and coefficient R^2 indicates the percentage of the variability of the response variable explained by the model (Table 14).

From the information gathered in these tables, the k-observer was chosen from among the selected observers, since it would make the best predictions.

Correlations

		DPU001	DP051_1	DP052	DP054	DP055
DPU001	Pearson Correlation	1,000	,226	,343	,236	,347
	Sig. (2-tailed)		,419	,211	,397	,205
	N	15	15	15	15	15
DP051_1	Pearson Correlation	,226	1,000	,320	,195	,188
	Sig. (2-tailed)	,419		,065	,270	,286
	N	15	34	34	34	34
DP052	Pearson Correlation	,343	,320	1,000	,746**	,766**
	Sig. (2-tailed)	,211	,065		,000	,000
	N	15	34	34	34	34
DP054	Pearson Correlation	,236	,195	,746**	1,000	,695**
	Sig. (2-tailed)	,397	,270	,000		,000
	N	15	34	34	34	34
DP055	Pearson Correlation	,347	,188	,766**	,695**	1,000
	Sig. (2-tailed)	,205	,286	,000	,000	
	N	15	34	34	34	34

**. Correlation is significant at the 0.01 level (2-tailed).

Table 13: Pearson correlation coefficents for the water body

regression model:		
modled sampling satation	station(s) used in the model	R^2
DP051	DP052	0,82
DP052	DP051, DP054	0,92
DP054	DP052	0,55
DP055	DP052	0,56
DPU001	.	.

Table 14: Predicting variables and the adjusted R squared

The optimal combination of sampling stations to be reviewed regularly consists of DPU001, DP051, DP054 and DP055. Using the regression model, virtual values for DP052 can be obtained.

Virtual values model for DP052

In the model to estimate the values of DP052, DP051 and DP054 are incorporated into the model as predictor variables. The model explains 92% of the variability of DP052.

The graph below shows the relationship between the predicted and real values for sampling station DP052.

Table 15 shows that the obtained virtual values for sampling station DP052 are consistent with the real ones. For two campaigns, the virtual concentrations have an important bias compared to the real values. As sampling station DP051 (crucial component in the model for DP052 - explains a significant part of the variability) is significantly influenced by the Albufera (hypereutrophic lagoon), the two extended biases probably result from directly influenced samples. Two samples analyzed at DP051 have chlorophyll *a* concentrations higher than 20 and 60 µg/l, which indicates that these samples contain a significant percentage of Albufera water. These values do not indicate the real trophic status of this part of the water body, but rather indicate the mixing process of coastal and Albufera waters. The fact that this mixing process is not detected by station DP052 (but is detected significantly on the predicted model value) is most likely caused by the change of the normal sea current conditions (direction North-South). Under such conditions, the Albufera influence decreased towards the South and increased towards the North. DPU001 was not active at this moment, but had it been active, it would likely have detected elevated chlorophyll a concentrations.

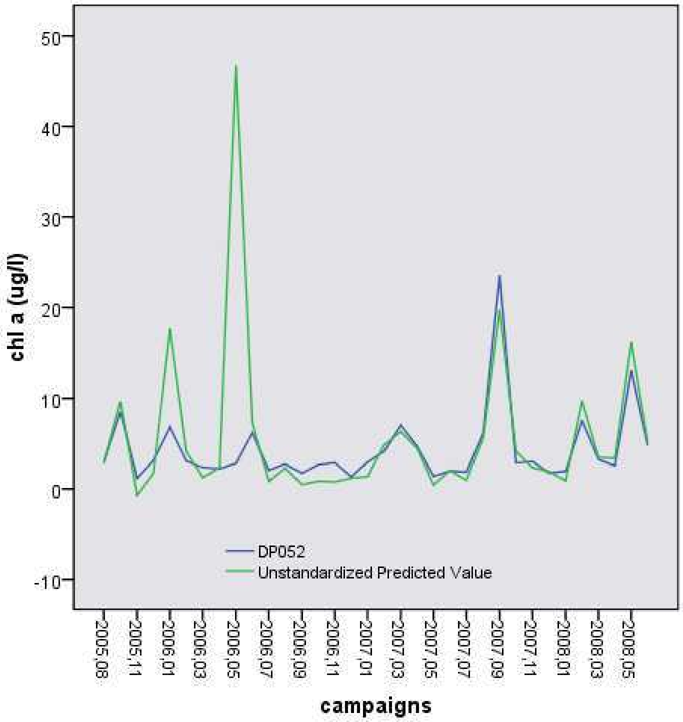


Figure 8: The chart above compares DP052's chlorophyll a series and the virtual series obtained using the model

Data used to obtain 90th percentile	90th percentile	BIAS (error)	Eco.class.
90th PCTL all sampling stations	7,960	-	Moderate
90th PCTL excluding DP052	8,088	0,148	Moderate
90th PCTL virtual values of DP034 and DP036	9,472	1,512	Moderate

Table 15: 90th percentiles and ecological classification calculated by using virtual data or excluding the sampling stations

Table 15 shows the 90th percentiles and the ecological classifications obtained using the reference data, selected k-observer and the superposition of the selected k-observer and the virtual values. All 90th percentiles reflect the same quality value (ecological class), but the superposition of the k-observer and the DP052 virtual value gives a 90th percentile with a significant error of 1.512 µg/l. This elevated 90th percentile error is caused by two overestimated virtual values (mentioned above) which are easy to identify as incorrect model values. Excluding these two values, the 90th percentile is closer to the reference value of 8.280 µg/l, with an acceptable error of 0.32 µg/l.

Therefore, since it is possible to obtain virtual values quite accurately for unrevised stations belonging to w.b.008, it is possible in this system to control for potential stable alterations in the anthropogenic characteristics of specific areas where unrevised stations are located by annually reviewing these stations and the comparison between virtual and actual values.

As mentioned above, the continuous presence of significant discrepancies may have implications for the spatial reduction, because the actual observer is potentially no longer the most suitable to provide a complete and stable view of the area, and performing a revision should be considered.

5.3.6.Water body 009

For water body 009, six alternatives were identified for spatial reduction. Analyzing the water body sampling stations, correlation coefficients were not obtained that indicate that a regression model can be functional for any of the alternatives. The only station for which virtual values can be obtained is DP061 which is included in most of the selected k-observers.

regression model:		
modled sampling satation	station(s) used in the model	R^2
DP056	DP061	0,33
DP057	DP061	0,17
DP059	DP061	0,35
DP061	DP059, DP056	0,55
DP062	DP061	0,27

Table 16: Predicting variables and the adjusted R squared

Correlations

		DP056	DP057	DP059_1	DP061	DP062_1
DP056	Pearson Correlation	1,000	,327	,245	,527**	,417*
	Sig. (2-tailed)		,059	,162	,001	,014
	N	34	34	34	34	34
DP057	Pearson Correlation	,327	1,000	,436**	,449**	,440**
	Sig. (2-tailed)	,059		,010	,008	,009
	N	34	34	34	34	34
DP059_1	Pearson Correlation	,245	,436**	1,000	,582**	,471**
	Sig. (2-tailed)	,162	,010		,000	,005
	N	34	34	34	34	34
DP061	Pearson Correlation	,527**	,449**	,582**	1,000	,548**
	Sig. (2-tailed)	,001	,008	,000		,001
	N	34	34	34	34	34
DP062_1	Pearson Correlation	,417*	,440**	,471**	,548**	1,000
	Sig. (2-tailed)	,014	,009	,005	,001	
	N	34	34	34	34	34

**. Correlation is significant at the 0.01 level (2-tailed).

*. Correlation is significant at the 0.05 level (2-tailed).

Table 17: Pearson correlation coefficents for the water body

5.3.7.Water body 010

In water body 010, ten alternatives for spatial reduction were selected.

Sampling stations	N. of reduced s.s.	90^{th} Percentile
DP066 DP068 DP070 DP072 DP073 DP074	**0**	**5,921**
DP066 DP068 DP072 DP073 DP074	1	6,001
DP066 DP068 DP070 DP072 DP073	1	5,980
DP066 DP068 DP070 DP073 DP074	1	5,995
DP068 DP070 DP073 DP074	2	5,744
DP068 DP070 DP072 DP074	2	5,921
DP068 DP072 DP073	3	5,975
DP068 DP073 DP074	3	5,980
DP068 DP070 DP074	3	5,995
DP068 DP070 DP073	3	5,973
DP068DP072DP074	3	6,001

Table 18: Selected k-observers and the 90th percentiles obtained by their data subsets

To study the possibility of estimating virtual values for each of the sampling stations in w.b. 010 using the other sampling stations of this system, we calculated the Pearson correlation coefficients between sampling stations (Table 19) and obtained stepwise regression models to predict the values of each of the sampling stations in w.b. 010 from the other sampling stations. For each station, stepwise regression has made a selection of the best predictors, and coefficient R^2 indicates the percentage of the variability of the response variable explained by the model (Table 20).

In water body 010, there was missing data in 13 of the 34 revised campaigns. Therefore, allocations have been made using the k-nearest neighbours method in these reviews, obtaining the estimated values for missing data from the average of the values of the 3 closest neighbours. Using the Hot-Deck technique, the data from the nine campaigns was recuperated.

From the information gathered in these tables, the k-observer was chosen from among the selected observers, since it would make the best predictions.

Correlations

		DP066_1	DP068_1	DP070	DP072_1	DP073_1	DP074_1
DP066_1	Pearson Correlation	1,000	,573**	,286	,335	,584**	,447*
	Sig. (2-tailed)		,000	,107	,065	,001	,012
	N	33	33	33	31	31	31
DP068_1	Pearson Correlation	,573**	1,000	,524**	,422*	,529**	,380*
	Sig. (2-tailed)	,000		,001	,016	,002	,035
	N	33	34	34	32	32	31
DP070	Pearson Correlation	,286	,524**	1,000	,531**	,248	,418*
	Sig. (2-tailed)	,107	,001		,002	,171	,019
	N	33	34	34	32	32	31
DP072_1	Pearson Correlation	,335	,422*	,531**	1,000	,502**	,370*
	Sig. (2-tailed)	,065	,016	,002		,004	,044
	N	31	32	32	32	31	30
DP073_1	Pearson Correlation	,584**	,529**	,248	,502**	1,000	,545**
	Sig. (2-tailed)	,001	,002	,171	,004		,002
	N	31	32	32	31	32	30
DP074_1	Pearson Correlation	,447*	,380*	,418*	,370*	,545**	1,000
	Sig. (2-tailed)	,012	,035	,019	,044	,002	
	N	31	31	31	30	30	31

**. Correlation is significant at the 0.01 level (2-tailed).

*. Correlation is significant at the 0.05 level (2-tailed).

Table 19: Pearson correlation coefficents for the water body

regression model:		
modled sampling satation	station(s) used in the model	R^2
DP066	DP068	0,50
DP068	DP066, DP073	0,57
DP070	DP072	0,30
DP072	DP073, DP070	0,40
DP073	DP068	0,37
DP074	DP066	0,28

Table 20: Predicting variables and the adjusted R square

The optimal combination of sampling stations to be reviewed regularly consists of DP068, DP070, DP073 and DP074. Using the regression model, virtual values for DP066 and DP072 can be obtained.

Virtual values model for DP066 and DP072

In the model to estimate the values of DP066, DP068 is incorporated into the model as a predictor variable. The model explains 50% of the variability of DP066. In the model to estimate the values of DP072, DP070 and DP073 are incorporated into the model as predictor variables. The model explains 40% of the variability of DP072.

The graph below shows the relationship between the predicted and real values for sampling stations DP066 and DP072.

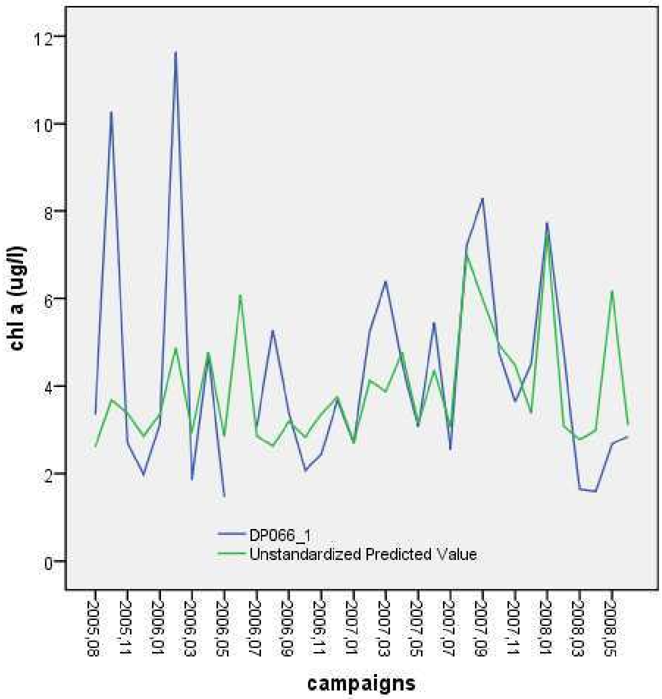


Figure 9: The chart above compares DP066's chlorophyll a series and the virtual series obtained using the model

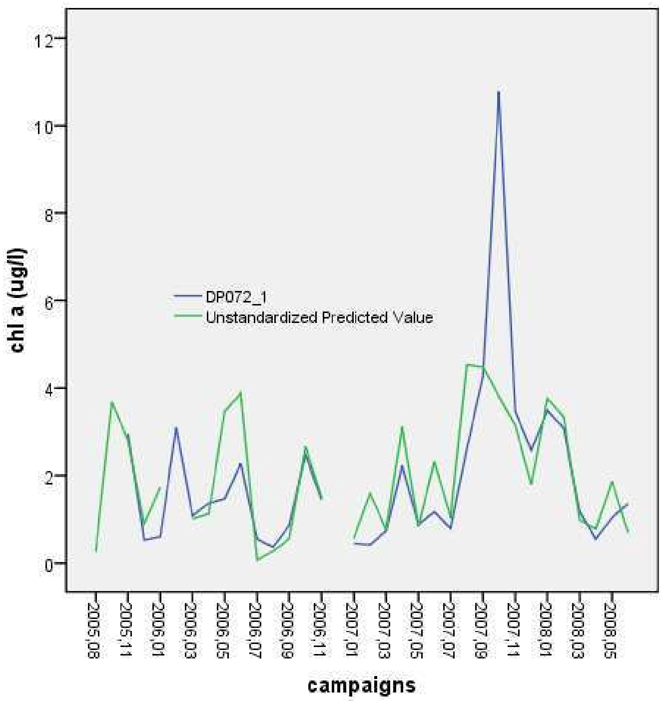


Figure 10: The chart above compares DP133's chlorophyll a series and the virtual series obtained using the model

Data used to obtain 90th percentile	90th percentile	BIAS (error)	Eco.class.
90th PCTL all sampling stations	5,921	-	Good
90th PCTL excluding DP066 and DP072	5,744	0177	Good
90th PCTL virtual values of DP076 and DP079	5,280	0,641	Good

Table 21: 90th percentiles and ecological classification calculated by using virtual data or excluding the sampling stations

Table 21 shows the 90th percentiles and the ecological classifications obtained using the reference data, chosen observer data and the joining of the chosen observer data and the virtual values of the unrevised sampling stations. The 90th percentiles obtained present negligible biases in respect to the 90th percentile obtained using reference data. All 90th percentiles reflect the same quality value (ecological class). These results confirm the accuracy of the virtual values obtained in this system for the unrevised sampling stations.

Therefore, since it is possible to obtain virtual values quite accurately for unrevised stations belonging to w.b.008, it is possible in this system to control for potential stable alterations in the anthropogenic characteristics of specific areas where unrevised stations are located by annually reviewing these stations and the comparison between virtual and actual values.

As mentioned above, the continuous presence of significant discrepancies may have implications for the spatial reduction, because the actual observer is potentially no longer the most suitable to provide a complete and stable view of the area, and performing a revision should be considered.

5.3.8.Water body 011

In water body 011, five alternatives for spatial reduction were selected.

Sampling stations	N. of reduced s.s	90^{th} Percentile
DP002a DP076 DP078 DP079 DP080 DP081	**0**	**2,2630**
DP076 DP078 DP079 DP080 DP081	1	2,2500
DP002a DP076 DP079 DP080 DP081	1	2,3420
DP076 DP078 DP079 DP080	2	2,3240
DP076 DP079 DP080 DP081	2	2,2690
DP002a DP078 DP080 DP081	2	2,2660

Table 22: Selected k-combinations and the 90th percentiles obtained by their data subsets

To study the possibility of estimating virtual values for each of the sampling stations in w.b. 011 using the other sampling stations of this system, we calculated the Pearson correlation coefficients between sampling stations (Table 23) and obtained stepwise regression models to predict the values of each of the sampling stations in w.b. 011 from the other sampling stations. For each station, stepwise regression has made a selection of the best predictors, and coefficient R^2 indicates the percentage of the variability of the response variable explained by the model (Table 24).

In water body 011, there was missing data in 20 of the 34 revised campaigns. Therefore, allocations have been made using the k-nearest neighbours method in these reviews, obtaining the estimated values for missing data from the average of the values of the 3 closest neighbours. Using the Hot-Deck technique, the data from the 13 campaigns was recuperated.

From the information gathered in these tables, the k-observer was chosen from among the selected observers, since it would make the best predictions.

Correlations

		DP002a_1	DP076	DP078	DP079_1	DP080_1	DP081_1
DP002a_1	Pearson Correlation	1,000	,529**	,411*	,315	,022	,216
	Sig. (2-tailed)		,005	,046	,144	,915	,311
	N	27	27	24	23	27	24
DP076	Pearson Correlation	,529**	1,000	,448*	,468*	,052	,310
	Sig. (2-tailed)	,005		,011	,024	,775	,131
	N	27	34	31	23	33	25
DP078	Pearson Correlation	,411*	,448*	1,000	,622**	,511**	,278
	Sig. (2-tailed)	,046	,011		,002	,004	,211
	N	24	31	31	22	30	22
DP079_1	Pearson Correlation	,315	,468*	,622**	1,000	,425*	,741**
	Sig. (2-tailed)	,144	,024	,002		,043	,000
	N	23	23	22	23	23	22
DP080_1	Pearson Correlation	,022	,052	,511**	,425*	1,000	,424*
	Sig. (2-tailed)	,915	,775	,004	,043		,039
	N	27	33	30	23	33	24
DP081_1	Pearson Correlation	,216	,310	,278	,741**	,424*	1,000
	Sig. (2-tailed)	,311	,131	,211	,000	,039	
	N	24	25	22	22	24	25

**. Correlation is significant at the 0.01 level (2-tailed).

*. Correlation is significant at the 0.05 level (2-tailed).

Table 23: Pearson correlation coefficents for the water body

regression model:		
modled sampling satation	station(s) used in the model	R^2
DP002a	DP076	0,70
DP076	DP002a, DP078	0,83
DP078	DP079, DP076	0,52
DP079	DP081	0,52
DP080	DP081	0,23
DP081	DP079	0,52

Table 24: Predicting variables and the adjusted R square

The optimal combination of sampling stations to be reviewed regularly consists of DP076, DP078, DP079 and DP080. Using the regression model, virtual values for DP0002a and DP081 can be obtained.

Virtual values model for DP002a and DP081

In the model to estimate the values of DP002a, DP076 is incorporated into the model as a predictor variable. The model explains 70% of the variability of DP076. In the model to estimate the values of DP081, DP079 is incorporated into the model as a predictor variable. The model explains 52% of the variability of DP081.

The graphs below show the relationship between the predicted and real values for sampling stations DP002a and DP081. Sampling station DP002a has a significant lack of data. Using the regression model, virtual values were obtained for all campaigns. This data can be used as a proxy to expand the information on the thropic state in the water body. Virtual values provide information that was not available before, however the accuracy of the model cannot be verified.

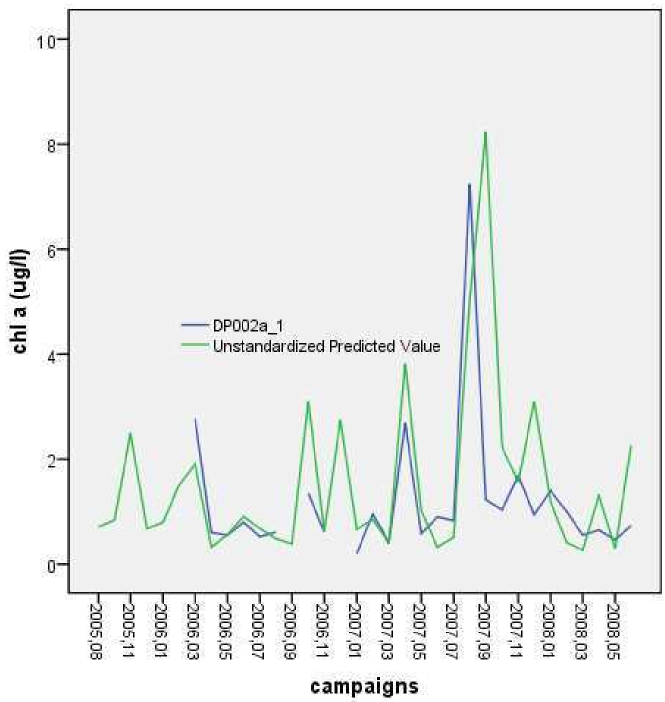


Figure 11: The chart above compares DP002a's chlorophyll a series and the virtual series obtained using the model

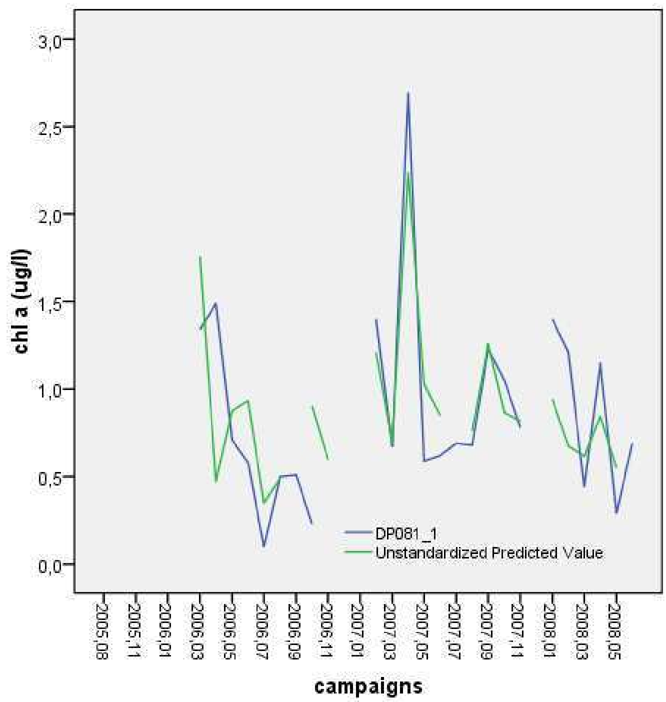


Figure 12: The chart above compares DP081's chlorophyll a series and the virtual series obtained using the model

Data used to obtain 90th percentile	90th percentile	BIAS (error)	Eco.class.
90th PCTL all sampling stations	2,263	-	Good
90th PCTL excluding DP076 and DP079	2,324	0,061	Good
90th PCTL virtual values of DP076 and DP079	2,450	0,187	Good

Table 25: 90th percentiles and ecological classification calculated by using virtual data or excluding the sampling stations

Table 25 shows the 90th percentiles and the ecological classifications obtained using the reference data, chosen observer data and the joining of the chosen observer data and the virtual values of the unrevised sampling stations. The 90th percentiles obtained present negligible biases in respect to the 90th percentile obtained using reference data. All 90th percentiles reflect the same quality value (ecological class). These results confirm the accuracy of the virtual values obtained in this system for the unrevised sampling stations.

Therefore, since it is possible to obtain virtual values quite accurately for unrevised stations belonging to w.b. 011, it is possible in this system to control for potential stable alterations in the anthropogenic characteristics of specific areas where unrevised stations are located by annually reviewing these stations and the comparison between virtual and actual values.

As mentioned above, the continuous presence of significant discrepancies may have implications for the spatial reduction, because the actual observer is potentially no longer the most suitable to provide a complete and stable view of the area, and performing a revision should be considered.

5.3.9.Water body 014

In water body 014, four alternatives for spatial reduction were selected.

Sampling stations	N. of reduced s.s.	90th Percentile
DP096 DP098 DP100 DP102	**0**	**1,791**
DP096 DP100 DP102	1	1,720
DP098 DP100 DP102	1	1,803
DP096 DP098 DP102	1	1,798
DP096 DP098 DP100	1	1,801

Table 26: Selected k-observers and the 90th percentiles obtained by their data subsets

To study the possibility of estimating virtual values for each of the sampling stations in w.b. 014 using the other sampling stations of this system, we calculated the Pearson correlation coefficients between sampling stations (Table 27) and obtained stepwise regression models to predict the values of each of the sampling stations in w.b. 014 from the other sampling stations. For each station, stepwise regression has made a selection of the best predictors, and coefficient R^2 indicates the percentage of the variability of the response variable explained by the model (Table 28).

In water body 014, there was missing data in 7 of the 34 revised campaigns. Therefore, allocations have been made using the k-nearest neighbours method in these reviews, obtaining the estimated values for missing data from the average of the values of the 3 closest neighbours. Using the Hot-Deck technique, the data from the five campaigns was recuperated.

From the information gathered in these tables, the k-observer was chosen from among the selected observers, since it would make the best predictions.

Correlations

		DP096_1	DP098	DP100_1	DP102_1
DP096_1	Pearson Correlation	1,000	,377*	,380*	,389*
	Sig. (2-tailed)		,030	,029	,025
	N	33	33	33	33
DP098	Pearson Correlation	,377*	1,000	,343	,229
	Sig. (2-tailed)	,030		,051	,201
	N	33	33	33	33
DP100_1	Pearson Correlation	,380*	,343	1,000	,060
	Sig. (2-tailed)	,029	,051		,735
	N	33	33	34	34
DP102_1	Pearson Correlation	,389*	,229	,060	1,000
	Sig. (2-tailed)	,025	,201	,735	
	N	33	33	34	34

*. Correlation is significant at the 0.05 level (2-tailed).

Table 27: Pearson correlation coefficents for the water body

regression model:		
modled sampling satation	station(s) used in the model	R^2
DP096	DP100,DP102	0,22
DP098	DP096	0,12
DP100	DP096	0,12
DP102	DP096	0,12

Table 28: Predicting variables and the adjusted R squared

The optimal combination of sampling stations to be reviewed regularly consists of DP098, DP100 and DP102. Using the regression model, virtual values for DP096 can be obtained with low accuracy. The virtual values, due to the low correlation, should be used as a proxy of the tendency of the trophic state in the area of station DP096.

Virtual values model for DP096

In the model to estimate the values of DP096, DP100 and DP102 are incorporated into the model as predictor variables. The model explains 22% of the variability of DP096.

The graph below shows the relationship between the predicted and real values for sampling station DP096.

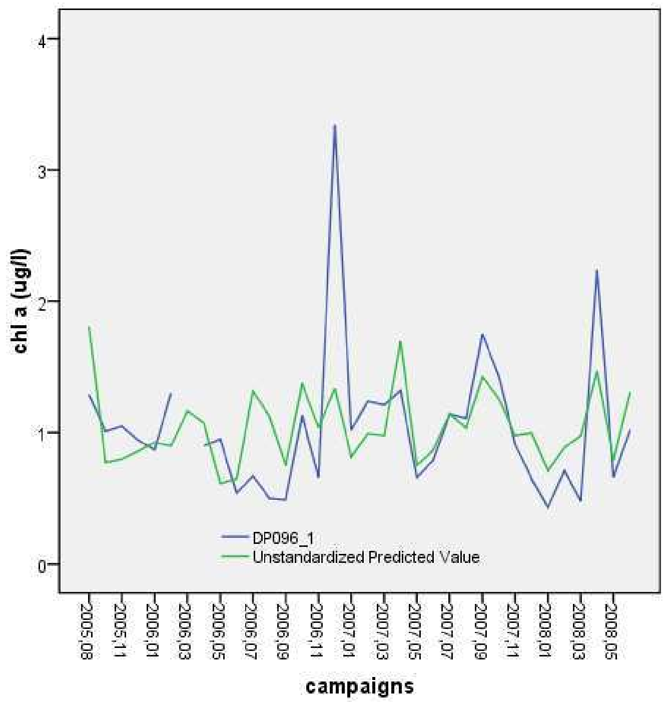


Figure 13: The chart above compares DP096's chlorophyll a series and the virtual series obtained using the model

Data used to obtain 90th percentile	90th percentile	BIAS (error)	Eco.class.
90th PCTL all sampling stations	1,791	-	High
90th PCTL excluding DP076 and DP079	1,803	0,012	High
90th PCTL virtual values of DP076 and DP079	1,699	0,092	High

Table 29: 90th percentiles and ecological classification calculated by using virtual data or excluding the sampling stations

Table 29 shows the 90th percentiles and the ecological classifications obtained using the reference data, chosen observer data and the joining of the chosen observer data and the virtual values of the unrevised sampling stations. The 90th percentiles obtained present negligible biases in respect to the 90th percentile obtained using reference data. All 90th percentiles reflect the same quality value (ecological class). These results confirm the accuracy of the virtual values obtained in this system for the unrevised sampling stations.

Therefore, since it is possible to obtain virtual values quite accurately for unrevised stations belonging to w.b.014, it is possible in this system to control for potential stable alterations in the anthropogenic characteristics of specific areas where unrevised stations are located by annually reviewing these stations and the comparison between virtual and actual values.

As mentioned above, the continuous presence of significant discrepancies may have implications for the spatial reduction, because the actual observer is potentially no longer the most suitable to provide a complete and stable view of the area, and performing a revision should be considered.

5.3.10.Water body 019

In water body 019, three alternatives for spatial reduction were selected.

Sampling stations	N. of reduced s.s.	90th Percentile
DP125 DP127 DP129 DP131	**0**	**3,1000**
DP125 DP129 DP131	1	3,3190
DP125 DP127 DP129	1	3,1000
DP129 DP127 DP131	1	3,2080

Table 30: Selected k-observers and the 90th percentiles obtained by their data subsets

To study the possibility of estimating virtual values for each of the sampling stations in w.b. 019 using the other sampling stations of this system, we calculated the Pearson correlation coefficients between sampling stations (Table 31) and obtained stepwise regression models to predict the values of each of the sampling stations in w.b. 019 from the other sampling stations. For each station, stepwise regression has made a selection of the best predictors, and coefficient R^2 indicates the percentage of the variability of the response variable explained by the model (Table 32).

In water body 019, there was missing data in 3 of the 34 revised campaigns. Therefore, allocations have been made using the k-nearest neighbours method in these reviews, obtaining the estimated values for missing data from the average of the values of the 3 closest neighbours. Using the Hot-Deck technique, the data from all campaigns was recuperated.

From the information gathered in these tables, the k-observer was chosen from among the selected observers, since it would make the best predictions.

Correlations

		DP125_1	DP127_1	DP129_1	DP131
DP125_1	Pearson Correlation	1,000	,770**	,518**	,385*
	Sig. (2-tailed)		,000	,002	,024
	N	34	34	34	34
DP127_1	Pearson Correlation	,770**	1,000	,734**	,507**
	Sig. (2-tailed)	,000		,000	,002
	N	34	34	34	34
DP129_1	Pearson Correlation	,518**	,734**	1,000	,154
	Sig. (2-tailed)	,002	,000		,385
	N	34	34	34	34
DP131	Pearson Correlation	,385*	,507**	,154	1,000
	Sig. (2-tailed)	,024	,002	,385	
	N	34	34	34	34

**. Correlation is significant at the 0.01 level (2-tailed).

*. Correlation is significant at the 0.05 level (2-tailed).

Table 31: Pearson correlation coefficents for the water body

regression model:		
modled sampling satation	station(s) used in the model	R^2
DP129	DP127, DP131	0,58
DP131	DP127, DP129	0,32
DP127	DP125, DP129, DP131	0,79
DP125	DP127	0,58

Table 32: Predicting variables and the adjusted R square

The optimal combination of sampling stations to be reviewed regularly is formed by DP125, DP129 and DP131. Using the regression model, virtual values for DP127 can be obtained.

Virtual values model for DP127

In the model to estimate the values of DP127, DP125 DP127 and DP131 are incorporated into the model as predictor variables. The model explains 79% of the variability of DP127.

The graph below shows the relationship between the predicted and real values for sampling station DP127.

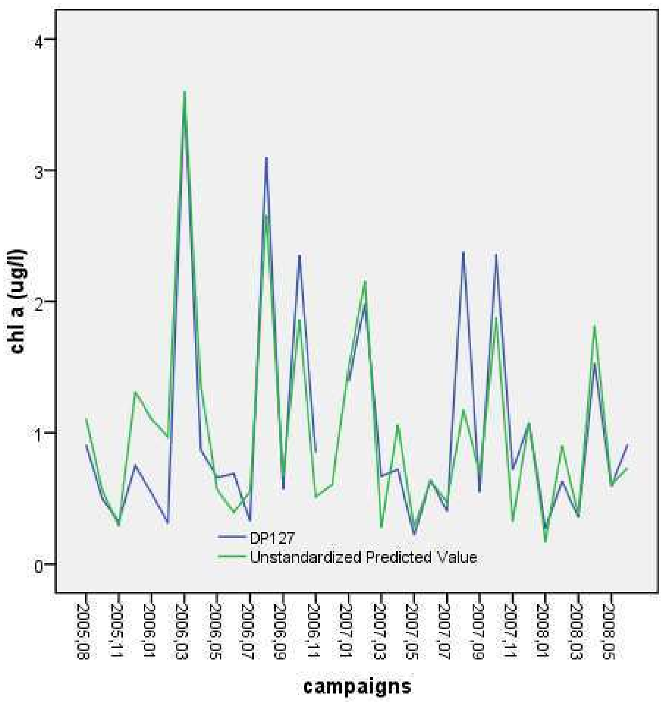


Figure 14: The chart above compares DP127's chlorophyll a series and the virtual series obtained using the model

Data used to obtain 90th percentile	90th percentile	BIAS (error)	Eco.class.
90th PCTL all sampling stations	3,100	-	Good
90th PCTL excluding DP076 and DP079	3,319	0,219	Good
90th PCTL virtual values of DP076 and DP079	2,878	0,222	Good

Table 33: 90th percentiles and ecological classification calculated by using virtual data or excluding the sampling stations

Table 33 shows the 90th percentiles and the ecological classifications obtained using the reference data, chosen observer data and the joining of the chosen observer data and the virtual values of the unrevised sampling stations. The 90th percentiles obtained present negligible biases in respect to the 90th percentile obtained using reference data. All 90th percentiles reflect the same quality value (ecological class). These results confirm the accuracy of the virtual values obtained in this system for the unrevised sampling stations.

Therefore, since it is possible to obtain virtual values quite accurately for unrevised stations belonging to w.b.019, it is possible in this system to control for potential stable alterations in the anthropogenic characteristics of specific areas where unrevised stations are located by annually reviewing these stations and the comparison between virtual and actual values.

As mentioned above, the continuous presence of significant discrepancies may have implications for the spatial reduction, because the actual observer is potentially no longer the most suitable to provide a complete and stable view of the area, and performing a revision should be considered.

5.3.11.Discussion

The comparative study of chlorophyll *a* concentration indicates that, although there is stability in the range of values obtained from different sampling stations, specific factors may be involved that alter the normal values of chlorophyll *a* concentration from a particular sampling station. This implies that spatial reduction should be supervised using a tool that can give notice if the relations between water body sampling stations are significantly changed and do not present the same environmental condition for which the observer has been selected.

The reduction decision made on the basis of historical data does not imply a possible future alteration in the water body, for which reason post-reduction management is necessary.

Because of this, regression models have been developed that can give us additional information for the parts of water bodies that are not surveyed after selecting the best observer. The chlorophyll *a* concentrations of unrevised sampling stations (virtual values) can be estimated by chlorophyll *a* values from observed stations using the regression models.

These virtual values should be used for post reduction management. It should be focused on the regression model if the model provides us with reliable information. The virtual values of chlorophyll *a* concentration should be compared with those values observed in occasional revisions of reduced sampling stations. In the event that, in subsequent revisions, the values observed in any of these stations strays significantly from the virtual values, more frequent monitoring of the sampling station and a possible review of the optimal combination of stations that will be reviewed regularly should be considered.

A continued discrepancy between observed and predicted values by the regression model in an unrevised station will indicate that there has been a possible alteration in the anthropogenic balance of the system, which should be investigated. Isolated discrepancies should be handled with caution, and actions should not be taken rapidly (virtual values should be treated as a proxy), especially in those systems in which the regression model explains the lower percentage of station variability.

Finally, accuracy in the regression models can be significantly enhanced by integrating environmental data such as currents, waves and rain. Significant variations in regular currents can cause a regression model to overestimate or underestimate the virtual values. For this reason, the virtual values should be combined with data on environmental factors (previously mentioned) so that over/underestimation can be detected.

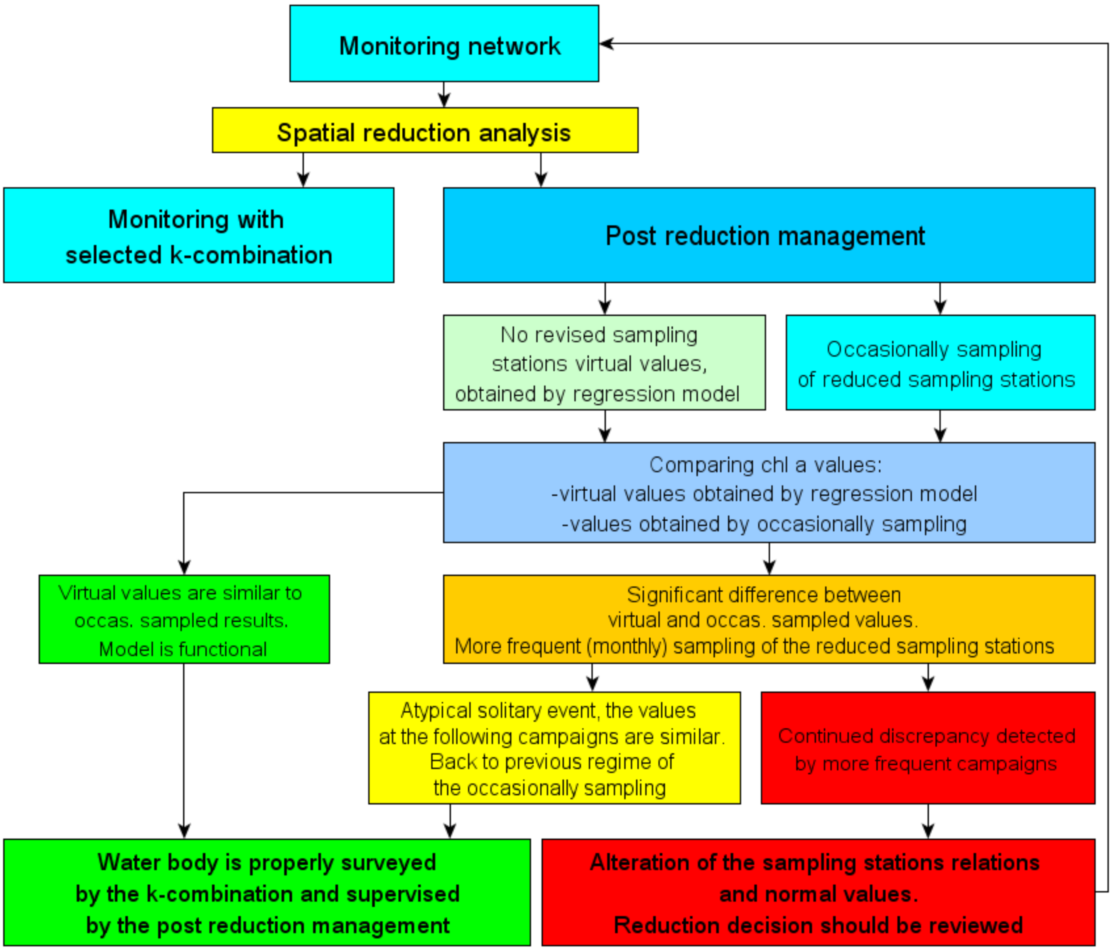


Figure 15: Post reduction management

6.Conclusions and future research

6.1.Conclusions

Three years of historical data allowed the detailed study of the monitoring network using temporal and spatial reduction. This study has determined the consequences for ecological classifications caused by reductions and has identified potential reductions in temporal frequency and spatial density.

Temporal reduction analysis provided results with the following conclusions:

1.Concerning the "Temporal reduction" exploratory study

- The percentage of unchanged classifications (for the three year period) does not descend below 74%, which can be considered a fairly good result, but changed classifications cannot be predicted and could have unexpected consequences for any of the bodies of water.

- The percentage of uncertainty (for bimonthly or trimestral campaigns) will not significantly decrease for the five year period (compared to the three year results), because it does not depend on the quantity of samples, but rather depends on the trend, regularity and patterns of the environmental parameters.

- These analysis results question the minimum survey frequency established by WFD and support one of the first decisions of the MEDGIG experts (at the first intercalibration reunion) to monitor coastal waters monthly in the scope of ecological classification.

2.Concerning the Seasonal pattern regularity

- Coastal ecosystems are complex and the level of eutrophic balance depends on anthropogenic influence and environmental factors that are determined by climate seasonality and irregularity.

- Environmental factors (precipitation and wave data) sometimes confirm seasonality, but not regularity. The climate has a generally seasonal pattern during the year (significant rain events and sea storms) but lacks regularity that would make it possible to know which months would have significant precipitation or sea storm events.
- Besides that, sea storms can have different effects on biomass growth:

  1. Waves can accelerate the mixing process of fresh water and superficial sea layers. This process can reduce nutrient concentration, and therefore biomass, during periods of significant continental loads.

2. During sea storms with elevated wave heights, vertical mixing of the sea column can occur. Deep, nutrient-rich sea water enriched with nutrients can accelerate phytoplankton biomass growth.

3.Concerning water body categories and temporal reduction results

- Based upon the temporal reduction analysis results, the following coastal ecosystem categorization has been established:

  1. Stable systems
     Water bodies with low chlorophyll *a* values that rarely or never exceed High/Good threshold. They have a low standard deviation value that is significantly lower than the “buffer” value and for all periods are classified as High.

  2. Unstable systems
     Water bodies that are classified as Moderate for all analyzed periods. The standard deviation value is extremely high and significantly exceeds the “buffer” values at the Good/Moderate threshold.

  3. Systems with ecological indeterminate classification
     All water bodies that are classified as neither stable nor unstable systems.

- All analyses indicate that temporal reduction is possible for stable ecosystems and the quality results will be unchanged for surveys with reduced frequency.

- Temporal reduction results suggest that reducing survey frequency is possible for unstable water bodies. Reduced monitoring frequency will not affect the ecological classification, but the precision of the time series will be seriously decreased

- Reducing the survey frequency for systems that are classified as neither stable nor unstable results in unpredictable consequences for ecological classification.

Spatial reduction analysis provided results with following conclusions:

4. Concerning the “Spatial reduction” exploratory study

   - The percentage of unchanged classifications (for the three year period), obtained by reducing the number of sampling stations in a system depends on the coastal ecosystem categorization:

   - In the stable systems with undoubted ecological classification for which we studied the possibility of spatial reduction (w.b. 001, 002 and 014), the percentage of classifications that were not altered by reducing one or two sampling stations was 100%.
   - In the systems with indeterminate ecological classification for which we studied the possibility of spatial reduction (w.b. 005, 009, 010, 011 and 019), the percentage of classifications that were not altered by reducing one or two sampling stations ranges from 66.7% to 100% and between 73.3% and 90%. It is possible in some of these systems to reduce by at least one station, but the selection of these stations should be carried out carefully.
   - In the only unstable system for which we studied the possibility of spatial reduction (w.b. 008), the percentage of classifications that were altered by reducing one or two sampling stations is 80%.

5. Concerning the selection of the sampling stations that should remain in monitoring programs

   - When making a reduction in the number of sampling stations reviewed in a system, it is especially relevant in the case of stable systems with indeterminate classification as well as unstable systems to make a proper selection of the stations which should remain in the monitoring program.

   - It is desirable that the sampling stations which remain within the monitoring program provide a complete and balanced view of the system, detecting abnormalities without biasing the information.

   - Statistical analysis of agreement between observers, adapted to our problem, allows us to make this selection properly.

   - MDS technique facilitates global assessment of the proximity between the information provided by the set of all sampling stations operating in a system and the information provided by different combinations of these stations. This makes the study of the cost effectiveness relationship for the selected space reductions easier.

6. Post-reduction management

- Spatial reduction should be supervised using a tool that can give notice if the relations between water body sampling stations are significantly changed and do not present the same environmental conditions for which the observer was selected. Reduction decisions made on the basis of historical data do not include the possibility of future changes in the water body, for which reason post-reduction management is a necessity.

- For this reason, we have developed regression models that can give us additional information for the parts of the water bodies that are not surveyed after selecting the best observer. Using these models, the chlorophyll *a* concentrations of unrevised sampling stations (virtual values) can be estimated using chlorophyll *a* values from observed stations.

- The virtual values of chlorophyll *a* concentration should be compared with values observed in occasional revisions of reduced sampling stations. A continued discrepancy between observed values and values predicted by the regression model in an unrevised station will indicate that there has been a possible alteration in the anthropogenic balance of the system, which should be investigated.

- Isolated discrepancies should be approached with caution and actions should not be taken hastily (virtual values should be treated as a proxy), especially in those systems in which the regression model explains the lower percentage of station variability.

6.2.Future research directions

- Using the same approach, analyze possible spatial and temporal reduction for five-year data.

- Apply the same approach to the phytoplankton composition systems indicator

- Apply the same approach to an analysis of phytoplankton bloom frequencies

- Apply the same approach to an analysis of the nutrients that determine chemical status at WFD

- Using five years of data, calibrate the described virtual value model and update with information provided from other environmental networks (waves, wind, rain…) to obtain more accurate results.

- Establish standardized algorithms for the percentiles that will be used for the classification of coastal water bodies in the WFD.

7.References

- Agostini VN, Bakun A. Ocean triads' in the Mediterranean Sea: physical mechanisms potentially structuring reproductive habitat suitability (with example application to European anchovy, Engraulis encrasicolus). *Fisheries Oceanography*. 2002;11(3):129-142.

- Andersen JH. Coastal eutrophication et al recent developments in definitions and implications for monitoring strategies. *Journal of Plankton Research*. 2006;28(7):621-628.

- Azov Y. Seasonal patterns phytoplankton productivity and abundance in nearshore oligotrophic waters of the Levant Basin (Mediterranean). *Journal of Plankton Research*. 1986;8:41-53.

- Ærtebjerg, G., Andersen, J.H. & Hansen, O.S. Nutrients and Eutrophication in Danish Marine Waters A Challenge for Science and Management. National Environmental Research Institute. 2003.

- Baretta-Bekker JG, Baretta JW, Ebenhöh W. Microbial dynamics in the marine ecosystem model ERSEM II with decoupled carbon assimilation and nutrient uptake. *Journal of Sea Research*. 1997;38(3-4):195-211.

- Batzias F, Siontorou CG. A novel system for environmental monitoring through a cooperative/synergistic scheme between bioindicators and biosensors. *Journal of Environmental Management*. 2007;82(2):221-239.

- Becacos, Kontos. Primary Production and Environmental Factors in an Oligotrophic. Biome in the *Aegean Sea.* Marine Biology.1977; 42(2). 93-98.

- Bethoux JP, Gentili B. Functioning of the Mediterranean Sea: past and present changes related to freshwater input and climate changes. *Journal of Marine Systems*. 1999;20(1-4):33-47.

- Bricker S. An integrated methodology for assessment of estuarine trophic status. *Ecological Modelling*. 2003;169(1):39-60.

- Bonin DJ, Bonin MC, Berman T. Experimental evidence of nutrients limiting the production of micro-nanoplankton and ultraplankton in the coastal waters of the eastern Mediterranean ocean, Haifa, Israel. Aqua Sci. 1989. 51, 129–152

- Cloern JE. Our evolving conceptual model of the coastal eutrophication problem. *Marine Ecology Progress Series*. 2001;210(3301):223-253.

•Conley DJ, Markager S, Andersen J, Ellermann T, Svendsen LM. Coastal eutrophication and the Danish National Aquatic Monitoring and Assessment Program. *Estuaries*. 2002;25(4B):848-861.

•Colpan S, Beken.Report on Existing or Planned Monitoring Activities and Suggestions For Optimal Modification. UNEP/MAP-MED POL, IASON. 2005.

•Jonge VN, Elliott M, Brauer VS. Marine monitoring: Its shortcomings and mismatch with the EU Water Framework Directive's objectives. *Marine Pollution Bulletin*. 2006;53(1-4):5-19

•Dubelaar GB, Geerders PJ, Jonker RR. High frequency monitoring reveals phytoplankton dynamics. *Journal of environmental monitoring JEM*. 2004;6(12):946-952.

•EU. *Council Directive 91/271/EEC of 21 May 1991 concerning urban waste-water treatment*. Official Journal of the European Communities, L135/40-52, 30 May 1991; 1991.

•European Commission , 2005. Towards a guidance document on eutrophication assessment in the context of European water policies. Interim report in the frame of the "Common implementation strategy for the Water Framework Directive, 133pp

•European Commission. Water Framework Directive.; 2000:1-72.

•European Commission. Directiva Nitratos (91/676/EEC). 1991.

•Ferreira JG, Vale C, Soares CV, et al. Monitoring of coastal and transitional waters under the E.U. Water Framework Directive. *Environmental Monitoring and Assessment*. 2007;135(1-3):195-216.

•Ferreira J, Nobre A, Simas T, et al. A methodology for defining homogeneous water bodies in estuaries – Application to the transitional systems of the EU Water Framework Directive. *Estuarine, Coastal and Shelf Science*. 2006;66(3-4):468-482.

•Giovanardi F, Vollenweider RA. Trophic conditions of marine coastal waters: experience in applying the Trophic Index TRIX to two areas of the Adriatic and Tyrrhenian seas. *Water Research*. 2004;63(2):199-218.

•Gómez M., Carretero J.C. Implementation of a high resolution third generation model on the Atlantic Coast of Spain. PIANC Bulletin. 1994; 83/84.

•Gunther,H.,Hasselmann,S.,Janssen,P.A.E.M. Wamodel Cycle 4.Technical Report No.4,Deutsche Klimarechenzentrum,Germany. 1991

•Hattermann F, Kundzewicz W: Water Framework Directive: Model supported Implementation - *A Water Manager's Guide. 2009.*

•Heiskanen A, Carstensen J, Gasiūnaitė Z, et al. Monitoring strategies for phytoplankton in the Baltic Sea coastal waters. *Environment*. 2005:45.

- Hojberg A, Refsgaard J, Van Geer F, Jorgensen L, Zsuffa I. Use of Models to Support the Monitoring Requirements in the Water Framework Directive. *Water Resources Management*. 2007;21(10):1649-1672.

- Huelsenbeck JP, Nielsen R, Bollback JP. Stochastic Mapping of Morphological Characters. *Systematic Biology*. 2003;52(2):131-158.

- Hunt CD, Rust SW, Sinnott L. Application of statistical modeling to optimize a coastal water quality monitoring program. *Environmental Monitoring and Assessment*. 2008;137(1-3):505-522.

- Hyndman RJ, Fan Y. Sample quantiles in statistical packages. *The American Statistician*. 1996;50(4):361-365.

- ICES. Report of the ICES/OSPAR Workshop on Integrated Monitoring of Contaminants and their Effects in Coastal and Open-sea Areas (WKIMON III), ICES Headquarters. ICES CM 2007. ACME:01. 209 pp.

- Jeffrey SW, Humphrey GF. New spectrophotometric equations for determining chlorophylls a, b, c1 and c2 in higher plants, algae and natural phytoplankton. *Biochem Physiol Pflanzen*. 1975;167(2):191-194.

- Nunes LM, Caeiro S, Cunha MC, Ribeiro L. Optimal estuarine sediment monitoring network design with simulated annealing. *Journal of Environmental Management*. 2006;78(3):294-304.

- Oros A, Mihnea V,: Report on the Anthropogenic Pressures on the Coastal Zone. UNEP/MAP-MED POL, IASON .2006.

- OSPAR. *Draft JAMP Eutrophication Monitoring Guidelines: Chlorophyll a*. Ad Hoc Working Group On Monitoring (MON); 1996.

- Pojed,S. Kveder. Investigation of nutrient limitation of phytoplankton production in the Northern Adriatic by enrichment experiments. Thalassia Jugosl. 1977; 13 , 13–24

- Rönnberg, C. & E. Bonsdorff,: Baltic Sea eutrophication: area-specific ecological consequences. – Hydrobiologia. 2004;514: 227-241.

- Smayda TJ. Patterns of variability characterizing marine phytoplankton, with examples from Narragansett Bay. *ICES Journal of Marine Science*. 1998;55(4):562-573.

- Vukadin, Stojanovski,. C:*N*:*Si*:*P Ratio in the waters of the middle and south Adriatic*, ICES Mar Sci Symp .1976; 23. 41-43.

- Wasson JG, Chandesris A, Pella H, Blanc L. *Typology and reference conditions for surface water bodies in France - the hydro-ecoregion approach*. CEMAGREF; 2002:5.

- Zavatarelli. Report on the design of a future coupled physical - ecosystem model of the system. UNEP/MAP-MED POL, IASON. 2005.

ANNEX 1

Annex "Exploratory study results"

Comparing the qualitative and quantitative results, obtained by the reduced number of sampling stations (k-combinations) with the results obtained by the complete data.

EXPLORATORY STUDY RESULTS WATER BODY: 001

k-combination spatial reduction analysis	error to original 90th percentile period 2005/08-2006/07 (µg/l)	error to original 90th percentile period 2006/08-2007/07 (µg/l)	error to original 90th percentile period 2007/08-2008/07 (µg/l)	error to original 90th percentile period 2005/08-2008/07 (µg/l)
DP001 DP003 DP005 DP007	0,480	0,214	-0,301	0,308
DP001 DP003 DP005 DP133	-0,805	-0,306	-0,745	-0,150
DP001 DP005 DP133 DP007	-0,805	-0,371	-0,517	-0,214
DP001 DP003 DP133 DP007	-0,094	-0,259	-0,073	-0,102
DP003 DP005 DP133 DP007	-0,079	0,044	0,269	0,154
DP005 DP133 DP007	-0,820	-0,049	0,170	-0,098
DP003 DP005 DP007	0,466	0,369	0,275	0,552
DP003 DP133 DP007	-0,188	-0,026	0,275	0,062
DP001 DP133 DP007	-1,056	-1,176	-0,289	-0,602
DP003 DP005 DP133	-0,820	0,042	0,167	-0,098
DP001 DP005 DP007	0,220	0,099	-2,569	0,134
DP001 DP003 DP133	-1,056	-1,176	-0,529	-0,584
DP001 DP003 DP007	0,474	0,000	-0,049	0,194
DP001 DP003 DP005	0,466	0,244	-2,581	0,278
DP001 DP005 DP133	-1,056	-1,386	-2,581	-0,794
DP005 DP133	-1,201	-0,581	-0,517	-0,653
DP005 DP007	0,165	0,487	0,221	0,532
DP003 DP007	0,497	0,164	0,281	0,374
DP133 DP007	-1,201	-0,483	0,221	-0,184
DP001 DP003	0,458	-0,921	-2,689	0,028
DP003 DP005	0,452	0,599	-0,004	0,662

Annex table: Comparing the qualitative (ecological status) and quantitative (90th percentile) results obtained by reduced number of sampling stations with results obtained by complete data. Red field – degradation of ecological status obtained by reduced data, green field –upgrade of ecological status, yellow field – preserve original ecological status for the all four periods.

EXPLORATORY STUDY RESULTS WATER BODY: 001

k-combination spatial reduction analysis	error to original 90th percentile period 2005/08-2006/07 (µg/l)	error to original 90th percentile period 2006/08-2007/07 (µg/l)	error to original 90th percentile period 2007/08-2008/07 (µg/l)	error to original 90th percentile period 2005/08-2008/07 (µg/l)
DP003 DP133	**-1,201**	**-0,444**	**0,224**	**-0,115**
DP001 DP133	**-1,607**	**-1,827**	**-2,689**	**-1,628**
DP001 DP007	**0,171**	**-0,986**	**-2,425**	**0,002**
DP001 DP005	**-0,388**	**-1,077**	**-2,689**	**-0,900**
DP001	**-1,102**	**-2,091**	**-3,283**	**-2,178**
DP133		**-1,596**		**-1,130**
DP007	**0,140**	**0,023**	**0,280**	**0,242**
DP005	**-0,522**	**1,019**		**0,746**
DP003	**0,564**	**0,162**		**0,472**

Annex table: Comparing the qualitative (ecological status) and quantitative (90th percentile) results obtained by reduced number of sampling stations with results obtained by complete data. Red field – degradation of ecological status obtained by reduced data, green field –upgrade of ecological status, yellow field – preserve original ecological status for the all four periods.

EXPLORATORY STUDY RESULTS WATER BODY: 002

k-combination spatial reduction analysis	error to original 90th percentile period 2005/08-2006/07 (µg/l)	error to original 90th percentile period 2006/08-2007/07 (µg/l)	error to original 90th percentile period 2007/08-2008/07 (µg/l)	error to original 90th percentile period 2005/08-2008/07 (µg/l)
DP134 DP135 DP136 DP010	**-0,240**	**-0,264**	**-0,613**	**-0,299**
DP135 DP136 DP137 DP010	**0,160**	**-0,042**	**0,879**	**0,307**
DP134 DP135 DP136 DP137	**0,340**	**0,548**	**0,774**	**0,388**
DP134 DP136 DP137 DP010	**-0,240**	**-0,264**	**-0,679**	**-0,326**
DP134 DP135 DP137 DP010	**-0,240**	**-0,237**	**-0,010**	**-0,072**
DP134 DP136 DP137	**-0,023**	**0,200**	**-0,224**	**0,143**
DP135 DP137 DP010	**-0,028**	**-0,597**	**1,404**	**0,373**
DP135 DP136 DP137	**0,484**	**0,652**	**1,444**	**0,753**
DP135 DP136 DP010	**-0,080**	**-0,726**	**-0,198**	**-0,088**
DP134 DP137 DP010	**-0,624**	**-0,786**	**-0,712**	**-0,677**
DP136 DP137 DP010	**-0,080**	**-0,726**	**-0,488**	**-0,092**
DP134 DP135 DP137	**0,301**	**0,517**	**1,082**	**0,580**
DP134 DP135 DP136	**-0,023**	**0,410**	**0,000**	**0,153**
DP134 DP136 DP010	**-0,624**	**-0,888**	**-0,856**	**-1,248**
DP134 DP135 DP010	**-0,624**	**-0,786**	**-0,646**	**-0,607**
DP136 DP010	**-0,368**	**-1,506**	**-0,869**	**-1,181**
DP137 DP010	**-0,368**	**-1,505**	**-0,564**	**-0,588**
DP136 DP137	**0,268**	**0,424**	**0,741**	**0,398**
DP134 DP010	**-1,038**	**-1,505**	**-1,996**	**-1,422**
DP135 DP137	**0,586**	**0,645**	**1,853**	**0,805**
DP134 DP135	**-0,944**	**0,350**	**-0,010**	**0,153**
DP134 DP136	**-1,121**	**0,020**	**-0,770**	**-0,330**
DP135 DP136	**0,337**	**0,680**	**1,002**	**0,794**
DP135 DP010	**-0,368**	**-1,505**	**-0,073**	**-0,464**

Annex table: Comparing the qualitative (ecological status) and quantitative (90th percentile) results obtained by reduced number of sampling stations with results obtained by complete data. Red field – degradation of ecological status obtained by reduced data, green field –upgrade of ecological status, yellow field – preserve original ecological status for the all four periods.

EXPLORATORY STUDY RESULTS WATER BODY: 002

k-combination spatial reduction analysis	error to original 90th percentile period 2005/08-2006/07 (µg/l)	error to original 90th percentile period 2006/08-2007/07 (µg/l)	error to original 90th percentile period 2007/08-2008/07 (µg/l)	error to original 90th percentile period 2005/08-2008/07 (µg/l)
DP134 DP137	**-1,013**	**0,175**	**-0,302**	**0,123**
DP136		**0,136**	**-0,592**	**0,074**
DP010	**-0,752**	**-2,175**	**-3,178**	**-1,582**
DP137		**0,393**		**0,634**
DP134		**-0,321**	**-2,104**	**-1,442**
DP135		**0,722**	**1,942**	**0,918**

Annex table: Comparing the qualitative (ecological status) and quantitative (90th percentile) results obtained by reduced number of sampling stations with results obtained by complete data. Red field – degradation of ecological status obtained by reduced data, green field –upgrade of ecological status, yellow field – preserve original ecological status for the all four periods.

Exploratory study results- WATER BODY 002

EXPLORATORY STUDY RESULTS WATER BODY: 005

k-combination spatial reduction analysis	error to original 90th percentile period 2005/08-2006/07 (µg/l)	error to original 90th percentile period 2006/08-2007/07 (µg/l)	error to original 90th percentile period 2007/08-2008/07 (µg/l)	error to original 90th percentile period 2005/08-2008/07 (µg/l)
DP034 DP036 DP038 DP040	0,138	-0,192	-0,319	-0,332
DP032 DP036 DP038 DP040	-0,557	-0,192	-0,074	-0,257
DP032 DP034 DP036 DP038	-0,557	1,109	-0,319	0,018
DP032 DP034 DP036 DP040	0,368	-0,018	0,726	0,439
DP032 DP034 DP038 DP040	0,348	-0,018	0,971	0,439
DP034 DP038 DP040	0,540	-0,202	-0,756	-0,171
DP034 DP036 DP038	-0,618	0,888	-0,776	-1,054
DP032 DP038 DP040	-0,468	-0,202	0,256	-0,021
DP032 DP036 DP040	-0,084	-0,202	0,686	0,009
DP036 DP038 DP040	-0,618	-0,317	-0,776	-0,723
DP032 DP036 DP038	-0,880	0,993	-0,776	-0,717
DP034 DP036 DP040	0,652	-0,202	0,328	-0,015
DP032 DP034 DP040	0,652	-0,036	2,958	1,325
DP032 DP034 DP036	-0,084	1,026	0,328	0,505
DP032 DP034 DP038	-0,468	1,026	-0,756	0,577
DP036 DP038	-2,787	-0,488	-1,631	-1,840
DP038 DP040	-2,370	-0,328	-1,631	-0,607
DP032 DP034	0,351	0,943	2,774	1,349
DP034 DP040	0,805	-0,212	2,705	1,219
DP034 DP038	-2,370	0,794	-1,631	-1,409
DP032 DP038	-2,796	0,972	-1,631	-0,271
DP032 DP036	-0,679	0,900	0,295	0,027
DP034 DP036	0,057	0,794	-0,147	-0,045
DP036 DP040	-0,006	-0,328	0,295	-0,317

Annex table: Comparing the qualitative (ecological status) and quantitative (90th percentile) results obtained by reduced number of sampling stations with results obtained by complete data. Red field – degradation of ecological status obtained by reduced data, green field –upgrade of ecological status, yellow field – preserve original ecological status for the all four periods.

EXPLORATORY STUDY RESULTS WATER BODY: 005

k-combination spatial reduction analysis	error to original 90th percentile period 2005/08-2006/07 (µg/l)	error to original 90th percentile period 2006/08-2007/07 (µg/l)	error to original 90th percentile period 2007/08-2008/07 (µg/l)	error to original 90th percentile period 2005/08-2008/07 (µg/l)
DP032 DP040	**0,288**	**-0,212**	**3,261**	**1,005**
DP040	**0,708**	**-0,339**	**3,082**	**-0,117**
DP032	**-0,680**	**0,861**	**3,206**	**1,328**
DP038	**-4,620**	**-0,652**	**-4,314**	**-2,981**
DP034	**0,796**	**0,700**	**0,292**	**1,333**
DP036	**-0,740**	**-1,542**	**-0,236**	**-0,955**

Annex table: Comparing the qualitative (ecological status) and quantitative (90th percentile) results obtained by reduced number of sampling stations with results obtained by complete data. Red field – degradation of ecological status obtained by reduced data, green field –upgrade of ecological status, yellow field – preserve original ecological status for the all four periods.

EXPLORATORY STUDY RESULTS WATER BODY: 008

k-combination spatial reduction analysis	error to original 90th percentile period 2005/08-2006/07 (µg/l)	error to original 90th percentile period 2006/08-2007/07 (µg/l)	error to original 90th percentile period 2007/08-2008/07 (µg/l)	error to original 90th percentile period 2005/08-2008/07 (µg/l)
DPU001 DP052 DP054 DP055	3,752	0,000	0,583	0,926
DPU001 DP051 DP054 DP055	-1,826	-0,063	0,583	-0,128
DP051 DP052 DP054 DP055	0,000	-0,143	-0,877	-0,160
DPU001 DP051 DP052 DP054	-1,826	-0,549	-0,442	-0,652
DPU001 DP051 DP052 DP055	-1,826	0,184	-0,877	-0,400
DP051 DP054 DP055	-1,826	-0,223	-0,174	-0,288
DPU001 DP051 DP052	-10,235	-0,360	-4,742	-1,722
DPU001 DP054 DP055	5,081	-0,063	1,690	1,350
DPU001 DP052 DP055	2,546	0,184	0,886	0,700
DP052 DP054 DP055	3,752	-0,143	-0,174	0,930
DPU001 DP051 DP055	-10,235	0,138	0,886	-0,744
DP051 DP052 DP055	-1,826	0,312	-4,742	-0,500
DPU001 DP051 DP054	-10,235	-0,801	0,536	-1,722
DPU001 DP052 DP054	2,546	-0,549	0,536	0,708
DP051 DP052 DP054	-1,826	-1,053	-4,742	-1,798
DP052 DP055	2,546	0,312	-1,019	0,862
DP054 DP055	5,081	-0,223	0,489	1,875
DP051 DP052	-10,235	-0,984	-7,481	-4,184
DPU001 DP052	1,040	-0,360	0,037	0,430
DP051 DP055	-10,235	0,205	-5,114	-2,498
DPU001 DP051	-45,368	-1,203	-4,058	-5,226
DPU001 DP055	5,120	0,138	3,157	0,910
DP052 DP054	2,546	-1,053	-0,758	0,880
DPU001 DP054	4,954	-0,801	1,472	0,930

Annex table: Comparing the qualitative (ecological status) and quantitative (90th percentile) results obtained by reduced number of sampling stations with results obtained by complete data. Red field – degradation of ecological status obtained by reduced data, green field –upgrade of ecological status, yellow field – preserve original ecological status for the all four periods.

EXPLORATORY STUDY RESULTS WATER BODY: 008

k-combination spatial reduction analysis	error to original 90th percentile period 2005/08-2006/07 (µg/l)	error to original 90th percentile period 2006/08-2007/07 (µg/l)	error to original 90th percentile period 2007/08-2008/07 (µg/l)	error to original 90th percentile period 2005/08-2008/07 (µg/l)
DP051 DP054	-10,235	-1,305	-4,853	-3,171
DPU001			3,280	-0,004
DP055	5,120	0,205	0,288	1,750
DP054	4,954	-1,305	0,442	1,455
DP051	-45,368	-2,234	-7,538	-11,329
DP052	1,040	-0,984	-9,632	-0,045

Annex table: Comparing the qualitative (ecological status) and quantitative (90th percentile) results obtained by reduced number of sampling stations with results obtained by complete data. Red field – degradation of ecological status obtained by reduced data, green field –upgrade of ecological status, yellow field – preserve original ecological status for the all four periods.

EXPLORATORY STUDY RESULTS WATER BODY: 009

k-combination spatial reduction analysis	error to original 90th percentile period 2005/08-2006/07 (µg/l)	error to original 90th percentile period 2006/08-2007/07 (µg/l)	error to original 90th percentile period 2007/08-2008/07 (µg/l)	error to original 90th percentile period 2005/08-2008/07 (µg/l)
DP056 DP057 DP059 DP061	0,000	0,922	-0,200	0,744
DP057 DP059 DP061 DP062	-1,720	-0,436	-0,009	-0,378
DP056 DP057 DP059 DP062	0,474	-0,006	0,309	1,360
DP056 DP059 DP061 DP062	0,725	-0,436	-0,009	-0,641
DP056 DP057 DP061 DP062	-1,512	-0,003	-0,195	-0,571
DP057 DP061 DP062	-3,457	-0,663	-0,235	-0,755
DP056 DP057 DP061	-1,512	1,102	-1,717	0,273
DP057 DP059 DP062	-0,132	-0,842	0,243	0,816
DP056 DP059 DP061	0,892	0,950	-0,231	0,187
DP056 DP059 DP062	0,890	-0,842	0,243	0,549
DP056 DP057 DP059	0,543	0,854	0,479	2,752
DP056 DP057 DP062	-0,066	-0,009	0,243	0,883
DP059 DP061 DP062	-3,016	-2,092	-0,018	-0,813
DP056 DP061 DP062	-2,557	-0,663	-0,235	-0,873
DP057 DP059 DP061	-1,720	0,390	-0,231	0,582
DP056 DP057	-0,066	1,070	-1,983	2,896
DP059 DP061	-5,330	-0,976	-0,262	-0,633
DP059 DP062	0,679	-2,900	0,177	-0,571
DP061 DP062	-6,294	-2,800	-0,275	-1,644
DP057 DP061	-5,522	0,495	-1,944	-0,435
DP056 DP061	-4,487	1,165	-1,944	-0,930
DP057 DP059	-0,132	0,260	1,047	2,342
DP056 DP059	1,060	0,850	0,348	2,531
DP057 DP062	-1,928	-1,745	0,177	-0,501

Annex table: Comparing the qualitative (ecological status) and quantitative (90th percentile) results obtained by reduced number of sampling stations with results obtained by complete data. Red field – degradation of ecological status obtained by reduced data, green field –upgrade of ecological status, yellow field – preserve original ecological status for the all four periods.

EXPLORATORY STUDY RESULTS WATER BODY: 009

k-combination spatial reduction analysis	error to original 90th percentile period 2005/08-2006/07 (μg/l)	error to original 90th percentile period 2006/08-2007/07 (μg/l)	error to original 90th percentile period 2007/08-2008/07 (μg/l)	error to original 90th percentile period 2005/08-2008/07 (μg/l)
DP056 DP062	**0,720**	**-1,745**	**0,177**	**-0,749**
DP059	**0,888**	**-1,518**	**0,683**	**1,453**
DP061	**-7,380**	**-1,126**	**-2,171**	**-3,106**
DP057	**-1,928**	**0,297**	**-1,723**	**2,724**
DP062	**-3,475**	**-6,519**	**0,111**	**-1,239**
DP056	**1,676**	**1,159**	**-11,223**	**3,129**

Annex table: Comparing the qualitative (ecological status) and quantitative (90th percentile) results obtained by reduced number of sampling stations with results obtained by complete data. Red field – degradation of ecological status obtained by reduced data, green field –upgrade of ecological status, yellow field – preserve original ecological status for the all four periods.

EXPLORATORY STUDY RESULTS WATER BODY: 010

k-combination spatial reduction analysis	error to original 90th percentile period 2005/08-2006/07 (µg/l)	error to original 90th percentile period 2006/08-2007/07 (µg/l)	error to original 90th percentile period 2007/08-2008/07 (µg/l)	error to original 90th percentile period 2005/08-2008/07 (µg/l)
DP066 DP068 DP070 DP073 DP074	-0,169	-0,306	-0,050	-0,074
DP068 DP070 DP072 DP073 DP074	0,108	0,580	0,224	0,653
DP066 DP068 DP070 DP072 DP073	-0,124	-0,004	-0,540	-0,059
DP066 DP068 DP072 DP073 DP074	-0,040	-0,375	-0,540	-0,080
DP066 DP070 DP072 DP073 DP074	1,156	-0,004	2,302	1,135
DP066 DP068 DP070 DP072 DP074	-0,169	-0,232	-0,450	-0,125
DP066 DP068 DP070 DP072	-0,174	-0,306	-0,940	-0,441
DP066 DP070 DP073 DP074	0,004	-0,380	2,229	0,645
DP068 DP070 DP072 DP073	0,068	0,570	-1,028	0,579
DP068 DP070 DP072 DP074	0,002	0,332	-0,940	0,000
DP066 DP068 DP072 DP074	-1,274	-0,760	-0,940	-0,471
DP066 DP068 DP072 DP073	-0,145	-0,450	-1,028	-0,467
DP068 DP070 DP073 DP074	0,002	0,285	-0,012	0,177
DP066 DP068 DP070 DP073	-0,174	-0,380	-0,585	-0,315
DP066 DP068 DP073 DP074	-1,274	-0,762	-0,585	-0,440
DP066 DP072 DP073 DP074	1,899	-0,450	1,163	0,661
DP066 DP070 DP072 DP073	0,296	-0,005	1,382	0,990
DP066 DP070 DP072 DP074	0,004	-0,306	1,405	0,651
DP068 DP072 DP073 DP074	1,304	0,624	-1,028	0,457
DP070 DP072 DP073 DP074	1,921	0,965	3,233	1,482
DP066 DP068 DP070 DP074	-1,004	-0,758	-0,495	-0,351
DP068 DP070 DP074	-0,168	0,020	-0,984	-0,074
DP068 DP072 DP073	0,404	0,623	-2,440	-0,054
DP066 DP072 DP074	-0,781	-0,762	0,488	0,434

Annex table: Comparing the qualitative (ecological status) and quantitative (90th percentile) results obtained by reduced number of sampling stations with results obtained by complete data. Red field – degradation of ecological status obtained by reduced data, green field –upgrade of ecological status, yellow field – preserve original ecological status for the all four periods.

EXPLORATORY STUDY RESULTS WATER BODY: 010

k-combination spatial reduction analysis	error to original 90th percentile period 2005/08-2006/07 (µg/l)	error to original 90th percentile period 2006/08-2007/07 (µg/l)	error to original 90th percentile period 2007/08-2008/07 (µg/l)	error to original 90th percentile period 2005/08-2008/07 (µg/l)
DP068 DP070 DP073	**-0,103**	**0,238**	**-1,072**	**-0,052**
DP068 DP072 DP074	**-0,019**	**0,278**	**-2,120**	**-0,080**
DP068 DP070 DP072	**-0,103**	**0,285**	**-2,120**	**-0,059**
DP070 DP073 DP074	**1,328**	**0,564**	**3,232**	**1,316**
DP066 DP072 DP073	**1,844**	**-0,525**	**0,384**	**0,647**
DP066 DP070 DP073	**-0,146**	**-0,454**	**1,188**	**0,641**
DP068 DP073 DP074	**-0,019**	**0,221**	**-1,072**	**-0,059**
DP066 DP068 DP073	**-2,659**	**-0,764**	**-1,072**	**-1,611**
DP070 DP072 DP074	**2,084**	**0,910**	**3,234**	**1,377**
DP066 DP068 DP072	**-2,659**	**-0,762**	**-2,120**	**-1,984**
DP066 DP068 DP070	**-2,384**	**-0,760**	**-0,984**	**-1,351**
DP066 DP068 DP074	**-3,456**	**-0,908**	**-0,984**	**-1,715**
DP066 DP070 DP074	**-1,800**	**-0,760**	**1,284**	**0,138**
DP066 DP073 DP074	**-1,196**	**-0,764**	**1,042**	**0,595**
DP072 DP073 DP074	**2,536**	**1,904**	**2,874**	**2,022**
DP070 DP072 DP073	**1,866**	**0,898**	**3,232**	**1,385**
DP066 DP070 DP072	**-0,146**	**-0,380**	**0,488**	**0,517**
DP068 DP072	**-0,124**	**0,221**	**-2,956**	**-1,403**
DP070 DP074	**0,034**	**0,202**	**3,233**	**1,159**
DP072 DP073	**2,051**	**2,080**	**2,807**	**1,889**
DP070 DP073	**0,468**	**0,553**	**3,231**	**1,263**
DP068 DP073	**-0,124**	**0,164**	**-2,852**	**-0,402**
DP072 DP074	**3,008**	**1,904**	**2,538**	**2,431**
DP068 DP070	**-0,173**	**0,014**	**-2,476**	**-0,315**

Annex table: Comparing the qualitative (ecological status) and quantitative (90th percentile) results obtained by reduced number of sampling stations with results obtained by complete data. Red field – degradation of ecological status obtained by reduced data, green field –upgrade of ecological status, yellow field – preserve original ecological status for the all four periods.

EXPLORATORY STUDY RESULTS WATER BODY: 010

k-combination spatial reduction analysis	error to original 90th percentile period 2005/08-2006/07 (µg/l)	error to original 90th percentile period 2006/08-2007/07 (µg/l)	error to original 90th percentile period 2007/08-2008/07 (µg/l)	error to original 90th percentile period 2005/08-2008/07 (µg/l)
DP068 DP074	-0,997	-0,685	-2,476	-0,440
DP066 DP068	-4,106	-1,253	-2,476	-2,929
DP066 DP072	-4,236	-0,764	-0,105	-0,971
DP066 DP070	-3,830	-0,762	0,436	-0,359
DP066 DP074	-4,921	-0,931	0,436	-0,561
DP066 DP073	-4,236	-0,766	0,332	-0,367
DP070 DP072	0,684	0,832	3,233	1,322
DP073 DP074	1,910	0,874	2,807	1,818
DP070	-0,116	0,149	3,232	1,136
DP073	1,855	0,820		1,741
DP074		-0,430	2,508	2,003
DP066		-1,666	-0,160	-2,259
DP068	-2,382	-1,158	-3,904	-3,171
DP072	2,938	2,080	-1,454	2,437

Annex table: Comparing the qualitative (ecological status) and quantitative (90th percentile) results obtained by reduced number of sampling stations with results obtained by complete data. Red field – degradation of ecological status obtained by reduced data, green field –upgrade of ecological status, yellow field – preserve original ecological status for the all four periods.

EXPLORATORY STUDY RESULTS WATER BODY: 011

k-combination spatial reduction analysis	error to original 90th percentile period 2005/08-2006/07 (µg/l)	error to original 90th percentile period 2006/08-2007/07 (µg/l)	error to original 90th percentile period 2007/08-2008/07 (µg/l)	error to original 90th percentile period 2005/08-2008/07 (µg/l)
DP002a DP076 DP079 DP080 DP081	-0,080	0,000	-0,306	-0,079
DP002a DP076 DP078 DP079 DP081	-0,010	0,190	-0,051	0,208
DP002a DP076 DP078 DP079 DP080	-0,045	-0,136	-0,324	-0,151
DP002a DP078 DP079 DP080 DP081	0,070	-0,017	0,386	0,298
DP002a DP076 DP078 DP080 DP081	-0,005	-0,119	-0,192	-0,187
DP076 DP078 DP079 DP080 DP081	0,038	-0,136	-0,036	0,013
DP002a DP076 DP078 DP079	-0,052	0,195	-0,846	0,053
DP002a DP076 DP078 DP080	-0,120	-0,218	-0,414	-0,222
DP002a DP076 DP079 DP080	-0,320	-0,225	-0,714	-0,187
DP002a DP076 DP079 DP081	-0,120	0,216	-0,714	0,109
DP076 DP078 DP080 DP081	0,070	-0,218	-0,231	-0,133
DP002a DP076 DP078 DP081	-0,020	0,038	-0,336	-0,043
DP076 DP078 DP079 DP081	0,030	0,195	-0,108	0,223
DP002a DP076 DP080 DP081	-0,320	-0,200	-0,396	-0,259
DP002a DP079 DP080 DP081	-0,130	-0,144	0,397	0,041
DP076 DP079 DP080 DP081	-0,045	-0,225	-0,324	-0,006
DP002a DP078 DP079 DP080	0,026	-0,170	0,320	0,001
DP076 DP078 DP079 DP080	-0,010	-0,226	-0,342	-0,061
DP002a DP078 DP080 DP081	0,074	-0,153	0,331	-0,003
DP002a DP078 DP079 DP081	0,062	0,152	0,488	0,603
DP078 DP079 DP080 DP081	0,075	-0,170	0,500	0,473
DP076 DP078 DP079	-0,017	0,200	-0,978	0,123
DP076 DP078 DP080	-0,016	-0,485	-0,432	-0,197
DP076 DP078 DP081	0,055	0,057	-0,393	0,070

Annex table: Comparing the qualitative (ecological status) and quantitative (90th percentile) results obtained by reduced number of sampling stations with results obtained by complete data. Red field – degradation of ecological status obtained by reduced data, green field –upgrade of ecological status, yellow field – preserve original ecological status for the all four periods.

EXPLORATORY STUDY RESULTS WATER BODY: 011

k-combination spatial reduction analysis	error to original 90th percentile period 2005/08-2006/07 (µg/l)	error to original 90th percentile period 2006/08-2007/07 (µg/l)	error to original 90th percentile period 2007/08-2008/07 (µg/l)	error to original 90th percentile period 2005/08-2008/07 (µg/l)
DP076 DP079 DP080	-0,120	-0,626	-0,846	-0,187
DP076 DP080 DP081	-0,120	-0,579	-0,414	-0,187
DP078 DP079 DP081	0,130	0,171	0,685	0,741
DP078 DP079 DP080	0,071	-0,665	0,375	0,060
DP078 DP080 DP081	0,079	-0,610	0,430	0,002
DP079 DP080 DP081	0,022	-0,749	0,514	0,136
DP076 DP079 DP081	-0,052	0,240	-0,846	0,179
DP002a DP076 DP079	-0,360	0,240	-2,206	-0,187
DP002a DP080 DP081	-0,826	-0,720	0,342	-0,301
DP002a DP076 DP078	-0,160	0,057	-1,506	-0,187
DP002a DP078 DP081	0,115	0,000	0,348	0,585
DP002a DP076 DP080	-0,626	-0,579	-1,374	-0,491
DP002a DP078 DP079	0,104	0,171	0,320	0,583
DP002a DP078 DP080	-0,178	-0,610	0,320	-0,256
DP002a DP079 DP080	-0,850	-0,749	0,143	-0,133
DP002a DP079 DP081	-0,250	0,434	0,516	0,693
DP002a DP076 DP081	-0,360	0,024	-1,374	-0,187
DP078 DP081	0,190	0,019	0,600	0,719
DP078 DP080	0,092	-0,923	0,320	-0,145
DP002a DP080	-1,594	-0,981	-0,152	-0,635
DP079 DP080	-0,506	-1,010	0,150	-0,005
DP002a DP081	-0,954	0,000	0,376	0,602
DP078 DP079	0,294	0,190	0,635	0,603
DP076 DP078	0,004	-0,026	-1,638	-0,187

Annex table: Comparing the qualitative (ecological status) and quantitative (90th percentile) results obtained by reduced number of sampling stations with results obtained by complete data. Red field – degradation of ecological status obtained by reduced data, green field –upgrade of ecological status, yellow field – preserve original ecological status for the all four periods.

EXPLORATORY STUDY RESULTS WATER BODY: 011

k-combination spatial reduction analysis	error to original 90th percentile period 2005/08-2006/07 (µg/l)	error to original 90th percentile period 2006/08-2007/07 (µg/l)	error to original 90th percentile period 2007/08-2008/07 (µg/l)	error to original 90th percentile period 2005/08-2008/07 (µg/l)
DP002a DP078	-0,104	0,019	0,320	0,337
DP079 DP081	0,014	0,496	0,702	0,818
DP002a DP079	-0,970	0,496	-1,348	0,388
DP080 DP081	-0,442	-0,981	0,444	-0,091
DP002a DP076	-0,699	0,048	-3,296	-0,427
DP076 DP079	-0,160	0,264	-2,424	-0,091
DP076 DP081	-0,160	0,048	-1,506	-0,187
DP076 DP080	-0,360	-0,894	-1,506	-0,818
DP080	-1,976	-2,450	-0,200	-1,350
DP076	-0,400	-0,103	-3,514	-0,432
DP079				0,693
DP002a			-4,128	-0,459
DP078	0,372	-0,098		0,585
DP081			0,617	0,827

Annex table: Comparing the qualitative (ecological status) and quantitative (90th percentile) results obtained by reduced number of sampling stations with results obtained by complete data. Red field – degradation of ecological status obtained by reduced data, green field –upgrade of ecological status, yellow field – preserve original ecological status for the all four periods.

EXPLORATORY STUDY RESULTS WATER BODY: 014

k-combination spatial reduction analysis	error to original 90th percentile period 2005/08-2006/07 (µg/l)	error to original 90th percentile period 2006/08-2007/07 (µg/l)	error to original 90th percentile period 2007/08-2008/07 (µg/l)	error to original 90th percentile period 2005/08-2008/07 (µg/l)
DP096 DP098 DP100	**-0,011**	**0,000**	**-0,075**	**-0,010**
DP096 DP098 DP102	**0,332**	**-0,132**	**-0,092**	**-0,007**
DP100 DP098 DP102	**-0,120**	**-0,012**	**-0,002**	**-0,012**
DP096 DP100 DP102	**-0,134**	**0,173**	**0,081**	**0,071**
DP100 DP102	**-0,428**	**0,111**	**0,126**	**0,037**
DP102 DP098	**0,144**	**-0,168**	**-0,103**	**-0,031**
DP100 DP098	**-0,438**	**-0,012**	**-0,088**	**-0,034**
DP096 DP102	**0,344**	**-0,934**	**0,004**	**0,092**
DP096 DP100	**-0,490**	**0,173**	**0,086**	**0,071**
DP096 DP098	**0,328**	**-0,132**	**-0,121**	**-0,039**
DP100	**-0,590**	**0,111**	**0,151**	**0,031**
DP096		**-0,916**	**-0,316**	**0,137**
DP102	**-0,032**	**-1,443**		**0,022**
DP098	**0,116**	**-0,168**	**-0,122**	**-0,117**

Annex table: Comparing the qualitative (ecological status) and quantitative (90th percentile) results obtained by reduced number of sampling stations with results obtained by complete data. Red field – degradation of ecological status obtained by reduced data, green field –upgrade of ecological status, yellow field – preserve original ecological status for the all four periods.

EXPLORATORY STUDY RESULTS WATER BODY: 019

k-combination spatial reduction analysis	error to original 90th percentile period 2005/08-2006/07 (µg/l)	error to original 90th percentile period 2006/08-2007/07 (µg/l)	error to original 90th percentile period 2007/08-2008/07 (µg/l)	error to original 90th percentile period 2005/08-2008/07 (µg/l)
DP125 DP127 DP129	0,457	-0,154	-0,011	0,000
DP129 DP127 DP131	-0,375	0,154	-0,588	-0,108
DP125 DP129 DP131	-0,243	-0,649	-0,705	-0,219
DP125 DP127 DP131	0,259	-0,154	0,083	0,610
DP125 DP127	1,123	-0,308	0,081	0,687
DP125 DP131	-0,092	-0,671	0,011	0,318
DP129 DP127	-0,066	0,154	-0,683	-0,036
DP131 DP127	-0,410	0,390	0,083	0,472
DP129 DP131	-0,830	-0,796	-1,249	-0,720
DP125 DP129	0,384	-0,671	-0,789	-0,153
DP125	1,047	-0,693	-0,011	0,610
DP127	0,363	0,304	0,081	0,728
DP129	-8,707	-1,176	-1,296	-0,512
DP131	-0,875	-1,566	-1,413	-0,985

Annex table: Comparing the qualitative (ecological status) and quantitative (90th percentile) results obtained by reduced number of sampling stations with results obtained by complete data. Red field – degradation of ecological status obtained by reduced data, green field –upgrade of ecological status, yellow field – preserve original ecological status for the all four periods.

ANNEX 2

Annex “Coastal monitoring network”

At this annex are presented Valencian coastal water bodies and the coastal monitoring network.

On the maps can be observed:

- Valencian coastal water bodies
- Actual coastal environmental monitoring network
- Position of the sampling stations in the each water body
- Quality number of the analyzed samples in each sampling station for the period August 2005 – July 2008

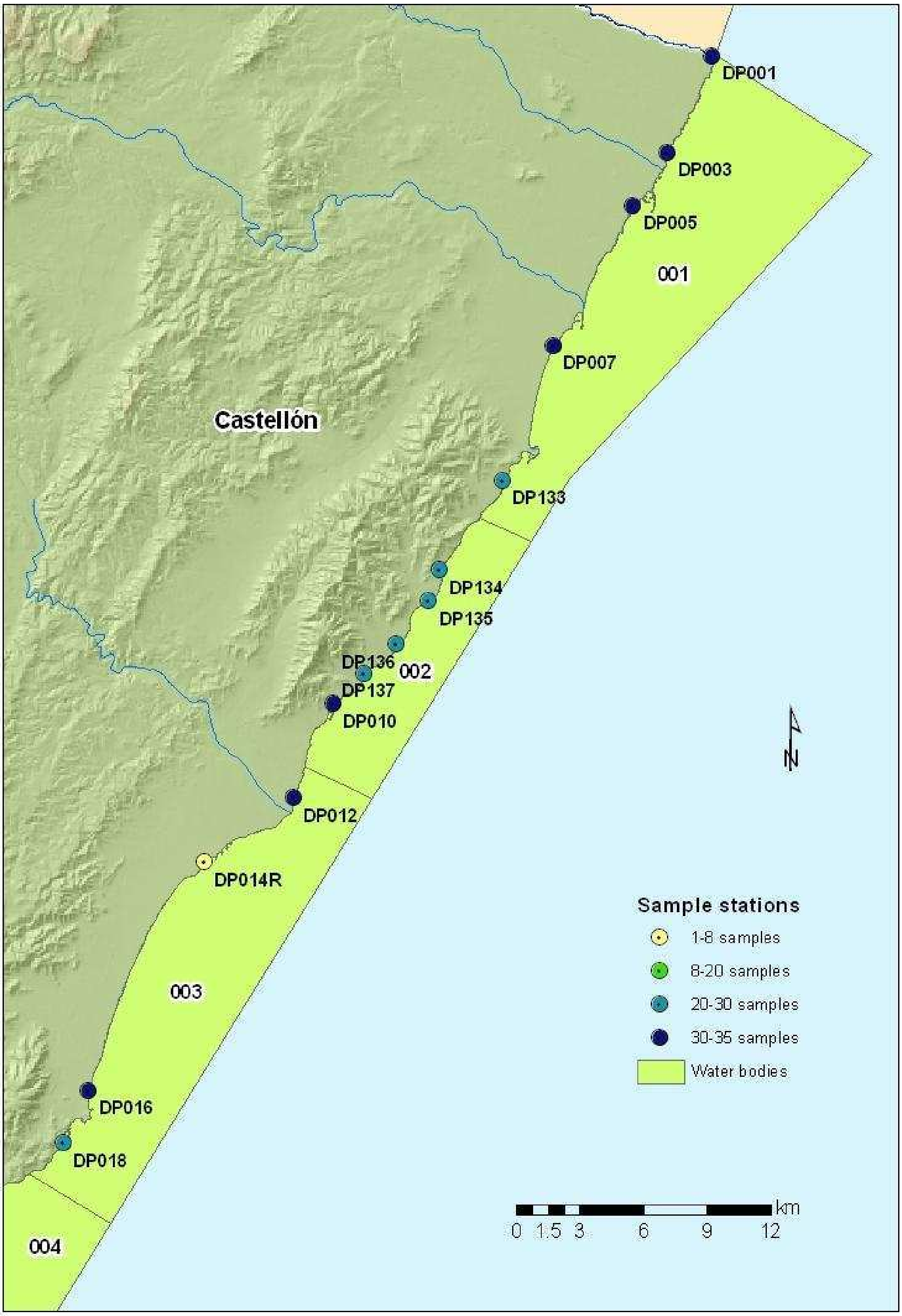


Figure 1: Coastal monitoring network; Water bodies 001, 002 and 003

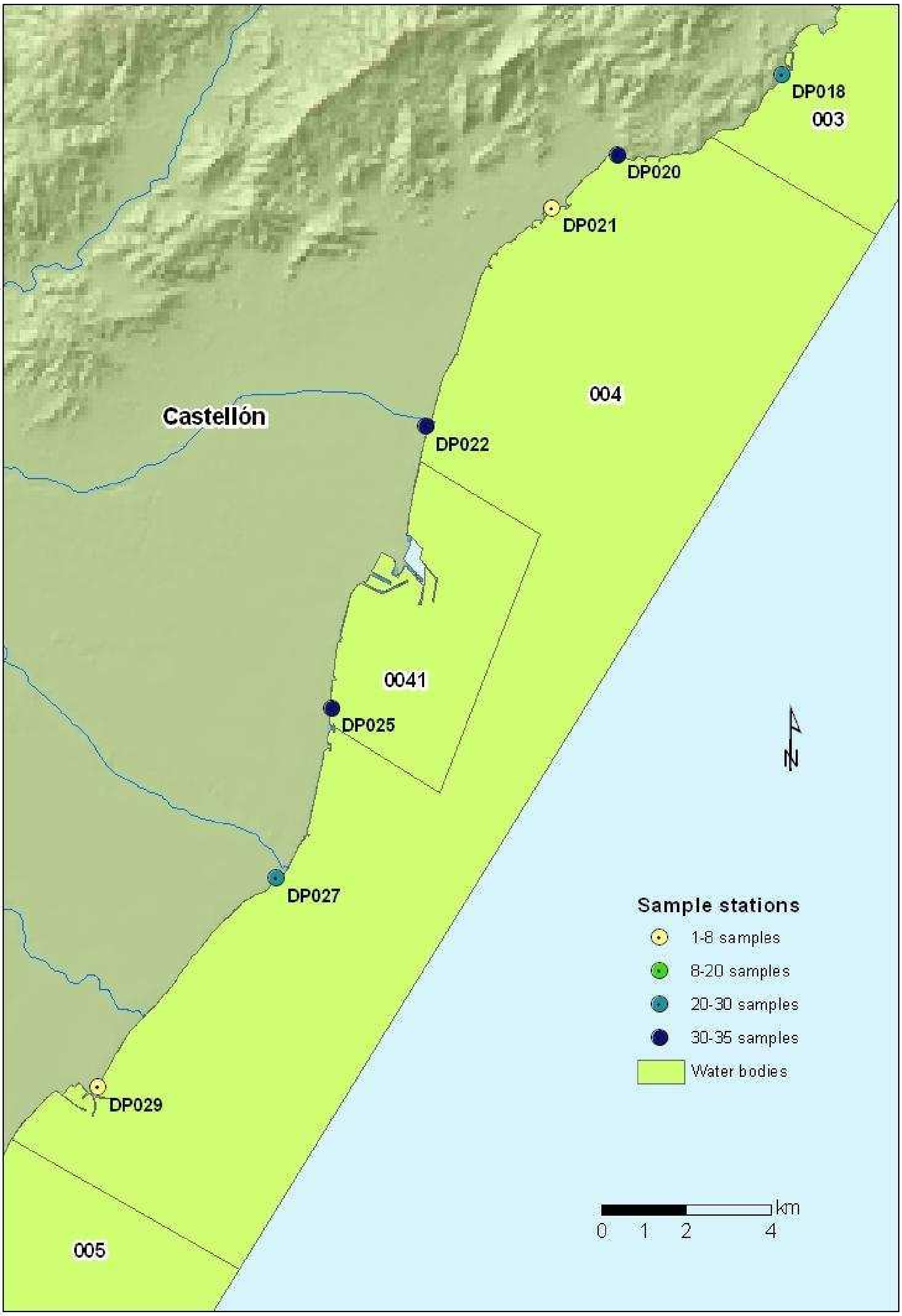


Figure 2: Coastal monitoring network; Water bodies 004, and 0041

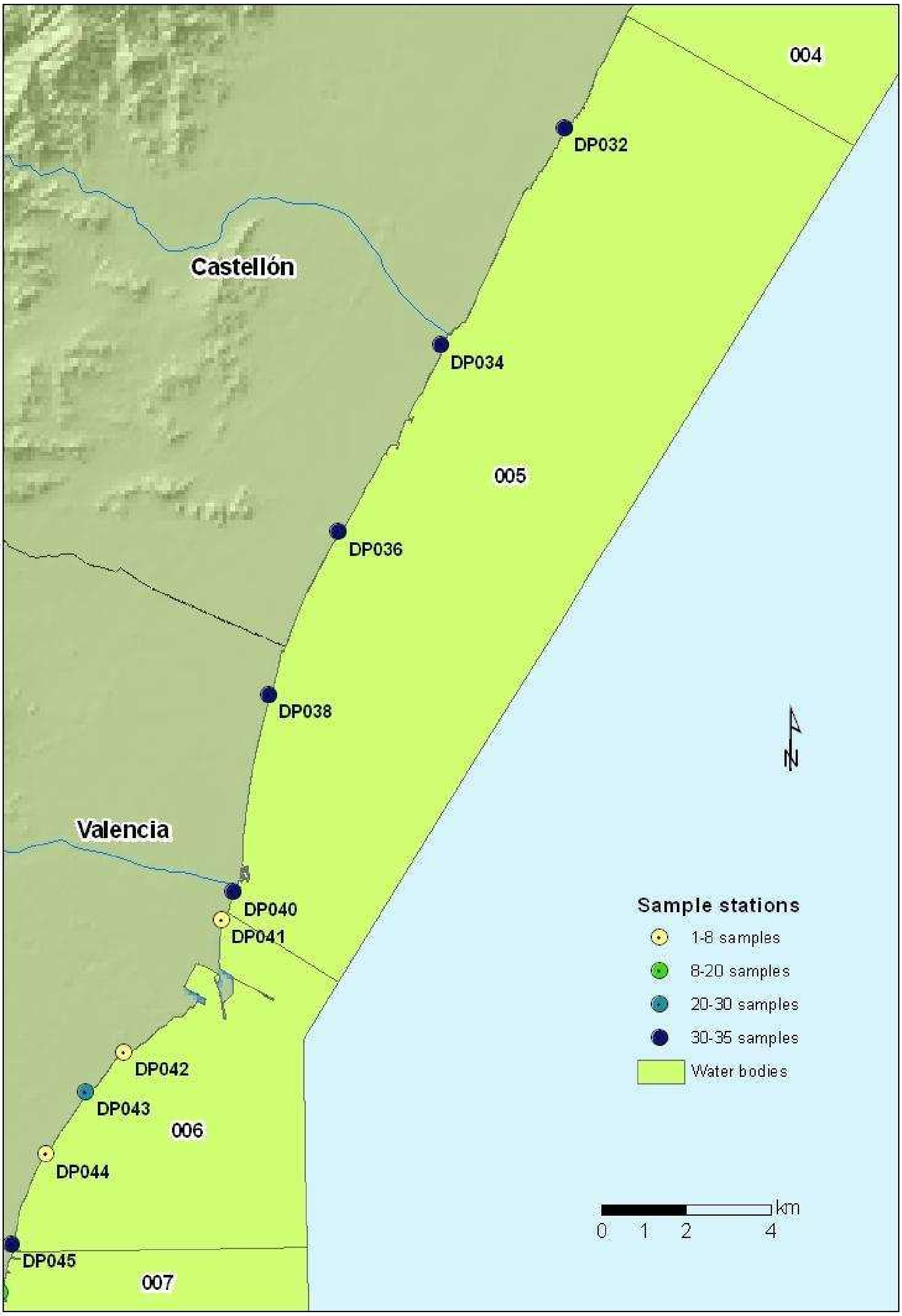


Figure 3: Coastal monitoring network; Water bodies 005 and 006

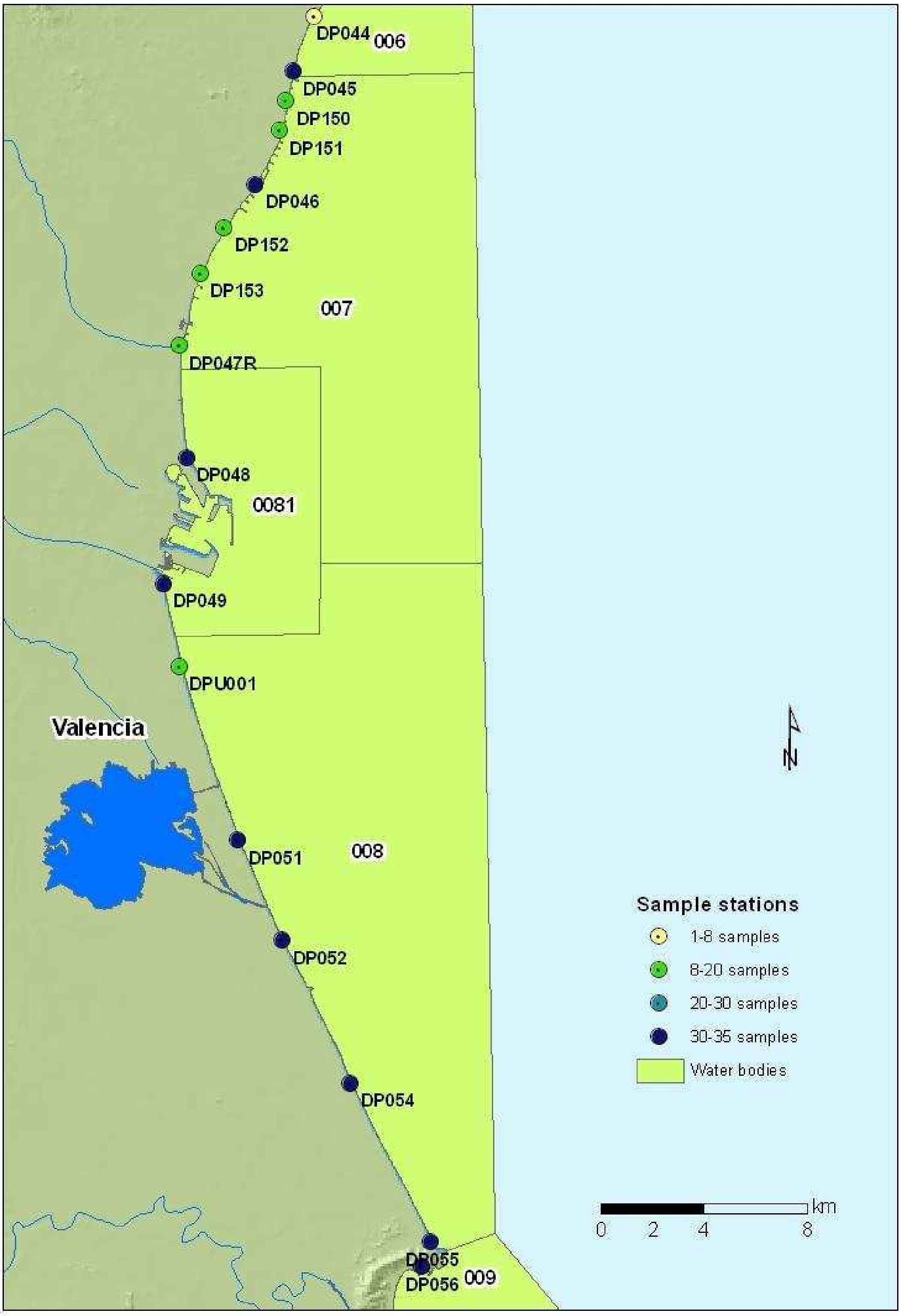


Figure 4: Coastal monitoring network; Water bodies 007, 0081 and 008

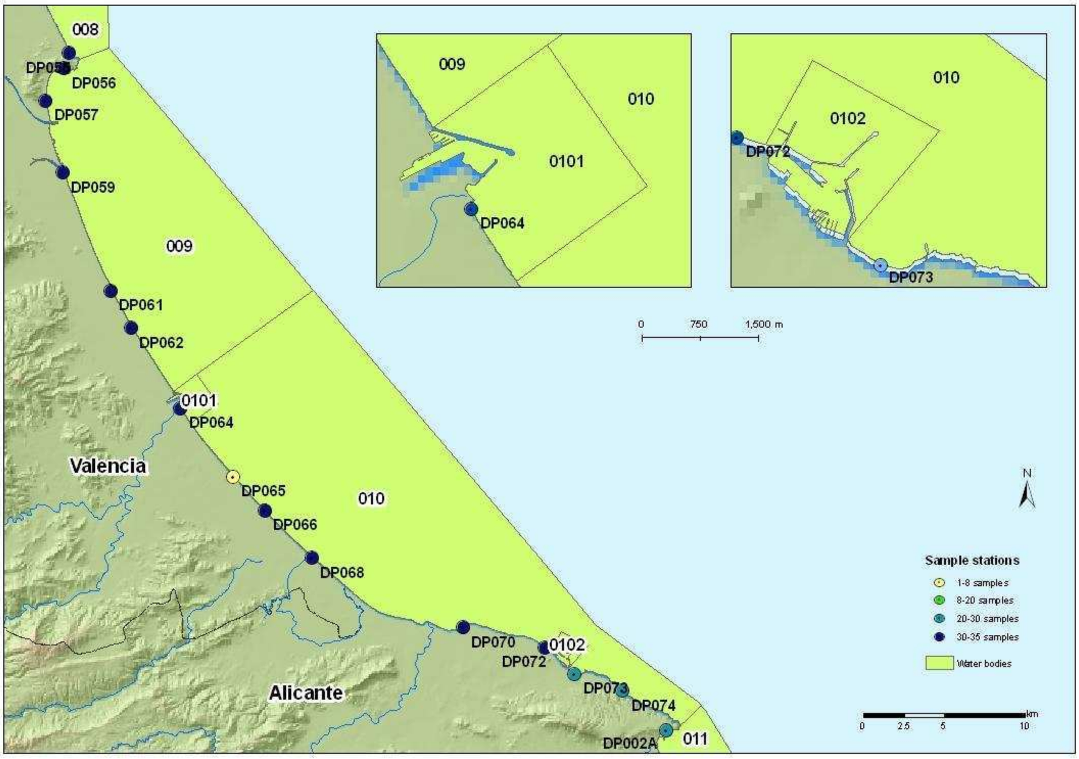


Figure 5: Coastal monitoring network; Water bodies 009, 0010, 0101 and 0102

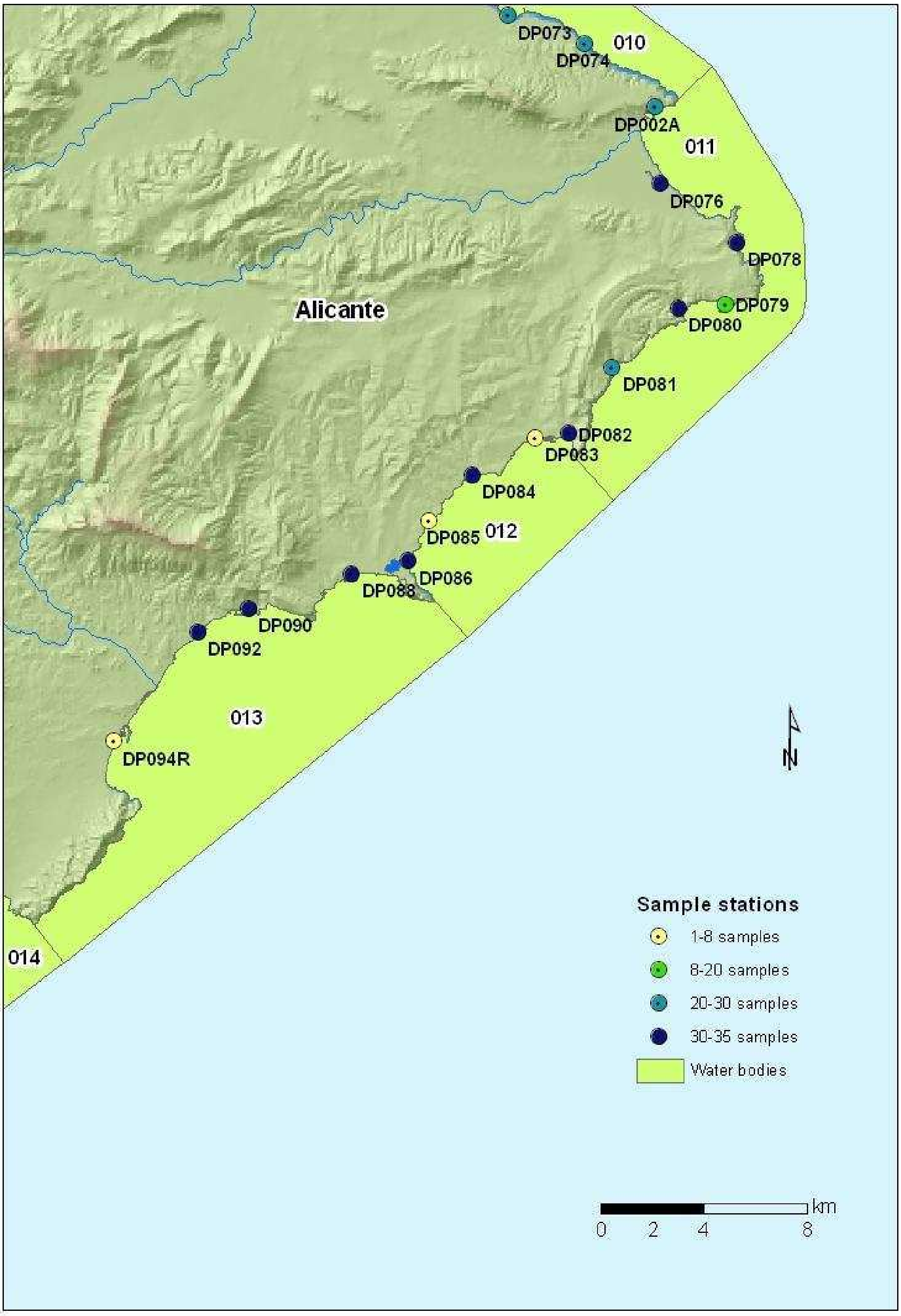


Figure 6: Coastal monitoring network; Water bodies 011, 012 and 013

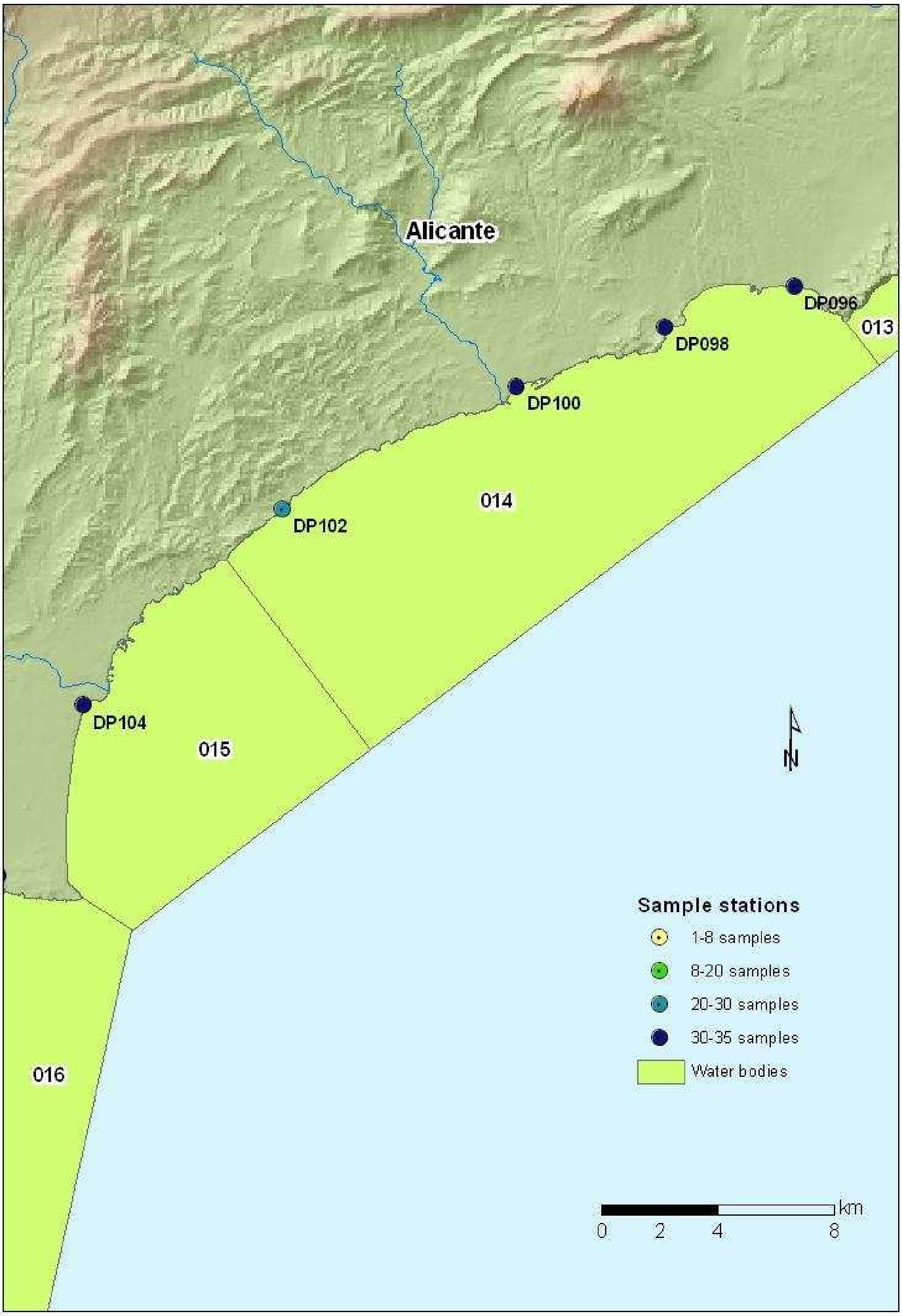


Figure 7: Coastal monitoring network; Water bodies 014 and 015

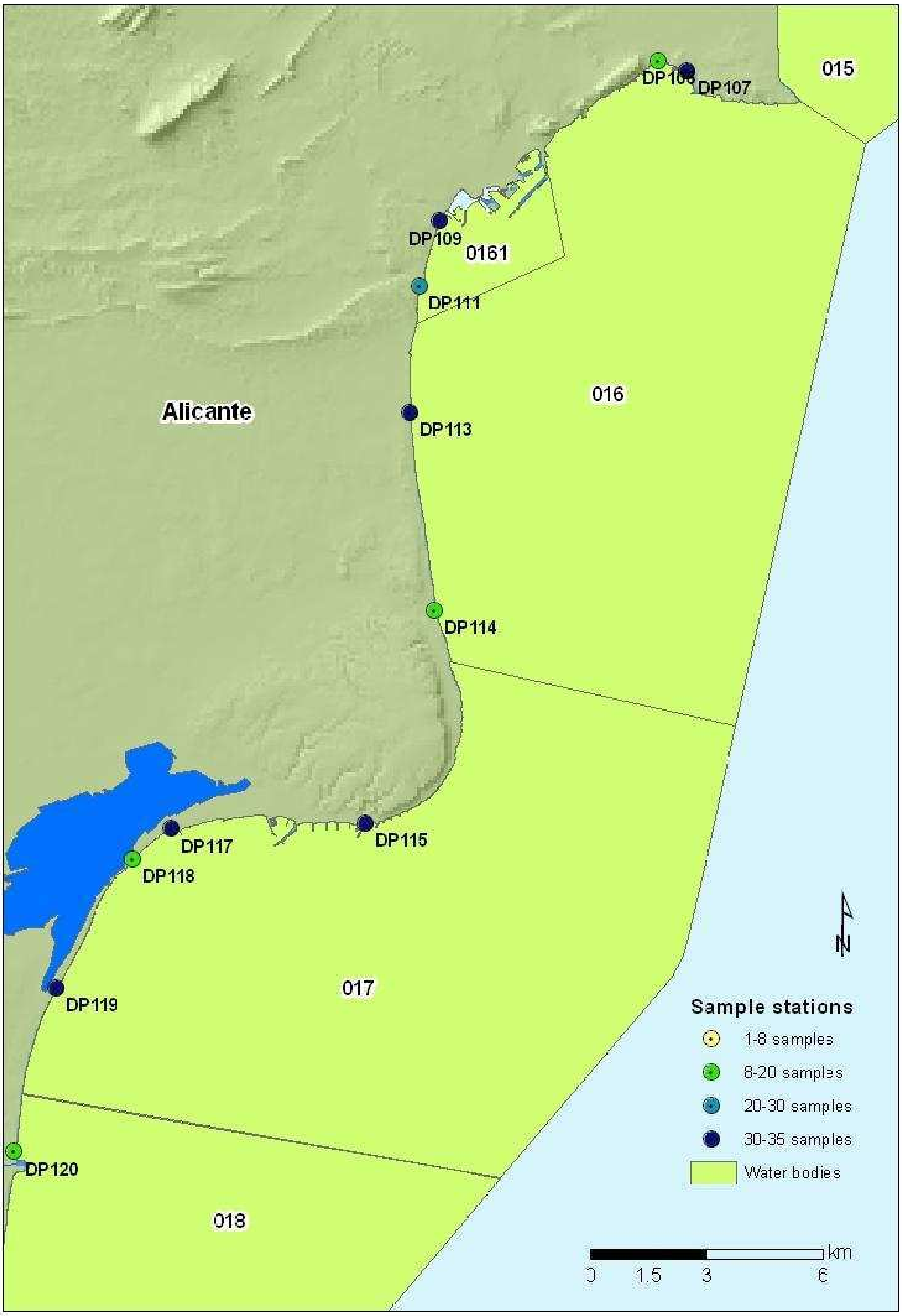


Figure 8: Coastal monitoring network; Water bodies 016, 0161 and 017

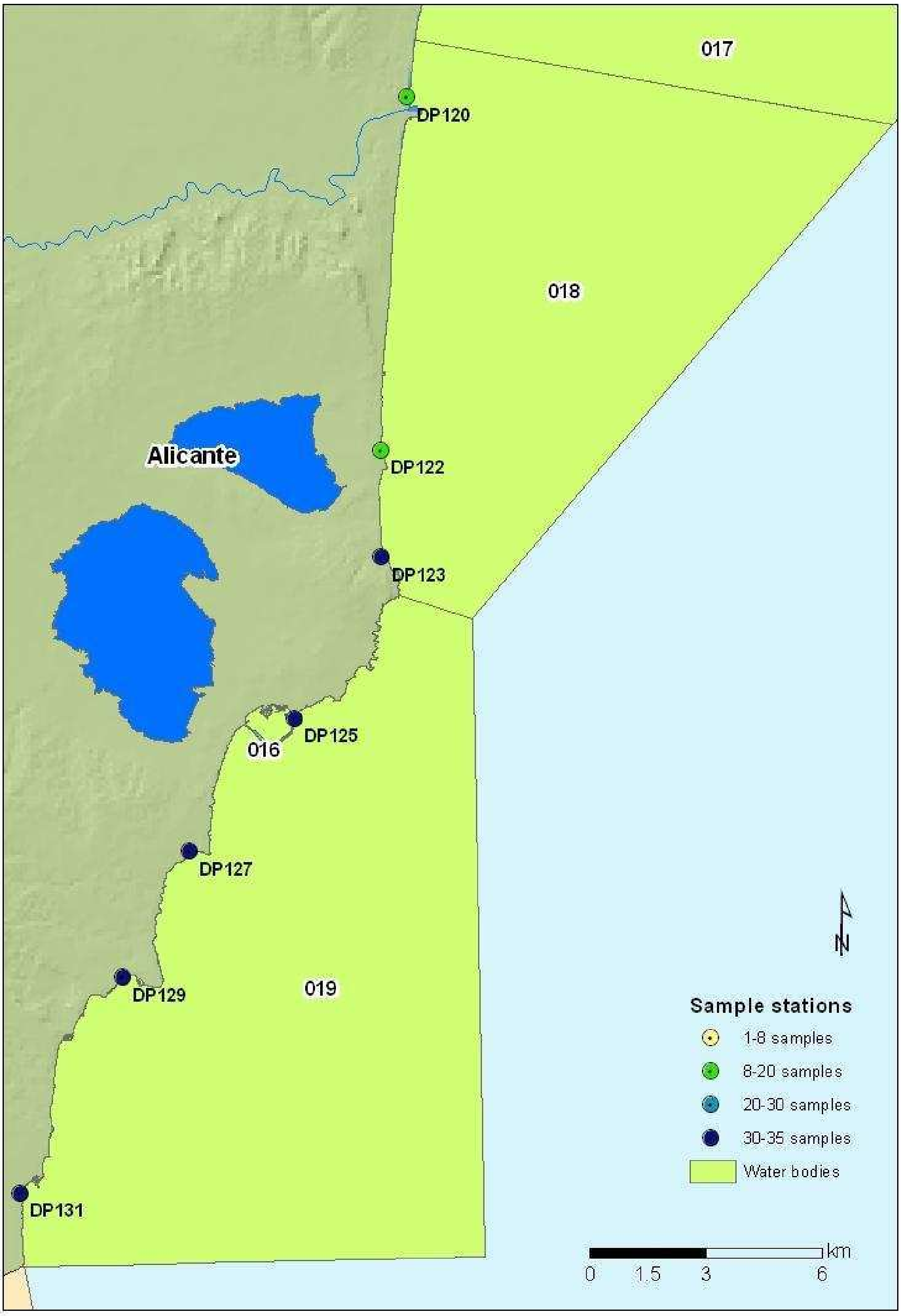


Figure 9: Coastal monitoring network; Water bodies 018 and 019