ARTICLE IN PRESS
JID: PATREC [m5G;August 12, 2017;23:33]
Pattern Recognition Letters 000 (2017) 1–10
Contents lists available at ScienceDirect
Pattern Recognition Letters
journal homepage: www.elsevier.com/locate/patrec

Data-driven approaches for meteorological time series prediction: A comparative study of the state-of-the-art computational intelligence techniques

Monidipa Das*, Soumya K. Ghosh

Department of Computer Science and Engineering, Indian Institute of Technology, Kharagpur 721302, India

A R T I C L E I N F O

Article history:
Available online xxx

Keywords:
Computational intelligence
Data-driven modeling
Bayesian network
Time series prediction
Meteorology

A B S T R A C T

With the proliferation of sensor generated weather data, the data-driven modeling for prediction of meteorological time series has gained increasing research interest in current years. The recent advancement in machine learning and artificial intelligence paradigm has made such data analysis process more effective, flexible and sound. This paper attempts to provide a comparative study of the state-of-the art computational intelligence (CI) techniques, which have been successfully applied for meteorological time series prediction purpose. The study has been carried out considering eleven distinct variants of CI techniques, especially based on artificial neural network (ANN), fuzzy logic, Bayesian network (BN) and other probabilistic models. Further, one more hybrid CI technique (SpaFBN), derived from the existing approaches, has been proposed in the present work. All these CI techniques have been empirically studied with respect to a multivariate meteorological time series prediction problem, in comparison with three benchmark statistical approaches. Overall, the experimental results demonstrate the superiority of the BN-based models in meteorological prediction. The presently proposed *spatial fuzzy Bayesian network* (SpaFBN) is also found to be an effective tool, especially for predicting humidity and precipitation rate time series. Moreover, the proposed SpaFBN is a generic CI technique which can be applied for predicting spatial time series from the domains other than meteorology as well.

1. Introduction

Meteorological time series prediction is important not only for our day-to-day planning, but also for long term decision making, which can have grave influence on the economic development of any country. Traditionally, the meteorological predictions are made mainly based on physics driven approaches as followed by various global circulation models or numerical weather prediction (NWP) models [2]. However, the recent data explosion has led to the emergence of a new paradigm, termed as *data driven modeling* or DDM [29], which aims at extensively analyzing historical data for generating insights, and utilizing those in further studies. The mathematical equations underlying such approaches are not derived from physical processes. Rather, these approaches are mainly based on various *computational intelligence* (CI) techniques, like artificial neural network (ANN), Bayesian/Belief network (BN), fuzzy logic (FL), genetic algorithms (GAs) and so on.

In this work, an attempt has been made to provide a comprehensive study of the various CI techniques that have been applied for *meteorological time series prediction*. The focus is kept on prediction of spatial time series data obtained from the spatially distributed sensors, and on the CI techniques which have been either recently proposed or most widely used in the recent past.

A number of works reporting the application of CI techniques in time series prediction can be found in literature. Sapankevych and Sankar [28] have presented an exhaustive survey on time series prediction using support vector machines (SVMs). SVM based prediction of air quality, rainfall, environment pollution etc. has been studied here along with other applications. A survey of the wind speed and wind power prediction has been provided by Lei et al. [16]. In their work, Thissen et al. [31] have found that the SVMs perform better than recurrent neural network (RNN) and statistical autoregressive model, in predicting nonlinear chaotic time series. The existing prediction approaches on SVM, ANN, fuzzy logic or combined neuro-fuzzy techniques have been reported here. Prediction of daily precipitation time series, considering variants of ANN models, has been discussed in the work by Partal et al. [24]. Similar study on rainfall prediction can be found in the work by Wu et al. [33]. A comparative study of traditional statistical autoregres-

* Corresponding author.
E-mail addresses: monidipadas@hotmail.com, monidipa42.das@rediffmail.com (M. Das).

http://dx.doi.org/10.1016/j.patrec.2017.08.009
0167-8655/

Please cite this article as: M. Das, S.K. Ghosh, Data-driven approaches for meteorological time series prediction: A comparative study of the state-of-the-art computational intelligence techniques, Pattern Recognition Letters (2017), http://dx.doi.org/10.1016/j.patrec.2017.08.009

sive models and autoregressive NN model, with respect to univariate prediction of rainfall time series, has been reported in the work of Chattopadhyay and Chattopadhyay [7]. The study shows superiority of NN model over the traditional statistical approaches.

However, most of the above-referred works either focus on a single computational intelligence family or consider a single meteorological/environmental parameter. Moreover, none of these works has studied on the Bayesian network-based approaches, which also belong to the computational intelligence family and have proved to show encouraging performance in environmental modeling [1].

In contrast, the present work provides a comparative study considering ANN, SVM, hybrid fuzzy-based model, BN and other probabilistic models together. The novelty lies here in accounting for the Bayesian network and its recently proposed variants. Moreover, one of the major contributions in this work is to propose SpaFBN as new CI technique, extending *spatial Bayesian network with incorporated fuzziness*. Further, in the present work, the comparative study has been made with respect to prediction of *three* primary meteorological variables, namely *temperature, relative humidity* and *precipitation rate*, from two separate climate regions.

In the present context of meteorological time series prediction, the overall prediction problem and the associated challenges have been discussed in the subsequent part of this section. Incidentally, the proposed SpaFBN and the other CI techniques discussed in this paper are applicable for predicting not only the meteorological time series but also the spatial time series from various other disciplines.

1.1. Problem statement and challenges

The problem of meteorological time series prediction, with respect to which the comparative empirical study has been made, can be formally stated as follows:

- Given, the historical daily time series data set over n meteorological parameters in $Z = \{z_1, z_2, \cdots, z_n\}$, corresponding to a set of L locations: $Loc = \{l_1, l_2, \cdots, l_L\}$ for previous t years: $\{y_1, y_2, \cdots, y_t\}$. Also given, the spatial attributes $SA = \{sa_1^{l_i}, sa_2^{l_i}, \cdots, sa_p^{l_i}\}$ for each location $l_i \in Loc$. The problem is to determine the daily time series of the variables in Z for any location $x \in Loc$ for future q years $\{y_{(t+1)}, y_{(t+2)}, \ldots, y_{(t+q)}\}$, when the spatial attributes of x is observed as $\{sa_1^x, sa_2^x, \cdots, sa_p^x\}$. Here, q is a positive integer, i.e. $q \in \{1, 2, 3, \cdots\}$.

The key challenges in such meteorological prediction mainly arise due to the spatio-temporal nature of the data. Unlike the classical data, the spatio-temporal data are highly autocorrelated, that means, the data from nearby locations are more likely to have similar values than those from locations that are far apart. Besides, such data are not independent; rather these are dependent on various co-located variables. Therefore, the conventional statistical methods, which assume that the data are independent and identically distributed, are not very suitable for analyzing such kind of data. Moreover, the meteorological data are non-linear, inherently chaotic, and full of uncertainty.

Research efforts have been made to extend existing traditional statistical and artificial intelligence techniques to cope up with these special properties of meteorological time series data. The present paper aims at summarizing the well-used and recently proposed variants of computational intelligence (CI) techniques, along with their pros and cons in respect of *meteorological time series prediction*. Comparative study has also been carried out empirically, with consideration to three meteorological variables.

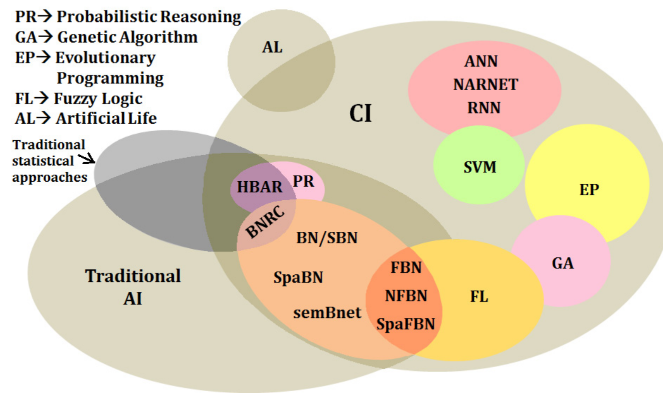

Fig. 1. Relationship between the various CI techniques.

1.2. Contributions

The key contributions of the present study are as follows:

- providing a compact discussion on the various CI techniques (including ANN, SVM, fuzzy logic, BN etc.) used for meteorological time series prediction;
- taking into account the recently proposed variants of Bayesian networks (BNs) which has not been covered by the earlier studies;
- proposing a *new extension of spatial Bayesian network*, namely SpaFBN, that can aid in meteorological prediction with reduced parameter uncertainty;
- performing a comparative empirical study of all the discussed CI techniques, in the light of predicting *temperature*, *humidity* and *precipitation* time series for *two* separate climate regions in India.

1.3. Organization of the paper

The rest of the paper is organized as follows. A comprehensive overview of all the considered CI techniques has been provided in Section 2. The theoretical foundation of the proposed hybrid CI technique (SpaFBN) has been presented in Section 3. The comparative study of all the considered and proposed CI techniques has been extensively discussed in Section 4, with respect to a multivariate time series prediction problem. Finally, the concluding remarks have been made in Section 5.

2. Overview of the state-of-the-art computational intelligence (CI) techniques

Computational intelligence (CI) is a set of nature-inspired computational methods to address the real world problems. It is based on the hypothesis that 'reasoning is computation'. According to Konar [15], the CI family consists of *granular computing* (fuzzy sets, probabilistic reasoning etc.), *neural computing* (e.g. artificial neural network or ANN), *evolutionary computing* (genetic algorithm, genetic programming etc.) and their *interaction* with artificial life, chaos theory and others.

The recent advancement of CI has greatly influenced the data-driven modeling, since CI techniques are capable of modeling the complex relationships among the parameters without knowing actual natural processes. A brief overview of all the CI techniques, which have been dealt with in the present study, is presented subsequently. The relationships among these CI techniques are depicted in Fig. 1.

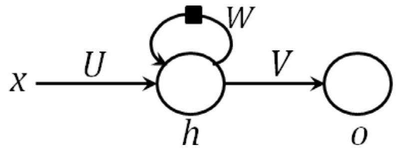

Fig. 2. A typical structure of RNN.

2.1. Feed forward back propagation neural network (FFBP)

The feed forward back propagation (FFBP) neural network is the most popular ANN architecture which has widely been applied especially for precipitation/rainfall prediction [12,23,24].

Levenberg–Marquardt backpropagation (LMBP) algorithm is one of the fastest methods used for NN training in this regard. The LMBP algorithm uses approximate Hessian matrix for weight update process in the following manner:

$$w_{i+1} = w_i - [J^T J + \mu I]^{-1} J^T e \quad (1)$$

where, I is the identity matrix, J is the Jacobian matrix, $J^T J$ approximates to Hessian matrix, $J^T e$ computes the gradient, μ is a scalar value controlling the learning rate, e is residual error.

Though FFBP is self-adaptive, the training performance of the FFBP is highly dependent on the random weight assignment at the beginning of each training simulation, number of hidden layers, and on the rate of learning. Therefore, to achieve the best FFBP performance, excessive FFBP simulations are needed [24].

2.2. Recurrent neural network (RNN)

A recurrent neural network (RNN) is a class of artificial neural network with at least one feed-back connection allowing it to exhibit dynamic temporal behavior and to learn sequences. In order to process the arbitrary sequences of inputs, RNN has the advantage of using their internal memory. A number of applications of RNN, especially for short term wind speed prediction, can be found in literature [3,4].

There exists several models for RNN. The Elman networks, Jordon networks, echo state network etc. are some of the commonly used RNN models. A typical structure of RNN is shown in Fig. 2.

Let, x_t is the input at time step t and h_t be the hidden state at time step t. Then h_t is calculated based on the previous hidden state and the current input, as follows:

$$h_t = f(Ux_t + Wh_{t-1}) \quad (2)$$

Where, f is a nonlinear function; W and U denotes the weight matrices.

2.3. Nonlinear autoregressive neural network (NARNET)

In contrast to the RNN, the nonlinear autoregressive neural network (NARNET) uses the past values of the time series to predict future values. The NARNET model, to predict the value of a data series y at time t using the past p values of the series, can be represented as follows:

$$y(t) = f(y(t-1), y(t-2), \cdots, y(t-p)) + \epsilon(t) \quad (3)$$

The function $f(.)$ is approximately determined during training of the neural network by updating weights and bias. $\epsilon(t)$ denotes the error of approximation.

The work by Benmouiza and Cheknane [5], Huang et al. [14] etc. are some examples where NARNET has been used for time series prediction. However, the major drawback of NARNET is that it needs high delay horizon and several neurons to achieve better prediction accuracy.

2.4. Support vector machine (SVM)

SVM based prediction, also called support vector regression (SVR), is a mechanism by which a function is estimated using observed data that in turn trains the SVM [28]. There can be linear and non-linear versions of SVM based regression. Given a time series $x(t)$, the prediction functions for linear and non-linear SVR can be represented as follows:

$$f_l(x) = (w.x) + b \quad (4)$$

$$f_{nl}(x) = (w.\phi(x)) + b \quad (5)$$

where, $\phi(x)$ is called the 'Kernel function' with maps a non-linear data to a higher dimensional feature set. w and b are the weights and threshold respectively, the optimal values for which are determined by solving the following minimization problem:

$$\textit{minimize} \qquad \frac{1}{2}||w||^2 + C\sum_{i=1}^{n} L(y(i), f(x(i), w))$$

where, C is a constant, y denotes the truth data, and L indicates a loss function. In case the loss function is a quadratic function, the method is called LS-SVM.

In a number of cases, the SVM/LS-SVM based prediction has shown encouraging result, outperforming the ANN [31,32]. An exhaustive study of SVMs can be found in the work of Sapankevych and Sankar [28].

2.5. Hierarchical Bayesian autoregressive (HBAR) model

The hierarchical Bayesian autoregressive (HBAR) model has been developed by Sahu and Bakar [27] for space-time modeling of large scale data. This is achieved by defining the autoregressive Gaussian predictive process approximation method within hierarchical Bayesian framework.

Let $Z_y(l_i, d)$ represents the observed value of a variable, at location l_i, $(i = 1, \cdots, L)$, on day d $(d = 1, \cdots, D)$ within year y for $y = y_1, \cdots, y_t$. Also let $c_{yj}(l, d)$ denotes the value of the jth covariate, $j = 1, \cdots, n$, on day d in year y and $C_{yd} = (c'_y(l_1, d), \cdots, c'_y(l_L, d))'$. Then, as per the HBAR model, the variable value at new location x and at day $D+1$ becomes:

$$Z_y(x, D+1) = c'_y(x, D+1)\beta + \tilde{\eta}_y(x, D+1) + \epsilon_y(x, D+1) \quad (6)$$

where, β denotes the prior distribution, and $\tilde{\eta}_y$ and ϵ_y are determined through autoregressive models with Gaussian predictive process approximations. The details of HBAR model can be found in the work of Sahu and Bakar [27].

Though HBAR is feasible for simultaneous modeling and analysis of large data sets, this is still time intensive because of non-availability of appropriate software.

2.6. Classical/standard Bayesian network (SBN)

The standard Bayesian network (SBN), also called Bayes network or belief network, is a probabilistic graphical model that allows us to represent and reason about uncertain domain. It is essentially a directed but acyclic graph, the nodes of which represent different random variables, and the edges represents the dependency between the variables. Each node X_i in the network is associated with conditional probability distribution $P(X_i|Parents(X_i))$, quantifying the effect of the parents on the node. The dependency structure in Bayesian network can be simply represented as joint probability density function *(PDF)* of the variables, by using factorization as a product of conditional/marginal probability distributions as presented below:

$$P(x_1, x_2, \cdots, x_i, \cdots, x_n) = \prod_{i=1}^{n} P(x_i \mid parents(x_i)) \tag{7}$$

where, x_i is a specific value for variable X_i and $parents(x_i)$ denotes the specific values of the variables in $Parents(X_i)$. This helps to easily get solutions of complex problems.

As per [1], Bayesian networks are highly suitable for environmental modeling. The work by Nandar [22], Madadgar and Moradkhani [18] etc. are some examples in this regard.

2.7. Fuzzy Bayesian network (FBN)

Fuzzy probability theory is an extension of classical probability theory that can better handle the uncertainty or imprecision present in data. A number of Bayesian network models with incorporated fuzzy logic [17,25,30] have been proposed till date. Among these, the most widely used fuzzy Bayesian network (FBN) is the one proposed by Tang and Liu [30].

Let $X = \{X_1, X_2, \cdots, X_p\}$ and $Y = \{Y_1, Y_2, \cdots, Y_q\}$ be two sets of events. Also let $\tilde{X}$ and $\tilde{Y}$ be two corresponding fuzzy events. Then according to this system of FBN,

$$P(\tilde{Y}|\tilde{X}) = \frac{\sum_{j=1}^{q}\sum_{i=1}^{p} \mu_{\tilde{Y}}(Y_j) \cdot \mu_{\tilde{X}}(X_i) \cdot P(X_i|Y_j) \cdot P(Y_j)}{P(\tilde{X})} \tag{8}$$

where, $P(\tilde{X})$ is fuzzy marginal probability, estimated as follows:

$$P(\tilde{X}) = \sum_{i=1}^{p} \mu_{\tilde{X}}(X_i) \cdot P(X_i) \tag{9}$$

In the meteorological prediction [9], FBN has shown better performance than ANN, SBN and the traditional statistical prediction models. However, the FBN suffers from high computational complexity.

2.8. New fuzzy Bayesian network (NFBN)

The NFBN [9] is a variant of FBN [30], and it produces more precise parameter estimates, considering the fuzzy membership of each individual observed values into the other ranges. Moreover, NFBN reduces the time requirement by using more simplistic computation involving only the observed values having non-zero membership in the considered range.

Let x and y be two variables, and $\{X_1, \cdots, X_p\}$ and $\{Y_1, \cdots, Y_q\}$ be the two sets of events corresponding to x and y respectively. Here, $X_1, \cdots, X_p$ and $Y_1, \cdots, Y_q$ are in the form of range of values achieved by x and y; $p, q \in I^+$, where I^+ is the set of positive integers. Also let $\tilde{X}$ and $\tilde{Y}$ be any two corresponding fuzzy events. Then according to *NFBN* [9],

$$P'(\tilde{Y}/\tilde{X}) = \frac{|\{m_i|\mu_{\tilde{Y}}(y_{m_i}) > 0, \quad \mu_{\tilde{X}}(x_{m_i}) > 0\}|}{N.P'(\tilde{X})} \tag{10}$$

where, $\{m_1, m_2, \cdots, m_N\}$ is a set of all the observations for the variable x and y; N is the total number of such observations; x_{m_i}= Value of the variable x in the ith observation (m_i); y_{m_i}= Value of the variable y in the ith observation (m_i); $\mu_{\tilde{X}}(x_{m_i})$ = Membership of the value x_{m_i} in the fuzzy set $\tilde{X}$; $\mu_{\tilde{Y}}(y_{m_i})$ = Membership of the value y_{m_i} in $\tilde{Y}$; '||' denotes set cardinality.

In *NFBN*, the fuzzy marginal probability $P'(\tilde{X})$ is defined as:

$$P'(\tilde{X}) = \frac{|\{m_i|\mu_{\tilde{X}}(x_{m_i}) > 0, m_i \in \{m_1, m_2, \cdots, m_N\}\}|}{N} \tag{11}$$

where, $\{m_1, \cdots, m_N\}$ is a set of all observations for the variable x; N is the total number of observations for x; $\mu_{\tilde{X}}(x_{m_i})$ = Membership of the value x_{m_i} in the fuzzy set $\tilde{X}$. '||' denotes set cardinality.

In meteorological time series prediction [9], NFBN has been able to show better performance than ANN, SBN, FBN and traditional statistical models.

2.9. Hybrid BN with residual correction (BNRC)

In the work of Das et al. [11], a hybrid structure of Bayesian network with residual correction module has been proposed to handle the situation of scarce data availability. As the number of variables increases, the parameter estimation in SBN requires more and more data to maintain the accuracy. To overcome this limitation, during the inference generation in BNRC, the inferred values are tuned/ corrected in an exponential manner with the help of error in previous state.

Let, at the end of training with the data of past t years, the final value of residual is ϵ_t, and the inferred value of the prediction variable for the year $y_{(t+1)}$ is $I_{(t+1)}$. Then, as per BNRC, the final predicted value of the variable becomes:

$$P_{(t+1)} = I_{(t+1)} + \epsilon_t \tag{12}$$

The residual value ϵ_t is determined as follows:

$$\epsilon_t = (\alpha E_{t-1}) + (1-\alpha)\epsilon_{t-1} \tag{13}$$

where, $\alpha \in [0, 1]$; E_{t-1} is the error corresponding to the year $y_{(t-1)}$ and is calculated as follows:

$$E_{t-1} = ActualValue - P_{t-1} \tag{14}$$

In comparison with the ANN (having FFBP architecture) and standard Bayesian network (SBN), the BNRC has been found to show better performance in hydrological time series prediction [11].

2.10. Semantic Bayesian network (semBnet)

The *semBnet* is an extension of SBN, which is able to incorporate domain knowledge or semantics during data analysis and prediction. It has been recently proposed by Das and Ghosh [10] and has been able to show encouraging performance in meteorological time series prediction.

The causal dependency graph of *semBnet* consists of two types of nodes: one set of nodes represents random variables with no semantic information available for themselves, and the other set of nodes represents random variables having semantic information available for themselves. The semantic information or domain knowledge is represented in terms of some semantic hierarchy or concept hierarchy. Then, during the semantic Bayesian learning, the marginal and conditional probabilities (P†) are determined considering semantic similarity (SS) between all relevant concepts. It can be represented as follows:

$$P^{\dagger} = f(P, SS) \tag{15}$$

where, P is the probability distribution obtained through SBN analysis. The semantic similarity $SS(c_1, c_2)$, between any two concepts c_1 and c_2, is determined based on the length of the shortest path between the two concepts and the depth of the subsume, i.e. the nearest common concept. The detailed working principle of *semBnet* can be found in the work by Das and Ghosh [10].

Though *semBnet* provides better prediction accuracy than SBN, ANN, FBN etc., it is time intensive like the SBN.

2.11. Spatial Bayesian network (SpaBN)

SpaBN has been proposed by Das et al. [8] to model the spatial variability of the influencing variables and thereby improve spatio-temporal prediction. Unlike the standard Bayesian network, SpaBN contains composite nodes along with the standard/classical nodes in the directed acyclic graph (DAG). Each composite node

Please cite this article as: M. Das, S.K. Ghosh, Data-driven approaches for meteorological time series prediction: A comparative study of the state-of-the-art computational intelligence techniques, Pattern Recognition Letters (2017), http://dx.doi.org/10.1016/j.patrec.2017.08.009

Table 1
Comparative study of the various computational intelligence (CI) techniques.

CI Family	CI Technique	Pros	Cons
ANN	FFBP	Efficient alternative for traditional methods for modeling the nonlinear time series; Self adaptive, having advantage of learning from previous error;	Performance is sensitive to the initial weight assignments, the number of hidden layers, and proper setting of the learning rate; Can not implicitly cope up with spatial/ spatio-temporal properties of data;
	RNN	Able to learn temporal sequences; Takes into account some kind of long-term dependencies;	High complexity; Can not implicitly cope up with spatial property of data;
	NARNET	If properly trained, it can provide suitable prediction accuracy for non-linear time series prediction;	Needs more complex model to get a good predictor; Can not utilize external information;
SVM	SVM	Convergence to the optimal solution is guaranteed; Less parameters involved;	Finding suitable heuristic to determine the free parameters is challenging; Can not implicitly cope up with spatial/ spatio-temporal properties of data;
Bayesian	HBAR	Fit for spatio-temporal modelling; Shows good performance for out-of-sample predictions as well;	Time intensive; Can not deal with the spatial semantics and any other kind of domain knowledge;
	SBN	Can reason with risk and uncertainty; Can automatically capture probabilistic information;	Can not implicitly cope up with spatial/ spatio-temporal properties of data; Sometimes needs expert knowledge for structuring; Exponential time and space complexity;
	BNRC	Retains all the advantages of BN/SBN; Self adaptive, having advantage of learning from previous error;	Needs expert knowledge for structuring; Can not implicitly cope up with spatial/ spatio-temporal properties of data;
	FBN	Retains all the advantages of BN/SBN; Better deals with ambiguity due to lack of information/knowledge;	Needs expert knowledge for structuring; Can not implicitly cope up with spatial/ spatio-temporal properties of data;
Bayesian Networks (BNs)	NFBN	Can deal with the problem of data discretization; Retains all the advantages of BN/SBN and FBN; Better alternative for FBN; More precise and less complex than FBN;	Extremely high time and space complexity; Needs expert knowledge for structuring; Can not implicitly cope up with spatial/ spatio-temporal properties of data;
	semBnet	Retains all the advantages of BN/SBN; Can incorporate the domain knowledge for providing better insights;	Needs expert knowledge for structuring; Time intensive like the SBN;
	SpaBN	Retains all the advantages of BN/SBN; Implicitly cope up with spatial variability and auto-correlation property of data; Drastically reduces the time complexity;	Needs expert knowledge for structuring; Suffers from parameter uncertainty due to sampling of discretized data; Can not deal with the spatial semantics;

is a composition of a number of standard/classical nodes associated with the same but spatially distributed variable. In SpaBN, the marginal and conditional probabilities are measured considering the classical probability distribution and also the spatial importance of each representative location associated with each constituting node within a composite node. Therefore, in SpaBN, the probability distribution ($P^{\dagger\dagger}$) can be represented as follows.

$$P^{\dagger\dagger} = f(P, SW) \tag{16}$$

Where, P is the probability distribution obtained through SBN analysis. SW denotes some spatial weight, reflecting the relevant importance of the associated spatial locations. The details of SpaBN can be found in the work by Das et al. [8].

The pros and cons of the discussed CI techniques are summarized in Table 1.

3. Proposed spatial fuzzy Bayesian network (SpaFBN)

Though the SpaBN has proved itself to be a superior extension of standard Bayesian network by outperforming the other models, it considers all the variables to be discrete, and therefore, suffers from the generic problem of data discretization to deal with the continuous variables. Whenever, a variable is discretized into ranges, problem arises for the boundary values. Each value, at the range boundary or adjacent to it, is treated to be strictly within one range, though it has enough relevance with its subsequent

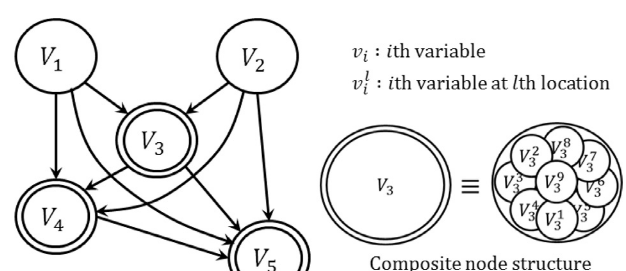

Fig. 3. A typical causal dependency graph (CDG) for SpaFBN.

range. Therefore, during parameter learning in SpaBN, a substantial number of records are skipped, leading to increased parameter uncertainty.

In order to overcome this limitation, in the present paper, a new variant of SpaBN, termed as SpaFBN, has been proposed. The proposed SpaFBN is a fuzzy extension of the spatial Bayesian network (SpaBN), which is able to reduce parameter uncertainty with the incorporated fuzziness. It can be treated as a hybrid CI technique, taking advantages of both NFBN and SpaBN.

To explain the working principle of proposed SpaFBN, let's consider the example causal dependency graph, shown in Fig. 3.

According to the scenario depicted through Fig. 3, the variables V_3, V_4 and V_5 are spatially distributed over $L = 9$ number of locations, and therefore have been represented with composite nodes, denoted by double-lined circles. The internal structure of a composite node is shown in Fig. 3. Variables V_3, V_4, and V_5 are also influenced by V_1 and V_2, which are not distributed spatially. Therefore, V_1 and V_2 have been represented by standard/ classical nodes.

Now, as per SpaFBN, the marginal probability distribution for any composite node, say V_3, is determined as follows:

$$P''(V_3) = \alpha . \left[\sum_{i=1}^{L} P'(V_3^i).SW_i \right] \tag{17}$$

where, α is normalization constant and SW_i is the spatial weight/ importance for the i-th location. $P'(V_3^i)$ is the marginal probability of the singular component V_3^i of composite node V_3, obtained in similar manner as that of NFBN (refer Eq. (11)). This helps to deal with the problem of data discretization, as faced by the SpaBN.

The marginal probability of other (standard) nodes (e.g. V_1, V_2 etc.) in the proposed SpaFBN are determined in a similar way as that of NFBN (refer Eq. (11)).

In the proposed SpaFBN, the conditional probability involving composite nodes is also determined with consideration to the spatial importance of the associated locations. For example, the conditional probability $P''(V_4|V_1, V_2, V_3)$ is calculated as follows:

$$P''(V_4|V_1, V_2, V_3) = \alpha . \left[\sum_{i=1}^{L} P'(V_4^i|V_1, V_2, V_3^i).SW_i \right] \tag{18}$$

where, α is normalization constant; SW_i is spatial weight/ importance for the ith location; and $P'(V_4^i|V_1, V_2, V_3^i)$ is the conditional probability involving singular component V_3^i and V_4^i of composite nodes V_3 and V_4, respectively. The $P'(V_4^i|V_1, V_2, V_3^i)$ is estimated in the same way as that of NFBN (refer Eq. 10). It helps to overcome the problem of data discretization and leads to reduced parameter uncertainty.

In the proposed SpaFBN, the conditional probability involving only standard nodes are determined in a similar manner as that of NFBN (refer Eq. (10)).

Now, once the parameter learning is over, the inference in SpaFBN is generated by utilizing the spatial weights (SW_i). For example, let the observed/ evidence variables are V_1 and V_2, from which the value of V_4 is to be inferred.

Then, as per SpaFBN, the inferred value of V_4 becomes:

$$V_4^{inferred} = \sum_{i=1}^{L} P'(V_4^i|V_1, V_2) \cdot SW_i \tag{19}$$

$$= \sum_{i=1}^{L} \sum_{V_3} P'(V_4^i|V_1, V_2, V_3^i) \cdot P'(V_3^i|V_1, V_2) \cdot SW_i \tag{20}$$

where the value for $P'(V_4^i|V_1, V_2, V_3^i)$ and $P'(V_3^i|V_1, V_2)$ can be determined from the conditional probability table for the variable V_4 and V_3, respectively. Among these inferred values, the predicted value becomes the one corresponding to the maximum probability estimate.

Spatial Weight/Importance calculation:

In the work of Das et al. [8], the spatial weight (SW_i) estimation has been illustrated with respect to hydrological prediction. In the present study, the weight estimation is described with respect to meteorological prediction.

Let the normalized inverse spatial distance between the ith location and the prediction location is $NISD_i$. Also let the *normalized correlation* between the time series of kth meteorological variable in the ith neighborhood location and that in the prediction location is $NCorr_{V_k}^i$. Then the spatial weight of the ith location is determined as follows:

$$SW_i = \frac{\sum_{k=1}^{n} NCorr_{V_k}^i + NISD_i}{\sum_{j=1}^{L} (\sum_{k=1}^{n} NCorr_{V_k}^j + NISD_j)} \tag{21}$$

where, $i = 1, 2, \cdots, L$. The estimation of SW_i can be varied based on the area of application.

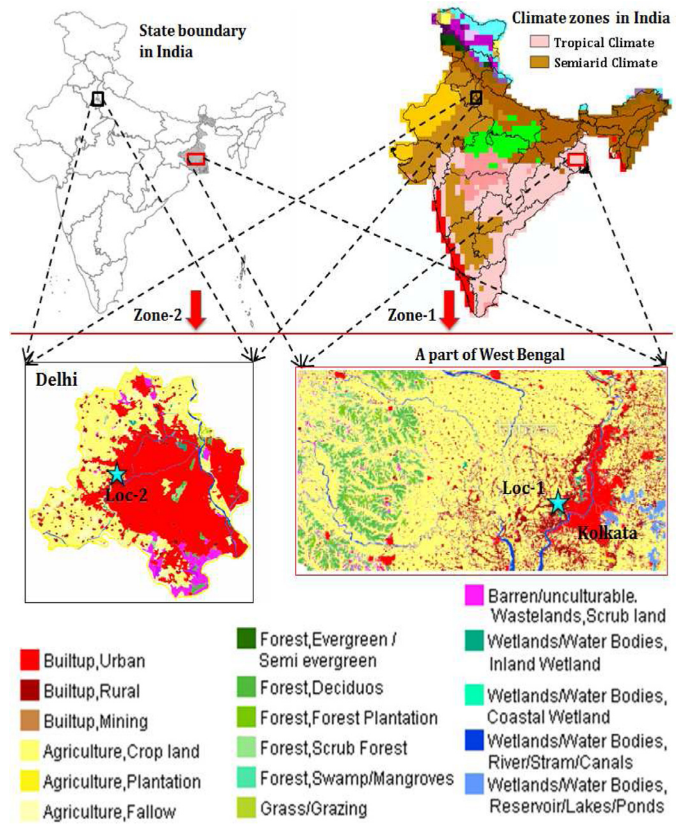

Fig. 4. Study zones and prediction locations: Loc-1 (22.82°N, 88.29°E) in West Bengal, India; Loc-2 (28.66°N, 77.07°E) in Delhi, India.

4. Performance evaluation using empirical study

This section provides a comparative performance analysis of all the discussed CI techniques (including the proposed SpaFBN), with respect to a case study on meteorological time series prediction in *India*. The details of study area, data sets, experimental setup, and prediction results have been thoroughly described in the subsequent part of this section.

4.1. Data set and study area

The experimentation has been carried out to predict three primary meteorological variables, namely *temperature, humidity*, and *precipitation* time series in two separate study zones in India. The *zone-1* belongs to tropical climate region in Eastern India, whereas the *zone-2* belongs to semi-arid climate in North-Western India (refer Fig. 4). From each study zone, a set of 10 locations have been randomly chosen as the sensor locations. The historical time series data on the considered meteorological variables for each selected location in the study zones have been collected from the Fetch-Climate Explorer [21]. Prediction has been made for two locations, one from each study zone, for the year 2015 and 2016 respectively. The details of the experimental data sets are given in Table 2.

Table 2
Details of data sets used in the comparative study.

Data sets	Prediction location details					Time series details	
	Locations	Latitude	Longitude	Land-cover	Zone details	Training Yr.	Prediction Yr.
Data set-1 (West Bengal)	Loc-1	22.82°N	88.29°E	Rural	zone-1 [Tropical climate]	2001–2014	2015
Data set-2 (Delhi)	Loc-2	28.66°N	77.07°E	Urban	zone-2 [Semi-arid climate]	2011–2015	2016

4.2. Experimental setup

The comparative study of proposed SpaFBN and all the above-discussed CI techniques has been carried out with respect to three benchmark time series prediction techniques, namely *Automated Autoregressive Integrated Moving Average* (A-ARIMA), *Vector Auto-Regressive Moving Average* (VARMA), and *Generalized Autoregressive Heteroskedasticity* (GARCH) model. The standards provided by R-tool packages [26] have been used to generate predicted time series from each of these models. R-tool has also been used to get forecast results from HBAR model [27].

Besides, the proposed SpaFBN have been implemented using MATLAB 7.12.0 (R2011a) in Windows 7 (64 bit OS, 3.10 GHz CPU, 4.00GB RAM). MATLAB has also been used to achieve predicted series from SpaBN [8], SBN, NFBN [9], FBN [30], BNRC [11], and semBnet [10] model. In order to get prediction results from FFBP [24], RNN, NARNET and SVM, the NNToolbox of MATLAB [19] has been utilized.

For all the considered CI techniques, which can explicitly deal with the spatial information (e.g. SBN, FBN, NFBN, BNRC etc.), the *land elevation, latitude*, and *land-cover* have been used as the spatial attributes. Moreover, the domain knowledge on land-cover, as utilized by semBnet [10], has been collected from the Bhuvan geo-portal [6].

4.3. Performance metrics

The performance has been measured in terms of four popular statistical measures: *normalized root mean square deviation (NRMSD), mean absolute error* (MAE), *mean absolute percentage error* (MAPE) [20], and *Coefficient of determination or R-squared* (R^2). NRMSD, also called Normalized Root Mean Square Error (NRMSE), is the normalized value of RMSE. The formal definition for each of these metrics are given below:

$$NRMSD = \frac{1}{(O_{max} - O_{min})} \sqrt{\frac{1}{N} \sum_{i=1}^{N} (Z_{o_i} - Z_{p_i})^2} \tag{22}$$

$$MAE = \frac{1}{N} \sum_{i=1}^{N} |Z_{o_i} - Z_{p_i}| \tag{23}$$

$$MAPE = \frac{|Z^s_{mo} - Z^s_{mp}|}{|Z^s_{mo}|} \times 100 \tag{24}$$

$$R^2 = \frac{\left[\sum_{i=1}^{N} (Z_{o_i} - Z^s_{mo})(Z_{p_i} - Z^s_{mp})\right]^2}{\sum_{i=1}^{N} (Z_{o_i} - Z^s_{mo})^2 . \sum_{i=1}^{N} (Z_{p_i} - Z^s_{mp})^2} \tag{25}$$

where $O_{\max}$ is the maximum value in the observed series; $O_{\min}$ is the minimum value in the observed series; Z_{o_i} and Z_{p_i} denote the observed value and the corresponding predicted value of the variable for the ith observation; Z^s_{mo} is the mean value of the observed series; Z^s_{mp} is the mean value of the predicted series; and N is the total number of observations in the series.

The best-fit between observed and predicted series leads to NRMSD = 0, MAE = 0, MAPE = 0, and $R^2 = 1$.

Moreover, in order to quantify the prediction uncertainty for each variable, the *Dawid–Sebastiani scores* [13] corresponding to all the considered CI techniques have been estimated. The Dawid–Sebastiani score (*DSS*) is measured as follows:

$$DSS(p, Z_o) = \left(\frac{Z_o - Z^s_{mp}}{Z^s_{sp}}\right)^2 + 2 \log Z^s_{sp} \tag{26}$$

where, Z_o is the observed value, p is the predicted time series, Z^s_{mp} is the mean prediction value, and Z^s_{sp} is the standard deviation of the predicted time series.

4.4. Results and discussions

The prediction performances in terms of NRMSD, MAE, MAPE and R^2 have been presented in Tables 3 and 4, for the *Data set-1* and *Data set-2* respectively. Further, the *DSS* scores for each CI technique have been graphically plotted in Fig. 5 (a)–(f), in comparison with the ideal prediction scenario. It is assumed that, in an ideal scenario, the predicted time series is exactly same as the observed time series.

Analyzing the results, summarized in Tables 3–4 and in Fig. 5, the following inferences can be drawn:

- Tables 3–4 show that, for both the data sets all the CI techniques have outperformed the traditional statistical models (ARIMA, VARMA and GARCH) and thereby proved themselves to be effective tools for data-driven modeling of meteorological time series prediction.
- Among the ANN models, the FFBP with Levenberg–Marquardt backpropagation algorithm is found to show comparatively better performance in case of predicting temperature. The FFBP also tends to produce lesser MAPE in case of predicting precipitation. However, the overall performance of NARNET in predicting precipitation is far better than FFBP. NARNET shows its efficacy in predicting humidity as well. The DSS values in Fig. 5 indicate that the prediction uncertainty corresponding to NARNET model is also lesser than the other ANN models.
- On the other side, it is evident from the Tables 3 to 4 that the HBAR, being a hybrid statistical space-time model, is able to outperform the pure statistical benchmark prediction techniques. However, the performance of ANN and BN based prediction is far better than that of HBAR.
- It can be noted from Tables 3 to 4 and Fig. 5 that the BN based prediction techniques have a generic tendency to outperform the ANN-based models, from both the perspective of prediction accuracy and prediction uncertainty. In case of predicting temperature, NFBN and semBnet are found to be most effective among the fuzzy BN and classical BN models, respectively. However, for the variable precipitation and humidity, the SpaBN and the newly proposed SpaFBN outperforms the others.
- On analyzing the performance of the proposed SpaFBN (refer Tables 3 and 4), it can be found that in most of the cases of prediction (especially for precipitation and humidity), the SpaFBN outperforms the other CI techniques by producing least NRMSD and least MAE values. The high values of R^2 (≈ 1) in all cases also indicate that the series predicted by the SpaFBN have the best match with the observed time series.The performance of SpaFBN in case of temperature prediction is also comparable to that of the NFBN and semBnet.

Please cite this article as: M. Das, S.K. Ghosh, Data-driven approaches for meteorological time series prediction: A comparative study of the state-of-the-art computational intelligence techniques, Pattern Recognition Letters (2017), http://dx.doi.org/10.1016/j.patrec.2017.08.009

Table 3
Prediction performance in case of the Data set-1: Comparative study of the proposed SpaFBN and other CI techniques.

Prediction Techniques	Prediction variables											
	Temperature (T)				Humidity (H)				Precipitation (P)			
	NRMSD	MAE	MAPE	R^2	NRMSD	MAE	MAPE	R^2	NRMSD	MAE	MAPE	$\boldsymbol{R^2}$
A-ARIMA	0.252	3.064	7.278	0.019	0.325	8.873	2.093	0.367	0.389	86.448	37.718	0.150
GARCH	0.214	2.592	0.993	0.000	0.355	8.612	5.137	0.000	0.526	110.975	88.258	0.000
VARMA	0.248	3.293	10.320	0.412	0.407	9.425	11.030	0.137	0.429	101.408	59.257	0.151
SVM	0.222	2.314	4.920	0.130	0.310	6.337	8.832	0.530	0.078	21.026	16.846	0.991
NARNET	0.145	2.093	6.813	0.832	0.137	3.601	5.303	**0.960**	0.066	20.007	16.133	**0.994**
FFBP	0.097	0.761	2.529	0.864	0.252	6.077	6.524	0.625	0.146	24.462	15.913	0.884
RNN	0.104	1.315	0.619	0.783	0.266	6.799	4.332	0.491	0.495	119.292	24.845	0.000
HBAR	0.170	2.371	3.899	0.472	0.133	3.386	2.420	0.874	0.355	88.518	71.203	0.744
SBN	0.080	1.028	0.085	0.862	0.101	2.374	3.239	0.958	0.093	21.457	9.430	0.953
FBN	0.065	0.815	0.101	0.908	0.101	2.374	3.239	0.958	0.080	17.540	6.271	0.962
NFBN	**0.061**	**0.712**	0.649	**0.921**	0.156	4.102	4.436	0.954	0.060	13.410	0.796	0.975
BNRC	0.078	1.026	1.070	0.874	0.100	2.344	3.175	0.958	0.092	21.191	8.771	0.953
semBnet	0.067	0.967	**0.025**	0.875	0.096	**2.245**	2.639	0.958	0.083	19.273	7.668	0.961
SpaBN	0.078	0.997	0.072	0.876	0.100	2.487	2.635	**0.960**	0.036	6.819	1.449	0.991
SpaFBN (proposed)	0.077	0.977	0.060	0.876	**0.091**	2.247	**1.702**	**0.960**	**0.029**	**4.555**	**0.281**	**0.994**

The boldface value indicates the best performance with respect to a given evaluation criteria.

Table 4
Prediction performance in case of the Data set-2: Comparative study of the proposed SpaFBN and other CI techniques.

Prediction Techniques	Prediction variables											
	Temperature (T)				Humidity (H)				Precipitation (P)			
	NRMSD	MAE	MAPE	R^2	NRMSD	MAE	MAPE	R^2	NRMSD	MAE	MAPE	R^2
A-ARIMA	0.245	6.341	7.078	0.217	0.316	10.366	2.954	0.014	0.321	46.364	20.468	0.108
GARCH	0.297	7.148	15.635	0.000	0.305	10.003	0.679	0.000	0.395	43.859	92.881	0.000
VARMA	0.326	7.997	24.982	0.233	0.341	11.645	5.709	0.032	0.356	41.903	68.009	0.089
SVM	0.106	2.430	2.784	0.848	0.322	7.808	11.069	0.130	0.290	29.398	53.675	0.571
NARNET	0.134	3.140	6.536	0.818	0.111	3.595	5.462	0.927	0.095	10.007	14.720	**0.927**
FFBP	0.124	3.016	1.054	0.798	0.219	6.076	7.737	0.690	0.144	15.884	3.581	0.817
RNN	0.490	11.981	5.775	0.009	0.257	8.552	0.646	0.316	0.349	46.170	38.353	0.067
HBAR	0.207	5.379	5.363	0.481	0.258	7.989	1.384	0.336	0.340	47.247	62.597	0.414
SBN	0.100	2.375	0.848	0.865	0.093	1.994	1.595	0.920	0.161	15.844	10.105	0.846
FBN	0.095	2.148	**0.092**	0.877	0.093	1.994	1.595	0.920	0.157	17.551	12.748	0.845
NFBN	0.088	**1.932**	0.144	0.893	0.061	1.335	0.948	0.963	0.103	8.313	**1.435**	0.904
BNRC	0.100	2.366	0.528	0.865	0.064	1.597	1.833	**0.965**	0.135	14.150	4.433	0.882
semBnet	**0.083**	1.970	0.153	**0.903**	0.063	1.584	1.685	**0.965**	0.135	14.246	6.645	0.882
SpaBN	0.106	2.388	0.639	0.874	0.063	1.584	1.685	**0.965**	0.123	13.478	8.061	0.910
SpaFBN (proposed)	0.097	2.169	0.117	0.888	**0.059**	**1.247**	**0.510**	**0.965**	**0.091**	**8.015**	3.141	**0.927**

The boldface value indicates the best performance with respect to a given evaluation criteria.

- Moreover, it is evident from Tables 3 and 4 that in every case the presently proposed SpaFBN outperforms the SpaBN, from which it has been derived. The prediction uncertainty (refer Fig. 5) corresponding to SpaFBN is also lesser than both SpaBN and NFBN in all cases, thus indicating the effectiveness of incorporating fuzziness in spatial BN based prediction.

Overall, all the variants of Bayesian networks (BNs) have shown high degree of potentiality in dealing with prediction uncertainty and providing comparatively better prediction accuracy from every respect. Further, the encouraging performance of the presently proposed SpaFBN fulfills our motivation of incorporating fuzziness in spatial Bayesian analysis to overcome the problem of discretized data and thereby reducing parameter uncertainty. The comparative study reveals that the proposed spatial fuzzy Bayesian network (SpaFBN) can be used as an effective CI technique for predicting meteorological time series, especially humidity and precipitation rate.

5. Conclusions

This paper provides a comparative study of the state-of-the-art computational intelligence (CI) techniques, which have extensively promoted the progress of data-driven modeling for meteorological time series prediction. The major contributions in this work lies in: (i) comprehensive discussion of eleven variants of CI techniques, including ANN, NARNET, RNN, SVM, HBAR, SBN, FBN, NFBN, BNRC, semBnet, and SpaBN; (ii) proposing a new, fuzzy extension of spatial Bayesian network, namely SpaFBN; and (iii) comparative empirical study with respect to prediction of *temperature, humidity* and *precipitation rate* time series for *two* locations from different climate regions in *India*. The study reveals that the CI-based meteorological prediction is more promising than that based on traditional pure statistical methods. More specifically, the experimental results show better performance of BN-based (especially the proposed SpaFBN based) models, compared to the others.

This study may help researchers to investigate advanced CI based models for meteorological time series prediction. In future, the work can be extended by considering other families of CI techniques, like genetic algorithm, genetic programming, hybrid neuro-fuzzy approach etc., and other kinds of spatial time series from diverse domains like hydrology, ecology, biology, medicine, remote sensing, and so on. Further, though as per the present objective in the paper, the performance of the proposed SpaFBN has been studied with respect to meteorological prediction, the SpaFBN is a

Please cite this article as: M. Das, S.K. Ghosh, Data-driven approaches for meteorological time series prediction: A comparative study of the state-of-the-art computational intelligence techniques, Pattern Recognition Letters (2017), http://dx.doi.org/10.1016/j.patrec.2017.08.009

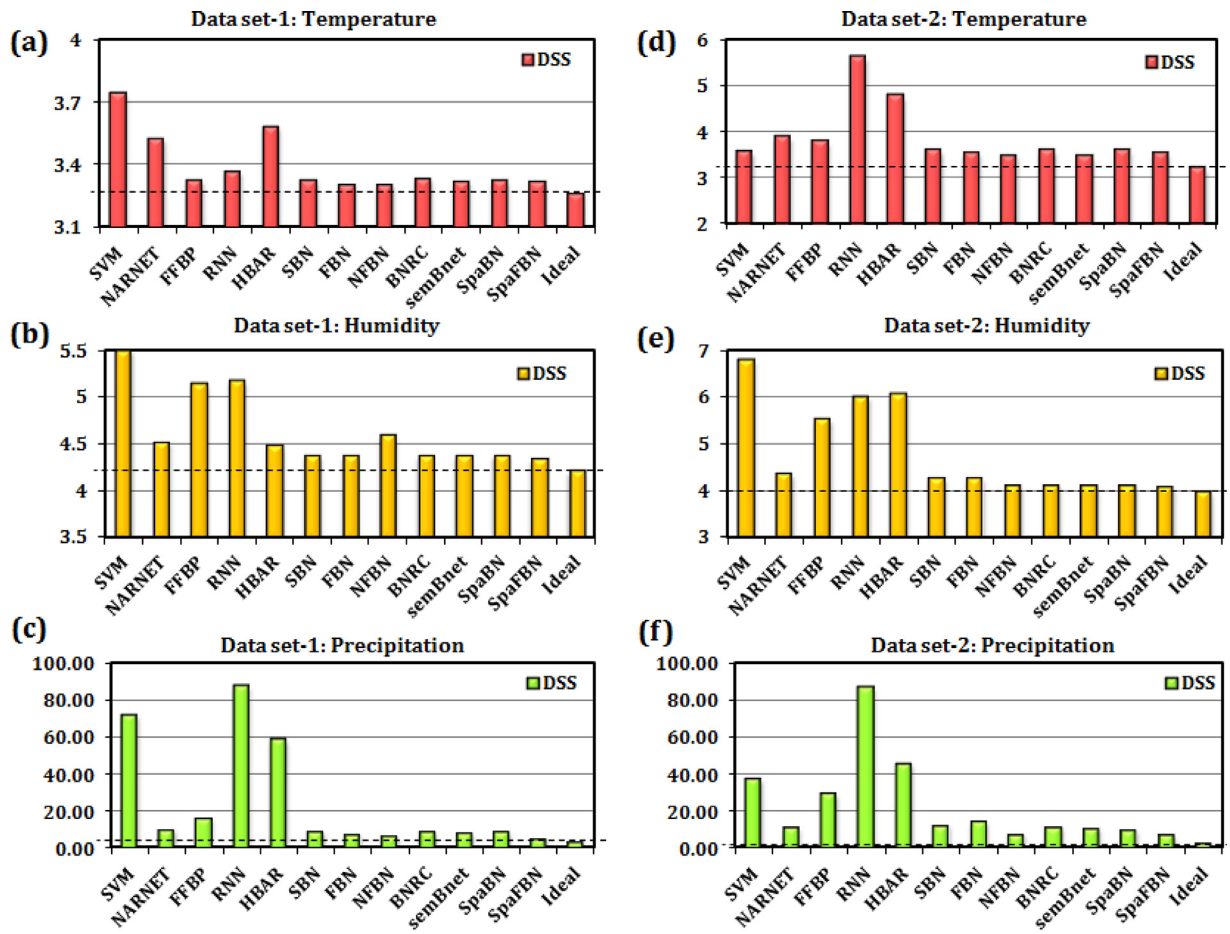

Fig. 5. Dawid–Sebastiani scores (DSS) for the CI techniques: (a)–(c) Data set-1; (d)–(f) Data set-2.

generic CI/machine learning technique, and therefore, ample scope remains in exploring SpaFBN for analyzing spatio-temporal data from other disciplines as well.

References

[1] P. Aguilera, A. Fernández, R. Fernández, R. Rumí, A. Salmerón, Bayesian networks in environmental modelling, Environ. Modell. Softw. 26 (12) (2011) 1376–1388.
[2] S. Al-Yahyai, Y. Charabi, A. Gastli, Review of the use of numerical weather prediction (nwp) models for wind energy assessment, Renewable Sustainable Energy Rev. 14 (9) (2010) 3192–3198.
[3] T. Barbounis, J.B. Theocharis, Locally recurrent neural networks for wind speed prediction using spatial correlation, Inf. Sci. 177 (24) (2007) 5775–5797.
[4] T.G. Barbounis, J.B. Theocharis, M.C. Alexiadis, P.S. Dokopoulos, Long-term wind speed and power forecasting using local recurrent neural network models, IEEE Trans. Energy Convers. 21 (1) (2006) 273–284.
[5] K. Benmouiza, A. Cheknane, Forecasting hourly global solar radiation using hybrid k-means and nonlinear autoregressive neural network models, Energy Convers. Manage. 75 (2013) 561–569.
[6] Bhuvan, *Indian Geo-Platform of ISRO*, 2016, (http://bhuvan.nrsc.gov.in/bhuvan_links.php#). [Online; Accessed 09-Dec-2016].
[7] S. Chattopadhyay, G. Chattopadhyay, Univariate modelling of summer-monsoon rainfall time series: comparison between arima and arnn, C.R. Geosci. 342 (2) (2010) 100–107.
[8] M. Das, S. Ghosh, P. Gupta, V. Chowdary, R. Nagaraja, V. Dadhwal, Forward: a model for forecasting reservoir water dynamics using spatial bayesian network (spabn), IEEE Trans. Knowl. Data Eng. 29 (4) (2017) 842–855, doi:10.1109/TKDE.2016.2647240.
[9] M. Das, S.K. Ghosh, A probabilistic approach for weather forecast using spatio-temporal inter-relationships among climate variables, in: Industrial and Information Systems (ICIIS), 2014 9th International Conference on, IEEE, 2014, pp. 1–6.
[10] M. Das, S.K. Ghosh, semBnet: A Semantic Bayesian Network for Multivariate Prediction of Meteorological Time Series Data, Pattern Recognit. Lett. 93 (2017) 192–201, doi:10.1016/j.patrec.2017.01.002.
[11] M. Das, S.K. Ghosh, V. Chowdary, A. Saikrishnaveni, R. Sharma, A probabilistic nonlinear model for forecasting daily water level in reservoir, Water Resour. Manage. 30 (9) (2016) 3107–3122.
[12] J. Farajzadeh, A.F. Fard, S. Lotfi, Modeling of monthly rainfall and runoff of urmia lake basin using feed-forward neural network and time series analysis model, Water Resour. Ind. 7 (2014) 38–48.
[13] T. Gneiting, M. Katzfuss, Probabilistic forecasting, Annu. Rev. Stat. Appl. 1 (2014) 125–151.
[14] J. Huang, M. Korolkiewicz, M. Agrawal, J. Boland, Forecasting solar radiation on an hourly time scale using a coupled autoregressive and dynamical system (cards) model, Sol. Energy 87 (2013) 136–149.
[15] A. Konar, Computational Intelligence: Principles, Techniques and Applications, Springer Science & Business Media, 2006.
[16] M. Lei, L. Shiyan, J. Chuanwen, L. Hongling, Z. Yan, A review on the forecasting of wind speed and generated power, Renewable Sustainable Energy Rev. 13 (4) (2009) 915–920.
[17] P.-c. Li, G.-h. Chen, L.-c. Dai, L. Zhang, A fuzzy bayesian network approach to improve the quantification of organizational influences in hra frameworks, Saf. Sci. 50 (7) (2012) 1569–1583.
[18] S. Madadgar, H. Moradkhani, Spatio-temporal drought forecasting within bayesian networks, J. Hydrol. 512 (2014) 134–146.
[19] MATLAB, Mathworks, 2016, (http://in.mathworks.com/products/matlab/?requestedDomain=www.mathworks.com). [Online; Accessed 15-Sep-2016].
[20] A. Mellit, A.M. Pavan, M. Benghanem, Least squares support vector machine for short-term prediction of meteorological time series, Theor. Appl. Climatol. 111 (1–2) (2013) 297–307.
[21] Microsoft-Research, FetchClimate, 2016, (http://research.microsoft.com/en-us/um/cambridge/projects/fetchclimate/app/). [Online; Jan-2016].
[22] A. Nandar, Bayesian network probability model for weather prediction, in: Current Trends in Information Technology (CTIT), 2009 International Conference on the, IEEE, 2009, pp. 1–5.
[23] D.R. Nayak, A. Mahapatra, P. Mishra, A survey on rainfall prediction using artificial neural network, Int. J. Comput. Appl. 72 (16) (2013).
[24] T. Partal, H.K. Cigizoglu, E. Kahya, Daily precipitation predictions using three different wavelet neural network algorithms by meteorological data, Stochastic Environ. Res. Risk Assess. 29 (5) (2015) 1317–1329.

[25] C.A. Penz, C.A. Flesch, S.M. Nassar, R.C. Flesch, M.A. De Oliveira, Fuzzy–bayesian network for refrigeration compressor performance prediction and test time reduction, Expert Syst. Appl. 39 (4) (2012) 4268–4273.

[26] R, *R-3.2.2 for Windows (32/64 bit)*, 2016, (https://cran.r-project.org/bin/windows/base/old/3.2.2/). [Online; Accessed 21-Dec-2016].

[27] S.K. Sahu, K.S. Bakar, Hierarchical bayesian autoregressive models for large space–time data with applications to ozone concentration modelling, Appl. Stoch. Models Bus. Ind. 28 (5) (2012) 395–415.

[28] N.I. Sapankevych, R. Sankar, Time series prediction using support vector machines: a survey, IEEE Comput. Intell. Mag. 4 (2) (2009).

[29] D. Solomatine, L.M. See, R. Abrahart, Data-driven modelling: concepts, approaches and experiences, in: Practical hydroinformatics, Springer, 2009, pp. 17–30.

[30] H. Tang, S. Liu, Basic theory of fuzzy bayesian networks and its application in machinery fault diagnosis, in: Fuzzy Systems and Knowledge Discovery, 2007. FSKD 2007. Fourth International Conference on, 4, IEEE, 2007, pp. 132–137.

[31] U. Thissen, R. Van Brakel, A. De Weijer, W. Melssen, L. Buydens, Using support vector machines for time series prediction, Chemom. Intell. Lab. Syst. 69 (1) (2003) 35–49.

[32] T.B. Trafalis, B. Santosa, M.B. Richman, Prediction of rainfall from wsr-88d radar using kernel-based methods, Int. J. Smart Eng. Syst. Des. 5 (4) (2003) 429–438.

[33] C. Wu, K. Chau, C. Fan, Prediction of rainfall time series using modular artificial neural networks coupled with data-preprocessing techniques, J. Hydrol. 389 (1) (2010) 146–167.

Please cite this article as: M. Das, S.K. Ghosh, Data-driven approaches for meteorological time series prediction: A comparative study of the state-of-the-art computational intelligence techniques, Pattern Recognition Letters (2017), http://dx.doi.org/10.1016/j.patrec.2017.08.009