RAIN PATTERN ANALYSIS AND FORECAST MODEL BASED ON GPS ESTIMATED ATMOSPHERIC WATER VAPOR CONTENT

A. Seco[(1)(*)], F. Ramírez[(1)], E. Serna[(1)], E. Prieto[(1)], R. García[(1)], A. Moreno[(2)], L. Miqueleiz[(1)], E. Priego[(3)]

[(1)]Dept. of Projects and Rural Engineering. Public University of Navarre. 31006 Pamplona, Spain. (E-mail: andres.seco@unavarra.es. Phone: 34948169682; Fax: 34948169148)

[(2)]Sistemas Inteligentes de Control y Gestión. Instituto Ibérmatica de Innovación (I3B) (E-mail: ai.moreno@ibermatica.com. Phone: 34948169682; Fax: 34948169148)

[(3)]Universidad Politécnica de Valencia. Camino de Vera, s/n - 46022 Valencia, Spain (E-mail: epriego@cgf.upv.es; Phone: 96-3877007 ext75595)

[(*)]Corresponding author

RAIN PATTERN ANALYSIS AND FORECAST MODEL BASED ON GPS ESTIMATED ATMOSPHERIC WATER VAPOR CONTENT

ABSTRACT

Rain is one of the fundamental processes of the hydrologic cycle as it can be the source of wealth or natural hazards. This experiment focuses in the relationship between rain occurrence and atmospheric pressure (Patm) and atmospheric water vapor content (PW), GPS estimated. The available nine years time series of each variable were analyzed. It allowed to state the existence of three rain patterns and monthly differences in the Patm-PW combinations. In spite of rain episodes take place only for some of the Patm-PW combinations, only these variables are unable to explain the rain occurrences because of not always they take place. This because a forecast sliding windows model with neural network was developed, to capture nonlinear relations that can not to be fully reflected by the lineal probabilistic ones based on the observed rains, Patm and PW series. This model stated a good correlation between the observed rains and the forecast, with a positive impact of the PW but negative of Patm. This model was able to predict the rain precipitation with a reasonable precision and reliable accuracy up to a 56 hours horizon.

KEYWORDS

GPS, water vapor, rain, forecast model, meteorology.

1. INTRODUCTION

From an environmental and a human point of view, rain is one of the fundamental processes of the hydrologic cycle. Rain is the water source of the natural vegetation and crops. It is also the origin of the majority of the domestic consumption water as well as in the industry, services, etc. Finally, it is possible to mention to rain like origin of natural hazards, both by the large periods of water absence and by the occurrence of torrential phenomena. In the Mediterranean climate zones this situation is especially evident: these zones are usually water needy, with little and irregular precipitations and with more or less frequent torrential episodes. These phenomena cause damages in environment and provoke great economic and human losses. For this reason, it is fundamental to advance in the knowledge of rain as a natural process, key for the management of this valuable resource and for the forecast of risks associated to extreme events. One of the key variables in the precipitation occurrence is the atmospheric water vapor content. Several studies have established the existence of large water vapor contents in the atmosphere previous to the intense precipitation occurrence in the Mediterranean area (Champollion et al., 2004; Cucurull et al., 2004; Brenot et al., 2006). Nevertheless aspects as the time between the atmospheric water vapor peak and the occurrence of rain or its intensity, they are not already satisfactorily stated, partly had to the complexity of the process and partly to the difficulty to determine the atmospheric water vapor content. GPS during this last decade became an instrument of great interest in meteorology. This is due to its proved effectiveness for the estimation of the water vapor content of the atmosphere, comparable in exactitude to the classic instruments in meteorology as radiosondes (RS) and the water vapor radiometers, without their limitations (Haase et al., 2003; Champollion et al., 2004; Vedel et al., 2004; Jade et al., 2005; Jin et al., 2007). Nowadays there are a great amount of GPS reference stations, which allows collecting atmospheric water vapor data, in all those zones in which this infrastructure exists. Where the series available are sufficiently long it is possible in addition the analysis of temporary patterns, long term trends, etc. (Gradanarsky et al., 2002; Jade et al., 2005; Jin et al., 2007).

This article shows an experimental analysis that establishes the relation between the variations of the water vapor content and the atmospheric pressure, with rains observed during the period 2002-2010 in the city of Pamplona, located in the north of Spain. Three patterns of precipitation occurrence have been observed and a model of rain forecasting, based on the time analysis of the atmospheric GPS water vapor content, has been developed.

2. MATERIALS AND METHODS

2.1. Location and local climatology

The selected location for this experiment is the city of Pamplona, in the north part of Spain (see figure 1), located at 440 meter above sea level.

FIGURE 1

From the climatic point of view, this city is located in a transition zone between the Atlantic climate and the Mediterranean one. During the winter period, several Atlantic disturbances from polar front, coinciding when the Azores anticyclone moves away, cause precipitation in a water or snow way. The appearance of winter anticyclones cause dry and cold days. During summer time temperature has a strong rise caused by the daytime warming and the weakly and the moving away of the draught stream due to the Azores anticyclone situated in a northern position. The high temperatures and the lack of rain during summer are only modified for the appearance of cold air at high altitude causing isolated storms which are the only precipitations of this period. The transition between winter and summer takes place gradually, with alternation of raining days caused by the Atlantic influence and with dry days caused by dorsal anticyclones (Gobierno de Navarra, 2001). Table 1 shows the more significant climatic parameters of Pamplona.

TABLE 1

2.2. Meteorological data.

As it has been stated in 2.1, the precipitation occurrences in Pamplona are related to changes in the Atmospheric Pressure (Patm), so both atmospheric variables were considered. RS atmospheric water vapor content data, expressed as Precipitable Water (PW), were considered as reference values. The nearest RS sites are Santander and Zaragoza cities, placed at 190 and 143 km respectively from Pamplona. All the meteorological data were supplied for the Spanish Meteorology Agency (AEMET). Precipitation and Patm were available with an hourly frequency meanwhile RS PW data are registered with an 12 hours interval (0000 UTC, 1200 UTC).

For the development of this experiment the following sets of data were available: (I) RS PW data for the period 2001-2010 at Santander, (II) 2006-2010 RS PW values at Zaragoza and (III) Hourly precipitation and atmospheric pressure data for the period 2001-2010 at Pamplona.

2.3. GPS atmospheric water vapor content determination

The effect of the troposphere causes a delay in the propagation of microwave signals that, in mid latitudes, results in a displacement in the zenith direction of about 2.4 m (Boehm et al., 2005). Total tropospheric delay is the sum of two components: first one is the hydrostatic component caused by the atmosphere dry gases, named Zenital Hydrostatic Delay (ZHD). This component is very steady and contributes with more than 90% of the total tropospheric effect. It is directly proportional to the ground atmospheric pressure and it can be determined with accuracy better than 1% using the Saastimoinen model (Saastamoinen, 1972). The second component is the wet one known as Zenital Wet Delay (ZWD). It is caused mainly by the atmospheric water vapor content and to a lesser extent to the liquid water in the atmosphere. It is responsible for most of the variations on the ZTD values and because of the heterogeneity of the water vapor content in the atmosphere it isn't possible its modelation with the required accuracy (Haase et al., 2003).

$$\mathrm{ZTD} = \mathrm{ZHD} + \mathrm{ZWD} \qquad [1]$$

The relation between Precipitable Water (PW) and ZWD is given by the formula

$$PW = \Pi \cdot ZWD \qquad [2]$$

where PW is the atmospheric water vapor content expressed as Precipitable Water in millimeters of water column and Π is an empirical term that is dependent on the average temperature of the atmospheric cross section (Haase et al., 2003; Champollion et al., 2004).

The GPS data used to monitor the PW content were 24-hour rinex files with a sampling rate of 30 s from the GPS reference stations available in Pamplona, located 500 m far from the meteorological station meteorological and Santander and Zaragoza GPS sites. Data were processed using the Gamit v-10.40 program (Herring et al., 2006). Gamit parameterizes the ZTD as a stochastic variation of the Saastimoinen model with piecewise linear interpolation in between the solution epochs. The variation is constrained to a Gauss-Markov process with an a priori power density. We used three 12-hour slicing window in each day to get hourly PW values, removing the first and the last values from each slicing window to avoid the edge effect of the Gauss-Markov process (Jin et al., 2007). The computation strategy was the usual in this kind of calculations, and it is shown in table 2 (Champollion et al., 2004; Cucurull et al., 2004; Brenot et al., 2006; Jin et al., 2007).

TABLE 2

3. RESULTS

3.1. GPS and radiosounding PW analysis

To verify the calculation's strategy goodness, RS PW registered series and the GPS PW estimated values were analyzed. In the case of Santander, the available series was 2001-2010 instead for Zaragoza it was 2006-2010.

Figure 2 shows the correlation between the GPS and the RS PW values, where R^2=0.83 for Santander and R^2=0.81 for Zaragoza were obtained.

FIGURE 2

These results evidence the goodness of the GPS used calculation strategy and the capacity of the GPS as atmospheric water vapor content estimator.

3.2. Rain, PW and atmospheric pressure relationship

As previously shown in 2.1, atmospherically speaking, unstable periods in Pamplona are associated to low pressures which make easy the coming in of the cool fronts from the north, causing precipitations. For this reason a study during the rains phenomena of the Patm and the water vapor content was carried out with the available data series. Three different patterns of these variables were established in this study during the rains phenomena. Figure 3 shows some examples of each type of the precipitation observed patterns.

FIGURE 3

TYPE 1

This type of rains takes place mainly in autumn and spring, during Patm decreasing phases associated to the coming in of fronts. The rains begin when Patm values are under 970 MBar. The mean rain intensities take values around 1.5 mm/h. The end of this type of rains coincides with a Patm relative minimum of around 960 Mbar. The usual duration of these kinds of rains is around 20-40 hours. During this period the atmospheric water vapor content remains stable with light variations mainly caused by the coming in of wet air from the Atlantic Ocean.

TYPE II

The second type of rains takes place mainly during the winter months. PW values are lower than in the case of type I rains because of winter period is cooler than type I months. In this case rains coinciding with a Patm increasing, beginning the rain around 950 Mbar, maintaining this situation until more or less 965 Mbar. The mean rain intensities also reach values around 1.5 mm/h and the duration of this kind of rains is also around 20-40 hours. During this time PW increases from values lower than 5 mm up to 10-15 mm because of the coming in of wet air from the sea, as in the case of type I.

TYPE III

The third type of rains appears coinciding with a minimum relative value of Patm and a maximum relative of PW. This type of rains occurs in any time and provokes rains with duration up to 10 hours and with larger intensities than in the types I and II, with values between 1.7 and 3.4 mm/h. Whereas in rains type I PW did not vary more than 5 mm and in type II showed an increase between 5-10 mm, in type III the precipitation had associated a PW decrease below 15 mm because the loss of humidity in the atmosphere during the rain it is not compensated with an entrance of humidity from the sea associated to the entrance of a new front.

The representativity of these types of precipitation patterns was analyzed with a joint representation of Patm, PW and hourly rain data series as it is showed in figure 4.

FIGURE 4

Figure 4 shows the characteristic triangular shape of the yearly PW-Patm data. Low PW values can be observed for all range of Patm, but maximum PW values are only observed with Patm around 965 Mbar. Between the whole set of data it can be observed that the PW-Patm records of the each month are distributed predominately in specific zones of the graph, being only some of these combinations which register rain. That way, the corresponding data from December and January are concentrated in the left zone of the graph. February, March, April and May show more dispersion in their register with a gradual movement to mean Patm and largest PW values zone, which are clearly observed during the month of July, August and September. Since October the dispersion of the data is going up and they start to move to the left zone of the graph as winter is arriving. At the same time, the precipitation data are concentrated in specific zones of PW and Patm combinations for each month. This situation is especially clear for the months of June, July and August where rain episodes, over all the more intenses, are located in determinate zones. In this time of the year precipitations correspond to type III. Types II and III, take place mainly during spring and autumn months, have smaller intensities and show a strong dispersion across the Patm-PW combinations. The concentration of the rains in the Patm-PW graph has been characterized by the circles that, with a minimum section, include 50 % of the total amount of rain for each month. Graphical representation of these circles is showed in figure 4 and numerical data are showed in table 3.

TABLE 3

This table shows the Patm-PW coordinates of the circles's centers, their radius and the observed probability of rain for the circles' area. The changes of the position of the circles's centers along the year agree with the observed rain pattern. The smaller radius of the circles takes place in the summer months, demonstrating a bigger concentration of the summer rains in specific Patm-PW combinations. The biggest radius are those corresponding to spring and autumn months, that also agree with the rains' dispersion observed for these epochs. Finally, the rain probability of the occurrence of rain in the circles takes values from 3.3% for September to 33.2% for December. These values are small, demonstrating that, although rains show a relationship with Patm and PW and their occurrence is few probable except for their specific values, these variables are not able to explain by themselves the occurrence of rain episodes.

4. TIME SERIES BASED RAIN FORECAST MODEL

As it was stated in the previous section, Patm and PW combinations were not able to explain completely the occurrence or not of rain events. This because, and having nine data years, rain time series and its relationship with the considered variables were analyzed. The choosed method was sliding windows with neural network to capture nonlinear relations that can not to be fully reflected by the lineal probabilistic ones (Johnson and Padilla, 2004; Daza, 2008; Mellit, 2008; Parisi, 2008; Roldao, 2008).

In this study, a 24 hours sliding window was analyzed. For each window, a neural net classifier runs in order to predict the next window. This neural net operator learned a model by means of a feed-forward neural network trained by a backpropagation algorithm (multi-layer perceptron). The used activation function was the usual sigmoid function. The values ranges of the attributes were scaled to -1 and +1 by this operator in order to normalize the neural inputs. With this classifier method, as shown in the figure 5, it is possible to follow the object series (precipitation), and its prediction based in Patm and PW series.

FIGURE 5

Visually, it is clear a correlation between the trends of the observed rain and the prediction model, but not in the absolute amount of the picks. So, it was studied the direct correlation between the real and predict series, and the no-lineal impact of the Patm, and PW values with the rain precipitation. For a 10 hours horizon window, the correlation model between observed rain and its prediction showed the following results: with both Patm and PW R^2 = 0.686, without Patm and with PW R^2 = 0.764 and with Patm and without PW R^2 = 0,568. (see figure 6)

FIGURE 6

This shows, from a mathematical point of view, a negative impact of Patm in the prediction developed model. The model is a time series of a linear component (temporal sequence) and other non linear more or less pronounced (relation between PW and the precipitation), and it makes sense the use of neural network techniques in order to capture nonlinear relations that have not been fully reflected by order methods.

Checking different scenarios of accuracy, the best fits correlation trends without the Patm attribute were studied for different horizon windows. Horizon window is the data series that we want to predict, so, in a larger window, there will be more data to discover, and the accuracy will be worst.

With this premise, there are takings trends correlations with different horizons, specifically, 10 hours, 24, 48 and 56, and on the whole case, with all attributes (year, moth, hour, PW and rain precipitations), except Patm.

FIGURE 7

Results of the correlation between the target series and the prediction ones are in figure 7. In the worst cases, with the biggest horizon window (56 hours), the correlation R^2 between the forecast rain precipitation and the real precipitation is greater than 0.563, and than the regression coefficient, $B_0 = 0.9$, which is statistically significant. It means that the rain trend can be estimated from its corresponding temporal behaviour and the no-linear relationship between the others attributes, PW, principally. So we can predict when it will be a strong variation in the temporal series and take account when will be a peak in the time series, in order to predict the strength of that variation, not in absolute terms, but the pulse power.

5. CONCLUSIONS

Across this experiment, the relationship between rain and GPS estimated water vapor content has been analyzed for the nine years available data series in Pamplona (Nort of Spain). The analysis of the Pamplona's Patm-PW-Rain series agreed with the fact that from a meteorological point of view, the occurrence of rain events in the study area is usually lied to atmospheric low pressures and water vapor entries, caused by Atlantic disturbances. Thus, three precipitation occurrence patterns were stated, associated to the different Patm-PW behavior across the year's epochs. The plot of the Patm-PW and rains combinations for each month showed differences between the different seasons across the year. Winter and summer month's demonstrated to have quite clear characteristic Patm-PW patterns, meanwhile during the spring and autumn months, Patm-PW patterns are gradually changing from winter to summer or vice versa. Not in all the possible Patm-PW combinations, rains take place. Mainly on winter and summer months rains appear at specific Patm-PW combinations. During spring and autumn months Patm-PW and rain patterns evidence the winter-summer transitions with not so clear Patm-PW-rain pattern. In spite of the three Patm-PW rain patterns observed and the monthly Patm-PW patterns, Patm and PW combinations were unable to explain by themselves the rain episodes occurrence because of, although rains only take place on determined Patm-PW combinations, not always at those combinations rains take place. To analyze the Patm-PW-rains relationship, a time series forecast model based on the observed rains, Patm and PW series was developed. This model stated a good correlation between the observed rains and the forecast, with a positive impact of the PW but negative of Patm. This unexpected result could have its origin in the yearly variability of the Patm-PW combinations' values when rains take place. The model based on the PW values, the temporal behaviour of the rain occurrences, and the temporal information (year, month, day), was able to predict the rain precipitation with a reasonable precision and reliable accuracy up to a 56 hours horizon.

REFERENCES

Brenot, H., Ducrocq, V., Walpersdorf, A., Champollion, C., Caumont, O. GPS zenith delay sensitivity evaluated from high-resolution numerical weather prediction simulations of the 8-9 September 2002 flash flood over southeastern France. Journal of Geophysical Research, Vol. 111, 2006.

Boehm, J., Cerveira, P.J.M., Schuh, H., Tregoning, P. The impact of tropospheric mapping functions based on numerical weather models on the determination of geodetic parameters. IAG Proceedings, 2005.

Champollion, C., Masson, F., Van Baelen, J., Walpersdorf, A., Chéry, J. GPS monitoring of the tropospheric water vapour distribution and variation during the 9 September 2002 torrential precipitation episode in the Cévennes (southern France). Journal of Geophysical Research, Vol. 109, 2004.

Cucurull, L., Vanderberghe, F., Barker, D., Vilaclara, E., Rius, A. Three-Dimensional Variational Data Assimilation of Ground-Based GPS ZTD and Meteorological Observations during the 14 December 2001 Storm Event over the Western Mediterranean Sea. Monthly weather review. Vol 132, pp: 749-763, 2004.

Daza F., 2008, Demanda De Agua En Zonas Urbanas En Andalucía, Departamento De Estadística, Econometría, I.O Y Organización De Empresas (Área De Organización De Empresas), Servicio De Publicaciones De La Universidad De Córdoba, ISBN-13: 978-84-7801-928-1

Gobierno de Navarra, Estudio Agroclimático de Navarra, Gobierno de Navarra, 2001.

Gradinarsky, L.P., Johansson, J.M., Bouma, H.R., Scherneck, H.G., Elgered, G. Climate monitoring using GPS. Physics and Chemistry of the Earth, N.27, pp:335-340, 2002.

Haase, J., Ge, M., Vedel, H., Calais, E. Accuracy and Variability of GPS Tropospheric Delay Measurements of Water Vapour in the Western Mediterranean. Journal of Applied Meteorology, Vol. 42, pp:1547-1568, 2003.

Herring, T.A., King, R.W., McClusky, S.C. GAMIT Reference Manual. Department of Earth, Atmospheric, and Planetary Sciences, 182 pp, 2006.

Jade, S. Vijayan, M.S.M., Gaur, V.K., Prabhu, T.P., Sahu, S.C. Estimates of precipitable water vapour from GPS data over the Indian subcontinent. Journal of Atmospheric and Solar-Terrestrial Physics, vol. 67, pp:623-635, 2005.

Jin, S., Park, J.-U., Cho, J.-H., Park, P.-H. Seasonal variability of GPS-derived zenith tropospheric delay (1994-2006) and climate implications. Journal of Geophysical Research, Vol. 112, 2007.

Johnson C.A., Padilla M.A., 2004, Regularidades no lineales en índices accionarios: una aproximación con redes neuronales. Escuela de Negocios, Universidad Adolfo Ibáñez, Santiago, Chile. Instituto de Ciencias Humanísticas y Económicas de la Escuela Superior Politécnica del Litoral (ESPOL), Guayaquil, Ecuador.

Mellit, S.A. Kalogirou, S. Shaari, H. Salhi, A. Hadj Arab. Methodology for predicting sequences of mean monthly clearness index and daily solar radiation data in remote areas. Renewable energy 33 (2008) 1570-1590.

Nansen, J.V.; Nelson, R.D. (2002). Data mining of time series using stacked generalizers. Neurocomputing, 43, 173-184

Parisi A., 2008, Estabilidad De Modelos De Redes Neuronales Artificiales Como Herramienta De Predicción Del Signo De La Variación Del Ipsa. Universidad de Chile. Escuela de Postgrado Economía y Negocios.

Roldao, A.M., 2008, Comparative study of artificial neural network and Box-Jekins Arima for Stock Price Indexes. Instituto Superior de Ciencias do Trabalho e da Empresa.

Saastamoinen, J. Atmospheric correction for the troposphere and stratosphere in radio ranging of satellites, in The Use of Artificial Satellites for Geodesy, Geophys. Monogr. Ser., Am. Geophys. Un., Washington.D.C., 15, 245-251, 1972.

Sfetsos, A.; Siriopulos, C. (2004). Combinational time series forecasting based on clustering algorithms and neural networks. Neural Computation and Applications 13

Vedel, H., Huang, X.-Y., Haase, J., Ge, M., Calais, E. Impact of GPS Zenith Tropospheric Delay data on precipitation forecasts in Mediterranean France and Spain. Geophysical Research Letters, vol. 31, 2004.

Figure

TABLE 1. Pamplona climatic station location and climatic parameters used in this study. (Gobierno de Navarra, 2001).

Latitude (UTM)	4736171
Longitude (UTM)	611323
Elevation	440
Annual Average Precipitation (mm)	701.3
Average Temperature (degrees)	12.6
Maximum Continental Index (Gorzinsky)	18.9
Köppen Index	Cfb
Gaussen Index	Sb
Potential evapotranspiration (PET), Thornthwaite Index	707.0
Agroclimate classification by Papadakis.	Wet Mediterranean (ME)

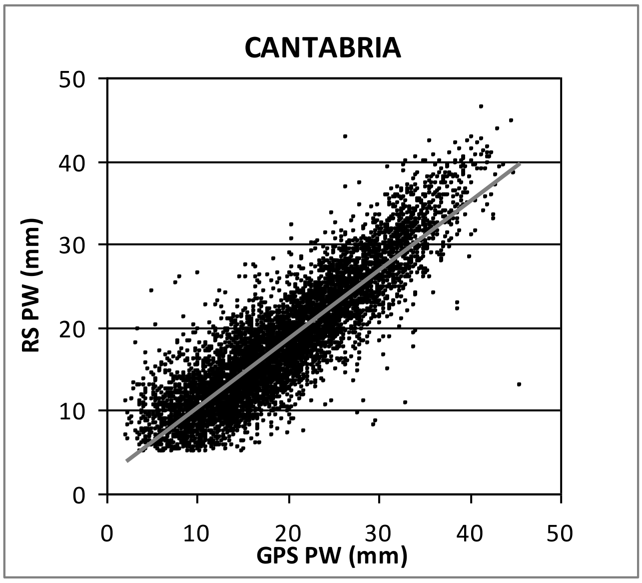
TABLE 2. Defining parameters of the computation strategy.

GPS observable	LC
ZTD interval	1 hour
Gauss-Markov process	Constraints 0.5 m Power density 2cm/hour$^{1/2}$ Correlation time 100 hours
Ephemeris	IGS Final
Earth Orientation Parameters	USNO
Reference Frame	ITRF 2000
Cutoff angle	10º
Sampling rate	30 s
Mapping functions	Global Mapping Functions
Coordinates constraints	Loose
Variation Model of the antenna phase center.	Absolute antenna phase centers

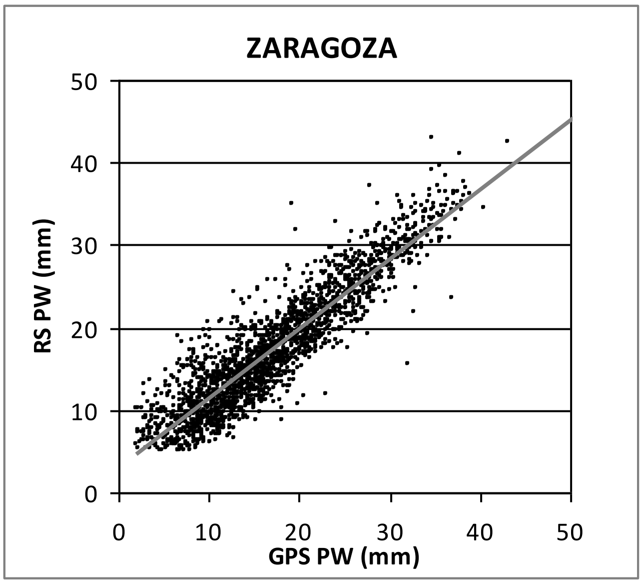
TABLE 3. Numerical parameters of the circles containing 50% of the monthly rains.

Month	Center		Radius	Rain Probability
	X (Pw)	Y (Patm)		
January	10.0	956.9	7.959	23.8
February	8.9	954.2	9.250	21.2
March	12.7	959.0	7.897	19.9
April	15.8	955.6	6.960	11.1
May	20.3	957.5	6.021	13.1
June	28.2	959.3	5.395	12.1
July	34.8	961.8	5.719	7.9
August	29.4	960.8	6.834	4.8
September	29.7	956.9	5.909	3.3
October	21.1	955.3	8.192	13.1
November	13.6	956.6	6.140	11.2
December	11.7	954.9	8.605	27.1

FIGURE 1. Location of the city of Pamplona, Santander and Zaragoza.



a)



b)

FIGURE 2. Comparison of Radiosounding and GPS PW values during the available period. a) Santander 2001-2010 and b) Zaragoza 2006-2010

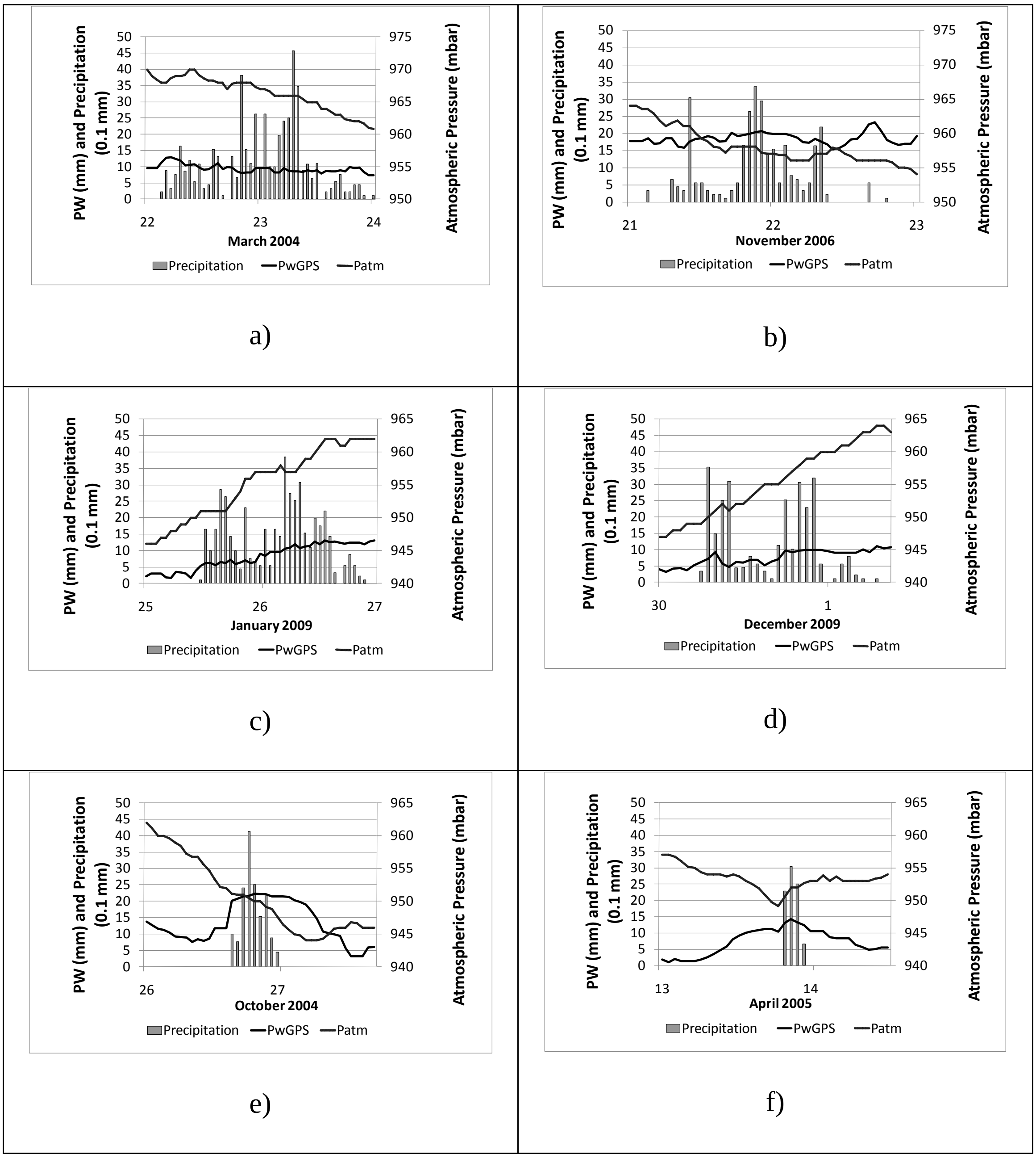


FIGURE 3. Different behavior of Patm-PW-Rain in Pamplona City. a) and b) Type I, c) and d) Type II, e) and f) Type III .

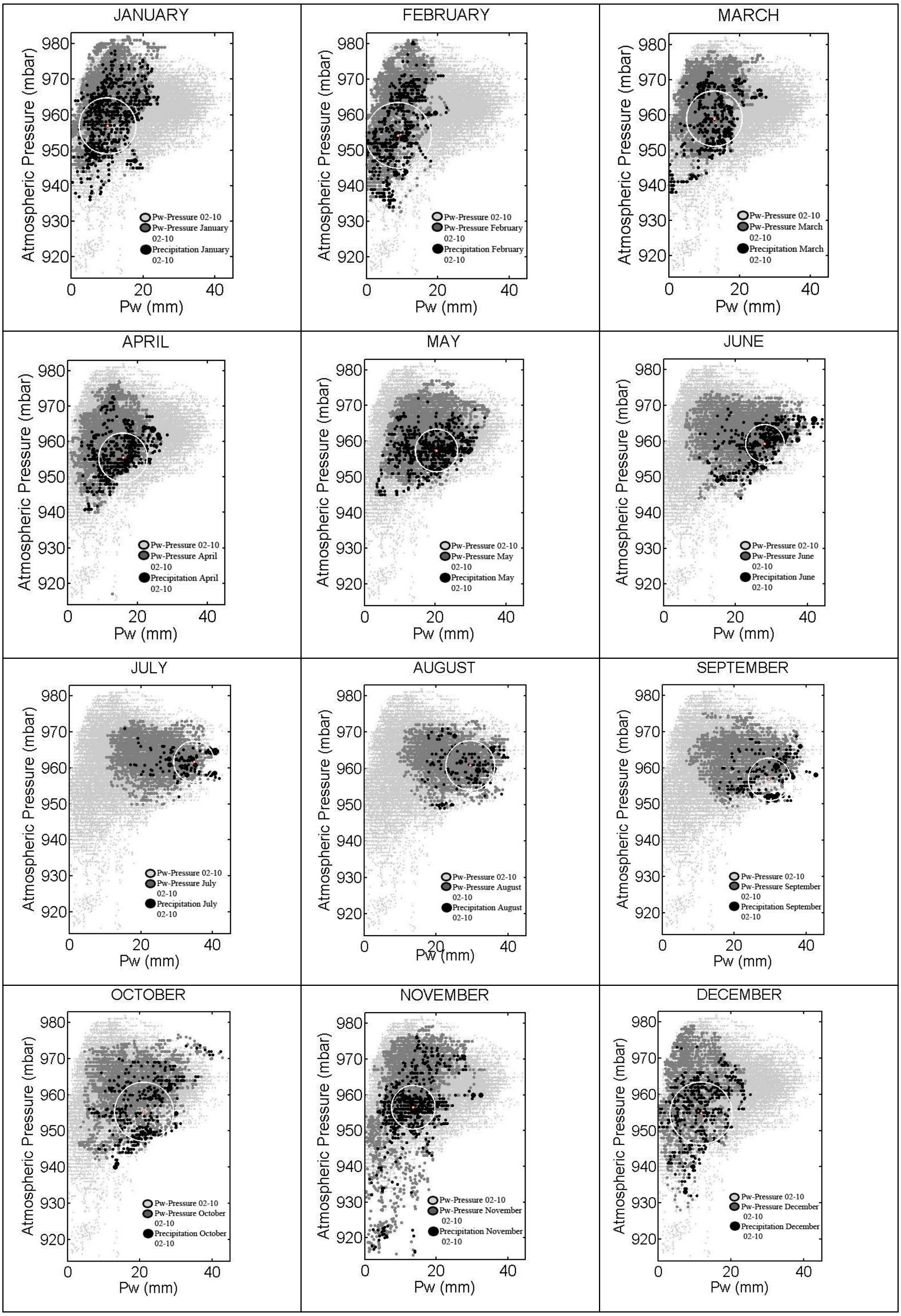
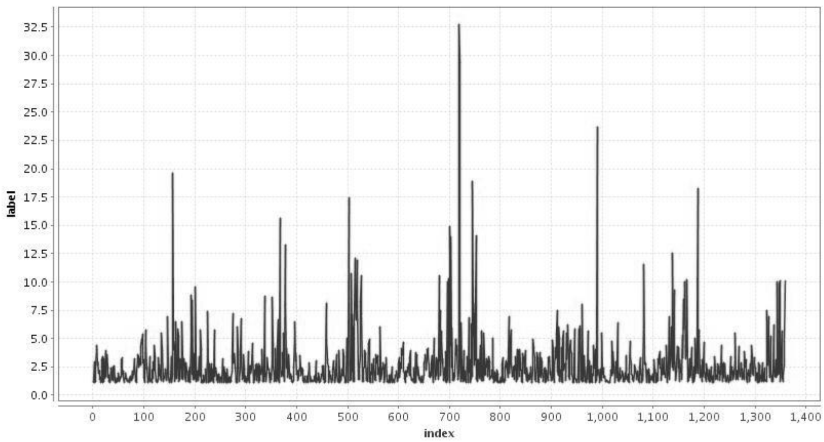
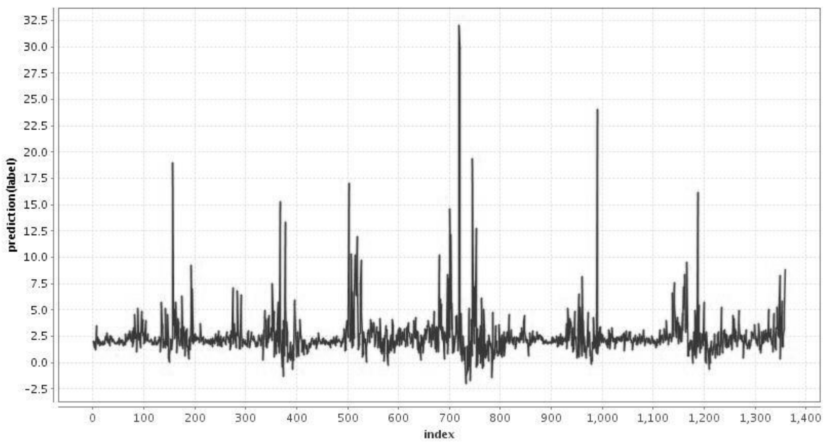


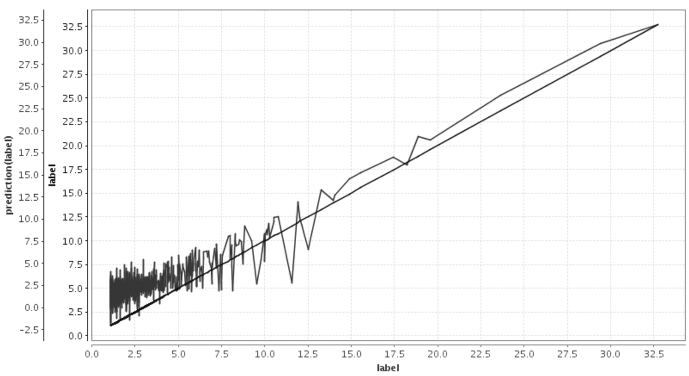
FIGURE 4. Hourly Patm-PW values distribution for the available serie (2001-2010). Over the whole set of data the values of each month and the observed rains are showed. Circles represent the minimum area including the 50% of the monthly observed rains.



a)

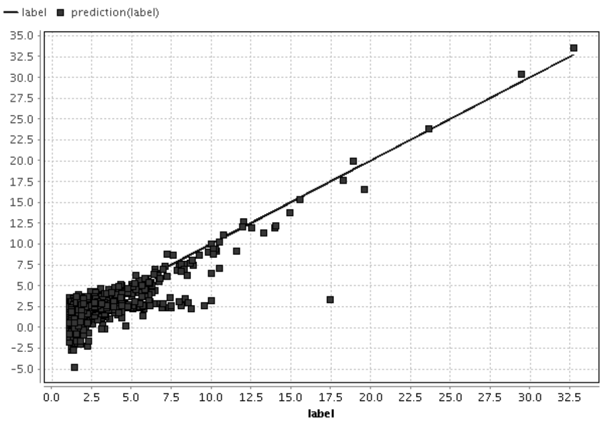


b)

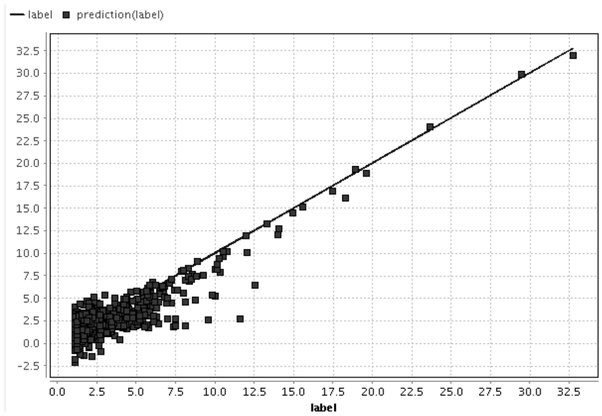


c)

FIGURE 5. Temporal trends and correlation between them.



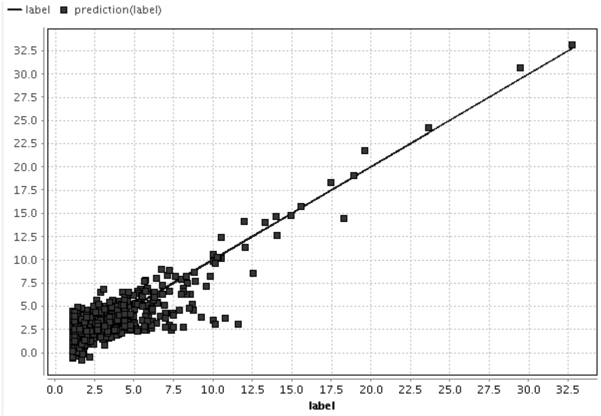
a)



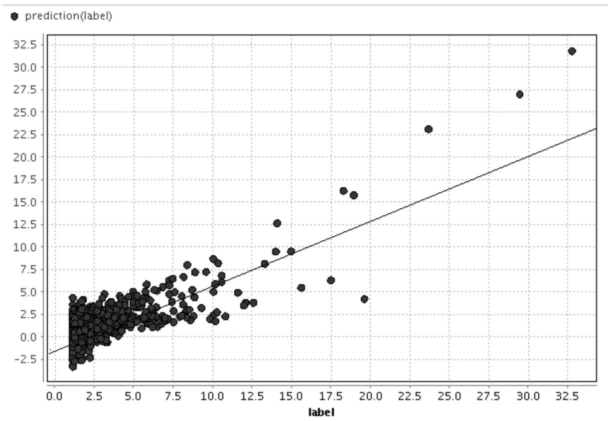
b)

c)

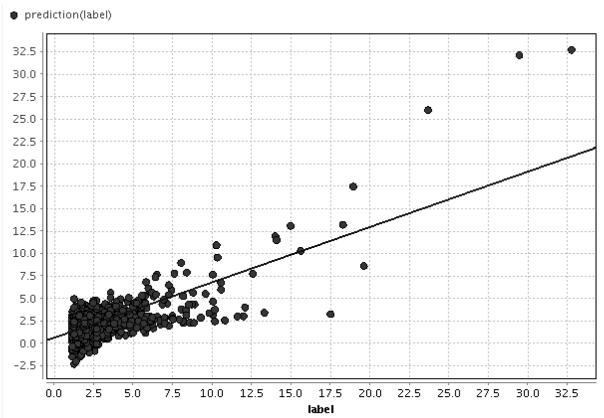
FIGURE 6. Trend series correlation, with Patm or without Patm.



a)



b)



c)

FIGURE 7. Correlations coefficients in different predict horizons.