Simultaneous Modelling of Rainfall Occurrence and Amount using a Hierarchical Nominal-Ordinal Support Vector Classifier

P. A. Gutierrez [a], S. Salcedo-Sanz* [b], J. Sánchez-Monedero [a] ,

C. Casanova-Mateo [c] and C. Hervás [a]

[a] *Department of Informatics and Numerical Analysis, Universidad de Córdoba.*

[b] *Department of Signal Processing and Communications, Universidad de Alcalá.*

[c] *Department of Applied Physics, Universidad de Valladolid, Valladolid, Spain.*

Abstract

In this paper we propose a novel computational system for simultaneous modelling of rainfall occurrence and amount. The proposed system is based on a hierarchical system of Nominal-Ordinal Support Vector Classifiers, the former to set the rainfall occurrence, and the latter to obtain the expected rainfall amount from a set of four different ordinal classes. In addition to the proposed model, we use a novel set of predictive meteorological variables, which improve the classifiers performance in this problem. We evaluate the proposed system in a real problem of rainfall forecast at Santiago de Compostela airport, Spain, where we have shown that the system is able to obtain an accurate prediction of occurrence and rainfall amount, and we discuss the usefulness of the proposed system as part of the airport weather forecast and warning system, in order to improve airport operations.

Key words: Rainfall occurrence; rainfall amount; nominal and ordinal classifiers;

Preprint submitted to Elsevier Science *18 December 2013*

ordinal regression.

* Corresponding author: Sancho Salcedo-Sanz. Department of Signal Theory and Communications, Universidad de Alcalá. 28871 Alcalá de Henares. Madrid. Spain.

Tlph: +34 91 885 6731

Fax: +34 91 885 6699

Email address: sancho.salcedo@uah.es (J. Sánchez-Monedero).

1 Introduction

Rainfall modelling is a very important problem that arises in many applications such in Agriculture [1], water resources management [2,3] or facilities maintenance and control [4], among others [5]. Currently, numerical weather prediction models have improved their performance, but they are still unable to provide accurate models for expected precipitation amount at high spacial and time resolutions. Different previous works have applied Soft-Computing approaches to overcome this difficulty, mainly based on neural networks and related approaches. These approaches have several advantages over global numerical models: they are much more easy and fast to train, can be applied to data from a specific point of measurement, and their performance is really competitive compared to global techniques.

Neural computation models for precipitation prediction started to be applied about twenty years ago [6–8]. Some of these first works applied multi-layer perceptrons to a set of predictive variables, carefully chosen to be related to rainfall, and with data from precipitation gauges (pluviometers) to obtain rainfall quantity [8,10,11]. The majority of these approaches considered short-term precipitation prediction, from 6 hours to 24 hours time-horizons, obtaining good results in the prediction [12]. There are other approaches focused on long-term rainfall prediction and precipitation trends in a given zone, such as [13], where the rainfall trend in the southern part of Indian Peninsula is analyzed by using an Adaptive Basis Function Neural Network with a backpropagation training algorithm. In [14] a multi-layer perceptron is applied to a problem of long-term precipitation prediction in California. More recently, in [15] an artificial neural network has been applied to model and forecast precipitation in Athens, Greece. In [16] a neural network was applied to forecast precipitation during the summer Monsoon station in India, using El Niño South Oscillation (ENSO) indices. In [17] a neural computation approach is applied to the short-term forecasting of thunderstorms rainfall.

Alternative classification and regression techniques have also been applied to problem of rainfall prediction and modelling. In [18] a comparison of machine learning algorithms (decision trees (DT), neural networks (ANN) and Support Vector Machines (SVMs)) has been carried out for a short-term precipitation prediction problem in Thailand. In [19] a hybrid SVM for regression with particle swarm optimization was applied to a problem of rainfall prediction. In [20] a SVM approach with different kernel functions is presented to predict monthly rainfall in a region of China. In [21] an novel wavelet-SVM approach was applied to precipitation forecasting from past data. SVMs have also been recently applied to precipitation related studies, such as precipitation downscaling [22,23] or streamflow prediction [37].

In spite of this huge work on rainfall prediction, there are not many papers focussed on the modelling and forecast of precipitation occurrence and amount together. There are two main articles dealing with this problem. In [24] several types of neural network models are applied to solve a problem of rainfall occurrence and amount modelling in northwest and southeast of England. The input data of this study are different measurement stations and also some large-scale climate predictors such as atmospheric circulation, thickness or moisture content at the surface, 850 and 500 hPa. More recently, in [25] a simple model for modelling rainfall occurrence and amount simultaneously has been proposed. It is based on a tweedy generalized linear modelling and the authors show that it performs well in modelling both occurrence and precipitation amount in Australia. Data from over 200 measurement stations spread all over Austria are used are inputs to the model. The use of joint models for simultaneous modelling of rainfall occurrence and amount is a hot topic in hydrology, since it provides information that can then be used in agriculture production systems and other applications.

In this paper we propose a novel system for simultaneous modelling of rainfall occurrence and amount, based on a hierarchical classifier, composed of a nominal and ordinal SVM classifier. First, a nominal SVM is used to set the rainfall occurrence model. A second ordinal SVM is then hybridized with the previous nominal classifier, in order to obtain the expected rainfall amount from a set of four different ordinal classes. In addition to the proposed model, we use a novel set of predictive variables, which improve the classifiers performance in this problem. First, we consider significant meteorological variables from atmospheric soundings. We also include as predictive variable the synoptic configuration of the atmosphere (synoptic situation using Hess-Brezowsky classification), that, to our knowledge, has not been either considered in precipitation prediction studies with machine learning techniques, in spite of its significance to establish precipitation regimes in mid-latitude regions [26]. We also evaluate the importance of other predictive variables such as humidity and Equivalent Potential Temperature (both measured in vertical soundings), and groups of these variables in the proposed hierarchical SVM performance. Regarding the objective variables, real rainfall data from a measurement station at Santiago de Compostela (Airport), Spain, are considered to establish the performance of the proposed system.

The rest of this paper is structured as follows: next section presents a review of the main predictive variables and precipitation data used in the study.

We also estate the exact modelling carried out, which includes the estimation/forecasting of rainfall occurrence and amount in the next 12 hours. Section 4 presents the proposed nominal and ordinal SVM bank for rainfall modelling. Section 5 presents the experimental part of the paper. Finally, we give some concluding remarks for closing the paper in Section 6.

2 Predictive and objective variables used

Rainfall requires the existence of adequate clouds to produce precipitation. Therefore in order for precipitation to occur, three basic factors should be combined in an adequate way: condensation nuclei, enough water vapor (moist) and vertical movements (updrafts and downdrafts as well as the atmospheric stability). As a consequence, data selection should cover all these three elements so as to obtain a robust group of predictive meteorological variables related to the physical processes involved in the production of precipitation. Fortunately, an adequate number of condensation nuclei (such as smoke from industrial, particles of salt, etc.) on which water vapor undergoes condensation to form water droplets or deposition to form ice crystals are almost always present in the atmosphere. Then, it is only necessary to select meteorological variables related to the presence of enough water vapor and vertical movements.

As has been shown in some studies [8,10], it is difficult to determine the criteria that should be followed to select the best set of meteorological variables to use in machine learning classifiers, based solely on an understanding of the physical mechanism of precipitation. Moreover, because precipitation is highly dependent on small-scale processes and local geography [27] a standardized pool of meteorological variables to forecast precipitation would be difficult to set. Nevertheless, considering the satisfactory results obtained in [8] using neural networks and those obtained in [10] using a neural approach with back-propagation training, it is possible to choose a reasonable group of meteorological variables following similar criteria.

In our study we combine observed variables, taken from an upper air sounding station, and meteorological variables derived from a numerical weather prediction model, plus the observed precipitation.

As mentioned before, observed precipitation (target variable) data was obtained from Santiago de Compostela Airport ground automatic station (latitude: 42.89; longitude: -8.41; altitude: 370 m). We chose this target area because Santiago de Compostela is located in one of the rainiest area of the Iberian Peninsula, without a dry season and with an average annual precipitation of 1886 mm [9]. This station is part of the State Meteorological Agency of Spain (AEMET) surface observing network and reports all meteorological data every 10 min (it calculates the average value for each meteorological variable every 10 min). Although the data are available on a ten-minute basis, we consider the rainfall prediction in a time horizon of 6 hours. Thus, the precipitation data's temporal resolution selected for this study is 6 hours. The meteorological data and variables employed for this study span the dates from 1st September 2009 to 31st August 2010, i.e., this study covers the 2009-2010 hydrological year.

We have used different predictive variables in order to predict precipitation occurrence and amount. Data from La Coruña (latitude: 43.36; longitude: -8.41; altitude: 67 m) radiosonde station, which is the nearest upperair station to our study area. This station belongs to AEMET and its data are freely available on the Internet [28]. The second set is formed by data from the medium-range global prediction model GFS (Global Forecast System) maintained by the National Center for Environmental Prediction (USA) [29]. In this case, the variables were taken at the grid point closest to the ground station used in this study. Likewise, the reason for using this numerical weather model is that data from GFS are freely available on the Internet. In addition, in this study we have used a novel predictive variable, trying to get better results in the forecast precipitation model proposed: the synoptic situation. As it is well known, some atmospheric circulation patterns promote precipitation whereas others make it difficult. In fact, some recent studies have been devoted to determine the more probable weather patterns that cause rainfall as well as the possible changing influence of the atmospheric circulation on surface precipitation [30,31]. In line with this idea, we have selected the subjective Hess-Brezowsky classification [32] of large scale circulation patterns as another predictive variable. This classification has shown its ability to improve the skill of a predictive model in a problem of daily maximum temperature prediction using a support vector regression algorithm [33]. In order to take into account water vapor content in the atmosphere, we have selected as predictors the meteorological variables shown in Table 1, whereas the meteorological variables chosen to determine updrafts and downdrafts as well as the atmospheric stability are shown in Table 2.

The target variable is the observed precipitation, which in this work is considered as a continuous variable describing the amount of rainfall in mm within a 6 hours interval. This variable has been discretized in four classes in order to transform the problem into an ordinal classification problem. It can be argued that the problem can be tackled as a standard regression problem, however the large amount of rain values equal to zero is a handicap for applying a regressor algorithm. The rainfall amount is mapped to different classes according to Table 3.

Briefly, an ordinal classification problem, also known as ordinal regression, is a supervised classification problem in which there is an order arrangement between categories. That order is often induced by the problem nature, as it is the case since $\{C_1 \prec C_2 \prec C_3 \prec C_4\}$ (see Table 3). Ordinal classifiers exploit this relationship of the data with the goal of improving performance. However, this performance cannot be measured as in nominal classification tasks, here, in addition to the error rate, the magnitude of the error should be considered. For instance, if we have a new unseen pattern of class C_3, an error classifying it as C_1 is more severe than classifying the pattern as C_4. For this reason specific performance metrics should be used (see Experimental Section).

3 Background

This section briefly introduces computational intelligence methods that are necessary to understand the paper proposal.

3.1 Support Vector Machine for Nominal Classification

The SVM [35,36] is perhaps the most common kernel learning method for statistical pattern recognition. The basic idea behind SVMs is to find a hyperplane that separates two different classes – positive and negative classes. This hyperplane, $b + \mathbf{w} \cdot \mathbf{x}$, is specified by its normal vector $\mathbf{w}$ and the bias b. The

SVMs overcome the limitations of the linear models by working with the patterns via a mapping function ϕ which transforms the patterns representation in the attributes or input space $\mathcal{X}$ to a high dimensional Reproducing Kernel Hilbert Space (RKHS). The reproducing kernel function is used, defined as $k(\mathbf{x}, \mathbf{x}') = \langle \phi(\mathbf{x}) \cdot \phi(\mathbf{x}') \rangle$, where $\langle \cdot \rangle$ denotes inner product in the RKHS.

Then, the hyperplane can be given as $\langle \mathbf{w} \cdot \phi(\mathbf{x}) \rangle + b = 0$, what yields the corresponding decision function:

$$f(\mathbf{x}) = y^* = \operatorname{sgn}\left(\langle \mathbf{w} \cdot \phi(\mathbf{x}) \rangle + b\right), \tag{1}$$

where $y^* = +1$ if $\mathbf{x}$ belongs to the corresponding class and $y^* = -1$ otherwise.

SVMs are linear models, based on a linear combination of a kernel function evaluated at the training data points. The solution to the problem of finding the maximum separating hyperplane is proven to be a convex optimization problem with a single global optimum. This optimization process implicitly selects a subset of patterns for building the model, which are know as *support vectors*. The initial formulation of SVMs is known as the hard-margin approach, which tends to suffer overfitting. Latter approaches included the concept of softmargin in order to better genereralize in the presence of noise, outliers or pre-labeling errors, which are common in real world problems. The soft margin is achieved with the inclusion of slack-variables ξ_i in the optimization process [36].

As Vapnik [36] shows, the optimal separating hyperplane is the one which maximizes the distance between the hyperplane and the nearest points of both classes (called margin) and results in the best prediction for unseen data. In this way, the optimal separating hyperplane with maximal margin can be formulated as the following Quadratic Programming (QP) problem:

$$\min_{\mathbf{w}\in\mathbb{R}^n,\boldsymbol{\xi}\in\mathbb{R}^n} L(\mathbf{w},\boldsymbol{\xi}) = \frac{1}{2}\|\mathbf{w}\|^2 + C\sum_{i=1}^{n}\xi_i, \tag{2}$$

subject to:

$$y_i \cdot (\langle \mathbf{w} \cdot \phi(\mathbf{x})\rangle + b) \geq 1 - \xi_i,\ \xi_i \geq 0,\ \forall i = 1, \cdots, n, \tag{3}$$

where y_i is the class of the input pattern $\mathbf{x}_i$.

In order to deal with the multiclass case, a "1-versus-1" approach can be considered, following the recommendations of Hsu and Lin [42]. The idea is to construct a binary classifier per each pair of classes and joining their multiple responses to obtain a final prediction.

Finally, among other specific issues, SVMs can have problems when dealing with imbalanced data (i.e. the number of patterns of each class significantly differs). This can lead to models that tend to ignore minority populated classes. The rainfall prediction problem is a clear example of imbalanced dataset, where the non-rain case is much more frequent than the rain cases. For dealing with imbalanced datasets, recently, the Cost Support Vector Classifier (CSVC) has been proposed [41]. In this case, different missclassification costs are assigned to each class, so the total misclassification cost $C\sum_{i=1}^{n}\xi_i$ is replaced with two terms:

$$C\sum_{i=1}^{n}\xi_i \rightarrow C_+\sum_{i\in I_+}\xi_i + C_-\sum_{i\in I_-}\xi_i, \tag{4}$$

being C_+ and C_- the soft-margin constants for positive and negative samples and I_+ and I_- the sets of positive and negative samples. This constants are set in such a way that the total penalty for each class should be equal [43], this is:

$$C_+ n_+ = C_- n_-, \tag{5}$$

where n_+ and n_- are the number of positive and examples.

3.2 Support Vector Machines for Ordinal Regression (SVOR).

The SVM formulation has been ported to the ordinal classification case (SVOR). In this case, classes are separated by different thresholds b_j and the QP problem is adapted [39]. In contrast to the binary case, where the class of the pattern is determined by the sign of the projection $\mathbf{w}^T \cdot \mathbf{x}$, the corresponding real line will be split into different intervals by using a threshold vector $\mathbf{b}$. This defines a set of parallel hyperplanes with the same $\mathbf{w}$ and different thresholds b_j.

In this paper we will work with SVOR with Implicit constraints of Chu and Keerthi (SVORIM) [40]. In contrast to the binary case, where only a pair of classes contributes to the error when finding the separating hyperplane, SVORIM redefines the QP problem for considering errors from the samples of all the categories when defining each hyperplane. In this way, the ordinal inequalities on the thresholds are *implicitly* satisfied at the optimal solution.

4 Proposed Hierarchical Nominal-Ordinal SVM

This paper proposes to address the rainfall prediction problem as an ordinal regression problem that will be tackled by using a hierarchical classifier.

This hierarchical classifier is composed of a binary classifier and an ordinal classifier. The binary one determines whether or not rain can occur, and an ordinal classification model is applied to perform a finer classification of the predicted rain cases. We call this method BInary and ORdinal Kernel classifier (BIORK).

The training process consist on simultaneously training the binary and the ordinal model with different subsets of the training patterns. For the binary model $f(\mathbf{x})$, rainfall classes (C_2, C_3, C_4) are grouped as the positive class $(y = +1)$ whereas the no-rain class (C_1) is the negative class $(y = -1)$ of the binary problem. The ordinal model $g(\mathbf{x})$ is trained only with rain classes so the model predicts $z \in \{1, 2, 3\}$ with $C_2 = 1, C_3 = 2, C_4 = 3$. Hyper-parameters of binary and ordinal models are adjusted independently with the purpose of getting a better fit of the models to the data. In addition, since the current data set is highly imbalanced regarding non-rain and rain patterns (see Table 3), we have selected the CSVC classifier for the binary model, where the cost C_+ is weighted according to the criteria shown in Eq. 5.

The prediction phase consist on first getting the binary prediction, and then perform a second classification of the positive class patterns with the ordinal model. Figure 1 shows the two models decision flow.

5 Experiments

5.1 Performance evaluation metrics

In this section experimental results are measured in terms of three metrics to observe different features of the models regarding classification performance of predicted labels $\{y*_1, y*_2, \ldots, y*_N\}$, with respect to the true targets $\{y_1, y_2, \ldots, y_N\}$:

- Acc: the accuracy (Acc), also known as Correct Classification Rate, is the rate of correctly classified patterns:

$$Acc = \frac{1}{N}\sum_{i=1}^{N} [\![y*_i = y_i]\!],$$

  where y_i is the true label, $y*_i$ is the predicted label and $[\![c]\!]$ is the indicator function, being equal to 1 if c is true, and to 0 otherwise. Acc values range from 0 to 100 and they represent a global performance on the classification task being not suitable for imbalanced datasets [44].

- GM: The geometric mean of the Sensitivity or precision for each class is typically used to evaluate performance in imbalanced problems [45]:

$$GM = \sqrt[J]{\prod_{j=1}^{J} S_j},$$

  where J is the number of classes and S_j is the accuracy of the classifier for patterns of class j. GM varies from 0 to 100. In the case $GM = 0$ this means that the classifier is not correctly labelling any pattern of one or more classes.

- $AMAE$: This measure evaluates the mean of the Mean Absolute Error (MAE) across classes [46]. It has been proposed as a more robust alternative to MAE (the most extended measure in ordinal regression) for imbalanced datasets. $AMAE$ is defined as:

$$AMAE = \frac{1}{J}\sum_{j=1}^{J} MAE_j = \frac{1}{J}\sum_{j=1}^{J}\frac{1}{n_j}\sum_{i=1}^{n_j} e(\mathbf{x}_i),$$

where n_j is the number of patterns in class j and MAE is defined as:

$$MAE = \frac{1}{N}\sum_{i=1}^{N} e(\mathbf{x}_i),$$

where $e(\mathbf{x}_i) = |\mathcal{O}(y_i) - \mathcal{O}(y*_i)|$. $AMAE$ values range from J to $J-1$.

As previously mentioned, ordinal regression problems need specific performance metrics.

5.2 Comparison methods

A wide selection of computational intelligence methods has been done for the experiments including ordinal classification state-of-the-art SVMs methods and artificial neural network methods. The nominal SVM classifier is included also as a reference method.

- BInary and ORdinal classification Kernel method (BIORK), that is the proposal of the paper. The method is implemented in Matlab by using Cost SVC available in LibSVM 3.0 [47] for the binary model and SVORIM for the ordinal model.
- Evolutionary extreme learning machine for ordinal regression (EELMOR) [48]. This algorithm applies differential evolution to improve neural network models trained with the extreme learning machine algorithm.
- Kernel Discriminant Learning for Ordinal Regression (KDLOR) [49] extends the Kernel Discriminant Analysis (KDA) using a rank constraint.

- ONN adapts the data replication method proposed in [50] to neural networks.
- The Proportional Odds Model (POM) is one of the first models specifically designed for ordinal regression [51], and it adapts the standard logistic regression to the ordinal case. For the POM model, the `mnrfit` function of Matlab software has been used.
- RED-SVM [1], by [52], applies the reduction from cost-sensitive ordinal ranking to weighted binary classification (RED) framework to SVM.
- SVM classifier, SVC, implemented in LibSVM 3.0 [47]. The "1-versus-1" multiclass approach is applied [2].
- The SVM for ordinal regression with implicit constraints, SVORIM.
- Pairwise Class Distances for Ordinal Classification (PCDOC) [53] with the *epsilon* Support Vector Regression (SVR) as the underlying regressor (SVR-PCDOC).

5.3 Experimental results

Regarding the experimental procedure, 30 different random splits of the dataset have been considered, with 75% and 25% of the instances in the training and generalization sets respectively. The partitions were the same for all compared methods. All the variables were property standardized and the SVM hyper-parameters have been adjusted by using a grid search in the parameters values space. The grid search consisted on a 5-fold validation procedure (exclusively using training data) with $AMAE$ as the parameters selection criteria.

[1] Source code available at `http://home.caltech.edu/htlin/program/libsvm/`

[2] Source code available at `http://www.csie.ntu.edu.tw/~cjlin/libsvm/`

All the kernel methods were configured to use the Gaussian kernel. For the support vector algorithms, i.e. BIORK, SVC, RED-SVM, SVORIM and ϵ-SVR (for SVRPCDOC), the corresponding hyper-parameters (regularization parameter, C, and width of the Gaussian functions, γ), were adjusted using a grid search over each of the 30 training sets by a 5-fold nested cross-validation with the following ranges: $C \in \{10^{-3}, 10^{-2}, \ldots, 10^{3}\}$ and $\gamma \in \{10^{-3}, 10^{-2}, \ldots, 10^{3}\}$. Regarding ϵ-SVR, the additional ϵ parameter was adjusted considering the range $\epsilon \in \{10^0, 10^1, 10^2, 10^3\}$. For KDLOR, the width of the Gaussian kernel was adjusted by using the range $\gamma \in \{10^{-3}, 10^{-2}, \ldots, 10^{3}\}$, and the regularization parameter, u, for avoiding the singularity problem values were $u \in \{10^{-2}, 10^{-3}, \ldots, 10^{-5}\}$. For ONN, the number of neurons in the hidden layer was selected by considering the following values, $M \in \{5, 10, 15, 20, 30, 40\}$. In the case of EELMOR, M value was chosen from the set $\{5, 10, 15, 20, 25, 30, 35, 40, 45, 50\}$, and the number of iterations was fixed to 50, and the population size 40. Finally, POM does not have hyper-parameters.

Table 4 shows the generalization performance of the different algorithms in terms of mean and standard deviation in the 30 generalization partitions. The results correspond to the previous explained metrics: Accuracy (Acc), geometric mean of the Sensitivities (GM) and $AMAE$. For each metric, the best result is highlighted in bold face and the second best result is highlighted in italics. Note that Accuracy alone is not enough to assess the performance of a classifier. As an illustrative example, a trivial classifier labelling all the patterns as no-rain class (C_1) will obtain an Accuracy performance near 69.04%. Observe than EELMOR have the best Acc performance, however it is not able of classifying any pattern of one or more classes since GM result is zero. Regarding the proposed method, it obtained the best performance both in GM and $AMAE$, which are more suitable performance metrics for the rainfall problem nature.

In order to better compare the performance of the algorithms, each pair of algorithms are compared by means of the Wilcoxon test [54]. A level of significance of $\alpha = 0.05$ was considered, and the corresponding correction for the number of comparisons was also included. The results of these tests are shown in Table 5.

5.4 Discussion: system usefulness for improving airports operations

Among other meteorological phenomena, precipitation can seriously affect airport operations. When heavy or very heavy rainfall rates are expected, raindrops impacting airplane windscreens can lead to a reduction of the visibility, and depending on the atmospheric conditions, windscreen wipers may not be able to fully cope with the rainfall rate. Not to mention that light, non-pressurised aircraft may find the heaviest rain rates allow water ingestion into the cabin, the cockpit or the engine compartments with subsequent risks to electronic equipment. On the other hand, precipitation can lead to runway flooding, what may directly affect take-off and landing performances.

Aiming at getting improved meteorological information for each airport, the International Civil Aviation Organization, in collaboration with the World Meteorological Organization (WMO), regulates the provision of meteorological services in support of airport operations. Specifically, the Annex 3 to the Convention on International Civil Aviation states that it is necessary to deliver specific weather forecasts and warnings to meet the needs of flight operations at each aerodrome. Thus, aeronautical meteorological service providers prepare and disseminate specific aeronautical weather forecasts for airports, such as TAF as well as Aerodrome Warnings:

1. TAF is the name of the code for reporting weather forecast information ("TAF" is an acronym of Terminal Aerodrome Forecast). The TAF describes weather conditions that are expected to occur over a specific period of time, that can range from 9 up to 30 hours. The TAF is one of the most valuable sources for the predicted weather at a specific airport. Among others, TAF specifies the occurrence of precipitation.
2. Aerodrome Warnings give concise information of meteorological conditions which could adversely affect aircraft on the ground, including parked aircraft, and the aerodrome facilities and services. An aerodrome warning is issued when a specific weather phenomena is observed or forecasted. Among others, accumulated precipitation is one of them.

Therefore, it is clear that aeronautical meteorological services providers need specific tools to accurate forecast precipitation occurrence and amount at each specific airport in order to deliver weather forecasts and warnings appropriated to contribute towards the safety, regularity and efficiency of airport operations. Thus, the BIORK system proposed in this paper could be useful as complementary system to obtain TAF reports and aerodrome warnings of rain occurrence and expected rainfall amount. The good performance in terms of accuracy and probability of error exhibited by BIORK system makes it a very interesting tool in rainfall prediction (which is one of the most difficult meteorological variables to be forecasted).

In future works we plan to improve the performance of the system by including specific predictive and objective data from convective – non-convective precipitation and extreme events, with a larger range of applications in alternative facilities or cases.

6 Conclusions

In this paper we have proposed a system for simultaneous prediction of rainfall occurrence and amount. The proposed system is based on a hierarchical system of nominal and ordinal Support Vector Classifiers, so called BInary and ORdinal classification Kernel method (BIORK), and we have also used a novel set of predictive meteorological variables, which improve the classifiers performance in this problem. We have evaluated the proposed system in a real problem of rainfall forecast at Santiago de Compostela airport, Spain, comparing the BIORK system against several alternative computational intelligence methods in the literature. We have shown that the BIORK approach is able to obtain the best results in terms of different metrics and according to the Wilcoxon test, these results are significant. This system can be used as part of the airport weather forecast and warning system, in order to improve airport operational performance.

Acknowledgement

This work has been partially supported by Spanish Ministry of Science and Innovation, under project number ECO2010-22065-C03-02. This work has been partially subsidized by the TIN2011-22794 project of the Spanish Ministerial Commission of Science and Technology (MICYT), FEDER funds and the P11-TIC-7508 project of the "Junta de Andalucía" (Spain). We would like to thank AEMET for providing the observed meteorological data used in this work.

References

[1] H. Wei, J. L. Li and T. G. Liang, "Study on the estimation of precipitation resources for rainwater harvesting agriculture in semi-arid land of China," *Agricultural Water Management*, vol. 71, no. 1, 33-45, 2005.

[2] P. S. Yu, S. T. Chen and I. F. Chang, "Support vector regression for real-time flood stage forecasting," *Journal of Hydrology*, vol. 328, no. 3-4, pp. 704-716, 2006.

[3] O. Kisi and M. Cimen, "A wavelet-support vector machine conjunction model for monthly streamflow forecasting," *Journal of Hydrology*, vol. 399, no. 1-2, pp. 132-140, 2011.

[4] A. Benedetto, "A decision support system for the safety of airport runways: the case of heavy rainstorms," *Transportation Research Part A: Policy and Practice*, vol. 36, no. 8, pp. 665-682, 2002.

[5] F. J. Tapiador, F. J. Turk, W. Petersen, A. Y. Hou, E. García-Ortega, L. A. Machado, C. F. Angelis, P. Salio, C. Kidd, G. J. Huffman and M. de Castro. "Global precipitation measurement: Methods, datasets and applications," *Atmospheric Research*, vol. 104-105, pp. 70-97, 2012.

[6] R. J. Kuligowski and A. P. Barros, "Experiments in short-term precipitation forecasting using artificial neural networks," *Monthly Weather Review*, vol. 126, pp. 470-482, 1997.

[7] H. D. Navone and H. A. Ceccatto, "Predicting Indian monsoon rainfall: a neural network approach," *Climate Dynamics*, vol. 10, no. 6-7, pp. 305-312, 1994.

[8] T. Hall, H. E. Brooks and C. A. Doswell, "Precipitation forecasting using a neural network," *Weather and Forecasting*, vol. 14, pp. 338-345, 1999.

[9] AEMET, "Guía resumida del clima en España 1971-2000," Ministerio de Medio Ambiente, Spain, 2004.

[10] R. J. Kuligowski and A. P. Barros, "Localized precipitation forecasts from a numerical weather prediction model using artificial neural networks," *Weather and Forecasting*, vol. 13, no. 4, pp. 1194-1204, 1998.

[11] N. Q. Hung, M. S. Babel, S. Weesakul and N. K. Tripathi, "An artificial neural network model for rainfall forescasting in Bangkok, Thailand," *Hydrology and Earth System Science*, vol. 13, pp. 1413-1425, 2009.

[12] Yuval and W. W. Hsieh, "An adaptive nonlinear MOS scheme for precipitation forecasts using neural networks," *Weather and Forecasting*, vol. 18, pp. 303-310, 2003.

[13] N. S. Philip and K. B. Joseph, "A neural network tool for analyzing trends in rainfall," *Computers & Geosciences*, vol. 29, pp. 215-223, 2003.

[14] D. Silverman and J. A. Dracup, "Artificial neural networks and long-range precipitation prediction in California," *Journal of Applied Meteorology*, vol. 39, pp. 57-66, 2000.

[15] P.T. Nastos, K.P. Moustris, I.K. Larissi, A.G. Paliatsos, "Rain intensity forecast using artificial neural networks in Athens, Greece," *Atmospheric Research*, vol. 119, pp. 153-160, 2013.

[16] R. P. Shukla, K. C. Tripathi, A. C. Pandey and I.M. Das, "Prediction of Indian summer monsoon rainfall using Niño indices: A neural network approach," *Atmospheric Research*, vol. 102, no. 1-2, pp. 99-109, 2011.

[17] A. Manzato, "Sounding-derived indices for neural network based short-term thunderstorm and rainfall forecasts," *Atmospheric Research*, vol. 83, no. 2-4, pp. 349-365, 2007.

[18] L. Ingsrisawang, S. Ingsriswang, S. Somchit, P. Aungsuratana and W. Khantiyanan, "Machine learning techniques for short-term rain forecasting system in the northeastern part of Thailand," *World Academy of Science, Engineering and Technology*, vol. 41, pp. 248-253, 2008.

[19] J. Wu, M. Liu and L. Jin, "A hybrid support vector regression approach for rainfall forecasting using particle swarm optimization and projection pursuit technology," *International Journal of Computational Intelligence and Applications*, vol. 9, no. 2, pp. 87-104, 2010.

[20] K. Lu and L. Wang, "A novel nonlinear combination model based on support vector machine for rainfall prediction," In *Proc. of the IEEE 4th International Joint Conference on Computational Sciences and Optimization*, pp. 1343-1347, 2011.

[21] O. Kisi and M. Cimen, "Precipitation forecasting by using wavelet-support vector machine," *Engineering Applications of Artificial Intelligence*, vol. 25, pp. 783-792, 2012.

[22] B. Tomassetti, M. Verdecchia and F. Giorgi, "NN5: A neural network based approach for the downscaling of precipitation fields - Model description and preliminary results," *Journal of Hydrology*, vol. 367, no. 1-2, pp. 14-26, 2009.

[23] S. T. Chen, P. S, Yu and Y. H. Tang, "Statistical downscaling of daily precipitation using support vector machines and multivariate analysis," *Journal of Hydrology*, vol. 385, no. 1-4, pp. 13-22, 2010.

[24] C. Harpham and R. L. Wilby, "Multi-site downscaling of heavy daily precipitation occurrence and amounts," *Journal of Hydrology*, vol. 312, no. 1-4, pp. 235-255, 2005.

[25] M. M. Hasan and P. K. Dunn, "A simple Poisson-gamma model for modelling rainfall occurrence and amount simultaneously," *Agricultural and*

Forest Meteorology, vol. 150, no. 10, pp. 1319-1330, 2010.

[26] R. M. Trigo and C. C. DaCamara, "Circulation weather types and their influence on the precipitation regime in Portugal," *International Journal of Climatology*, vol. 20, pp. 1559-1581, 2000.

[27] S. Applequist, G. E. Gahrs, R. L. Pfeffer, X. F. Niu, "Comparison of methodologies for probabilistic quantitative precipitation forecasting," *Weather and Forecasting*, vol. 17, no 4, pp. 783-799, 2002.

[28] http://weather.uwyo.edu/upperair/sounding.html

[29] M. Kanamitsu, J. C. Alpert, K. A. Campana, P. M. Caplan, D.G. Deaven, et al., "Recent changes implemented into the Global Forecast System at NMC," *Weather and Forecasting*, vol. 6, no. 3, pp 425-436, 1991.

[30] O. Planchon, H. Quénol, N. Dupont and S. Corgne, "Application of the Hess-Brezowsky classification to the identification of weather patterns causing heavy winter rainfall in Brittany (France)," *Natural Hazards and Earth System Sciences*, vol. 9, pp. 1161-1173, 2009.

[31] P. D. Jones, D. H. Lister, "The influence of the circulation on surface temperature and precipitation patterns over Europe," *Climate of the Past*, vol. 5, pp. 259-267, 2009.

[32] J. Kysely and R. Huth, "Changes in atmospheric circulation over Europe detected by objective and subjective methods," *Theoretical and Applied Climatology* vol. 85, pp 19-36, 2006.

[33] A. Paniagua-Tineo, S. Salcedo-Sanz, C. Casanova-Mateo, E. G. Ortiz-García, M. A. Cony and E. Hernández-Martín, "Prediction of daily maximum temperature using a support vector regression algorithm," *Renewable Energy*, vol. 36, no. 11, pp. 3054-3060, 2011.

[34] C.L. Wu and K.W. Chau, "Prediction of rainfall time series using modular soft computing methods," *Engineering Applications of Artificial Intelligence*, vol. 26, no. 3, pp. 997-1007, 2013.

[35] B. Boser, I. Guyon, and V. Vapnik, "A training algorithm for optimal margin classifiers," in *Proceedings of the Fifth Annual ACM Workshop on Computational Learning Theory*, pp. 144-152, 1992.

[36] C. Cortes and V. Vapnik, "Support-vector networks," *Machine Learning*, vol. 20, no. 3, pp. 273–297, 1995.

[37] O. Kisi, and M. Cimen, "Precipitation forecasting by using wavelet-support vector machine conjunction model," *Engineering Applications of Artificial Intelligence*, vol. 25, no. 4, pp. 783-792, 2012.

[38] A. Zarnani, P. Musilek, X. Shi, X. Ke, H. He, and R. Greiner, "Learning to predict ice accretion on electric power lines," *Engineering Applications of Artificial Intelligence*, vol. 25, no. 3, pp. 609-617, 2012.

[39] A. Shashua and A. Levin, "Ranking with large margin principle: Two approaches," *Advances in Neural Information Processing Systems*, vol. 15, pp. 937-944, MIT Press, Cambridge, 2003.

[40] W. Chu and S. S. Keerthi, "Support vector ordinal regression," *Neural Computation*, vol. 19, no. 3, pp. 792-815, 2007.

[41] N. Thai-Nghe, Z. Gantner, and L. Schmidt-Thieme. "Cost-sensitive learning methods for imbalanced data," In *Proc. of the 2010 International Joint Conference on Neural Networks* (IJCNN), pages 1-8, 2010.

[42] C.-W. Hsu and C.-J. Lin, "A comparison of methods for multi-class support vector machines," *IEEE Transaction on Neural Networks*, vol. 13, no. 2, pp. 415-425, 2002.

[43] A. Ben-Hur and J. Weston. “A User’s Guide to Support Vector Machines,” In Data Mining Techniques for the Life Sciences Methods in Molecular Biology, vol. 609, pp. 223-239, 2010.

[44] J. Sánchez-Monedero, P. A. Gutiérrez, F. Fernández-Navarro, and C. Hervás-Martínez. “Weighting efficient accuracy and minimum sensitivity for evolving multi-class classifiers,” *Neural Processing Letters*, vol. 34, no. 2, pp. 101-116, 2011.

[45] M. Kubat and S. Matwin. “Addressing the curse of imbalanced training sets: one-sided selection,” In *Proc. of the 14th International Conference on Machine Learning*, pp 179-186, Morgan Kaufmann, 1997.

[46] S. Baccianella, A. Esuli and F. Sebastiani. “Evaluation measures for ordinal regression,” In *Proc. of the Ninth International Conference on Intelligent Systems Design and Applications* (ISDA’09), pp. 283-287, 2009.

[47] C.C. Chang and C.J Lin. “LIBSVM: A library for support vector machines,” *ACM Trans. Intell. Syst. Technol.* vol. 2, pp. 27:1–27:27, 2011.

[48] J. Sánchez-Monedero, P.A. Gutiérrez and C. Hervás-Martínez, “Evolutionary ordinal extreme learning machine,” In *Proc. of the International Conference on Hybrid Artificial Intelligence Systems* (HAIS2013), Lecture Notes in Computer Science, vol. 8073, pp. 500-509, 2013.

[49] B. Y. Sun, J. Li, D. Dash Wu, X.M. Zhang and W.B. Li. “Kernel Discriminant Learning for Ordinal Regression,” *IEEE Transactions on Knowledge and Data Engineering*, vol. 22, pp. 906-910, 2010.

[50] J. S. Cardoso and J. F. Pinto da Costa, “Learning to Classify Ordinal Data: The Data Replication Method,” *Journal of Machine Learning Research*, vol. 8, pp. 1393-1429, 2007.

[51] P. McCullagh, "Regression models for ordinal data," *Journal of the Royal Statistical Society*, Series B (Methodological), vol. 42, 109-142, 1980.

[52] H.T. Lin and L. Li. "Reduction from cost-sensitive ordinal ranking to weighted binary classification," *Neural Computation*, vol. 24, pp. 1329–1367, 2012.

[53] J. Sánchez-Monedero, P. A. Gutiérrez, P. Tiňo and C. Hervás-Martínez. "Exploitation of pairwise class distances for ordinal classification," *Neural Computation*, vol. 25, pp. 2450-2485, 2013.

[54] F. Wilcoxon, "Individual comparisons by ranking methods," *Biometrics Bulletin*, vol. 1, no. 6, pp. 80-83, 1945.

List of Tables

List of Figures

Table 1

Variables selected to take into account the content of water vapor present in the atmosphere.

Variable	Measurement units	Pressure level (hPa)	Source
Total Precipitable Water	mm	Entire column	upper air sounding
Equivalent Potential Temperature	K	950, 850, 700, 500	upper air sounding
Humidity	%	950, 850, 700, 500	upper air sounding

Table 2

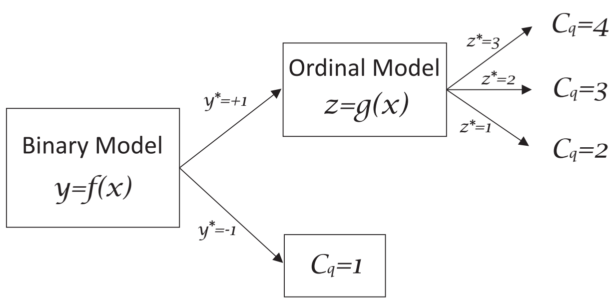
Variables selected to determine updrafts, downdrafts and the atmospheric stability.

Variable	Measurement units	Pressure level (hPa)	Source
Temperature	K	950, 850, 700, 500	upper air sounding
Wind Speed	m/s	950, 850, 700, 500, 300	upper air sounding
Wind direction	Degrees	950, 850, 700, 500, 300	upper air sounding
CAPE	J/kg	Entire column	upper air sounding
CIN	J/kg	Entire column	upper air sounding
ω	m/s	850, 500	GFS numerical model

Table 3

Observed rainfall in mm/6 h mapping to class labels.

Observed rainfall mm/6 h (w)	Label	Class number	Number of patterns
$w = 0.0$	class C_1 (no rain)	1	899
$w > 0.0$ and $w <= 0.2$	class C_2	2	329
$w > 0.2$ and $w <= 0.4$	class C_3	3	51
$w > 0.4$	class C_4	4	23



(a)

Fig. 1. Hierarchical classifier prediction process.

Table 4

Mean and standard deviation (SD) of the generalization performance of the proposed method and state-of-the-art methods for different performance metrics.

Method/DataSet	Accuracy Mean_{SD}	GM Mean_{SD}	AMAE Mean_{SD}
BIORK	$76.500_{2.210}$	$\mathbf{35.470_{20.560}}$	$\mathbf{0.710_{0.090}}$
EELMOR	$\mathbf{80.440_{1.450}}$	$0.000_{0.000}$	$0.900_{0.040}$
KDLOR	$73.290_{3.750}$	$30.840_{21.620}$	*$0.770_{0.100}$*
ONN	$70.900_{2.450}$	$9.310_{16.060}$	$1.170_{0.290}$
POM	$78.840_{1.450}$	$0.000_{0.000}$	$0.890_{0.060}$
REDSVM	$77.920_{2.660}$	$33.230_{19.390}$	*$0.770_{0.100}$*
SVC	*$79.220_{2.120}$*	$29.960_{20.420}$	$0.780_{0.090}$
SVORIM	$77.940_{2.650}$	*$33.330_{19.470}$*	*$0.770_{0.100}$*
SVRPCDOC	$75.710_{2.620}$	$22.640_{20.490}$	$0.870_{0.110}$

Table 5

Wilcoxon tests over different performance metrics.

	Acc			*GM*			*AMAE*		
Method	Wins	Draws	Loses	Wins	Draws	Loses	Wins	Draws	Loses
BIORK	2	3	3	4	4	0	7	1	0
EELMOR	7	1	0	0	2	6	1	2	5
KDLOR	0	2	6	3	5	0	4	4	0
ONN	0	1	7	0	3	5	0	0	8
POM	4	3	1	0	2	6	1	2	5
REDSVM	2	5	1	3	5	0	4	3	1
SVC	4	4	0	3	5	0	4	3	1
SVORIM	2	5	1	3	5	0	4	3	1
SVRPCDOC	1	4	3	2	5	1	1	2	5