http://ansinet.com/itj

ITJ

ISSN 1812-5638

INFORMATION TECHNOLOGY JOURNAL

ANSInet
Asian Network for Scientific Information
308 Lasani Town, Sargodha Road, Faisalabad - Pakistan

Information Technology Journal 12 (21): 6031-6037, 2013
ISSN 1812-5638 / DOI: 10.3923/itj.2013.6031.6037

Application Study of Machine Learning in Lightning Forecasting

Taorong Qiu, Shanshan Zhang, Hou Zhou, Xiaoming Bai and Ping Liu
Department of Computer, Nanchang University, Jiangxi, Nanchang, 330031, China

Abstract: Aiming at the lack of highresolution and short-term (5, 5km and next 3 h) thunderstorm forecast model, the three thunderstorm forecast models based on Rough Sets (RS), Support Vector Machine (SVM) and RS-SVM are presented in this study through analyzing Rough set, Support machine vector and the characteristic of lightning forecasting. With the real data set, the models are tested and the experimental results are analyzed, which show that the RS-SVM based model is effective and moreover, acquires a higher forecast precision.

Key words: Lightning forecasting, machine learning, rough set, support vector machine

INTRODUCTION

Lightning is a powerful natural electrostatic discharge produced within thunderstorms (cumulonimbus), cloud-to-ground and cloud-to-air. Lighting disaster was suggested as one of ten kinds of nature disasters by the U.N. departments. It named as "a popular disaster in electron era" by China Electrician Council (Ma *et al*., 2008). Due to this, meteorologists have done much research on lightning forecasting andmuchprogresshas beenmade (Sun *et al*., 2009; Yang and Tang, 2009). With the development of information technology, electronic products are essential to people's daily life and economic activities. However, large-scale meteorological fails to meet the need of electronic age. To research short-term (forecast scale: 5, 5 km) and immediate (next 3 h) meteorological weather is of great significant to the theoretical research and practical applications.

The traditional forecasting mainly includes multiple regressionequation, discriminatoryanalysis, indices superposition and so on. However, owing to the complexity of the mechanism of lightning generation, it is characterized by nonlinear obviously. Therefore, the methods mentioned above have some difficulties in the thunderstorm prediction. Many scholars have introduced some soft computing methods, such as Rough set (Liu, 2005), Neural Network and SVM, into the thunderstorm forecast recently, also achieved certain effect (Feng and Chen, 2004; Guo, 2008; Kong and Jin, 2007; Lu *et al*., 2008; Ma *et al*., 2009; Qian and Sun, 2009; Wang *et al*., 2011). There is still a virgin area in high resolution and very short range forecast, though some meteorologists have taken rough set method into the analysis of meteorological data. Support Vector Machine(SVM),which has a profound mathematical foundation, is a novel machine learning method based on statistical theory (Li *et al*., 2004; Zeng and Zhang, 2011). It has been successfully used in text classification, handwriting recognition, image classification, biological information and other fields with a good classification accuracy and generalization ability achieved by balancing the VC dimension and the empirical risk. Nevertheless, it remains in the preliminary stage to apply SVM to the weather forecast.

Rough set, as a newly developed mathematical tool for dealing with uncertain knowledge, has many advantages. However, it is vulnerable in fault tolerance and generalization, therefore, during the data classification, there will likely be a phenomenon of rule rejection. In most of these cases, SVM generalization performance either matches or is better than competing methods. It turns out that every global solution is also unique. Whereas in training for a mass dataset, SVM always faces the problem of large memory usage. In addition, SVM usually biases to a class of large proportion when training unbalanced data. Thence it follows that SVM and RS theory has a strong complementary. For the sake of enhancing the training efficiency, we first adopt rough set theory to process data reduction and then take the reductions as the training sets of SVM. It is feasible to establish the thunderstorm forecast model combining rough set theory and SVM method.

The article is organized as follows: After a brief overview of the research meaning and the status about the relevant subject, Section 2 discusses the related theories needed in this study and Feasibility Analysis to both SVM and RS applied in Thunderstorm Forecast; In Section 3, the thunderstorm forecast model based on RS-SVM along with the implementation technique of the classifier are introduced in details; Section 4 gives the results of experiments performed on the real datasets by

Corresponding Author: Taorong Qiu, Department of Computer, Nanchang University, Jiangxi, Nanchang, 330031, China

three models and an analysis for different methods as well; finally, some conclusions including also considerations on the future are presented in Section 5.

INTRODUCTION TO RELATED THEORIES AND FEASIBILITY ANALYSIS

Basic concepts of rough sets: Knowledge representation system (The symbolic expression of knowledge is essential to handle intelligent data, while the primary section of KRS is to research the set of objects, so the KRS can be formalized as follows (Liu, 2005; Pawlak, 1991):

$$S = (U, A, V, f)$$

where, U is a non-empty finite set of objects called the universe; A is the set of attributes, $A = C \cup D$ and $C \cap D = \phi$, along which C is the set of conditional attributes and D is the set of decision attributes; $V = \cup_{\alpha \in A} V_\alpha$ denotes the value set of attributes such that a: $U \rightarrow V_\alpha$, for every $\alpha \in A$; f represents the map of $U \times A \rightarrow V$.

Indiscernibility relation: For x, y, $\in U$, $P \subseteq A$, if $\forall q \in P$: $f_q(x) = f_q(y)$ holds, then objects x and y are indiscernible from each other. Ind (P) is called the P-indiscernibility relation.

Set approximation: Suppose $P \subseteq A$, $Y \subset U$, $x \subset U$, $[x]_p = \{y \subset U\}|x$ ind (P) y}, the lower and upper approximations can be represented respectively as:

$$\underline{PY} = pos_P(Y) = \{x \subset U \mid [x]_P \subseteq Y\} \tag{1}$$

$$\overline{PY} = \{x \subset U \mid [x]_P \cap Y \neq \varnothing\} \tag{2}$$

The set $bn_R(Y) = \overline{PY} - \underline{PY}$ is called the boundary of Y. Obviously, if $bn_R(Y) \neq \varnothing$ or $\overline{PY} \neq \underline{PY}$ holds, Y is roughly definable. The objects in $\underline{PY}$ can be with certainty classified as members of Y on the basis of knowledge in P, while the objects in $\overline{PY}$ can be only classified as possible members of Y on the basis of knowledge in P. The P-boundary region of Y is the difference set of $\underline{PY}$ and $\overline{PY}$, consisted of those objects that we cannot decisively classify into Y. There are also some other definitions in Rough set: $pos_R(Y) = PY$ is called the R-positive region of Y, while inversely $neg_R(Y) = U - \overline{PY}$ is the R-negative region of Y.

Knowledge reduction: There are two basic notions in knowledge reduction: reduction (reduct) and core.

Let R be the indiscernibility relation, $r \in R$, if Ind (R) = ind (R-{r} then r is dispensable in R; or r is indispensable in R.

If for any $r \in R$ is indispensable, then R is independent; or else R is dependent.

If $Q \subseteq P$ is independent and ind (Q) = ind (P), Q is called a reduction of P. The set of indispensable attributes in P is called the core of P, denoted as core (P). Suppose red (P) denotes the reductions of P, then core (P) = $\cap$red (P) holds.

Suppose P and Q are the sets of equal relationship on U, we can normalize and define P-positive region of Q as:

$$pos_P(Q) = \bigcup_{Y \in U/Q} \underline{PY} \tag{3}$$

If $a \in P$ and $pos_P(Q) = pos_{P-\{\alpha\}}(Q)$ hold, we say that the relationship α is Q-dispensable in P, otherwise, α is Q-indispensable. If for every α in P is Q-indispensable, then P is independent on Q, or we say that P depends on Q. If there is a single Q-reduction of P, P is exactly certain. On the contrary, there are various expressions of and the uncertainty will aggravate on condition P vthat core (P) = $\varnothing$.

Decision rule: Another important issue is the rule extraction after reduction. Let X_i and Y_j be the equivalence class of U/C and U/D respectively, des (X_i) denotes the specific value of every conditional attribute, similarly, des (Y_j) denotes the specific values of every decision attribute. Then the decision rule can be defined as $r_{ij:des}(X_i) \rightarrow des(Y_j)$, $X_i \cap Y_j \neq \varnothing$.

Formally the certainty factor of rules can be computed in the following way:

$$\mu(X_i \cap Y_j) = |X_i \cap Y_j| / |X_i|,\ 0 < \mu(X_i, Y_j) < 1 \tag{4}$$

When $\mu(X_i \cap Y_j) = 1$, we say that r_{ij} is certain; while $0 < \mu(X_i, Y_j) < 1$ we say that r_{ij} is uncertain.

Data processing step in rough set:

- Organize the collecting data into the decision table. The data must fully reflect the features of knowledge to be dealt with. So, it is necessary to eliminate the redundant attributes to reduce the data dimension and finally determine conditional attributes and decision attributes
- Data discretization. Because rough set is less favorable in dealing with the continuous data, it is desirable to process data discretization. A successful

approach to discretization is to construct a set of cuts with a minimal number of elements on the condition that the discerned data must remain the original feature or not introduce the extra noise

- Reduction of decision table. The reduction of decision table can be divided into attribute reduction and value reduction
- Classifier construction. Apply the rules gained by step 3 to the decision algorithm to process data classification

Introduction to support vector machine: Support Vector Machine (SVM) is a supervised learning model with associated learning algorithm that analyzes data and recognizes patterns, used for classification and regression analysis. SVM can be extended into a nonlinear classifier by mapping the space of objects into a high-dimensional (possibly infinite- dimensional) space. By choosing an adequate mapping, the data become linearly separable or mostly linearly separable in the high-dimensional space. In general, the whole procedure is to make the data dimension raising and linearization. Usually, the dimension raising sample space may lead to the complexity of problems, even "dimension curse". Because of this, researchers tend to prefer reducing the dimension, e.g., PCA method. However, if it is just used in classification or regression analysis, this can yet be regarded as a great choice. In that case, the set of samples which cannot be handled in the low dimension can be separated by a linear hyperplane simply. SVM ingeniously figures out an increased computational complexity caused by dimension raising via the kernel function expansion theorem, through which a linear learning machine is built in the high dimension instead of an explicit formula transforming the low dimension to a high dimension.

Frequently used kernel functions for SVM include:

Linear kernel function:

$$K(X_i, X_j) = X_i^T X_j \tag{5}$$

Polynomial kernel function:

$$K(X_i, X_j) = (\gamma X_i^T X_j + r)^d, \gamma > 0 \tag{6}$$

Radial basis function(RBF) kernel:

$$K(X_i, X_j) = \exp(-\gamma \| X_i - X_j \|^2), \gamma > 0 \tag{7}$$

Sigmoid kernel:

$$K(X_i, X_j) = \tanh(\gamma X_i^T X_j + r) \tag{8}$$

where, γ and d are kernel-specific parameters.

Feasibility analysis

Features of thunderstorm forecast: Thunderstorm forecast, as it is, is a binary classification which divides the meteorological data into thunderstorm and non-thunderstorm class. It is through mining the relationship between predictors and the thunderstorm to make the prediction. The lightning forecast is characterized by the followings:

- Conditions for lightning occurrence are diverse in different scales, areas and predictable time. The prior knowledge about thunderstorm cannot be expanded infinitely, at the same time, there are few literatures evaluating on the small scale forecast. Therefore, the prior knowledge for the lightning forecast research of special area and scale is not enough
- There are numbers of predictors describing weather conditions, while not all of these have a relationship with thunderstorms, which are redundant mean to lightning prediction. The redundance can lead to a large training data scale and a more complicated model if they are not rejected. Hence, it is of great significant for improving the accuracy and efficiency to select the key factors
- The lightning happens non-linearly with the predictors. It is necessary to illustrate the relationship between them
- The historical data of meteorology are enormous, which are difficult for whatever model to handle all of them. That means the model must have a better ability to deal with the incomplete information

Advantages to the combination of SVM and rough set: Rough set, as a newly developed mathematical tool for dealing with uncertain knowledge, has many advantages. However, it is vulnerable in fault tolerance and generalization. Therefore, during the data classification, there will likely be a phenomenon of rule rejection. In most of these cases, the generalization performance of SVM either matches or is better than competing methods. It turns out that every global solution is also unique. Whereas in training for every large datasets, SVM always faces the problem of large memory usage. It can also import noise to the model which may infect the precision without reducing protectors. In addition, SVM usually biases to a class of large proportion when training the unbalanced data. Thence it follows that SVM and RS theory has a strong complementary. For the sake of enhancing the training efficiency, we first adopt rough set theory to process data reduction and then take the reductions as the training sets of SVM. So, it is feasible to improve the performance of SVM in learning unbalanced data via associating SVM with RS and reduce the error data of SVM by RS.

THUNDERSTORM FORECAST MODEL BASED ON RS-SVM

In this study, we combined the Rough Set with SVM to construct the thunderstorm forecast model. The frame of the thunderstorm forecast is shown as Fig. 1.

As shown in Fig. 1, the model is made up of two modules: The classify structure and the thunderstorm forecast. Data discretization, predictor extraction, classifier training and some other sub-modules consist of the classify structure. Among which, the undersampling as well as the data discretization belong to data pretreatments; the predictor extraction which aims to improve the efficiency and precision creates the reduced meteorological decision table via reducing the dimension of the discretized data; after that, SVM decision functions will be gained by training the corresponding data using the rules extracted from the reduced decision table.

In the thunderstorm-predict module, the rule base and SVM decision function will be applied into the prediction of new data. Firstly, it processes the identification by using rough set and this will produce three probable outcomes: Thunderstorm, non-thunderstorm and rejection. Then, with regard to the rejections, it will be rejudged by SVM.

A more detailed description of the classifier construction is now presented in the following.

Classifier structure based on RS (RS model): The classifier should analyze the real-time acquisitions and predict whether they would lead to the occurrence of thunderstorms. The practical classify process is composed of three steps:

- **Step 1:** Choose the corresponding data fields from the predict data according to the attribute reduction results; Perform the discretization using the cuts got from the discretization algorithm
- **Step 2:** Compare the forecast data processed in Step 1 with rules in the rule base. If the attributes (expect the "*") of the rule have the same value with the data set, that means the rule is satisfied. Then, the support of this decision attribute is computed according to the rule's decision values. If the data matches numbers of rules in the base, the support is the summation of each support. Besides, the supports of the thunderstorm and the non-thunderstorm are also calculated
- **Step 3:** Move the predict results to the decision class with largest support values which is gained through Step 2

Classifier structure based on SVM (SVM model): In this study, a SVM with RBF kernel is taken to construct the classifier:

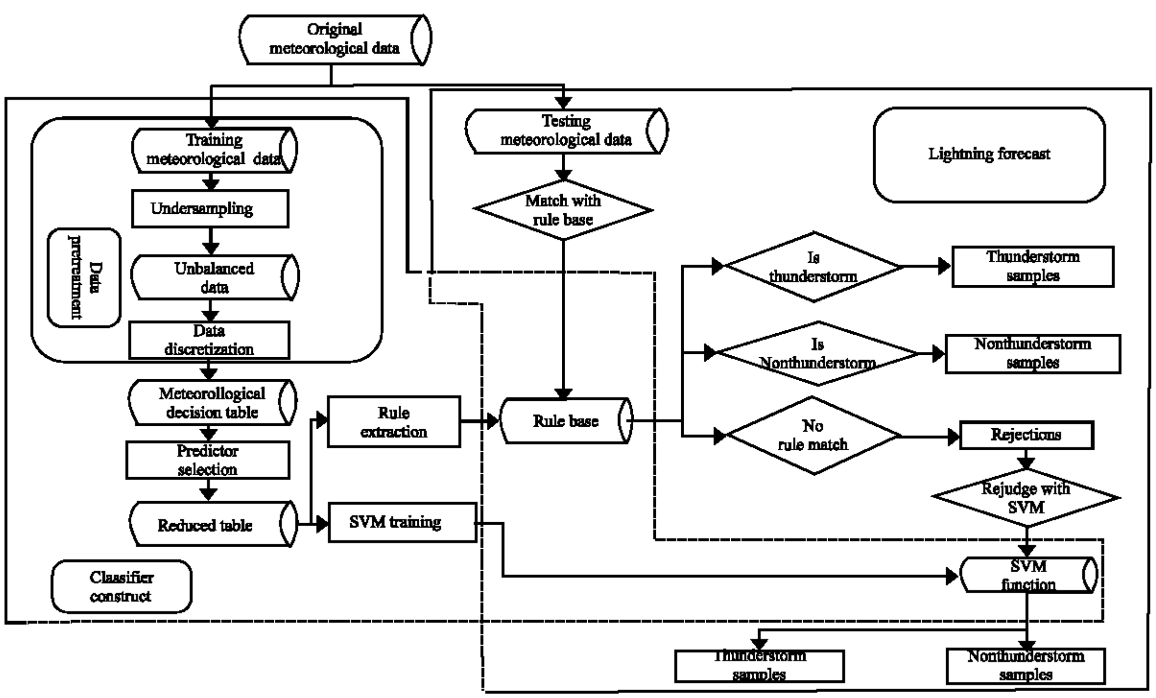

Fig. 1: Frame of thunder forecast model

Table 1: Table test result of SVM grade

	C				
σ^2	2^7	2^9	2^{11}	2^{13}	2^{15}
2^{-3}	0.65	0.65	0.8	0.85	0.7
2^{-5}	0.65	0.65	0.6	0.7	0.7
2^{-7}	0.8	0.8	0.6	0.7	0.7
2^{-9}	0.7	0.7	0.6	0.7	0.65
2^{-11}	0.7	0.55	0.55	0.55	0.75

$$K(\vec{x},\vec{x}_i)=\exp\left(\frac{-\|\vec{x}-\vec{x}\|^2}{2\sigma^2}\right) \tag{9}$$

The recognition rate of the classifier can be optimally achieved through adjusting the parameters of the kernel such as c and σ^2. There are lots of solutions to get the optimization (c, σ^2), one is the grid method, in which, c takes N values and σ^2 takes M values, respectively. For every (c, σ^2) in N×M group, SVM is trained respectively. Through evaluating all the generalized recognitions, the optimal parameter is the group with the highest recognition rate. The benefit of this method is its parallel processing for each SVM is trained independently and irrespective of others. In this study, c takes values of [2^7, 2^9, 2^{11}, 2^{13}, 2^{15}] , while σ^2 takes [2^{-3}, 3^{-5}, 2^{-7}, 2^{-9}, 2^{-11}].

In this study, we take the positive region data samples generated by RS and the reductive attributes as the training data of SVM .The training results of the grid method are shown in Table 1.

Table 1 indicates that the maximum of prediction accuracy takes 0.85 on the condition that $c = 2^{13}$ and $\sigma^2 = 2^{-3}$. In this case, $c = 2^{13}$ and $\sigma^2 = 2^{-3}$ are chosen as parameters in the present study.

Forecast model based on RS-SVM (RS-SVM model): Firstly, predicate the forecasting data with RS based classifier. If it can be determined, the prediction result is obtained.

Secondly, for the data which cannot get a definite result via RS based classifier, SVM based classifier is used to acquire the prediction result.

EXPERIMENTS AND RESULTS ANALYSIS

Experimental environment: The experiments are performed on a loaded Intel (R) Core (TM) 2 Dou CPU T5250 @1.50 GHz,250G of hard disk,2G of RAM, Windows XP, Matlab 7.1 and GrADS 1.9.

Original meteorological data preparation: The dataset is taken from LAPS system of Jiangxi Meteorological Administration. It provides an original meteorological data reanalysis with 8 times every day (each at time of 00, 03, 06, 09, 12, 15, 18, 21 o'clock). And GRADS software is used to interpolate the laps data into a latitude-longitude grid with specified spacing. Lightning positioning data with 0.045E×0.045N spacing during the longitude of 113.5-118.54 and latitude of 24.5-30.035 are collected from the historical meteorological information of Jiangxi province in 2010. The laps data are also interpolated in the same grid. Because the frequent thunderstorm period focuses on months 5-8, laps data with a high number of thunderstorm in that time are selected to form the experimental data, which consists of 201468 records, including 100734 thunderstorm samples and 100734 non-thunderstorm samples. We used a random selection of 40294 testing data.

STEPS

Data discretization: After sampling and discretization, the laps data are organized as the lightning decision table, which includes 313 conditional attributes and 1decision attribute. In the Table, 1 denotes non-lightening sample and 2 denotes lightening sample.

Predictor extraction: The attribute reduction method based on rough set is used on the decision table produced by step (1) to get predictors. And finally, 33 attributes are selected from 313 conditional attributes

Rules generate and classifier construct: Utilize the technique noted in the previous to construct the corresponding classifier.

Model test: Apply the classifier to classify the test data.

Result analysis: The 40294 samples are tested, respectively on the three models mentioned above. Table 2 records the test indexes.

In view of the unbalanced original data, the following evaluation indexes are introduced to get a better description of the performance of the classifier:

Forecast accuracy:

$$Acc=\frac{TP+TN}{TP+TN+FN+FP} \tag{10}$$

Non-thunderstorm forecast accuracy:

$$FPrate=\frac{TN}{TN+FP} \tag{11}$$

Thunderstorm forecast accuracy:

$$TPrate=\frac{TP}{TP+FN} \tag{12}$$

Table 2: Records of forecast results

	Prediction results	
Real results	Thunderstorm	Non-thunderstorm
Thunderstorm	Correct (TP)	Comitting (FN)
Non-thunderstorm	False (FP)	Correct (TN)

Table 3: Forecast accuracy on RS model

Non-thunderstorm forecast accuracy	Thunderstorm forecast accuracy	Forecast accuracy
0.72	0.74	0.47

Table 4: Forecast accuracy on SVM model

Non-thunderstorm forecast accuracy	Thunderstorm forecast accuracy	Forecast accuracy
0.67	0.63	0.65

Table 5: Forecast accuracy on RS-SVM model

Non-thunderstorm forecast accuracy	Thunderstorm forecast accuracy	Forecast accuracy
0.66	0.75	0.71

The forecast accuracy in RS is characterized by:

$$acc = \frac{TP + TN}{TP + TN + FN + UD} \tag{13}$$

where, UD denotes the number of samples (reject rate).

Experiment 1-thunderstorm forecast based on rough set: Test results about prediction accuracy and precision based on RS classifier fill in Table 3.

Table 3 reveals that at least one third of data samples in the majority of non thunderstorm samples are rejected in the prediction process. Two reasons contribute the phenomenon: one is the limited generation of rules, and the other is that it is difficult to extract rules which can be applied in all cases because of the weather's particularity as well as the variety relating to different seasons and years.

Experiment 2-thunderstorm forecast based on SVM: Positive region data and reductive attributes generated by RS are adopt as the training data. A RBF kernel has been taken with $c = 2^{13}$ and $\sigma^2 = 2^{-3}$ for the SVM model.

Table 4 shows the accuracy and precision of the forecast model in the case of the above SVM.

It is illustrated from Table 4 that SVM classifier has a relatively higher recognize rate in non-thunderstorm samples than that in thunderstorm samples. Though it has lower accuracy than that of the RS classifier comparing with Table 3, there is no rejection phenomenon.

Experiment 3-thunderstorm forecast based on RS-SVM: The accuracy and precision of the forecast model in the case of the RS-SVM based classifier are given in Table 5.

Table 5 demonstrates that the RS-SVM classifier achieved the highest forecast precision in all these three classifiers. Nevertheless, due to the inadequate non-thunderstorm data samples, there exists misinformation phenomenon. That ultimately led to a lower accuracy in non-thunderstorm forecast.

Compare with other forecast models based on SVM: In recent years, Support Vector Machine (SVM) is usually used in lightning forecast, e.g., a research on the thunderstorm forecast based on least-squares SVM is stated in the literature (Wang *et al*., 2011; Ma *et al*., 2009) applied the SVM technique based on data field to the thunderstorm forecast. Compared with the models discussed in the literatures (Ma *et al*., 2009; Wang *et al*., 2011), the thunderstorm forecast model studied in the present paper though has a similar forecast accuracy, differences are thoroughly aired on the following two points:

- Forecast scales are different. In the literature (Wang *et al*., 2011), it depicts a range of 1E×1N thunderstorm forecast and the literature (Ma *et al*., 2009) specialize a prediction on weather stations, while in this study the research is concerned essentially with a small-scale forecast in 0.045E×0.045N
- The method and technique route used to improve the forecast precision are different. The literature (Wang *et al*., 2011) chooses predictors by experience and adopts LS-SVM to improve the precision, while in the literature (Ma *et al*., 2009), in order to solve the imbalance problem of datasets, it introduces the concept of data field and, in the meanwhile, proposes a method which is combined with a weighted SVM to get a higher precision. The high efficiency and accuracy makes it especially suitable for non-linear meteorological field

However, it may be biased towards the majority class. In this study, with the full consideration of particular features and complement of RS and SVM, it puts forward a combined method to reduce the influence to the accuracy and precision caused by the imbalanced data, and the powerful combination produces a great advantage.

CONCLUSION

Through researching the process of model construction, a high resolution and short time nowcasting model is proposed combined RS with SVM. With the real data, experiments are performed respectively on these three models, that is, RS-based model, SVM-based model and RS-SVM based model. The results demonstrate that the new RS-SVM based model discussed in the present study is more suitable and operational.

The future assignments are arranged as follows: Firstly, expecting to raise the prediction accuracy, a model special for a certain regional will be built, secondly, it is necessary to increase the experimental data. The historical data of lightning is huge, however, for the limit of experimental installation and finite time, this paper just made the experiment with data of a given year; once more, improve and perfect the given algorithms is also requisite; finally, thunderstorm models using other soft computing methods associated with fusion approaches still need a further test and comparison.

ACKNOWLEDGMENTS

Supported by National Natural Science Foundation of China (No. 61070139), Natural Science Foundation of Jiangxi Province (No. 20114BAB201039) and the Science and Technology Support Planning Project of JiangXi Province (No. 20112BBG70087).

REFERENCES

Feng, H.Z. and Y.Y. Chen, 2004. A new method for non-linear classify and non-linear regression: Application of support vector machine to weather forecast. J. Applied Meteorol. Sci., 15: 355-365.

Guo, Q.M., 2008. Research and application of rough set theory in cumulonimbus forecast. Civil Aviation University of China.

Kong, Q.Y. and L. Jin, 2007. Application of rough set theory to regional mean rainfall forecast. J. Guangxi Acad. Sci., 23: 147-149.

Li, G.Z., M. Wang and H.J. Zeng, 2004. An Introduction to Support Vector Machines and other Kernel-based Learning Methods. China Electronic Industry Press, China.

Liu, Q., 2005. Rough Set and Rough Set Reasoning. 3rd Edn., China Science Press, Beijing, China.

Lu, Z.Y., Y.Y. Li, J. Lu and Z.C. Zhao, 2008. Parameters optimization of RBF-SVM sand-dust storm forecasting model based on PSO. J. Tianjin Univ., 41: 413-418.

Ma, J., W. Fan and H.Y. Yuan, 2009. Application of SVM technology based on data field in thunderstorm report. Comput. Eng., 35: 263-265.

Ma, M., W.T. Lv and Y.J. Zhang, 2008. Characteristics of lightning exposure in China from 1997-2006. J. Applied Meteorol. Sci., 19: 393-400.

Pawlak, Z., 1991. Rough Sets: Theoretical Aspects of Reasoning about Data. Kluwer Academic Publishers, Dordrecht, Boston, ISBN 0792314727, Pages: 229.

Qian, Y.Y. and M.Y. Sun, 2009. Analysis mmeteorological data based on rough set. Plateau Mountain Meteorol. Res., 29: 695-703.

Sun, L., J.J. Zhou and J. Yang, 2009. Advances in early-warning and forecasting of thunderstorms. Plateau Mountain Meteorol. Res., 29: 75-79.

Wang, Z.H., W. Zhang and J. Zhu, 2011. A preliminary study on thunderstorm forecast based on LS-SVM. J. Trop. Meteorol., 27: 161-165.

Yang, Z.J. and X.F. Tang, 2009. The application of step-wise decreasing FAR method to thunderstorm potential trend forecast. Scient. Meteorol. Sin., 29: 810-814.

Zeng, H.J. and Y.K. Zhang, 2011. Machine Learning. China Machine Press, China.