SNOWFALL PREDICTION TECHNIQUES - A STATE OF THE ART REVIEW

C Kishor Kumar Reddy
Research Scholar, Dept. of CSE
K L University, Guntur, India
Kishoar23@gmail.com

B Vijaya Babu
Professor, Dept. of CSE
K L University, Guntur, India
Vijaymtech28@gmail.com

Abstract: - Snow carries the imperative role in the matter of aquatic, animal and human life in complete etiquette of climate circumstances, considered to be one of the most natural climatic wonders whose nowcasting is arduous and challenging. In the mountainous areas, snow has a resilient impact on traffic, sewer systems. On the other hand, the snow hydrological cycle is one of the most amalgamated and challenging elements to distinguish and to model. This survey paper provides a diverse assortment of approaches which are epitomized by various academicians, researchers, scientists and meteorological departments in nowcasting snow. Hence the upshot of this paper is to show the available techniques and therefore comparison table exposes the accuracy and future scope of an individual methods used for nowcasting snowfall.

Key-Words: - Artificial Intelligence, Classification, Hydrology, Neural Networks, Nowcasting

1 Introduction

Since ancient times, forecasting of snow is technologically and scientifically exigent problems around the globe. Researchers and Scientists have tried to forecast snow using a number of approaches, some of these approaches is more precise than previous models. Since the oldest human civilization, people have attempted to Nowcast snow informally. The increasing accessibility of weather data from different sources: radar, raiosonde balloons, observational records and satellite maps, observations from ship and aircraft, proxy data, etc, during the last decades created an importance to find efficient and precise models to extract and analyze hidden data from the vast amount of data. At present, snow nowcasting is made through the application of science and technology [1] [22-23].

The weather forecasts are divided into the following categories:

Now casting: in which the particulars about the present weather and forecasts up to a few hours ahead are given.

Short range forecasts: in which the particulars about the weather between 1-3 days.

Medium range forecasts: in which the particulars about the weather between 4-10 days.

Long range /Extended Range forecasts: in which particulars about the weather above 10 days.

Worldwide, an ample range of snow nowcasting methods are employed. Basically there are two methods to Nowcast snow: Empirical approach and Dynamical approach. The Empirical method is completely based on the study of weather historical data and its association to a mixture of atmospheric variables: humidity, temperature, pressure, wind speed, dew point, visibility and so on. The most common empirical approaches used for snow nowcasting are Artificial Intelligence, Neural Networks, Classification, Regression, Fuzzy Logic etc [19][24]. In dynamical approach, physical models based on system of equations generate the predictions of future snow. Nowcasting of weather by processor using statistical formulae are known as numerical weather prediction. To predict the weather by numerical weather prediction, meteorologist has developed atmospheric models that approximate the changes in humidity, wind speed, temperature, pressure, dew point etc using mathematical equations [22-23].

During winter season, Snow precipitation around the globe makes the movement of civilians as well as army personnel complicated due to closure of roads, leads to the growth of perilous situations due to avalanches. In the mountainous areas, snowfall has a durable impact on traffic, sewer systems and also some more human undertakings. To make the best of fair weather days, precise snow and avalanche forecasts are required in advance, in order to ensure safe movement of army and civilians, necessary services and supplies along road tracks. On the other hand, the hydrological cycle of snowfall is one of the most complex and difficult elements to distinguish and to model. This is due to the intricacy of the various atmospheric processes: humidity, temperature, pressure, wind speed, dew point, visibility etc. which create snowfall and the momentous range of disparity over a wide range of scales mutually in time and space. Till date, many snowfall prediction techniques are proposed, but not fully satisfied because the snow has nonlinear nature. Hence, there is urgent need for developing a snow nowcasting model, predicts weather in terms of snow/no snow and likely amount of snow in advance [22-23].

For the past era, there are many machine learning approaches based on historical weather datasets, Satellite and Radar imagery have simplified for the growth of innovative methods to global precipitation observations. In recent times various machine learning precipitation algorithms have been established which produce precipitation products involving of higher spatial and temporal resolution that is useful for hydrologic researches and water resources applications [19][24].

The rest of the document is organized as follows: Section 2 depicts the previous approaches, section 3 provides the summary on previous approaches and section 4 concludes the paper followed by references.

2 PREVIOUS APPROACHES

2.1 Intelligent Visibility Meter [4]

It measures the intensity of forward scattered infrared rays in the atmosphere and converses MOR (Meteorological Optical Range) value. Furthermore, by means of the artificial intelligence methodologies, it observes category and intensity of precipitation. Chief functions of Intelligent Visibility Meter are given as follows: Calculation function of accumulated snow depth, amount of precipitation, intensity of precipitation, visibility and classification of precipitation [4].

The snowfall nowcasting precision about 74% is achieved by using precipitation type code of Intelligent Visibility Meter as a result of the analysis. This means that the very short term snowfall forecasting of the specific area can be achieved from the ground observation value with Intelligent Visibility Meter without using upper air layer observation data. [4].

2.2 Machine Learning Tools for Long-Lead Heavy Precipitation Prediction with Multi-sensor Data [2] [19][24]

Authors introduced two new concepts including the NEPCs and the EPCs and designed the NSC to form the balanced data set, and apply the Fast-OSFS and DFS to search for the most-relevant features. Such a machine learning framework enables to analyze a mix of semi structured and unstructured data in search of valuable information and insights for precipitation forecasting with a credible accuracy. Authors carried out the precipitation prediction in several word situations by using 2010, 2008, 2007 as testing data sets, and 2005~2009, 2003~2007, and 2002~2006 as training data sets, respectively. In these datasets, first they used the NSC to balance the training data, then the Fast-OSFS and the DLA to screen out the most related and proper features; finally, they opted the KNN for performing the final prediction. The performance measure in terms of recalls is 78.95%, 70.77%, and 87.27% for these three years [2].

2.3 Precipitation Estimation from Remotely Sensed Information Using Neural Networks (PERSIANN) [7]

A computerized system for precipitation evaluation from remotely sensed information using artificial neural networks, the PERSIANN has been proposed. The model mainly estimates the rainfall from geosynchronous satellite long wave infrared imagery. The basic algorithm of the PERSIANN system depends upon the neural network and when it becomes available it can be certainly modified to integrate the appropriate information [7].

2.4 Nearest Neighbour Model [19][24]

The model presents only the most similar situations corresponding to the current situation and the decision making itself depends on the forecaster. Further, the selection of weights for different parameters is based on the intuition of the expert, although attempts have been made for optimizing weights also. The basic concept of the nearest-neighbour model lies in the fact that similar situations will lead to similar outcomes. Thus the nearest-neighbour technique looks into the history of the events in the past data. The similarity of the present situation with the past ones is defined in terms of the similarity metric. For the development of the present model Euclidean metric has been taken as the similarity measure.

Based on this approach, a quantitative snowfall forecast model has been developed for a station in Jammu and Kashmir, using surface meteorological data of the past 12 winters (1991–92 to 2003–04, excluding the data of winter 1994–95, which was not available). The model predicts weather in terms of snow/no snow day and the amount of snowfall (snow height in cm) for three consecutive days in advance. The performance of the model has been tested for four winters for day-1, day-2 and day-3 forecasts. For qualitative snowfall forecast, the model performance for day-1, day-2 and day-3 forecasts turns out to be 80–90, 70–80 and 65–75%. The model estimates the expected snowfall amount at the station for day-1, day-2 and day-3 in advance, and based on the value of the estimated snowfall amount, it is categorized in the expected snowfall range based on the already established criterion. Quantitatively, the model predicts snowfall amount accurately for day-1 and the average accuracy of the model for different ranges of established categories varies from 25 to 55% for day-1 forecast. The model over-predicts the expected snowfall amount for day-2 and day-3 compared to day-1.

2.5 Real-Time Forecasting of Snowfall Using a Neural Network [5]

A neural network approach to snowfall forecasting, using only forecast information available on a real-time basis, shows substantial gains relative to the lookup and climatologically snow-ratio methods. In particular, in addition to reducing errors in the snow ratio, the network approach compensates for errors in QPFs, helping to constrain the overall error in forecast snow amount. Although the procedure employed in this study has proven effective based on a small sample collected over the course of two cold seasons (November–March), several elements could be further refined. First, the assignment of a snow-ratio number to the predicted class is somewhat ad hoc; it may be possible to make improvements by relating the specific within-class ratio to the forecast class probability, as was informally done for the light class (e.g., a high probability of a heavy snow ratio might indicate a lower snow ratio than a lower probability within that same class). Second, refinements based upon the relationship between the forecast vertical motion and the details of the sounding profile might allow for further refinement, particularly in the light class when maximum vertical motions may align with temperature conditions most conducive for dendrite growth [5] [20].

2.6 Bayesian networks [8]

Bayesian networks (BNs) belong to the family of probabilistic graphical models (GMs). These graphical structures are used to represent knowledge about an uncertain domain. In particular, each node in the graph represents a random variable, while the edges between the nodes represent probabilistic dependencies among the corresponding random variables. These conditional dependencies in the graph are often estimated by using known statistical and computational methods. They have been used for applications in various areas, such as machine learning, text mining, natural language processing, speech recognition, signal processing, bioinformatics, error-control codes, medical diagnosis, weather forecasting, and cellular networks [8][19][24].

2.7 Sequential Minimal Optimization (SMO) [9-19]

SMO implements John C. Platt's sequential minimal optimization algorithm for training a support vector classifier using polynomial or RBF kernels. This implementation globally replaces all missing values and transforms nominal attributes into binary ones. It also normalizes all attributes by default [19][24].

2.8 Naïve Bayes Classifier [9-19]

Naïve Bayes is a statistical learning algorithm that applies a simplified version of Bayes rule in order to compute the posterior probability of a category given the input attribute values of an example situation. Prior probabilities for categories and attribute values conditioned on categories are estimated from frequency counts computed from the training data. Naïve Bayes is a simple and fast learning algorithm that often outperforms more sophisticated methods. The Bayesian classification represents a supervised learning method as well as a statistical method for classification. Assumes an underlying probabilistic model and it allows us to capture uncertainty about the model in a principled way by determining probabilities of the outcomes [19][24].

2.9 Decision Stump [9-19]

It is a one level decision tree. It has one internal node which is connected directly to the terminating nodes. A single root node decides how to classify inputs based on a single feature. Each leaf represents possible feature value, the class label that should be assigned to inputs whose features have that value. For using this method, one must decide the feature and build the tree. It is the simplest method by which decision stumps can be build for each possible feature and which feature is giving highest accuracy on the training data can be checked [19][24].

2.10 J48 [9-19]

It is a new version of an earlier very popular algorithm C4.5 Decision trees developed by J. Ross Quinlan. It is providing variety of options which generates an unpruned or pruned C4.5 decision tree. A predictive machine-learning model which decides the target value of a new sample based on different attribute values of the available data is J48 decision tree. The different attributes denote by the internal nodes of a decision tree, the branches between the nodes tell us the possible values that these attributes can have in the experimental samples, while the terminal nodes tell us the final value of the dependent variable [19][24].

2.11 LMT [9-19]

A classification model with an associated supervised training algorithm that combines logistic prediction and decision tree learning is logistic model tree (LMT). Logistic model trees use a decision tree that has linear regression models at its leaves to provide a section wise linear regression model. The algorithm can deal with binary and multi-class target variables, numeric and nominal attributes and missing values. It does not require any tuning parameters. It is often more accurate than C4.4 decision trees and standalone logistic regression [19][24].

2.12 Random Forest [9-19]

Random forests are an ensemble learning method for classification, regression and other tasks, which operate by constructing a multitude of decision trees at training time and outputting the class that is the mode of the classification or mean prediction of the individual trees. Random forests correct for decision trees' habit of over fitting to their training set. Random forests are a way of averaging multiple deep decision trees, trained on different parts of the same training set, with the goal of reducing the variance. This comes at the expense of a small increase and some loss of interpretability, but generally greatly boosts the presentation of the final model [19][24].

2.13 REP Tree [9-19]

It is a fast decision tree learner, builds a decision tree using attribute selection measure as information gain and prunes it using reduced-error pruning. Only sorts values for numeric attributes once. Missing values are dealt with by splitting the corresponding instances into pieces [19][24].

2.14 Other Approaches

Irene Y.H Gu et al., put forward a full automatic image analysis system for detection and analysis of snow/ice coverage on electric insulators of power lines using images which were captured by visual cameras in a remote outdoor laboratory test bed [25]. Jinmei Pan et al., put forth a passive microwave remote sensing techniques that detected wet snow in the south of china [26]. Yajaira Mejia et al., gave an approach for estimating the snowfall using neural networks on multi source remote sensing observations and ground based meteorological measurements [27]. Melanie Wetzel et al., projected a technique that supports the snowfall forecast and for the verification of radar limited mountainous terrain that includes matching the output parameters and graphics from high resolution mesoscale models to surface mesonets [28].

Michael A. Rotondi illustrated a Markov chain models across eight national weather stations using historical data from the global historical climatology network to predict a ‘snow day’ [30]. Gail M. Skofronick Jackson et al., in their research interpreted how instruments like the W-band radar of Cloudsat, Global Precipitation Measurement Dual-Frequency Precipitation Radar ku- and Ka-bands, and the Microwave Imager can be used in the simulations of lake effect and synoptic snow events in order to determine the minimum amount of snow [31]. Gail M. Skofronick Jackson et al., demonstrated thresholds for detecting falling snow from satellite-borne active and passive sensors [29] [32].

3 SUMMARIZATION OF PREVIOUS APPROACHES

Table 1 illustrates the accuracy of various previously proposed algorithms. Table 2 illustrates the accuracy of non-decision tree algorithms. Table 3 illustrates the accuracy of decision trees. WEKA tool has been used for the analysis of snow datasets [21]. The main objective of this comparison is not to disparage which is the best technique, but to prove its usage and to create alertness in their fields.

TABLE 1
COMPARISON OF PREVIOUS PROPOSED METHODS FOR SNOW NOWCASTING

S No.	Proposed By	Model	Accuracy
1	Yahui Di et al. [2]	NSC+KNN	52.80
2	Yahui Di et al. [2]	NSC+Fast-OSFS+KNN	55.14
3	Yahui Di et al. [2]	NSC+Fast-OSFS+DLA+KNN	60.28
4	Yahui Di et al. [2]	KNN	77.53
5	Minoru Numata et al. [4]	Intelligent Visibility Meter	77.80
6	Josh Barnwell [6]	Cobb Method	79.90
7	Geetha A [3]	Rapid Miner	81.78
8	Dan Singh [33]	Nearest Neighbour Model	86.74

TABLE 2
COMPARISON OF NON-DECISION TREE APPROACHES FOR SNOW NOWCASTING

S No.	Model	Accuracy
1	Bayesian Networks	88.00
2	Naïve Bayes	86.72
3	SMO	88.66
4	Simple Logistic	88.58

TABLE 3
COMPARISON OF DECISION TREE APPROACHES FOR SNOW NOWCASTING

S No.	Model	Accuracy
1	Decision Stump	85.56
2	J48	88.42
3	LMT	88.21
4	Random Forest	88.54
5	REP Tree	88.19

4 CONCLUSION

From this survey a detailed report can be obtained for nowcasting snow by using quite a lot of techniques over fifteen years. This paper provides an idea that maximum researchers and scientists use the above techniques for nowcasting snowfall and also they attain substantial results. Hence the upshot of this paper is to show the available techniques and therefore comparison table exposes the accuracy and future scope of an individual methods used for nowcasting snowfall. This survey may inspire more researchers to use the following algorithms to solve many research problems put-forth by the available huge amounts of data for knowledge discovery. Moreover, in future this paper will lead a moral support for the researches to Nowcast snow accurately and efficiently.

5 ACKNOWLEDGEMENTS

The authors would like to thank the referees for their valuable comments and suggestions, which greatly enhanced the clarity of this paper. We are indebted to Dr. Tirupathi Rao, Dr V Chandra Prakash and Dr G Pradeepni for the valuable suggestions. We are also grateful to the Management and Department of Computer Science and Engineering of Stanley College of Engineering & Technology for Women, Hyderabad, K L University, Guntur for providing their maximum support during the paper write up.

References:

[1] Robert A. Houze: Cloud Dynamics. Academic Press (1994).

[2] Yahui Di, Wei Ding and Yang Mu, "Developing Machine Learning Tools for Long-Lead Heavy Precipitation Prediction with Multi-sensor Data", IEEE International Conference on Networking, Sensing and Control, 2015.

[3] A Geetha and G M Nasira, "Artificial Neural Networks' Application in Weather Forecasting – Using RapidMiner", International Journal of Computational Intelligence and Informatics, Vol. 4: No. 3, 2014.

[4] Minoru Numata Jun Okamoto Toru Tashiro and Takeshi Ito, " A Basic Study of Short Term Forecasting Methods of Snowfall and Road Surface Temperature Using Intelligent Visibility Meter", SIRWEC, 2015.

[5] Paul j. Roebber, melissa r. Butt, and sarah j. Reinke, " Real-Time Forecasting of Snowfall Using a Neural Network" AMS Weather and Forecasting, 2007

[6] Josh Barnwell, "Verification of the Cobb Snowfall Forecasting Algorithm", Dissertations & Theses in Earth and Atmospheric Sciences, University of Nebraska , 2011.

[7] S. Sarumathi, N. Shanthi and S. Vidhya, "Satellite Rainfall Prediction Techniques - A State of the Art Review", International Journal of Environmental, Chemical,

Ecological, Geological and Geophysical Engineering Vol:9, No:2, 2015

[8] Ben-Gal I., Bayesian Networks, in Ruggeri F., Faltin F. and Kenett R., “Encyclopedia of Statistics in Quality & Reliability”, Wiley & Sons, 2007.

[9] Shilpa Sonawani1 and Debajyoti Mukhopadhyay, “A Decision Tree Approach to Classify Web Services using Quality Parameters”, ICWA, 2013

[10] Lior Rokach and Oded Maimon, “Top-Down Induction of Decision Trees Classifiers—A Survey”, IEEE Transactions on Systems, Man, and Cybernetics: Applications And Reviews, 2005

[11] Layla Safwat Jamil, “Data Analysis based on Data Mining Algorithms using WEKAWorkbench”, International Journal of Engineering Sciences & Research Technology, 2016

[12] Purva Sewaiwar and Kamal Kant Verma, “Comparative Study of Various Decision Tree Classification Algorithm Using WEKA”, International Journal of Emerging Research in Management &Technology, 2015

[13] Kittipol Wisaeng, “A Comparison of Decision Tree Algorithms for UCI Repository Classification”, International Journal of Engineering Trends and Technology, 2013.

[14] Kishor Kumar Reddy C, Vijaya Babu B, Rupa C H: SLEAS: Supervised Learning using Entropy as Attribute Selection Measure. International Journal of Engineering and Technology (2014) 2053-2060.

[15] Kishor Kumar Reddy C, Rupa C H and Vijaya Babu B: A Pragmatic Methodology to Predict the Presence of Snow/No-Snow using Supervised Learning Methodologies. International Journal of Applied Engineering Research (2014) 11381-11394.

[16] Kishor Kumar Reddy C, Rupa C H and Vijaya Babu: SPM: A Fast and Scalable Model for Predicting Snow/No-Snow. World Applied Sciences Journal (2014) 1561-1570.

[17] Kishor Kumar Reddy C, Rupa C H and Vijaya Babu: SLGAS: Supervised Learning using Gain Ratio as Attribute Selection Measure to Nowcast Snow/No-Snow. International Review on Computers and Software (2015).

[18] Kishor Kumar Reddy C, Vijaya Babu and Rupa C H: ISIQ: Improved Supervised Learning using in Quest to Nowcast Snow/No-Snow. WSEAS Transactions on Computers (2015).

[19] Arun k pujari: Data Mining Techniques. Universities Press (2004).

[20] http://www.wunderground.com/.

[21] Nabilah Filzah Mohd Radzuan, Andi Putra, Zalinda Othman, Azuraliza Abu Bakar and Abdul Razak Hamdan: Comparative Study – Three Artificial Intelligence Techniques for Rain Domain in Precipitation Forecast. International Journal of Environmental, Ecological, Geological and Mining Engineering (2013) 898-903.

[22] http://en.wikipedia.org/wiki/Weather_forecasting.

[23] http://en.wikipedia.org/wiki/Quantitative_precipitation_forecast.

[24] J. Han: Data Mining: Concepts and Techniques. Morgan Kaufmann Publishers. (2001)

[25] Irene Y.H. Gu, Unai Sistiag and Anders Fahlstrom: Online Detection of Snow Coverage and Swing Angles of Electrical Insulators on Power Transmission Lines. IEEE International Conference on Image Processing (2009) 3249 – 3252.

[26] Jinmei Pan, Lingmei Jiang and Lixin Zhang: Wet Snow Detection in the South of China by Passive Microwave Remote Sensing. IEEE International Geo Science and Remote Sensing Symposium (2012) 4863-4866.

[27] Yajaira Mejia, Hosni Ghedira and Reza Khanbilvardi: A Neural Network Based Approach for Multi-Spectral Snowfall Detection and Estimation, IEEE International Geo Science and Remote Sensing Symposium (2007) 2276 – 2279.

[28] Melanie Wetzel, Michael Meyers, Randolph Borys, Ray Mcanelly, William Cotton, Andrew Rossi, Paul Frisbie, David Nadler, Douglas Lowenthal, Stephen Cohn, and William Brown: Mesoscale Snowfall Prediction And Verification In Mountainous Terrain. AMS Journal Of Weather And Forecasting (2004) 806-828.

[29] Pascal Sirguey, Renaud Mathieu Yves Arnaud, Muhammad M. Khan and Jocelyn

Chanussot: Improved Resolution For The Detection Of Snow With Modis Using Wavelet Fusion. IEEE International Geo Science and Remote Sensing Symposium (2007) 3975-3978.

[30] Michael A. Rotondi: Estimating Transition Matrices to Predict Tomorrow's Snowfall Using Real Data. Journal of Statistics Education (2010) 1-14.

[31] Gail M. Sko Fronick Jackson, Benjamin T. Johnson, and S. Joseph Munchak: Detection Thresholds of Falling Snow From Satellite-Borne Active And Passive Sensors, IEEE Transactions On Geo Science and Remote Sensing (2013) 4177-4189.

[32] Gail M. Skofronick Jackson and Benjamin T. Johnson: Thresholds Of Detection For Falling Snow From Satellite-Borne Active And Passive Sensors, IEEE International Geo Science And Remote Sensing Symposium (2007) 2637-2640.

[33] Dan Singh*, Ashwagosha Ganju and Amreek Singh, "Weather prediction using nearest-neighbour model", Research and evelopment Centre, Snow and Avalanche Study Establishment, Chandigarh India