

Article

A Methodology for Discriminant Time Series Analysis Applied to Microclimate Monitoring of Fresco Paintings

Sandra Ramírez ^{1,2}, Manuel Zarzo ^{1,*}, Angel Perles ³ and Fernando-Juan García-Diego ^{4,*}

¹ Department of Applied Statistics, Operations Research and Quality, Universitat Politècnica de València, Camino de Vera, s/n 46022 Valencia, Spain; smramirez@javerianacali.edu.co

² Department of Natural Sciences and Mathematics, Pontificia Universidad Javeriana Cali, 760031 Cali, Colombia

³ ITACA Institute, Universitat Politècnica de València, Camino de Vera, s/n 46022 Valencia, Spain; aperles@disca.upv.es

⁴ Department of Applied Physics (U.D. Agriculture Engineering), Universitat Politècnica de València, Camino de Vera, s/n 46022 Valencia, Spain

* Correspondence: mazarcas@eio.upv.es (M.Z.); fgarcid@upv.es (F.-J.G.-D.); Tel.: +34-96-387-4900 (M.Z.)

Abstract: The famous Renaissance frescoes in Valencia's Cathedral (Spain) have been kept under confined temperature and relative humidity (RH) conditions for about 300 years, until the removal of the baroque vault covering them in 2006. In the interest of longer-term preservation and in order to maintain these frescoes in good condition, a unique monitoring system was implemented to record both air temperature and RH. Sensors were installed at different points at the vault of the apse during the restoration process. The present study proposes a statistical methodology for analyzing a subset of RH data recorded by the sensors in 2008 and 2010. This methodology is based on fitting different functions and models to the time series, in order to classify the different sensors. The methodology proposed, computes classification variables and applies a discriminant technique to them. The classification variables correspond to estimates of model parameters of and features such as mean and maximum, among others. These features are computed using values of functions such as spectral density, sample autocorrelation (sample ACF), sample partial autocorrelation (sample PACF), and moving range (MR). The classification variables computed were structured as a matrix. Next, sparse partial least squares discriminant analysis (sPLS-DA) was applied in order to discriminate sensors according to their position in the vault. It was found that the classification of sensors derived from Seasonal ARIMA-TGARCH showed the best performance (i.e., lowest classification error rate). Based on these results, the methodology applied here could be useful for characterizing the differences in RH, measured at different positions in a historical building.

Keywords: ARIMA; art conservation; Holt–Winters; sensor diagnosis; sparse PLS-DA; TGARCH



Citation: Ramírez, S.; Zarzo, M.; Perles, A.; García-Diego, F.-J. A Methodology for Discriminant Time Series Analysis Applied to Microclimate Monitoring of Fresco Paintings. *Sensors* **2021**, *21*, 436. <https://doi.org/10.3390/s21020436>

Received: 18 November 2020

Accepted: 31 December 2020

Published: 9 January 2021

Publisher's Note: MDPI stays neutral with regard to jurisdictional claims in published maps and institutional affiliations.



Copyright: © 2021 by the authors. Licensee MDPI, Basel, Switzerland. This article is an open access article distributed under the terms and conditions of the Creative Commons Attribution (CC BY) license (<https://creativecommons.org/licenses/by/4.0/>).

1. Introduction

Over the past 300 years, the famed Renaissance frescoes in Valencia's cathedral were kept under confined conditions because they were covered by a baroque vault. However, this vault was removed in 2006 [1]. In the interest of longer-term preservation and in order to maintain these frescoes in good condition, a monitoring system was implemented to record both air temperature and relative humidity (RH). Sensors were located at different points in the apse vault. The approximate location of each sensor can be seen in Figure 1. The positions are: cornice \mathcal{C} , ribs \mathcal{R} , walls \mathcal{W} , and frescoes \mathcal{F} . Some sensors were inserted on the painting's surface itself. It is a unique system because sensors are rarely placed inside the frame or on the canvas of paintings. Details about the installation of probes in the frescoes can be seen in figures from [1,2]. A perspective of the upper part of the apse and terrace above the frescoes can be found in [1].

The system was intended to detect water entering from the roof at specific points, or excessive general humidity in the vault itself. Any indication of high levels of thermo-

hygrometric conditions would instigate corrective measures [2]. The data analysis carried out by Zarzo et al. [1] showed the advantages of using humidity sensors for the monitoring of frescoes, so as to maximize their protection and prevent deterioration. The microclimatic requirements for churches and cathedrals are similar to those of museums, which also contain valuable works of art [2]. As in the case of museums, the indoor thermo-hygrometric conditions should be maintained at optimal levels in order to conserve the artefacts. The risks constituted by ventilation systems, air-conditioning, central heating, and the presence of visitors should be assessed in order to prevent or slow down the process of deterioration. Ideally, the temperature of walls and their surfaces should be the same as the air in the immediate proximity because, otherwise, an airflow is generated along the wall surface that increments the aerodynamic deposition of airborne particles and wall soiling. Cultural heritage sites are subjected to climatic changes that put them at risk, which has been widely discussed in [3,4].



Figure 1. Approximate location of the probes and sensors at the apse vault of the cathedral of Valencia. Details of the installation of the probes and a scheme depicting the installation of probes in the frescoes can be seen in [1] and [2]. The image shows the position of the 29 probes for monitoring the relative humidity (RH) of the indoor atmosphere, displayed in different colors according to their position. Seven probes were located on the ribs (orange), two at the cornice (light orange), ten on the walls below the severies (purple), and ten probes on the frescoes (green).

The internal environment should be appropriate [5] because changes in air temperature and RH can affect the conservation of fresco paintings [6–8]. Different studies [9–12] have monitored thermo-hygrometric parameters inside museums in order to assess the potential risks related to temperature and humidity. Other authors, such as Camuffo et al. [8], have studied the interactions between the indoor atmosphere and walls supporting frescoes or mural paintings. Similar works have been carried out in churches [13–17]. Frasca et al. [13] performed a microclimatic monitoring of the historic church of Mogila Abbey to analyze the impact of the environmental parameters on the works of art. Among their results, they found that vulnerable objects were at a high risk of mechanical damage approximately 15% of the time. The main cause of the vulnerability was the RH variability.

The problems of deterioration due to high humidity identified in the Renaissance frescoes at the cathedral of Valencia were studied by Zarzo et al. [1]. The researchers suggested that these problems could be caused by the infiltration of rainwater through

the roof above the apse and, that maintenance or regular monitoring should therefore be conducted for the long-term preservation of the valuable frescoes. Bernardi et al. [18] studied the importance of waterproofing in the roof above frescoes in St. Stephan's church in Nessebar, Bulgaria.

In the same way, the European Standards [19–24] summarized in the standard [25] as well as Corgnati and Filippi [26] adopted the approach of the Italian Standard UNI 10829 (1999) for the monitoring, elaboration, and analysis of microclimatic data for the preservation of artefacts.

A big economic effort is being carried out by governments within the European Union to preserve artworks in museums. Several works have monitored the microclimate within museums to analyze its relationship with the degradation of materials from which works of art are made, for example, with the goal of preserving artwork and artefacts [27].

Concerning data analysis, García-Diego and Zarzo [2] used monthly principal components analysis (PCA) in their research for February, September, October, and November of 2007. Furthermore, Zarzo et al. [1] also fitted a PCA per year for the years 2007, 2008, and 2010. The resulting loading plots highlight the most relevant similarities and dissimilarities among sensors. Regarding RH recorded in 2007, researchers observed that the daily evolution versus time of the RH mean per hour (\overline{RH}) was rather parallel for all sensors. It was observed that sensors H, N, and R (inserted on the frescoes) recorded higher values of \overline{RH} than those installed on the walls. Interestingly, sensors H and R were located in the zone where there was a moisture problem after their installation. In 2008 and 2010, the correlation between \overline{RH} and the first principal component PC1 was very high (greater than 0.994) [1]. After computing the average of moving ranges with order 2 of RH (per hour H_{MV}, day DMV and month MMV), it was shown that PC2 for 2008 could be predicted as: $PC2 = 232.88 - 2.69\overline{RH} - 32.39DMV$. Regarding 2010, the estimated regression model was: $PC2 = 297.03 - 3.87\overline{RH} - 32.12DMV$. Based on the results, researchers concluded that PC1 could be interpreted as the yearly RH average, while PC2 provided basic information about daily mean variations. Furthermore, researchers detected an abnormal performance in one sensor that might correspond to a failure of the monitoring system [28] or a change in the microclimatic conditions surrounding that particular sensor. They also concluded that the use of humidity sensors and the interpretation of the first two principal components can be very useful when discussing the microclimatic air conditions surrounding fresco paintings. Hence, PCA is a powerful statistical method for characterizing the different performance among sensors of the same type, located at different positions [1]. The advantage of PCA for sensor diagnosis has also been reported by Dunia et al. [28] and Zhu et al. [29].

The present study re-analyzes time series of RH from sensors located at the apse vault of Valencia cathedral. The data sets used here correspond to subsets of the database used in the study conducted by Zarzo et al. [1]. The present work focuses on RH measurements recorded from 23 sensors in 2008 and from 20 sensors in 2010. These time series of RH do not contain missing values.

This research aims to bring forward a methodology for discriminating sensors according to their position. For this purpose, the approach applied in this study consists of three stages: (1) The different time series were divided according to climatic condition and changes of the slope and level of the time series; (2) Three methods (M1, M2, and M3) were applied to obtain the classification variables per part of the time series identified in stage 1; (3) Sparse partial least squares discriminant analysis sPLS-DA was applied three times (one per method) as a discriminant technique in order to classify sensors, by using the set of classification variables as predictors.

The methodology proposed in this research is new in the context of time series clustering, as well as in sensor classification, when applied with the aim of conserving works of art. This methodology is unique because it uses a Seasonal ARIMA-TGARCH model to extract information from the time series, for discrimination purposes. It is also singular because it employs sPLS-DA in order to classify the time series.

According to the results of this study, sPLS-DA together with ARIMA-TGARCH-Student has a high capability of classifying time series with very similar characteristics, which often occurs in museums or similar buildings. The proposed methodology is well-suited to monitoring the sensors in this type of building.

This approach can be very useful in defining how microclimatic measurements should be carried out for monitoring conditions in heritage buildings or similar sites. Furthermore, the methodology could be useful for reducing the number of sensors required to monitor the microclimate. In summary, this approach could help to better manage the preventive conservation of cultural heritage sites.

2. Materials and Methods

2.1. Materials: Description of the Data Sets

Regarding the frescoes in Valencia's cathedral, 29 probes were implemented to monitor the indoor air conditions. Each probe contains an integrated circuit model DS2438 (Maxim Integrated Products, Inc., San Jose, CA, USA) that incorporates an analogue-to-digital voltage converter. This converter measures the output voltage of a humidity sensor (HIH-4000, Honeywell International, Inc.) and a temperature sensor. The recorded values of RH have an accuracy of $\pm 3.5\%$. Details of the probes, RH sensors, functions of calibration, and their installation in the apse vault are described elsewhere [1,2]. Seven probes were placed on the ribs (\mathcal{R}), two at the cornice (\mathcal{C}), ten on the walls below the severies (\mathcal{W}), and ten on the frescoes (\mathcal{F}) (see Figure 1).

The data sets used here do not contain missing values and correspond to subsets of the database used by Zarzo et al. [1]. The electronics platform Arduino was used <https://www.arduino.cc/en/Guide/Introduction>. Such data sets correspond to the mean RH per hour or day (\mathbf{RH}_h or \mathbf{RH}_d), where RH_{h_i} is the average of measurements per hour, while RH_{d_i} corresponds to the average of measurements per day.

The RH datasets correspond to those sensors located in the cornice \mathcal{C} , ribs \mathcal{R} , walls \mathcal{W} , and frescoes \mathcal{F} . As the statistical analysis was performed separately for each season, sensors \mathcal{M} in \mathcal{W} and $\tilde{\mathcal{N}}$ at the \mathcal{R} were discarded because it was necessary to deal with time series comprising a time period of at least 300 observations without missing values. Thus, time series are available in 2008 for 23 sensors: two on the cornice (A and B), five at the ribs (C, D, I, J, and X), nine on the frescoes (E, H, K, O, R, T, W, Y, and AB), and seven on the walls (G, L, P, U, V, Z, and AA). In 2010, information from sensors H, Y, AB, G, and Z was not available, but there were two additional sensors (S and Q) located at the \mathcal{C} and \mathcal{W} , respectively. Hence, 20 sensors could be used for 2010: 8 at position \mathcal{RC} (\mathcal{R} or \mathcal{C}), 6 on the walls (\mathcal{W}), and 6 on the frescoes (\mathcal{F}) (see Figure 1).

For both years, the follow-up time spanned the seasons of winter W_r , spring Sp , and summer Sm . The data set \mathbf{RH}_h for 2008 consists of 3851 observations: 1430 for W_r , 2099 for Sp , and 322 for Sm . Regarding 2010, 3414 observations are available: 636 for W_r , 2178 for Sp , and 600 for Sm .

Autumn was not considered because the number of observations was not high enough according to the established conditions of the study (i.e., at least 300 observations). Data sets correspond to the periods between 15 January and 4 July 2008, and from 22 February to 18 July 2010. Sensors from the selected times—both in 2008 and 2010—did not experience electronic malfunction. In 2008, there was no evidence of conservation problems in the frescoes where these sensors were located. By contrast, in 2010 there was evidence of salt efflorescence found in the same zones, before the first restoration works in the apse vault [1].

The periods corresponding to each season were defined as follows: spring was considered as being between 19 March and 20 June, summer between 20 June and 22 September, and winter between 22 September and 22 January.

2.2. Statistical Methods

Three different methods (M) were applied to the RH data in order to extract estimates of parameters and features used subsequently as classification variables. These methods

consist of fitting the time series to different statistical models or functions. M1 included functions such as spectral density, sample autocorrelation function (ACF), sample partial ACF (PACF), and moving range (MR) [30–33]; M2 was the Additive Seasonal Holt–Winters method (Additive SH-W) [34,35]; and M3 was a Seasonal ARIMA with threshold generalized autoregressive conditional heteroskedastic (TGARCH) model considering the Student distribution for residuals (Seasonal ARIMA-TGARCH-Student) [36–39].

The three methods were carried out separately for various seasons of the year (Wr, Sp, and Sm) for both 2008 and 2010. Once the classification variables were computed from the three methods, sparse partial least square discriminant analysis (sPLS-DA) [40] was applied to classification variables three times—one time per method—using all classification variables per season. sPLS-DA was used to discriminate between sensors according to their three possible positions in the vault: \mathcal{R} (\mathcal{R} or \mathcal{C}), \mathcal{W} , and \mathcal{F} . The classes \mathcal{R} and \mathcal{C} were joined as a new class called \mathcal{RC} in order to ensure a similar number of sensors per group.

The statistical methodology applied consisted of different steps: Firstly, the identification of structural breaks in the time series (Section 2.2.1), which leads to the establishment of periods where the analyses were carried out. Secondly, calculation of classification variables using M1 (Section 2.2.2). Thirdly, the calculation of classification variables using M2 (Additive SH-W; Section 2.2.3). Fourthly, the calculation of classification variables using M3 (Seasonal ARIMA-TGARCH-Student; Section 2.2.4). Finally, sensor classification was done by means of sPLS-DA (Section 2.2.5).

R software [41] version 4.03 was used to carry out the analyses. The main R packages used were `aTSA` [42], `forecast` [43,44], `mixOmics` [45,46], `rugarch` [37,47], `strucchange` [48], `tseries` [49], and `QuantTools` [33].

2.2.1. Identification of Structural Breaks in the Time Series

Many time series models (e.g., ARMA [36], ARCH, and GARCH [37]) assume the lack of sudden changes due to external factors that might appear occasionally. However, when analyzing time series of real situations, it can be found that external factors produce dramatic shifts such as a change in the slope of a linear trend, which cannot be properly modeled. Such occasional events are known as structural breaks [38]. In order to detect such events, different tests can be applied from the generalized fluctuation test framework (e.g., CUSUM and MOSUM), which are based on empirical fluctuation processes [50]. Others like the Chow [51] test are based on checking sequences of F statistics [52–54], while the supF test [55] consists of applying the former at all possible structural breaks. The null hypothesis is “no structural change”, versus the alternative: “the vector of coefficients varies over time” [55].

By visually inspecting the evolution versus time of \mathbf{RH}_h for both 2008 and 2010 (see Figure 2), potential structural breaks were identified in at least two points. Their significance was assessed by means of the CUSUM and supF tests. Both tests were carried out with the logarithmic transformation [56], that is, $r_t = \ln(\mathbf{RH}_{h_t})$, which has been used in other works to stabilize the data variance [38,56,57]. It was observed that most of the daily time series undergo seasonal trends, which makes it necessary to apply regular differentiation (i.e., $w_t = r_t - r_{t-1}$) in order to remove any trend [56], which is a common pretreatment in time series analysis. In this paper, \mathbf{r} refers to transformed values using the logarithmic function, while \mathbf{W} refers to data that was subjected to a logarithmic transformation and one regular differentiation. Figure 3 displays the plot of the time series of RH from sensor Y (2008). Additionally, this figure shows the plots of the time series of the logarithmic transformation of RH as well as one regular differentiation of the previous time series.

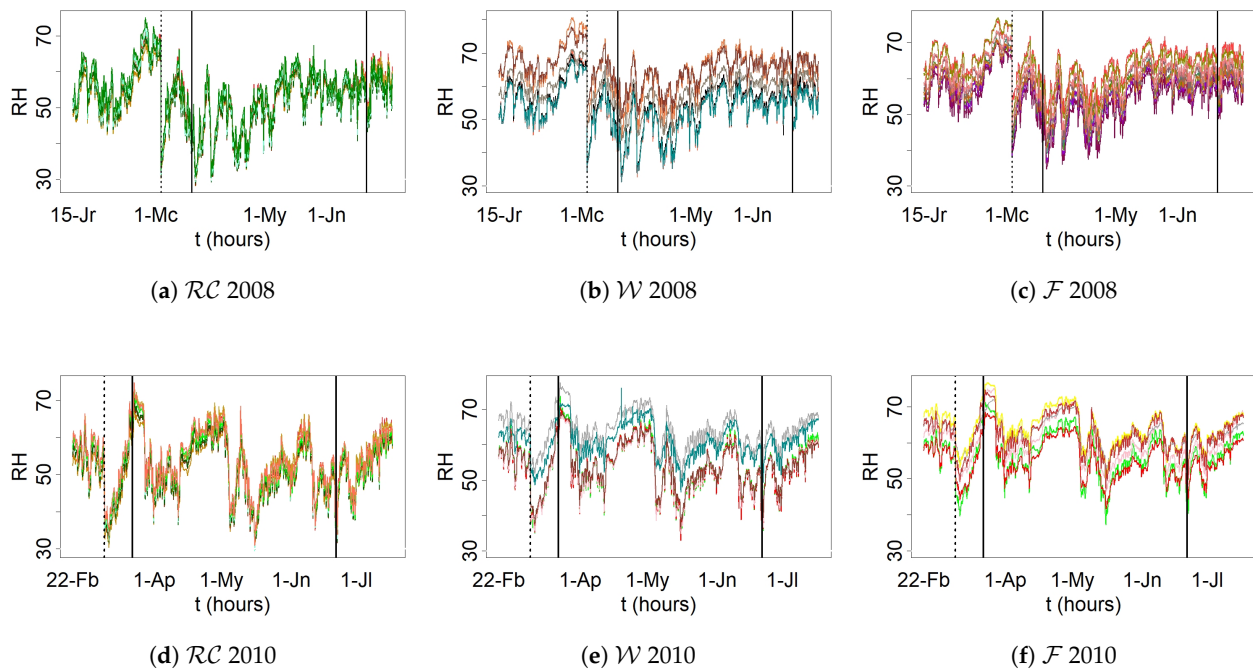


Figure 2. Evolution of RH_h . Trajectories of sensors located at equivalent positions in the apse vault are depicted in the same chart (data recorded between January 15 and July 4 2008): cornice and ribs (\mathcal{RC}) (a), walls (\mathcal{W}) (b), and frescoes (\mathcal{F}) (c). Likewise for data collected between February 22 and July 18 2010: \mathcal{RC} (d), \mathcal{W} (e), and \mathcal{F} (f). Separation by seasons (Wr, Sp and Sm) is indicated by means of vertical solid lines. Wr is divided into two periods (dashed line) because a structural break was identified according to the *supF* and *CUSUM* tests.

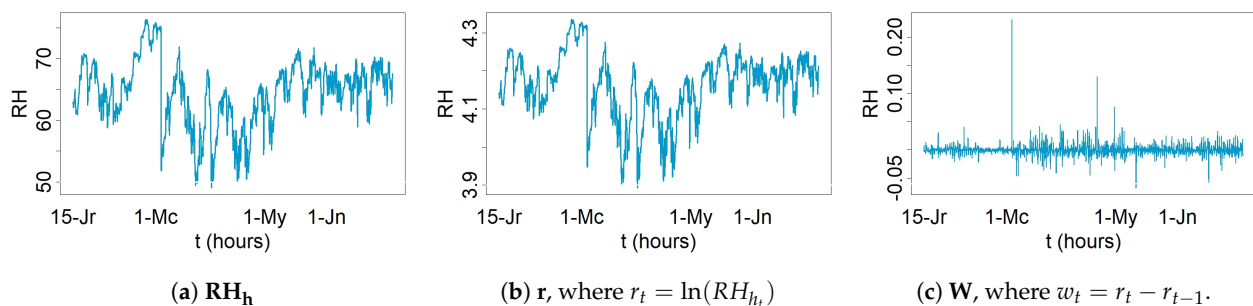


Figure 3. (a) Observed time series of RH from sensor Y (2008), (b) logarithmic transformation of the time series of (a), (c) one regular differentiation of the time series of (b).

The *supF* and *CUSUM* tests were applied to six groups: Wr 2008 (group 1, $n = 1429$), Sp 2008 (group 2, $n = 2098$), Sm 2008 (group 3, $n = 321$), Wr 2010 (group 4, $n = 635$), Sp 2010 (group 5, $n = 2177$), and Sm 2010 (group 6, $n = 559$).

According to the *supF* test, a structural break was identified in Wr 2008, after the 1058th observation (March 27 at 7:00 a.m., p -value = 0.01). Another break was found in 2010 at the 338th observation (March 8 at 1:00 p.m., p -value = 0.04). The *CUSUM* chart identified a significant shift at the same instant of time (1058th value in 2008 and 338th observation in 2010). The main reason for structural breaks could have been the strong changes of RH that occur in Valencia.

Ignoring structural breaks can lead to negative implications such as inconsistency of the parameter estimates and forecast failures [58]. Accordingly, for each structural break, it was decided to fit one model before this event, and another one after the structural break. On the other hand, in congruence with the physical characteristics of the data,

it might be convenient to split the statistical analysis per season and year. According to both considerations, the analysis was carried out separately in four periods, denoted as WrA, WrB, Sp, and Sm. WrA corresponds to the winter period before the structural break, while WrA refers to the following period (see Figure 2).

2.2.2. Calculation of Classification Variables—Method M1

This method is based on features using estimates of ACF (ρ_l at lag l), PACF (α_l at lag l), as well as features using mean (μ) and moving range (MR). Furthermore, features from spectral density were used, which was estimated using the periodogram ($I(w)$) of signals w . These features can help to characterize and reveal interesting properties of the underlying stochastic process without using any specific parametric model. Figure 4 shows a summary of the steps of M1.

RH values were used for estimating PACF, mean, and MR. By contrast, logarithm transformation and regular differencing were applied before estimating ACF and spectral density in order to stabilize the variances and remove the trend (i.e., **W** was employed). The objective of using both the ACF and spectral density with **W** was to focus on the seasonal component of the time series. These functions are briefly explained below:

Firstly, the mean of \mathbf{RH}_h was estimated for each period because this variable appeared as important for discrimination purposes in the preliminary study [1,2].

Secondly, MR with order n correspond to range values over n past values [33]. This function was applied to \mathbf{RH}_h and \mathbf{RH}_d . For each period, the mean and variance were computed for all MR values with order 2. These variables were calculated in order to estimate HMV and DMV, which were used in the preliminary investigations of this project [1,2]. However, MMV (i.e., MR of order 2 for \mathbf{RH}_m , estimated with the average RH per month) could not be calculated in this research because the number of observations per month was too low. For \mathbf{RH}_h , the mean of MR ($\hat{\mu}_{MR}$) corresponds to HMV and for \mathbf{RH}_d , the mean of MR is represented by DMV.

Thirdly, spectral density was estimated by means of the periodogram, which was calculated on the log scale using a spectrum function [30]. The periodogram displays information about the strengths of the various frequencies for explaining the seasonal components of a time series. The maximum peaks of spectral density and their corresponding frequencies were identified [30]. These functions were applied to \mathbf{RH}_h .

Finally, an estimation of ACF at lag l is the correlation (quantified by means of Pearson's correlation coefficient) between the values of a given time series, with the lagged values of the same time series at l time steps (l refers to lags) [30]. **W** values were used for this calculation. The values of sample ACF for the lags from 1 to 72 were used as classification variables because they showed greater variations, while for further lags they displayed lower values close to zero, comprised between the limits of a 95% confidence interval in the ACF correlogram.

Regarding sample PACF, according to Cowpertwait and Metcalfe [59], "the partial autocorrelation at lag l is the correlation that results after removing the effect of any correlations due to the terms at shorter lags". Sample ACF and sample PACF plots are commonly used in time series analysis and forecasting (e.g., autoregressive moving average (ARMA) models and their particular cases such as autoregressive (AR) and moving average (MA) models [59]). These plots, also called correlograms, illustrate the strength of a relationship between the values observed at a certain instant of time with those recorded in previous moments (with lag l) in the same time series. If sample ACF values decline exponentially and there are spikes in the first or more lags of sample PACF values, the time series can be modeled as an AR process. If sample PACF values decline exponentially and there are spikes in the first or more lags of the sample ACF values, the time series can be modeled as an MA process. If both sample PACF and sample ACF values decline exponentially, the time series can be modeled as a mixed ARMA process [59]. Sample PACF was computed for \mathbf{RH}_h values. The sample PACFs for the first four lags were calculated

for each period and were regarded as classification variables because they are usually the most important ones for capturing the relevant information in time series.

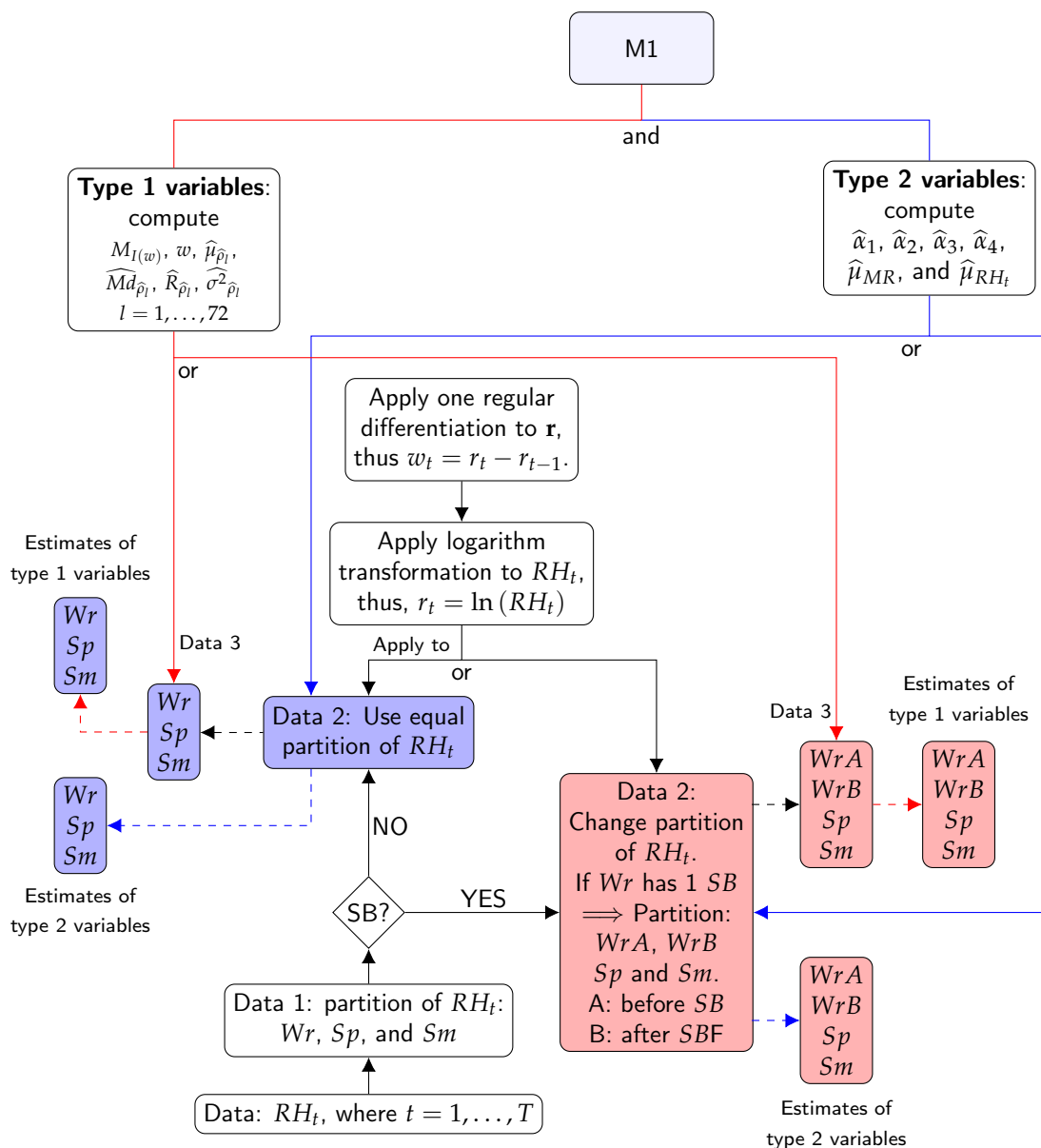


Figure 4. Flow chart for the steps of method 1: Blue lines indicate type 2 variables. Red lines indicate type 1 variables. Solid lines indicate processes. Dashed lines indicate results. The first step consists of dividing the different time series according to the climatic conditions: W_r , S_p , and S_m (Data 1). The second step consists of dividing the time series (Data 1) according to possible structural breaks (SBs) (Data 2). The third step applies a logarithm transformation and one regular differentiation to Data 2. The result is Data 3. The fourth step consists of applying the formulas of type 2 variables to Data 2. This is the first result. The fifth stage is carried out by applying the formula of type 1 variables to Data 3. The outcome produced is result 2. Different boxes contain symbols such as W_r , S_p , and S_m (or W_rA , W_rB , S_p , and S_m). They indicate that the results were computed for all different parts of the time series.

The features computed using the values of \mathbf{RH} were called type 1 variables and features calculated using values of \mathbf{W} were referred to as type 2 variables. The list of type 1 variables resulting from M1 are the estimates of the following parameters:

- Mean of \mathbf{RH}_h ($\hat{\mu}_{RH}$).
- Mean of MR ($\hat{\mu}_{MR}$) of order 2 for \mathbf{RH}_d and \mathbf{RH}_h .

- Variance of MR ($\hat{\sigma}_{MR}^2$) of order 2 for \mathbf{RH}_d and \mathbf{RH}_h .
- PACF for the first four lags ($\hat{\alpha}_1, \hat{\alpha}_2, \hat{\alpha}_3$, and $\hat{\alpha}_4$).

The list of type 2 variables resulting from M1 are the estimates of the following parameters:

- Maximum of spectral density ($M_{I(w)}$) and frequency corresponding to the maximum (w).
- Mean ($\hat{\mu}_{\hat{\rho}_1}$), Median ($\widehat{Md}_{\hat{\rho}_1}$), range ($\widehat{R}_{\hat{\rho}_1}$), and variance of the sample ACF ($\hat{\sigma}_{\hat{\rho}_1}^2$) for the first 72 lags.

2.2.3. Calculation of Classification Variables—Method M2: Additive SH-W

Winters [35] extended Holt's method [34] for capturing the seasonality of a time series. Hence, it was called Holt–Winters (H-W), which is a particular method of exponential smoothing aimed at forecasting [60].

The Seasonal H-W approach (SH-W) is based on three smoothing equations: for the level, trend, and seasonality. The parameter S denotes the number of values per season, while three additional parameters capture the information at time t : a_t denotes the time series level, b_t is the slope, and s_t is the seasonal component [60]. There are two different SH-W methods, depending on whether seasonality is modeled additively or multiplicatively [60]. The Seasonal H-W approach (SH-W) is based on three smoothing equations: for the level, trend, and for seasonality. The parameter S denotes the number of values per season, and three additional parameters capture the information at time t : a_t denotes the time series level, b_t is the slope, and s_t is the seasonal component [60].

In this research, the Additive SH-W method was fitted to both time series of RH and to their logarithmic transformations, but it turned out that the best outcomes were obtained with the transformed data. The period, the number of observations per season, was considered as $S = 24$ (i.e., 24 hourly values per day). Although this method does not require a residual analysis, one was carried out in an attempt to extract further information. Autocorrelation within the time series appeared in at least 10 out of the 22 lags for over 80% of the Ljung–Box Q (LBQ) tests [61] applied. Furthermore, the Kolmogorov–Smirnov normality (KS normality) [62] and Shapiro–Wilk (SW) tests [57,63,64] rejected the hypothesis of normality for at least 80% of the cases applied. The KS normality test compares the empirical distribution function with the cumulative distribution function. The test statistic is the maximum difference between the observed and theoretical values (normality). The statistic of the KS normality test was used as a classification variable in order to gather information about the distribution pattern of residuals and quantify departure from a normal distribution. The SW test detects deviations from normality due to either skewness or kurtosis, or both. The statistic of the SW test was also employed as a classification variable in order to identify lack of normality in the residuals according to skewness and kurtosis. Furthermore, given that data sets in this study are seasonal with a period of 24 h, where 72 is the maximum of lags, this maximum value was also considered for estimation of the mean, median, range, and variance of the sample ACF for the residuals. Figure 5 shows a summary of the steps of M2.

The features computed from residuals of the SH-W method were called type 3 variables. The other classification variables corresponded to the estimates of the method's parameters. The list of classification variables resulting from M2 are the following:

- Estimates of the parameters of the SH-W method: trend (\hat{b}), level (\hat{a}), and seasonal components ($\hat{S}_1, \dots, \hat{S}_{24}$).
- type 3 variables: sum of squared estimate of errors (SSE), maximum of spectral density ($\widehat{M}_{I(w)}$), frequency corresponding to maximum of spectral density (\hat{w}), and the mean ($\hat{\mu}_{\hat{\rho}_k}$), median ($\widehat{Md}_{\hat{\rho}_k}$), range ($\widehat{R}_{\hat{\rho}_k}$), and variance ($\hat{\sigma}_{\hat{\rho}_k}^2$) of sample ACF for 72 lags. The statistic of the SW test ($shap.t$), and the statistic of the KS normality test ($kolg.t$) are also included in this list.

These patterns can be studied using the GARCH family of models. Two of the most important ones for capturing such changing conditional variance are the ARCH and generalized ARCH (GARCH) models developed by Engle [69] and later extended by Bollerslev [70]. Engle and Bollerslev [71] were pioneers in the area of volatility modeling by introducing ARCH and, subsequently, GARCH models, which provide motion dynamics for the dependency in the conditional time variation of the distributional parameters of the mean and variance.

In recent years, different studies have applied hybrid forecasting models in various fields, and have shown a good performance for rainfall data [72], for the price of gold [73], for forecasting daily load patterns of energy [74], and for stock market prices [75].

According to Ghalanos [37], the family of GARCH models is broad, including the standard, integrated, and exponential models, as well as the GJR-GARCH, the asymmetric power ARCH, and the threshold GARCH (TGARCH) of Zakoian [76]. They capture the asymmetry of occasional impacts as well as abnormal distributions to account for the skewness and excess kurtosis. For GARCH models, error terms can sometimes be assumed from the Student distribution [77]. Bollerslev [78] described the GARCH-Student model as an alternative to the normal distribution for fitting the standardized time series. In particular, in the TGARCH-Student(s, r) model, s is the number of GARCH parameters β_i ($i = 1, \dots, s$), r is the number of ARCH and rotation parameters α_j and η_{1j} , respectively ($j = 1, \dots, r$), while ω is the variance intercept parameter. Error terms ϵ_t are assumed to be a white noise sequence following a Student distribution with degrees of freedom v [37].

Thus, instead of considering the standard ARIMA approach, whose focus is the conditional mean, it seems convenient to use here a hybrid approach based on ARIMA and GARCH models which can simultaneously deal with both the conditional mean and variance [38].

Given that data sets in this study are seasonal, it is necessary to use Seasonal ARIMA models, which are capable of modeling a wide range of seasonal data. A Seasonal ARIMA(p, d, q)(P, D, Q) $_S$ model is characterized by additional seasonal terms: P is the number of seasonal AR (SAR) terms, D is the number of differences taken, Q is the number of seasonal MA (SMA) terms, and S is the number of observations per period ($S = 24$ in this study). Furthermore, Φ_i ($i = 1, \dots, P$) are the parameters of the SAR part of the model, Θ_j ($j = 1, \dots, Q$) are the parameters of the SMA part, and the ϵ_t are error terms, which are assumed to be a white noise sequence [38]. In particular, in the Seasonal ARIMA(p, d, q)(P, D, Q) $_S$ -TGARCH(s, r)-Student model, the errors ϵ_t from Seasonal ARIMA(p, d, q)(P, D, Q) $_S$ follow a TGARCH(s, r)-Student process of orders s and r , so that their error terms ϵ_t are assumed to be a white noise sequence following a Student distribution, with degrees of freedom v .

Two steps were considered for the application of a hybrid approach based on Seasonal ARIMA and GARCH models, as briefly explained below.

Firstly, the most successful Seasonal ARIMA (or ARIMA) model was selected and the residuals were computed. Next, the most successful GARCH model was applied to fit these residuals. The following steps were carried out:

- The condition of stationarity was checked, that is, whether the statistical characteristics of the time series were preserved across the time period. The null hypothesis was that mean and variance do not depend on time t and the covariance between observations RH_t and RH_{t+1} does not depend on t [38]. To examine this null hypothesis, the augmented Dickey–Fuller (ADF) [79] and LBQ tests were applied for 48 lags. Furthermore, the sample ACF and sample PACF plots were also used.
- Transformation and differencing: the logarithmic transformation and regular differentiation were applied to \mathbf{RH}_t data before fitting ARMA in order to transform nonstationary data into stationary data [59]. The criterion for determining the values of d (differencing) is explained in the next step. The logarithmic transformation was preferred over other transformations because the variability of a time series becomes more homogeneous using logarithmic transformation, which leads to better forecasts [80].

- Identification of the most appropriate values for (p, d, q) and (P, D, Q) . Sample ACF and sample PACF plots were used to identify the appropriate values of (p, q) . Furthermore, the *corrected Akaike information criterion (AICc)* [60] was useful for evaluating how well a model fits the data and determining the values of both (P, D, Q) and (p, d, q) , taking into account the restriction that d and D should be 0 or 1. The most successful model for each time series was chosen according to the lowest *AICc* value. The *AICc* values were compared for models with the same orders of differencing, that is, equal values of d and D .

Secondly, the maximum likelihood estimation (MLE) method was used for estimating the parameters of the Seasonal ARIMA (or ARIMA) [38]. The models chosen were statistically examined in order to ensure that the resulting residuals do not contain useful information for forecasting. For this purpose, different tests were applied to determine whether all conditions and model assumptions were fulfilled. The analysis of residuals was carried out as follows:

- The condition of white noise was checked. Error terms can be regarded as white noise if their mean is zero and the sequence is not autocorrelated [38]. In order to check this issue, the ADF and LBQ tests were applied to the residuals and their squared values for 48 lags. Furthermore, the sample ACF plots were also used.
- To study the absence of Arch effects: for this purpose, the Lagrange multiplier test [69] and sample ACF plots [67] were applied to the residuals and their squared values [81].
- To check the distribution of residuals: by means of the Q–Q normal scores plots as well as the SW and KS normality tests.

The analysis of residuals revealed that error terms follow a GARCH process in all the different ARIMA models that were fitted. Therefore, it was necessary to fit a GARCH model to these residuals. The estimated model parameters were checked to determine if they were statistically significant, and their residuals were evaluated as described above. Finally, a hybrid model was fitted for each sensor and each period using the combined Seasonal ARIMA (or ARIMA) and TGARCH approach. After repeating the steps iteratively in order to select the most successful model, the normality tests applied to their residuals rejected the hypothesis of normality in all cases. Furthermore, all Q–Q normal scores plots showed that residuals were not falling close to the line at both extremes. Thus, a Student distribution was used to fit the residuals of the different TGARCH models.

For each period, a common model was applied to the hourly data of each sensor (one day corresponds to a sequence of 24 values).

- WrA (2008): seasonal ARIMA(1, 1, 0)(2, 0, 0)₂₄ – TGARCH(1,1)-Student.
- WrA (2010): ARIMA(1, 1, 2) – TGARCH(1,1)-Student.
- WrB (2008 and 2010): seasonal ARIMA(1, 1, 1)(2, 0, 0)₂₄ – TGARCH(1,1)-Student.
- Sp (2008 and 2010): seasonal ARIMA(1, 1, 2)(0, 0, 2)₂₄ – TGARCH(1,1)-Student.
- Sm (2008 and 2010): seasonal ARIMA(1, 1, 1)(1, 0, 0)₂₄ – TGARCH(1,1)-Student.

A seasonal model was not selected for WrA (2010) because the analysis of residuals of the selected model showed similar results to the best seasonal model, and the selected model was simpler.

When analyzing the residuals from Seasonal ARIMA-TGARCH-Student models for 2008, it turned out that in the case of WrA, the time series from 12 sensors out of the 23 available did not satisfy all expected conditions. The same happened for Sp: 14 out of the 20 models did not fulfill all requirements. Thus, in an attempt to extract further information not properly captured by these models, some features were calculated from the residuals. Figure 6 shows a summary of the steps of M3.

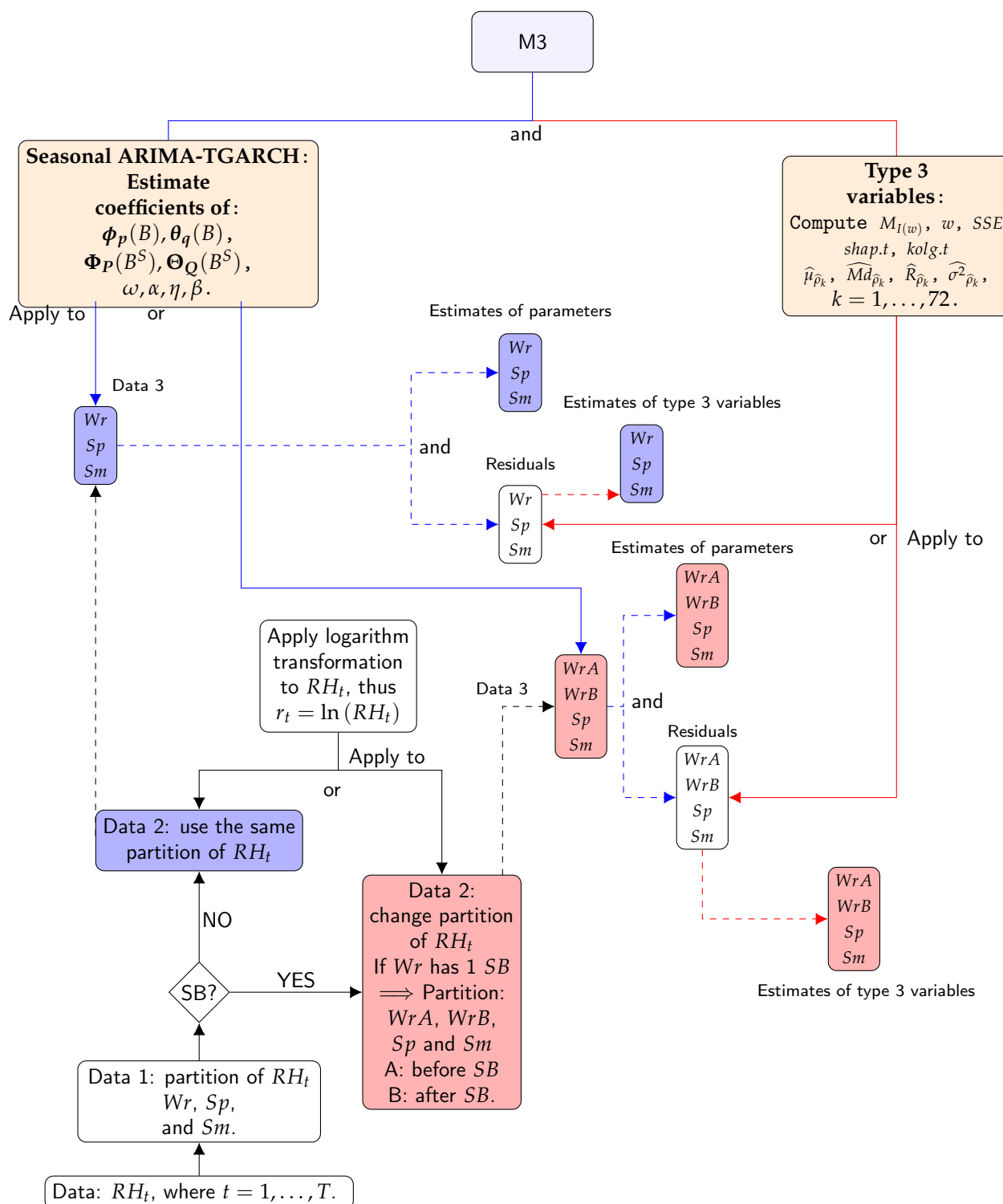


Figure 6. Flow chart for the steps of method 3: Blue lines indicate estimated parameters. Red lines indicate type 3 variables. Solid lines indicate processes. Dashed lines indicate results. The first step divides the different time series according to the climatic conditions (W_r , S_p , and S_m) (Data 1). The second step organises the time series according to possible structural breaks (SBs) (Data 2). The third step applies logarithmic transformation to Data 2, the result is Data . The fourth step applies the model to Data 3 in order to obtain the estimates of model parameters (first result) and then the residuals from the method. The fifth step consists of applying the formulas of type 3 variables to the residuals (second result). Different boxes display symbols W_r , S_p , and S_m (or W_rA , W_rB , S_p , and S_m). This indicates that the results correspond to all different parts of the time series.

In all cases, residuals from the ARIMA-GARCH-Student models displayed evidence of stationarity for 48 lags. However, in some cases, there was evidence of autocorrelation as well as the presence of ARCH effect. For the tests applied to residuals, 0.03 was the maximum p-value found to reject the null hypothesis. Regarding 2008, the number of time series (from the 23 sensors) that satisfied all tests in the residual analysis, is the following: 12 in WrA, 22 in WrB, 22 in Sp, and 20 in Sm. In 2010, out of the 20 sensors available, the values are: 18 in WrA, 19 in WrB, 14 in Sp, and 15 in Sm.

The features computed from residuals of the models were called type 3 variables. The other classification variables corresponded to the estimates of the parameters of the selected models. The estimates of the parameters are as follows:

- Estimated parameters from ARIMA of: (1) the regular autoregressive operator ($\phi_p(B)$) of order p and the regular moving average operator ($\theta_q(B)$) of order q : $\hat{\phi}_1, \hat{\phi}_2, \hat{\theta}_1, \hat{\theta}_2$, etc.; (2) the seasonal autoregressive operator ($\Phi_P(B^{24})$) of order P and the seasonal moving average operator ($\Theta_Q(B^{24})$) of order Q : $\hat{\Phi}_1, \hat{\Phi}_2, \hat{\Theta}_1, \hat{\Theta}_2$, etc.
- Estimated parameters from TGARCH(1,1): $\alpha_1, \eta_{11}, \beta_1, \omega$, and v (for Student distribution).

The estimate of type 3 variables:

- Variance of the residuals ($\hat{\sigma}^2$), maximum of spectral density of the residuals ($\hat{M}_{I(w)}$), frequency corresponding to maximum of spectral density (\hat{w}), mean ($\hat{\mu}_{\hat{\rho}_k}$), median ($\hat{Md}_{\hat{\rho}_k}$), range ($\hat{R}_{\hat{\rho}_k}$), and variance ($\hat{\sigma}_{\hat{\rho}_k}^2$) of sample ACF for 72 lags. The statistic of the SW test (*shap.t*) and the statistic of the KS normality test (*kolg.t*) are also included.

2.2.5. Sensor Classification by Means of sPLS-DA

Once all classification variables were computed as described above for the data from 2008, they were structured in three matrices, one per method (denoted as X_1, X_2 and X_3 , respectively), with 23 rows (sensors) where the classification variables are in columns. The total number of variables obtained from each method was 53 for X_1 , 141 for X_2 , and 59 for X_3 . Likewise, regarding 2010, classification variables were structured in three analogous matrices with 20 rows and with the same number of variables.

As the number of classification variables is much greater than the number of sensors, this scenario suggests a high degree of multicollinearity, and it might lead to severely ill-conditioned problems. Different approaches can be considered to deal with this problem. One solution is to perform variable selection, or to apply methods based on projection to latent structures like partial least squares discriminant analysis (PLS-DA).

One advantage of this multivariate tool is that it can handle many noisy and collinear classification variables, being computationally very efficient when the number of variables is much greater than the number of sensors. Even though PLS-DA is extremely efficient in a high-dimensional context, the interpretation of results can be complex in the case of a high number of variables. In such a case, sparse PLS-DA (sPLS-DA) has very satisfying predictive performance, and is able to select informative variables easily. Therefore, it was decided to apply sPLS-DA [40] here using the classification data sets mentioned above in order to identify a small subset of components and classification variables aimed at sensor clustering.

The algorithm of sPLS-DA used here was the one proposed by Rohart et al. [45], which corresponds to a modified version developed by Lê Cao et al. [40]. This new version uses the penalty ℓ_1 (lasso) on the loading vector of the regressor matrix by shrinking to zero the coefficient of some variables according to Rohart et al. [45].

With the aim of sensor clustering, sPLS-DA was applied to the previously mentioned classification data sets. Three categories of positions were considered for the sensors: \mathcal{RC} , \mathcal{W} , and \mathcal{F} . This method was applied to the three matrices (X_1, X_2 , and X_3) containing the classification variables, with dimension $n \times p$, where p is the number of classification variables and n is the number of sensors. Furthermore, Y is a vector of length n that indicates the class of each sensor, with values coded as 1 (for \mathcal{RC}), 2 (\mathcal{F}) and 3 (\mathcal{W}). This vector has to be converted into a dummy matrix (Z , i.e., with values either 0 or 1) with

dimension $n \times K$, where n is the number of sensors and $K = 3$ the number of classes or positions of sensors.

Before applying sPLS-DA, all anomalous values of each classification variable were removed and considered as missing data after being previously identified using normal probability plots and box plots for each variable. As a result, in 2008: 1.20% (M1), 1.04% (M2), and 1.06% (M3) were the percentages of missing values of the classification data sets. In 2010, the corresponding percentages were 0.40%, 1.39%, and 0.49%, respectively, for each method. These values are relatively low. Furthermore, all classification variables were normalized (i.e., centered and scaled to unitary variance). The package `mixOmics` [45] was used to perform sPLS-DA, which is able to handle missing values by using the NIPALS algorithm [45,82].

Three-fold cross-validation (three-fold CV, S1 supplementary information of [45]) was used to evaluate the performance (i.e., low classification error rate) of the PLS-DA. It was used to determine both the optimal number of components and the optimal number of variables. The three-fold CV was performed with stratified subsampling, where all positions (\mathcal{RC} , \mathcal{F} , and \mathcal{W}) are represented in each fold.

In order to select the optimal number of components, three-fold CV was applied for a maximum number of ten components, which was repeated 1000 times for each fold. With the objective of assessing the PLS-DA performance, the classification error rate (CER), the overall classification error rate (denoted as Overall), and the balanced classification error rate (BER) were computed [45]. Each BER value corresponds to the average proportion of wrongly classified sensors in each class, weighted by the number of sensors in each class. BER is less biased towards majority classes during the performance assessment when compared with the Overall criterion [45]. Thus, BER was considered instead of the latter.

The classification of sensors was determined according to different prediction distances (PD): maximum, centroid, and Mahalanobis [45]), which were computed for each sensor. Among the three distances calculated, it was found that the centroid one performed better in most cases for the classification, and hence it was selected. Regarding the centroid distance, the software computed the centroid (G) of the learning set of sensors (training data) belonging to the classes (\mathcal{RC} , \mathcal{F} , and \mathcal{W}). Each centroid G was based on the H latent components associated with X . The distances were calculated from the components of the trained model. The position of the new sensor was assigned according to the minimum distance between the predicted score and the centroids G calculated for the three classes considered.

The optimal number of components H was achieved by determining the best performance, based on the BER criterion and prediction distances according to the centroid distance. Once the optimal number of components was determined, repeated three-fold CV was carried out to establish the optimal number of variables according to the criteria of centroid and BER. Finally, once the optimal number of components and variables was decided, the final PLS-DA model was computed.

Figure 7 displays the results from the first three-fold CV for the three methods and for both years. For 2010, the values of BER and centroid distance suggested that one component is enough to classify the time series, while for 2008 the results indicate that one or two components are necessary. From this step, the centroid distance and BER were selected in order to determine the number of components.

Figure 8 shows the results from the second three-fold CV for the three methods and for both years. For 2008, the results suggested that the number of variables per one component were 15 (M1), 10 (M2), and 5 (M3). For 2010, the results suggest the number of variables per one component was 15 for all methods. The information in Figures 7 and 8 (centroid.dist, BER, and number of variables per component) was used to apply the final PLS-DA. Figure 9 describes the steps used to apply sPLS-DA, using the results from the three methods in the study.

Figure 9 shows the summary of steps of the sPLS-DA.

The main outputs from the analysis are: (1) a set of components (C) associated with X_1 , X_2 , and X_3 for the matrix Z ; (2) a set of loading (L) vectors containing the coefficients assigned to each variable that define each component; (3) a list of selected variables (V) from X_i ($i = 1, 2, 3$) associated with each component; (4) the values of BER for each component; and (5) the predicted class (PC) for each sensor. Coefficients in a given loading vector indicate the importance of each variable.

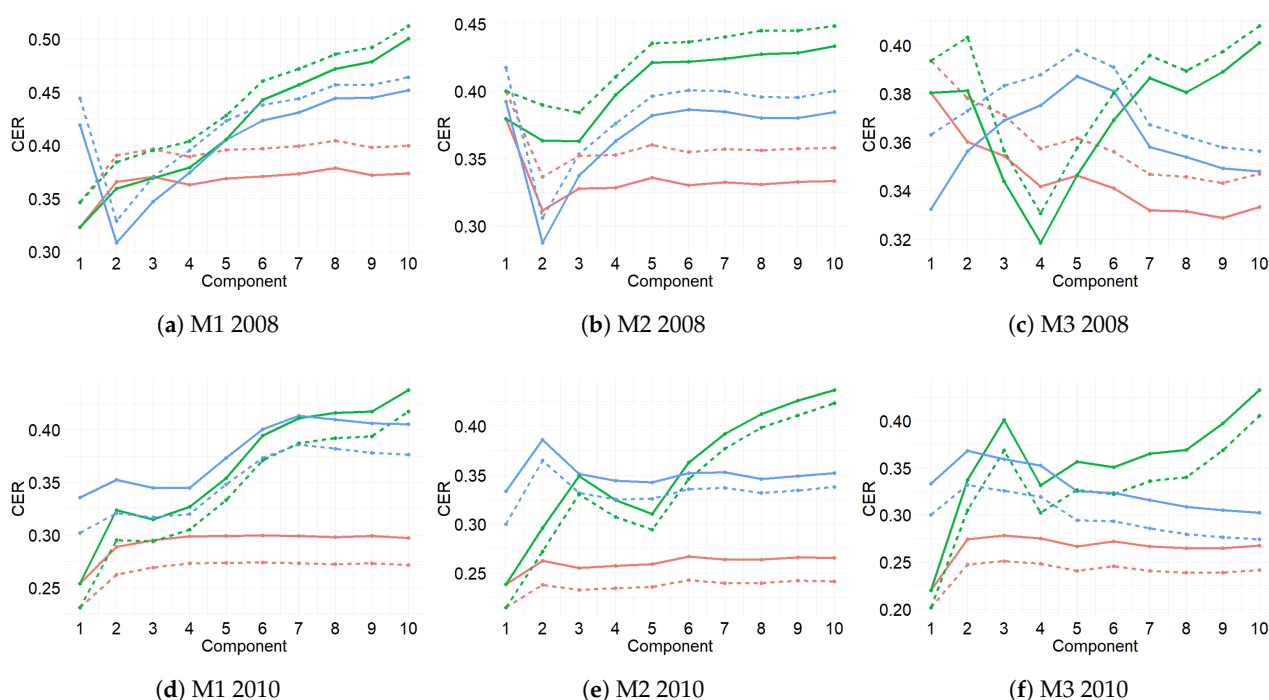


Figure 7. Classification error rate (CER) from PLS-DA for 10 components. The CER was computed for each prediction distance (maximum, centroid, and Mahalanobis) per method (M1, M2, and M3) for 2008 and 2010. Two types of error rate are indicated: balanced BER (dashed lines) or Overall (solid lines). Blue lines refer to maximum distance, red lines to centroid distance, and green lines to Mahalanobis distance. PLS-DA was carried out using repeated three-fold CV 1000 times. For 2010, BER and centroid distance showed the best performance achieved by one component. In 2008, M1 and M2 performed the best for maximum distance and Overall, while Mahalanobis distance performed the best in M3. The second-best distance for the three methods was the centroid distance.

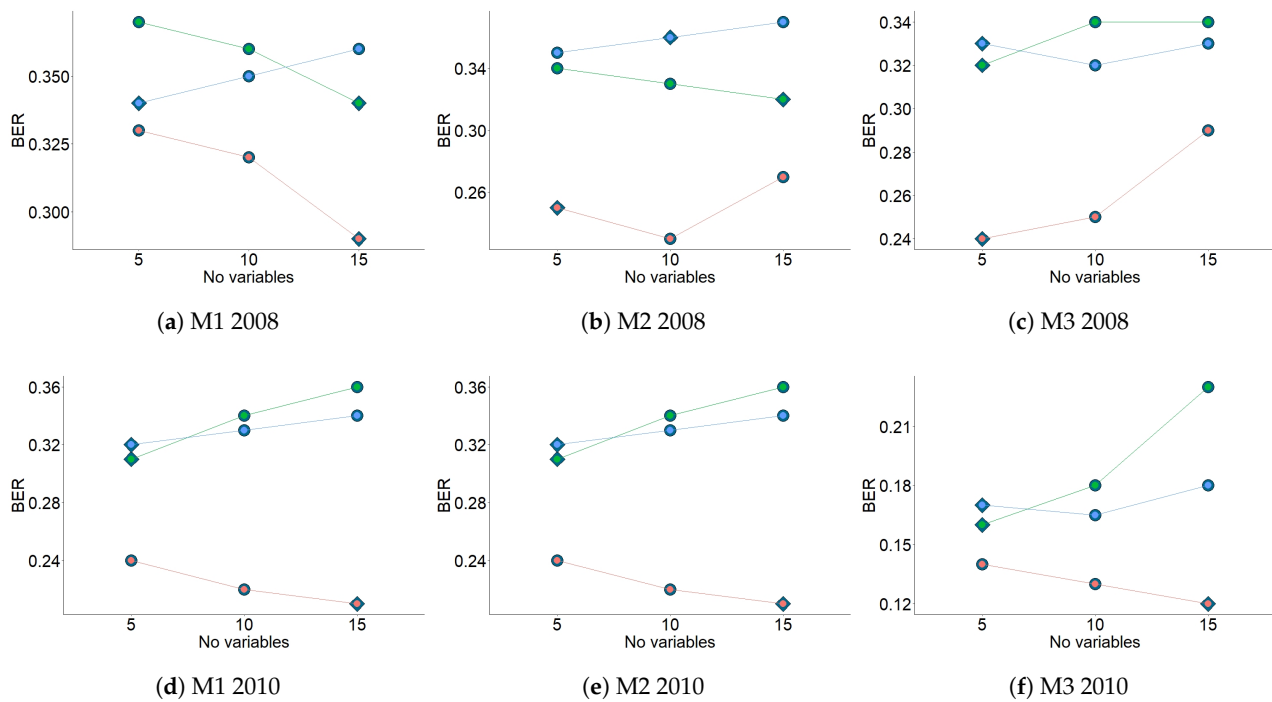


Figure 8. BER according to the number of variables (5, 10 or 15) and different number of components (1: orange dots, 2: blue dots, or 3: green dots) for each method, for 2008 and 2010. Three-fold CV was run 1000 times using centroid distance prediction. Diamonds indicate the optimal number of variables per component according to the lowest value of BER.

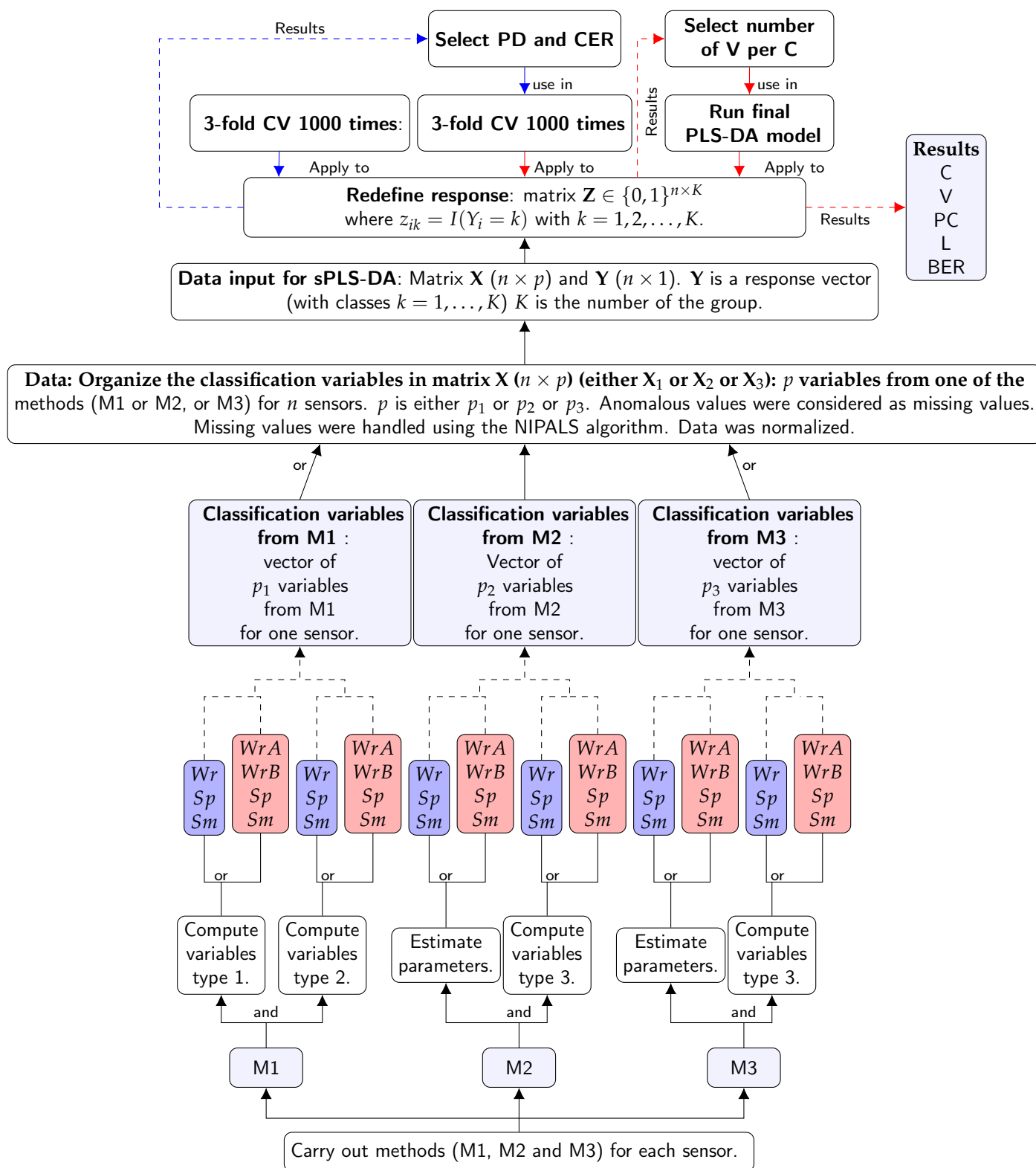


Figure 9. Flow chart for the stages used to apply sparse partial least squares discriminant analysis (sPLS-DA) to the results from the three methods. In the box titled “Data”, the information corresponds to the variables from one of the three methods. If the information is from M_i then $X=X_i$. The values were computed for all sensors. Thus, a matrix X was obtained. The values were treated before running the sPLS-DA algorithm. In the box titled “Data input for sPLS-DA”, the information corresponds to the response vector converted into a dummy matrix Z . In the following boxes the PLS-DA algorithm runs from left to right. The first three-fold CV was used to evaluate PLS-DA and the prediction distance PD, classification error rate (CER) with the optimal number of components selected. This information was used in the second three-fold CV to check PLS-DA in order to select the optimal number of variables V . The information obtained using both three-fold CVs was used to run the final PLS-DA.

3. Results

Components from sPLS-DA are linear combinations of variables that might correspond to WrA, WrB, Sp, or Sm. By applying sPLS-DA to X_1 , X_2 , and X_3 , only one component appeared to be relevant in all cases. The variables selected (per component) by the sPLS-DA algorithm are indicated in the following paragraph. The final model which used the classification variables from M1 (2008) is based on 15 selected variables. The selected model from M2 (2008) consists of 10 variables, while just 5 variables were considered for M3 (2008). The final model for M1, M2, and M3 (2010) comprises 1 component and 15 selected variables from each model (see Table 1).

The BER values are indicated in Table 1a,b for the three methods, using data from 2008 and 2010, respectively. For both years, the classification variables which turned out to be the most important for the first component were ordered according to the absolute value of their loading weights, from highest to lowest. The notation of the results is $\widehat{M}_{I(w)}$ (spec.mx); for \mathbf{RH}_h : $\widehat{\mu}_{MR}$ (rMh), $\widehat{\sigma}_{MR}^2$ (rVh); for \mathbf{RH}_d : $\widehat{\mu}_{MR}$ (rMd), $\widehat{\sigma}_{MR}^2$ (rVd). Also, SSE (sse), *kolg.t* (kolg.t), $\widehat{\sigma}^2$ (res.v), ω (omega), $\widehat{\alpha}_2$ (pacf2), S_1 (s1), S_{18} (s18), S_{20} (s20), S_{24} (s24), α (alpha), v (shape), $\widehat{\mu}_{\rho_1}$ (acf.m), \widehat{M}_{ρ_1} (acf.md).

An explanation about the most important variables, regardless of period and year, is presented below.

- M1: spec.mx, rMh, rMd, rVh, rVd, and pacf2 (see Table 1). The features rMh and rMd account for changes in the mean of the time series, while rVh and rVd are intended to explain changes in the variance. The rest of the features mentioned provide information about the dynamic structure of each time series. It was found that rMh, rMd, and rVh were important in the four periods considered, both in 2008 and 2010. rMd was relevant for WrA and WrB in 2008. The variable spec.mx was relevant in WrA and WrB for 2008 and 2010, as well as WrB. The variable pacf2 was found in WrB 2010. Hence, consistent results were derived from the two years under study.
- M2: sse, kolg.d, and spec.mx (computed from the residuals), as well as b, s1, s18, s19, s20, and s24 (from the models). From the residuals, sse accounts for the variance that is not explained by the models. This parameter appeared as important in all periods considered, except WrA 2010. kolg.d quantifies the deviation from normality for the residuals, and was relevant in all periods except Sp 2008 and WrB 2010. The third feature, spec.mx, which provides information about the dynamic structure of each time series, was relevant for all periods except WrB 2008, WrA 2010, and Sp 2010. Regarding the parameters computed from the models, b is related to the trend component of the time series, which was important in WrA 2010. The other variables mentioned are related to the seasonal components of the time series, which were shown to be important in Sm 2010.
- M3: res.v, shape, spec.mx, acf.m, and acf.md (computed from the residuals), as well as omega and alpha (from the model). From the residuals, res.v is aimed to explain the variance not explained by the models. It was relevant in all periods except Sp and Sm 2008. The variable shape provides information about the distribution of residuals, but it was only relevant in WrA 2010. The other features (i.e., spec.mx, acf.m, and acf.md) are intended to describe the dynamic structure of each time series. Spec.mx was important in all periods except Sp and Sm 2008, while the last two only appeared in Sp 2010. Regarding the parameters from the models, omega explains the changes in the mean of the conditional variance, while alpha quantifies the impact of the rotation on the conditional variance. The variable alpha only appeared in WrA 2010. Again, the fact that most variables were common in the three periods and in both years suggests strong consistency in the underlying phenomena explaining the discrimination between sensors.

In all cases, the classification variables corresponded to the different parts of the time series (WrA, WrB, Sp, and Sm), except for M3 in 2008 which only showed variables from winter (see Table 1).

Table 1. Results from sPLS-DA (2008 and 2010): variables selected for the first component per method (M) for each period (*WrA* stands for winter-A, *WrB* for winter-B, *Sp* for spring, and *Sm* for summer). With respect to the names of the selected classification variables, a prefix is used to indicate the period of the time series that was used to apply the three methods. For example, *WrA* was used as a prefix in the name of the classification variables to indicate the part of the time series where the variables came from. Thus, *WrA_{rMh}* indicates that the variable *rMh* corresponds to period *WrA*. For each component, variables are ordered according to the absolute value of their loading weights, from highest to lowest. Variables with negative weights are highlighted in bold. The balanced classification error rate (*BER*) is indicated for the first component per method.

(a) Results from sPLS-DA (2008).		
Method	Variables	BER
M1	WrA_{spec.mx} , WrA_{rMh} , WrB_{rMh} , Sp_{rMh} , Sm_{rMh} , WrA_{rVh} , WrB_{rVh} Sp_{rVh} , Sm_{rVh} , WrA_{rMd} , WrB_{rMd} , Sp_{rMd} , Sm_{rMd} , WrA_{rVd} , WrB_{rVd}	30.02%
M2	<i>WrA_{sse}</i> , <i>WrA_{spec.mx}</i> , <i>WrB_{sse}</i> , <i>Sp_{sse}</i> , WrA_{kolg.t} , <i>Sp_{spec.mx}</i> , WrB_{kolg.t} Sm_{kolg.t} , <i>Sm_{sse}</i> , <i>Sm_{spec.mx}</i>	24.05%
M3	WrA_{spec.mx} , WrA_{res.v} , WrB_{spec.mx} , WrB_{res.v} , WrA_{omega}	22.60%
(b) Results from sPLS-DA (2010).		
Method	Variables	BER
M1	<i>WrA_{spec.mx}</i> , <i>WrA_{rMd}</i> , <i>WrB_{rMd}</i> , <i>Sp_{rMd}</i> , <i>Sm_{rMd}</i> , <i>WrA_{rMh}</i> , <i>WrB_{rMh}</i> <i>Sp_{rMh}</i> , <i>Sm_{rMh}</i> , <i>WrA_{rVh}</i> , <i>WrB_{rVh}</i> , <i>Sp_{rVh}</i> , <i>Sm_{rVh}</i> , <i>WrB_{spec.mx}</i> , WrB_{pacf2}	24.08%
M2	<i>WrB_{spec.mx}</i> , <i>Sm_{sse}</i> , Sp_{kolg.t} , <i>Sp_{sse}</i> , WrA_{kolg.t} , <i>WrB_{s1}</i> , Sm_{kolg.t} , <i>WrB_{sse}</i> <i>WrB_{s24}</i> , <i>Sm_{spec.mx}</i> , Sm_{s19} , Sm_{s18} , WrA_b , Sm_{s20} , <i>Sp_{s24}</i>	21.17%
M3	<i>Sp_{res.v}</i> , <i>Sm_{res.v}</i> , <i>Sm_{spec.mx}</i> , <i>WrA_{res.v}</i> , <i>WrB_{res.v}</i> , <i>Sp_{spec.mx}</i> , <i>Sp_{omega}</i> <i>Sm_{omega}</i> , <i>WrA_{spec.mx}</i> , <i>WrB_{omega}</i> , WrB_{spec.mx} , WrA_{alpha} , WrA_{shape} , Sp_{acf.md} , Sp_{acf.m}	12.81%

The results shown in Figure 10 correspond to the score plots for the first two components from sPLS-DA applied to the classification of sensors. They depict their projection over the two principal latent structures that best discriminate sensors according to their position. In 2008, the first component for each method allowed a rather good discrimination of sensors at the \mathcal{RC} position with respect to the rest, though a poor discrimination was achieved between \mathcal{F} and \mathcal{W} (see Figure 10a–c).

In 2010, the first and second components for M3 displayed a clear discrimination between sensors located on the three positions. However, for M1 and M2, only \mathcal{RC} sensors appear far apart from those on the walls, while the \mathcal{F} group is located in between (see Figure 10d–f).

Regarding the performance of the three methods for achieving the classification of sensors, the best results were derived from M3 and the worst from M1. M3 yielded higher correct classification percentages: 77.40% in 2008 and 87.19% in 2010 (see Table 1b). For 2008, the final classification resulting from M3 variables displayed the following wrongly classified sensors: Y, AA, Z, and V (see Figure 11a). Three of them (Y, Z, and AA) were installed near the location where the salt efflorescence was found. For 2010, the final results from sPLS-DA for M3 showed that all sensors were classified correctly (see Figure 11b).

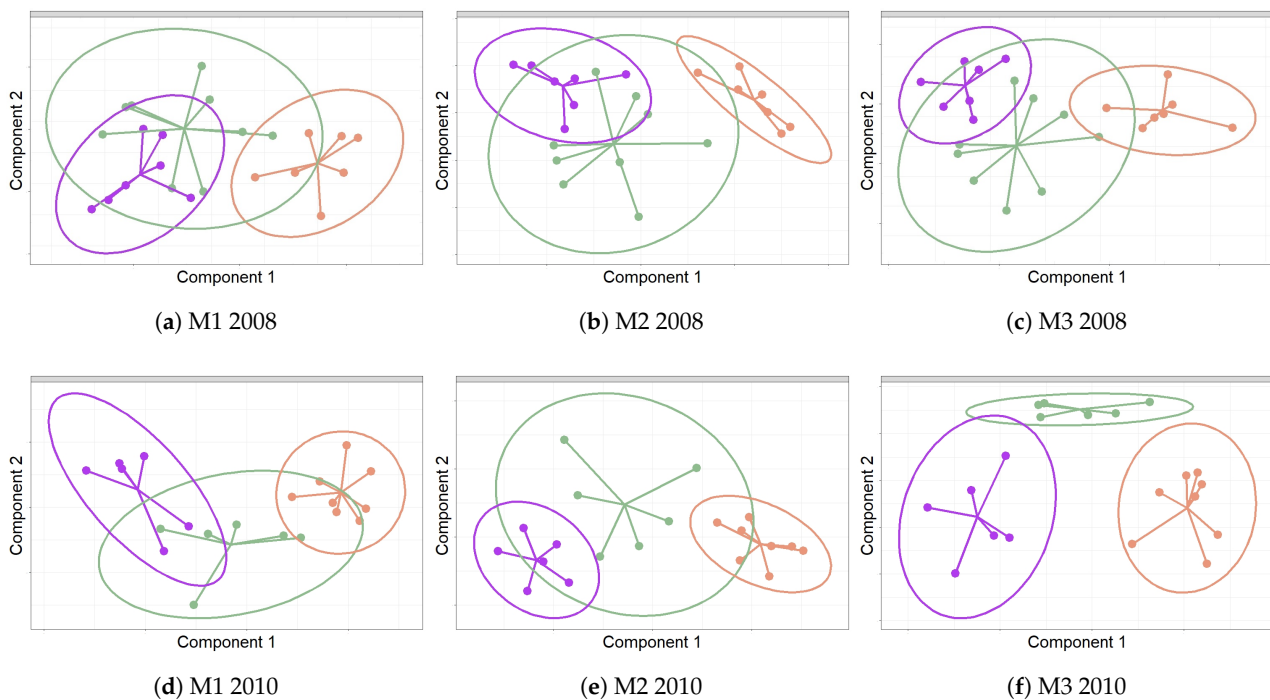


Figure 10. Discrimination of the time series of RH according to the position of sensors: Frescoes (\mathcal{F}), Cornice and Ribs (\mathcal{RC}), and Wall (\mathcal{W}). Color codes: \mathcal{F} sensors are shown in green, \mathcal{RC} in orange, and \mathcal{W} in purple. Graphics correspond to the projection of sensors over the first two components from sPLS-DA. Each graph shows confidence ellipses for each class to highlight the strength of the discrimination at a confidence level of 95%.

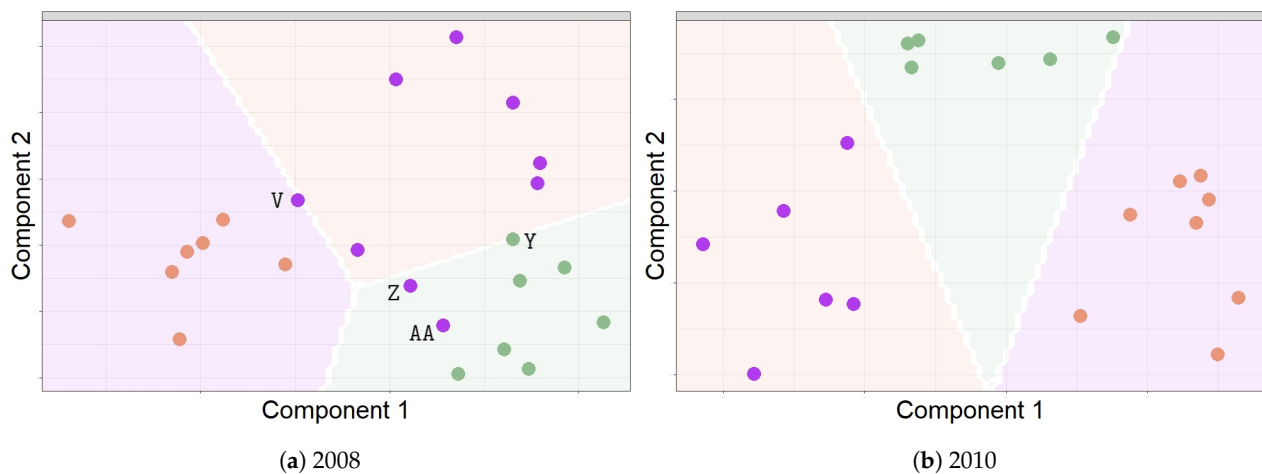


Figure 11. Prediction classes derived from M3: these plots display the predicted classes of each sensor located in different points in a simulated grid. The final classification prediction for the position of the sensors is displayed. The points in orange (\mathcal{RC}), purple (\mathcal{W}) and green (\mathcal{F}) represent the class prediction for the sensors in the study. (a) For 2008: according to the classification prediction, the wrongly classified sensors were as follows: V, Y, Z, and AA. (b) For 2010: all sensors were classified correctly.

4. Discussion

The methodology proposed here consists of using sPLS-DA to classify time series of RH according to classification variables that were computed from different functions (e.g., sample ACF, sample PACF, spectral density, and MR). Additionally, the Seasonal ARIMA-TGARCH-Student model and the Additive SH-W method were used. Furthermore, estimated parameters of the models, as well as the mean, variance, and maximum values of the functions (e.g., sample ACF, sample PACF, spectral density, and statistics of the

KS normality test, among others) were applied to the residuals derived from the models. The centroid distance was applied to classify the sensors, and the lasso penalty was used to select the optimal variables that determine the relevant components. Additionally, the BER parameter was employed to evaluate the performance of the classification methodology.

We used sPLS-DA because the classification data in this study are characterized by more variables than the number of time series (sensors), and in the interest of easily interpreting the results. This technique leads to underlying latent variables (components) that summarize the relevant information from the data for the purpose of discrimination. It performs variable selection for each component, which is an advantage. The key issue in time series clustering is how to characterize the similarities and dissimilarities between time series. Various metrics for measuring such similarity have been proposed, based on: parameters from models [83–87], serial features extracted from the original time series [88–91], the complexity of the time series [92–97], the properties of the predictions [98,99], and the comparison of raw data [100]. Regarding methods based on model parameters, the criterion most commonly considered is to assume that time series are properly explained by ARIMA processes. Piccolo [83] introduced the Euclidean distance between their corresponding AR expansion [38] as a metric and used a complete linkage clustering algorithm to construct a dendrogram. One problem of this metric is related to the numerical computations of AR coefficients. The same metric was also considered by Otranto [101] for dealing with GARCH processes. For ARMA models, Maharaj [102] developed an agglomerative hierarchical clustering procedure based on the p-value of a hypothesis test applied to every pair of stationary time series. Kalpakis et al. [103] studied the clustering of ARIMA time series by using the Euclidean distance between the linear predictive coding (LPC) cepstrum of two time series as their dissimilarity measure. Xiong and Yeung [104] classified univariate ARIMA time series by considering ARMA models. They derived an expectation maximization (EM) algorithm for estimating the coefficients and parameters of the models. However, if the underlying clusters are very close to each other, the clustering performance might diminish significantly. According to the review of the previously mentioned studies about the clustering of time series, it seems that the methodology applied here is rather unique because it uses a hybrid model comprising ARIMA and GARCH to calculate a distance for classifying time series. This is also probably the first using sPLS-DA in order to classify time series.

We found that the time series of RH, one per sensor, were very similar despite their different positions in the apse vault of the cathedral. When classifying the sensors, it turned out that few parameters appeared as relevant, most of which were features extracted from the residuals of models. This is most likely due to the similarity among the time series studied. As a consequence, the information that was not properly explained by the models was decisive for characterizing the differences between time series. The classification variables derived from the ARIMA-TGARCH-Student model yielded better performance than those from SH-W, which might suggest that the former model captures more information from the data than the latter. In fact, SH-W is an algorithm intended for producing point forecasts [60].

A comparison of the results from method M1 with those from the preliminary study of Zarzo et al. [1] indicates that a better classification was obtained here. The variables HMV and DMV (rMh and rMd) were relevant in both studies. Although the mean for the total observations was important in the preliminary project [1], this variable was not selected by the sPLS-DA. The classification variables selected per sPLS-DA explained the changes in the mean and variance of the time series with rMh , rMd , rVh , and rVd . Furthermore, the method obtains variables from sample PACF and spectral density which explain the autocorrelation of the time series. The research by Zarzo et al. [1] did not use variables related to the autocorrelation of the time series.

One disadvantage of sPLS-DA is the need to use the same number of classification variables for each sensor. As a consequence, a unique ARIMA-TGARCH-Student model was used for all sensors in the same part of the time series (WrA , WrB , Sp , and Sm). This

means that a better fit might result, as it considers a different model for each time series. Another disadvantage is that it is necessary to know a priori the number of classes of the time series (sensors) for their classification. According to the previous idea, the limitations of the statistical methodology proposed in this study are: (1) sPLS-DA needs to know the number of classes before implementing the algorithm. (2) When applying a unique ARIMA-TGARCH parametric model to all sensors, it is unlikely that the best values for the classification variables will be found. This can affect the classification error rate of the sensors.

One advantage of using both sPLS-DA and ARIMA-TGARCH-Student is the capability of classifying time series with very similar characteristics. Additionally, the functions and models utilized here can be easily implemented because different packages are available in R software. One such example is the `mixOmics` package of R, which has different functions that allow sPLS-DA to be implemented simply and makes it easy to display the different results for interpretation. Furthermore, this package can handle missing values using the NIPALS approach. It takes advantage of the PLS algorithm which performs local regressions on the latent components. There are two main advantages of using PLS—it both handles missing values and calculates the components sequentially. In this study, the anomalous values of classification variables were considered as missing values in order to avoid possible problems with the classification of the sensors. The percentage of values that were used as missing were lower than 2%.

In relation to future studies, alternative classification variables could be considered depending on the different scenarios and according to the characteristics of the time series. In order to obtain classification variables that capture more information from the data, flexible models can be proposed. Some options for calculating the classification variables might be the following:

- Cepstral coefficients: Ioannou et al. [105] studied several clustering techniques in the context of the semiparametric model: spectral density ratio. They found that the cepstral-based techniques performed better than all the other spectral-domain-based methods, even for relatively small subsequences.
- Structural time series model: the flexibility required from this model can be achieved by letting the regression coefficients change over time [106].
- A nonparametric approach of the GARCH [107,108].

Regarding classification techniques when there are fewer variables than time series, sPLS-DA can be extended by using the elastic net [109] as the penalization. Finally, a further study might be carried out in controlled scenarios, where time series can be computationally simulated by controlling different characteristics in order to identify the strengths and weaknesses of the proposed methodology. In alignment with the previous ideas for improving the methodology, future research will use sPLS-DA with two methods: a nonparametric Seasonal ARIMA-GARCH model and a structural time series model. Furthermore, several time series will be computationally simulated in controlled scenarios in order to evaluate the results when using sPLS-DA, together with one of the previously mentioned methods.

García-Diego and Zarzo [2] concluded that the environment surrounding the Renaissance frescoes was not the same at all points of the apse vault of the cathedral. Sensors located on the walls or on the paintings registered higher RH values than those in the vault ribs. Thus, the mean value of RH is related with the three previously mentioned classes. The ideal goal is obviously to achieve a correct classification of all sensors. However, a poor classification error rate might be caused by either the malfunctioning of some sensors or a poor performance of the classification technique, or if there is a problem related with the microclimate where the sensors are located. Those sensors incorrectly classified by the technique should be checked to identify possible moisture problems in the artworks. In this work, the main cause of sensor malfunction was the development of salt deposits around the probes as a consequence of fitting some of the probes inside the layer of plaster supporting the frescoes.

The results indicate that sPLS-DA could be implemented for the online monitoring of fresco paintings aimed at preventive conservation using the parameters and features previously extracted from the hybrid models based on GARCH and ARIMA as classification variables. This analysis might be carried out for each season of every year.

5. Conclusion

The methodology proposed here is useful for understanding the differences in thermohygro-metric conditions monitored inside large buildings or museums, which might provide a basis for better assessing the potential risks related to temperature and humidity on the artworks. Among the methods proposed, a hybrid approach based on ARIMA and GARCH models with sPLS-DA yielded the best performance. Parsimonious models with a small subset of components and classification variables were obtained using sPLS-DA, which offers satisfactory results with easy interpretation. Another advantage of sPLS-DA is that it can be implemented easily with `mixOmics`, which allows a focus on graphical representation in order to better understand the relationships between the different observations and variables. Furthermore, this package can deal with missing values. Finally, the use of a hybrid approach based on ARIMA and GARCH models as well as sPLS-DA is a novel proposal for classifying different time series. In order to improve the methodology proposed in this research, future research will use sPLS-DA with two methods that are more flexible than those applied in this study. This will capture more information from the data. Furthermore, a computational simulation will be carried out in order to evaluate the new methodology in different possible scenarios.

Author Contributions: Data curation, F.-J.G.-D.; Formal analysis, S.R. and M.Z.; Funding acquisition, A.P.; Methodology, S.R. and M.Z.; Project administration, F.-J.G.-D.; Writing—original draft, S.R.; Writing—review and editing, M.Z., A.P., and F.-J.G.-D. All authors have read and agreed to the published version of the manuscript

Funding: This project received funding from the European Union’s Horizon 2020 research and innovation programme under grant agreement No. 814624. Furthermore, the project was partially supported by Instituto Colombiano de Crédito Educativo y Estudios Técnicos en el Exterior (ICETEX) by means of Programa crédito Pasaporte a la Ciencia ID 3595089, and also by Pontificia Universidad Javeriana Cali (Nit 860013720-1) through the Convenio de Capacitación para Docentes O. J. 086/17.

Institutional Review Board Statement: Not applicable because humans or animals are not involved in this study.

Informed Consent Statement: Not applicable because humans or animals are not involved in this study.

Data Availability Statement: The data sets used are available at: <https://zenodo.org/record/4412126> DOI: 10.5281/zenodo.4412126.

Conflicts of Interest: The authors declare no conflicts of interest.

References

1. Zarzo, M.; Fernández-Navajas, A.; García-Diego, F.J. Long-term monitoring of fresco paintings in the cathedral of Valencia (Spain) through humidity and temperature sensors in various locations for preventive conservation. *Sensors* **2011**, *11*, 8685–8710. [[CrossRef](#)] [[PubMed](#)]
2. García-Diego, F.J.; Zarzo, M. Microclimate monitoring by multivariate statistical control: the renaissance frescoes of the cathedral of Valencia (Spain). *J. Cult. Herit.* **2010**, *11*, 339–344. [[CrossRef](#)]
3. Verticchio, E.; Frasca, F.; García-Diego, F.D.; Siani, A.M. Investigation on the use of passive microclimate frames in view of the climate change scenario. *Climate* **2019**, *7*, 98. [[CrossRef](#)]
4. Sesana, E.; Gagnon, A.S.; Bertolin, C.; Hughes, J. Adapting cultural heritage to climate change risks: perspectives of cultural heritage experts in Europe. *Geosciences* **2018**, *8*, 305. [[CrossRef](#)]
5. Camuffo, D. *Microclimate for Cultural Heritage*, 1st ed.; Elsevier: Amsterdam, The Netherlands, 1998.
6. Camuffo, D. Indoor dynamic climatology: investigations on the interactions between walls and indoor environment. *Atmos. Environ.* (1967) **1983**, *17*, 1803–1809. [[CrossRef](#)]

7. Camuffo, D.; Bernardi, A. Study of the microclimate of the hall of the giants in the carrara palace in padua. *Stud. Conserv.* **1995**, *40*, 237–249. [[CrossRef](#)]
8. Camuffo, D.; Pagan, E.; Bernardi, A.; Becherini, F. The impact of heating, lighting and people in re-using historical buildings: a case study. *J. Cult. Herit.* **2004**, *5*, 409–416. [[CrossRef](#)]
9. Bernardi, A. Microclimate in the british museum, London. *Mus. Manag. Curatorship* **1990**, *9*, 169 – 182. [[CrossRef](#)]
10. Bernardi, A.; Camuffo, D. Microclimate in the chiericati palace municipal museum Vicenza. *Mus. Manag. Curatorship* **1995**, *14*, 5–18. [[CrossRef](#)]
11. Camuffo, D.; Bernardi, A.; Sturaro, G.; Valentino, A. The microclimate inside the Pollaiolo and Botticelli rooms in the Uffizi Gallery. *J. Cult. Herit.* **2002**, *3*, 155–161. [[CrossRef](#)]
12. Merello, P.; García-Diego, F.D.; Beltrán, P.; Scatigno, C. High frequency data acquisition system for modelling the impact of visitors on the thermo-hygrometric conditions of archaeological sites: a Casa di Diana (Ostia Antica, Italy) case study. *Sensors* **2018**, *18*, 348. [[CrossRef](#)] [[PubMed](#)]
13. Frasca, F.; Siani, A.M.; Casale, G.R.; Pedone, M.; Bratasz, L.; Strojecki, M.; Mleczkowska, A. Assessment of indoor climate of Mogita Abbey in Krak Kraków (Poland) and the application of the analogues method to predict microclimate indoor conditions. *Environ. Sci. Pollut. Res.* **2017**, *24*, 13895–13907. [[CrossRef](#)] [[PubMed](#)]
14. Tabunschikov, Y.; Brodatch, M. Indoor air climate requirements for Russian churches and cathedrals. *Indoor Air* **2004**, *14* (Suppl. 7), 168–174. [[CrossRef](#)] [[PubMed](#)]
15. Camuffo, D.; Sturaro, G.; Valentino, A. Thermodynamic exchanges between the external boundary layer and the indoor microclimate at the Basilica of Santa Maria Maggiore, Rome, Italy: the problem of conservation of ancient works of art. *Bound.-Layer Meteorol.* **1999**, *92*, 243–262. [[CrossRef](#)]
16. Vuerich, E.; Malaspina, F.; Barazutti, M.; Georgiadis, T.; Nardino, M. Indoor measurements of microclimate variables and ozone in the church of San Vincenzo (Monastery of Bassano Romano Italy): a pilot study. *Microchem. J.* **2008**, *88*, 218–223. [[CrossRef](#)]
17. G, L.; Charpantidou, E.; Kioutsoukis, I.; Rapsomanikis, S. Indoor microclimate, ozone and nitrogen oxides in two medieval churches in Cyprus. *Atmos. Environ.* **2006**, *40*, 7457–7466. [[CrossRef](#)]
18. Bernardi, A.; Todorov, V.; Hristova, J. Microclimatic analysis in St. Stephan’s church, Nessebar, Bulgaria after interventions for the conservation of frescoes. *J. Cult. Herit.* **2000**, *1*, 281–286. [[CrossRef](#)]
19. EN16883. Conservation of Cultural Heritage. Guidelines for Improving the Energy Performance of Historic Buildings. 2017. Available online: https://standards.cen.eu/dyn/www/?p=204:110:0:::FSP_PROJECT:36576&cs=113EECDB855EBEF2097C9F626231290BE (accessed on 1 January 2021).
20. EN16141. Conservation of Cultural Heritage. Guidelines for Management of Environmental Conditions. Open Storage Facilities: Definitions and Characteristics of Collection Centres Dedicated to the Preservation and Management of Cultural Heritage. 2012. Available online: https://standards.cen.eu/dyn/www/?p=204:110:0:::FSP_PROJECT:30978&cs=19D7899D83F3E1FAFF740C53B6D9C068F (accessed on 1 January 2021).
21. EN16242. Conservation of Cultural Heritage. Procedures and Instruments for Measuring Humidity in the Air and Moisture Exchanges between Air and Cultural Property. 2012. Available online: https://standards.cen.eu/dyn/www/?p=204:110:0:::FSP_PROJECT:34048&cs=1000E6B80FEC23200296847848BFD8390 (accessed on 1 January 2021).
22. EN15898. Conservation of Cultural Property. Main General Terms and Definitions. 2019. Available online: https://standards.cen.eu/dyn/www/?p=204:110:0:::FSP_PROJECT:61301&cs=1BFEBDBB425EAF8FEDC7D300B26CB0F1E (accessed on 1 January 2021).
23. EN15758. Conservation of Cultural Property. Procedures and Instruments for Measuring Temperatures of the Air and the Surfaces of Objects. 2010. Available online: https://standards.cen.eu/dyn/www/?p=204:110:0:::FSP_PROJECT:28488&cs=18FBB4BE512863FD1F25E8415D5BA9ACE (accessed on 1 January 2021).
24. EN15757. Conservation of Cultural Property. Specifications for Temperature and Relative Humidity to Limit Climate-Induced Mechanical Damage in Organic Hygroscopic Materials. 2010. Available online: https://standards.cen.eu/dyn/www/?p=204:110:0:::FSP_PROJECT:28487&cs=1CA6AC7E107FE7F852A4F9C8D11CCD217 (accessed on 1 January 2021).
25. EN16893. Conservation of Cultural Heritage. Specifications for Location, Construction and Modification of Buildings or Rooms Intended for the Storage or Use of Heritage Collections. 2018. Available online: https://standards.cen.eu/dyn/www/?p=204:110:0:::FSP_PROJECT:35659&cs=15B18B1A035AB93FB36262CA746F5F7EC (accessed on 1 January 2021).
26. Corgnati, S.P.; Filippi, M. Assessment of thermo-hygrometric quality in museums: Method and in-field application to the Duccio di Buoninsegna exhibition at Santa Maria della Scala (Siena, Italy). *J. Cult. Herit.* **2010**, *11*, 345–349. [[CrossRef](#)]
27. Collectioncare.eu. Available online: <https://www.collectioncare.eu/> (accessed on 1 January 2021).
28. Dunia, R.; Qin, S.J.; Edgar, T.F.; McAvoy, T.J. Use of principal component analysis for sensor fault identification. *Comput. Chem. Eng.* **1996**, *20*, S713–S718. [[CrossRef](#)]
29. Zhu, D.; Bai, J.; Yang, S. A multi-fault diagnosis method for sensor systems based on principle component analysis. *Sensors* **2009**, *10*, 241–253. [[CrossRef](#)]
30. Venables, W.N.; Ripley, B.D. *Modern Applied Statistics with S*, 4th ed.; Springer: New York, NY, USA, 2002.
31. Box, G.E.P.; Jenkins, G.M. *Time Series Analysis: Forecasting and Control*, Revised ed.; Holden-Day: San Francisco, CA, USA 1976.
32. Brockwell, P.J.; Davis, R.A. *Time Series: Theory and Methods*, 2nd ed.; Springer: New York, NY, USA, 1987.

33. Kovalevsky, S. *Package 'QuantTools'*; R Core Team: Vienna, Austria, 2020. Available online: https://quanttools.bitbucket.io/_site/index.html (accessed on 1 January 2021).
34. Holt, C.C. Forecasting seasonals and trends by exponentially weighted moving averages. *J. Econ. Soc. Meas.* **2004**, *20*, 5–10. [[CrossRef](#)]
35. Winters, P.R. Forecasting sales by exponentially weighted moving averages. *Manag. Sci.* **1960**, *6*, 324–342. [[CrossRef](#)]
36. Hannan, E.J.; Rissanen, J. Recursive Estimation of Mixed Autoregressive-Moving Average Order. *Biometrika* **1982**, *69*, 81–94. [[CrossRef](#)]
37. Ghalanos, A. *Introduction to the Rugarch Package (Version 1.4-3)*; R Core Team: Vienna, Austria, 2020. Available online: https://cran.r-project.org/web/packages/rugarch/vignettes/Introduction_to_the_rugarch_package.pdf (accessed on 1 January 2021).
38. Palma, W. *Time Series Analysis. Wiley Series in Probability and Statistics*, Har/Psc ed.; John Wiley & Sons Inc.: Hoboken, NJ, USA, 2016.
39. Hamilton, J.D. *Time Series Analysis*, 5th ed.; Princeton University Press: Princeton, NJ, USA, 1994.
40. Lê Cao, K.A.; Boitard, S.; Besse, P. Sparse PLS discriminant analysis: biologically relevant feature selection and graphical displays for multiclass problems. *BiomMed Cent. Bioinform.* **2011**, *12*, 253. [[CrossRef](#)] [[PubMed](#)]
41. R Core Team. *R: A Language and Environment for Statistical Computing*; R Foundation for Statistical Computing: Vienna, Austria, 2014.
42. Qiu, D. *Package 'aTSA'*; R Core Team: Vienna, Austria, 2015. Available online: <https://cran.r-project.org/web/packages/aTSA/index.html> (accessed on 1 January 2021).
43. Hyndman, R.J.; Khandakar, Y. Automatic time series forecasting: the forecast package for R. *J. Stat. Softw.* **2008**, *26*, 1–23. [[CrossRef](#)]
44. Hyndman, R.; Athanasopoulos, G.; Bergmeir, C.; Caceres, G.; Chhay, L.; O'Hara-Wild, M.; Petropoulos, F.; Razbash, S.; Wang, E.; Yasmeen, F. *Package 'Forecast: Forecasting Functions for Time Series and Linear Models'*; R Core Team: Vienna, Austria, 2020. <https://cran.r-project.org/web/packages/forecast/> (accessed on 1 January 2021).
45. Rohart, F.; Gautier, B.; Singh, A.; Lê Cao, K.A. mixOmics: An R package for omics feature selection and multiple data integration. *PLoS Comput. Biol.* **2017**, *13*, 1–19. [[CrossRef](#)]
46. Le Cao, K.A.; Dejean, S. *mixOmics: Omics Data Integration Project*; Bioconductor: Vienna, Austria, 2020. Available online: <http://www.bioconductor.org/packages/release/bioc/html/mixOmics.html> (accessed on 1 January 2021).
47. Ghalanos, A.; Kley, T. *Package 'Rugarch'*; R Core Team: Vienna, Austria, 2020. Available online: <https://cran.r-project.org/web/packages/rugarch/index.html> (accessed on 1 January 2021).
48. Zeileis, A. Implementing a class of structural change tests: an econometric computing approach. *Comput. Stat. Data Anal.* **2006**, *50*, 2987–3008. [[CrossRef](#)]
49. Trapletti, A.; Hornik, K. *Package: 'tseries: Time Series Analysis and Computational Finance'*; R Core Team: Vienna, Austria, 2020. Available online: <https://cran.r-project.org/web/packages/tseries/> (accessed on 1 January 2021).
50. Leisch, F.; Hornik, K.; Kuan, C.M. Monitoring structural changes with the generalized fluctuation test. *Econom. Theory* **2000**, *16*, 835–854. [[CrossRef](#)]
51. Chow, G.C. Tests of equality between sets of coefficients in two linear regressions. *Econometrica* **1960**, *28*, 591–605. [[CrossRef](#)]
52. Hansen, B.E. Tests for parameter instability in regressions with I(1) processes. *J. Bus. Econ. Stat.* **2002**, *20*, 45–59. [[CrossRef](#)]
53. Andrews, D.W.K. Tests for parameter instability and structural change with unknown change point. *Econometrica* **1993**, *61*, 821–856. [[CrossRef](#)]
54. Andrews, D.W.K.; Ploberger, W. Optimal tests when a nuisance parameter is present only under the alternative. *Econometrica* **1994**, *62*, 1383–1414. [[CrossRef](#)]
55. Zeileis, A.; Leisch, F.; Hornik, K.; Kleiber, C. Strucchange: an R package for testing for structural change in linear regression models. *J. Stat. Softw.* **2002**, *7*, 1–38. [[CrossRef](#)]
56. Cryer, J.D.; Chan, K.S. *Time Series Analysis: With Applications in R*; Springer Texts in Statistics; Springer: New York, NY, USA, 2008.
57. Royston, J.P. Algorithm AS 181: the W test for normality. *J. R. Stat. Soc. Ser. C Appl. Stat.* **1982**, *31*, 176–180. [[CrossRef](#)]
58. Gaetano, D. Forecast combinations in the presence of structural breaks: evidence from U.S. equity markets. *Mathematics* **2018**, *6*, 34. [[CrossRef](#)]
59. Cowpertwait, P.S.P.; Metcalfe, A.V. *Introductory time Series with R*; Springer Series: Use R, Springer: New York, NY, USA, 2009.
60. Hyndman, R.; Koehler, A.B.; Ord, J.K.; Snyder, R.D. *Forecasting with Exponential Smoothing: The State Space Approach*; Springer: Berlin, Germany, 2008.
61. Ljung, G.M.; Box, G.E.P. On a measure of Lack of fit in time series models. *Biometrika* **1978**, *65*, 297–303. [[CrossRef](#)]
62. Conover, W.J. *Practical Nonparametric Statistics*, 1st ed.; John Wiley & Sons, Inc.: Hoboken, NJ, USA, 1999. ISBN:0-471-16068-7.
63. Royston, J.P. An extension of Shapiro and Wilk's W test for normality to large samples. *J. R. Stat. Soc. Ser. C Appl. Stat.* **1982**, *31*, 115–124. [[CrossRef](#)]
64. Royston, P. Remark AS R94: A remark on algorithm AS 181: the W-test for normality. *J. R. Stat. Soc. Ser. C Appl. Stat.* **1995**, *44*, 547–551. [[CrossRef](#)]
65. Anderson, O.D. *Time series analysis: forecasting and control: (Revised Edition)* by George E.P. Box and Gwilym M. Jenkins. *J. Frankl. Inst.* **1980**, *310*. [[CrossRef](#)]

66. Yaziz, S.R.; Azizan, N.A.; Zakaria, R.; Ahmad, M.H. The performance of hybrid ARIMA-GARCH modeling in forecasting gold price. In Proceedings of the 20th International Congress on Modelling and Simulation, Adelaide, Australia, 1–6 December 2013; pp. 1–6.
67. Tseng, J.J.; Li, S.P. Quantifying volatility clustering in financial time series. *Int. Rev. Financ. Anal.* **2012**, *23*, 11–19. [[CrossRef](#)]
68. Engle, R. Risk and volatility: econometric models and financial practice. *Am. Econ. Rev.* **2004**, *94*, 405–420. [[CrossRef](#)]
69. Engle, R.F. Autoregressive conditional heteroscedasticity with estimates of the variance of United Kingdom inflation. *Econometrica* **1982**, *50*, 987–1007. [[CrossRef](#)]
70. Bollerslev, T. Generalized autoregressive conditional heteroskedasticity. *J. Econom.* **1986**, *31*, 307–327. [[CrossRef](#)]
71. Engle, R.F.; Bollerslev, T. Modelling the persistence of conditional variances. *Econom. Rev.* **1986**, *5*, 1–50. [[CrossRef](#)]
72. Yusof, F.; Kane, I.L. Volatility modeling of rainfall time series. *Theor. Appl. Climatol.* **2013**, *113*, 247–258. [[CrossRef](#)]
73. Yaziz, R.S.; Azlinna, N.A.; Ahmad, M.; Zakaria, R. Modelling gold price using ARIMA-TGARCH. *Appl. Math. Sci.* **2016**, *10*, 1391–1402. [[CrossRef](#)]
74. Hor, C.L.; Watson, S.J.; Majithias, S. Daily load forecasting and maximum demand estimation using ARIMA and GARCH. In Proceedings of the International Conference on Probabilistic Methods Applied to Power Systems, Stockholm, Sweden, 11–15 June 2006; pp. 1–6. [[CrossRef](#)]
75. Xing, J. The research on stock market volatility in China based on the model of ARIMA-EARCH-M (1, 1) and ARIMA-TARCH-M (1, 1). In *International Conference on Advances in Education and Management*, 2011; Volume 210. [[CrossRef](#)]
76. Xing, J. The research on stock market volatility in China based on the model of ARIMA-EARCH-M (1, 1) and ARIMA-TARCH-M (1, 1). In Proceedings of the International Conference on Advances in Education and Management. ISAEBD 2011: Education and Management, Dalian, China, 6–7 August 2011; Volume 210. [[CrossRef](#)]
77. Johnston, K.; Scott, E. GARCH models and the stochastic process underlying exchange rate price change. *J. Financ. Strateg. Decis.* **2000**, *13*, 13–24.
78. Bollerslev, T. A conditionally heteroskedastic time series model for speculative prices and rates of return. *Rev. Econ. Stat.* **1987**, *69*, 542–547. [[CrossRef](#)]
79. Fuller, W.A. *Introduction to Statistical Time Series*, 2nd ed.; John Wiley and Sons: New York, NY, USA, 1996.
80. Lütkepohl, H.; Xu, F. The role of the log transformation in forecasting economic variables. *Empir. Econ.* **2012**, *42*, 619–638. [[CrossRef](#)]
81. McLeod, A.I.; Li, W.K. Diagnostic checking ARMA time series models using squared-residual autocorrelations. *J. Time Ser. Anal.* **1983**, *4*, 269–273. [[CrossRef](#)]
82. Wold, H. Path models with latent variables: the NIPALS approach. *Quantitative Sociology*; Elsevier: Amsterdam, The Netherlands, 1975. [[CrossRef](#)]
83. Piccolo, D. A distance measure for classifying arima models. *J. Time Ser. Anal.* **1990**, *11*, 153–164. [[CrossRef](#)]
84. Maharaj, E.A. A significance test for classifying ARMA models. *J. Stat. Comput. Simul.* **1996**, *54*, 305–331. [[CrossRef](#)]
85. Maharaj, E.A. Comparison of non-stationary time series in the frequency domain. *Comput. Stat. Data Anal.* **2002**, *40*, 131–141. [[CrossRef](#)]
86. Kakizawa, Y.; Shumway, R.H.; Taniguchi, M. Discrimination and clustering for multivariate time series. *J. Am. Stat. Assoc.* **1998**, *93*, 328–340. [[CrossRef](#)]
87. Vilar, J.A.; Pértega, S. Discriminant and cluster analysis for gaussian stationary processes: local linear fitting approach. *J. Nonparametric Stat.* **2004**, *16*, 443–462. [[CrossRef](#)]
88. Struzik, Z.R.; Siebes, A. The haar wavelet in the time series similarity paradigm. In Proceedings of the Principles of Data Mining and Knowledge Discovery. PKDD 1999. Lecture Notes in Computer Science, Prague, Czech Republic, 15–18 September 1999; Volume 1704, pp. 12–22. [[CrossRef](#)]
89. Galeano, P.; Peña, D. Multivariate analysis in vector time series. *Stat. Econom. Ser.* **2000**, *15*, 1–19.
90. Caiado, J.; Crato, N.; Peña, D. A periodogram-based metric for time series classification. *Comput. Stat. Data Anal.* **2006**, *50*, 2668–2684. [[CrossRef](#)]
91. Douzal, C.A.; Nagabhushan, P. Adaptive dissimilarity index for measuring time series proximity. *Adv. Data Anal. Classif.* **2007**, *1*, 5–21. [[CrossRef](#)]
92. Uspensky, V. A. and Shen, A. Ming Li and Paul Vitányi. An introduction to Kolmogorov complexity and its applications. Texts and monographs in computer science. *J. Symbolic Logic* **1995**, *60*, 1017–1020. [[CrossRef](#)]
93. Li, M.; Badger, J.H.; Chen, X.; Kwong, S.; Kearney, P.; Zhang, H. An information-based sequence distance and its application to whole mitochondrial genome phylogeny. *Bioinformatics* **2001**, *17*, 149–154. [[CrossRef](#)] [[PubMed](#)]
94. Li, M.; Chen, X.; Li, X.; Ma, B.; Vitanyi, P. The similarity metric. *IEEE Trans. Inf. Theory* **2004**, *50*, 3250–3264. [[CrossRef](#)]
95. Cilibrasi, R.C.; Vitányi, P.M. Clustering by compression. *IEEE Trans. Inf. Theory* **2005**, *51*, 1523–1545. [[CrossRef](#)]
96. Keogh, E.; Lonardi, S.; Ratanamahatana, C.A.; Wei, L.; Lee, S.H.; Handley, J. Compression based data mining of sequential data. *Data Min. Knowl. Discov.* **2007**, *14*, 99–129. [[CrossRef](#)]
97. Brandmaier, A.M. Permutation Distribution Clustering and Structural Equation Model Trees. Ph.D. Thesis, Universitat des Saarlandes: Saarbrücken, Germany, 2011. Available online: <https://publikationen.sulb.uni-saarland.de/handle/20.500.11880/26345> (accessed on 1 January 2021).

98. Alonso, A.M.; Berrendero, J.R.; Hernández, A.; Justel, A. Time series clustering based on forecast densities. *Comput. Stat. Data Anal.* **2006**, *51*, 762–776. [[CrossRef](#)]
99. Vilar, J.A.; Alonso, A.M.; Vilar, J.M. Non-linear time series clustering based on non-parametric forecast densities. *Comput. Stat. Data Anal.* **2010**, *54*, 2850–2865. [[CrossRef](#)]
100. Batista, G.; Wang, X.; Keogh, E. A complexity-invariant distance measure for time series. In Proceedings of the eleventh SIAM international conference on data mining, SDM11, Mesa, AZ, USA, 28–30 April 2011; pp. 699–710. [[CrossRef](#)]
101. Otranto, E. Clustering heteroskedastic time series by model-based procedures. *Comput. Stat. Data Anal.* **2008**, *52*, 4685–4698. [[CrossRef](#)]
102. Maharaj, E.A. Cluster of time series. *J. Classif.* **2000**, *17*, 297–314. [[CrossRef](#)]
103. Kalpakis, K.; Gada, D.; Puttagunta, V. Distance measures for effective clustering of arima time-series. In Proceedings of the Data Mining, 2001. ICDM 2001. IEEE International Conference, San Jose, CA, USA, 29 November–2 December 2001; pp. 273–280. [[CrossRef](#)]
104. Xiong, Y.; Yeung, D.Y. Mixtures of arma models for model-based time series clustering. In Proceedings of the 2002 IEEE International Conference on Data Mining, Maebashi City, Japan, 9–12 December 2002; pp. 717–720 [[CrossRef](#)]
105. Ioannou, A.; Fokianos, K.; Promponas, V.J. Spectral density ratio based clustering methods for the binary segmentation of protein sequences: A comparative study. *Biosystems* **2010**, *100*, 132–143. [[CrossRef](#)]
106. Harvey, A.C.; Shephard, N. 10 Structural time series models. In *Econometrics; Handbook of Statistics*; Elsevier Science Publishers B.V: Amsterdam, The Netherlands, 1993; Volume 11, pp. 261–302. [[CrossRef](#)]
107. Bühlmann, P.; McNeil, A.J. An algorithm for nonparametric GARCH modelling. *Comput. Stat. Data Anal.* **2002**, *40*, 665–683. [[CrossRef](#)]
108. Rohan, N.; Ramanathan, T.V. Nonparametric estimation of a time-varying GARCH model. *J. Nonparametric Stat.* **2013**, *25*, 33–52. [[CrossRef](#)]
109. Zou, H.; Hastie, T. Regularization and variable selection via the elastic net. *J. R. Stat. Soc. Ser. B Stat. Methodol.* **2005**, *67*, 301–320. [[CrossRef](#)]