CUMULONIMBUS PREDICTION USING ARTIFICIAL NEURAL NETWORK BACK PROPAGATION WITH RADIOSONDE INDECES

Agie Wandala Putra*), Chidchanok Lursinsap**)
*) Indonesian Meteorological Climatological and Geophysics Agency (BMKG), Jakarta, Indonesia
**) Department of Mathematics and Computer Science, Faculty of Science
Chulalongkorn University, Bangkok, Thailand
e-mail: agie.wandala@bmkg.go.id

Abstract

This research purpose an accurate quantitative forecasting of Cumulonimbus (*Cb*) Cloud development events. Neural network will use to develop a quantitative cumulonimbus events forecasting model that highly improve forecasting skill especially in Tropical Area. Cumulonimbus storm cells can produce torrential rain of a convective nature and flash flooding, as well as straight-line winds. Most storm cells die after about 20 minutes, when the precipitation causes more downdraft than updraft, causing the energy to dissipate. If there is enough solar energy in the atmosphere as in tropical are, the moisture from one storm cell can evaporate rapidly—resulting in a new cell forming just a few miles from the former one. This can cause thunderstorms to last for several hours. Cumulonimbus clouds can also bring dangerous storms which bring lightning, thunder, and torrential ice. The technique is Back propagation neural network (BPN), the advantages using this engine are more reasonably to estimated a large class of functions and efficient than numerical differentiation. The BPN inputs were not only raw sounding observation data also derived indices value, using Principle component analysis (PCA) as feature selection processing. PCA has been widely known as method to reduce the dimension of input to multivariate data by minimizing loss of information input. In this research PCA is used to reduce the dimension of input to the BPN and reconstruct the new input data. Clustering and initialization process centers on neural network done with the Self Organizing Map (SOM) technique and the determination of the weight of the hidden center during the learning process using the algorithm Recursive Orthogonal Least Square. The initial result show that the combination of dataset engine are workable, the proposed technique result in better accuracy of prediction and can implemented in operational used for early warning system to reduce destruction impact from weather hazard in tropical area.
Key Words: *MTSAT, Artificial Neural Network, PCA, Radiosonde, Upper Air*

Abstrak

Penelitian ini bertujuan untuk peramalan kuantitatif yang akurat peristiwa terbentuknya awan *Cumulonimbus* (*Cb*). Model neural network digunakan untuk mengembangkan model peramalan kuantitatif peristiwa terbentuknya cumulonimbus yang sangat meningkatkan kemampuan peramalan khususnya di area tropis. Sel badai cumulonimbus dapat menghasilkan hujan deras yang bersifat konvektif dan banjir bandang, serta angin garis lurus. Kebanyakan sel badai mati setelah sekitar 20 menit, ketika curah hujan menyebabkan lebih downdraft daripada updraft, menyebabkan energi menghilang. Jika ada energi surya yang cukup di atmosfer seperti di berada tropis, kelembaban dari satu sel badai bisa menguap dengan cepat - menghasilkan sel baru membentuk hanya beberapa mil dari sel sebelumnya. Hal ini dapat menyebabkan badai berlangsung selama beberapa jam. Awan cumulonimbus juga dapat membawa badai berbahaya yang membawa petir, guntur, dan hujan es lebat. Teknik yang digunakan adalah *Back Propagation Neural Network* (BPN), keuntungan menggunakan mesin ini adalah lebih masuk akal untuk memperkirakan kelas besar fungsi dan efisien daripada diferensiasi numerik. BPN masukan tidak hanya data observasi mentah juga berasal dari nilai indeks, menggunakan *Principle Component Analysis* (PCA) sebagai pengolahan seleksi fitur. PCA telah dikenal luas sebagai metode untuk mereduksi dimensi input data multivariat dengan meminimalkan hilangnya masukan informasi. Dalam penelitian ini PCA digunakan untuk mengurangi dimensi dari masukan kepada BPN dan merekonstruksi input data baru. *Clustering* dan inisialisasi proses berpusat pada jaringan saraf (*neural network*) dilakukan dengan *Self Organizing Map* (SOM) teknik dan penentuan berat pusat tersembunyi selama proses pembelajaran dengan menggunakan algoritma *Rekursif Orthogonal Least Square*. Hasil awal menunjukkan bahwa kombinasi dari mesin dataset dapat diterapkan, hasil teknik yang diusulkan dalam akurasi yang lebih baik dari prediksi dan dapat diimplementasikan dalam operasional digunakan untuk sistem peringatan dini untuk mengurangi dampak kerusakan dari bahaya cuaca di daerah tropis .
Kata Kunci: *MTSAT, Artificial Neural Network, PCA, Radiosonde, Upper Air*

1. Introduction

The artificial neural networks have been studied since the nineteen sixties (Rosenblatt, 1958), but their use for forecasting meteorological events appeared only in the last 20 years. Some “early” references regarding different meteo events could be: Lee et al. (1990) for cloud classification, Frankel et al. (1990) for lightning, Schizas et al. (1991) for minimum temperature, McCann (1992) for thunderstorms, French et al. (1992) or Allen and Le Marshall (1994) for rain, Pasini and Potestá (1995) for visibility and fog, Xiao and Chandrasekar (1996) for snow, Marzban and Stumpf (1996) for tornado, Pankiewicz (1997) for satellite convective cells, Marzban and Witt (2001), for hail. Manzato for Thunderstorm (2005) and Hail (2009). Weng (2010) for Ligtning. A General references to Feature Selection are Everitt and Dunn (2001), Krzanowski (2000), Krzanowski and Marriott (1994) and Rencher (1995, 1998),Joliffe(2006), and Preisendorfer and Mobley (1988), Monahan (1999) specific for meteorlogy and oceanograpphy, ANN are Masters (1993), Bishop (1996), Haykin (1998) and Marzban (2002).

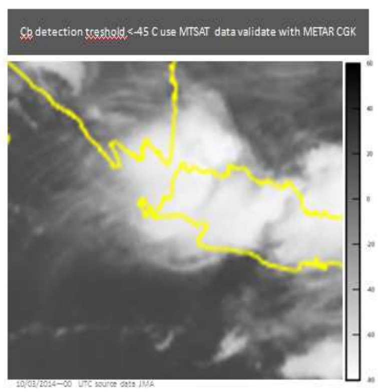

Figure 1-1. Cumulonimbus detection using MTSAT

The radiosonde is an expendable, balloon-borne device that measures the vertical profile of meteorological meteorological variables and transmits the data to a groundbased receiving and processing station. These profiles are typically obtained twice each day and are the core of the global weather observing system that provides inputs to numerical forecast models. The sensor package routinely measures the variation with altitude of temperature, humidity, and pressure as the balloon ascends from the land or ocean surface to heights up to about 30km (a pressure altitude of about 11 hectopascals,hPa). When the device also measures winds, it is more properly called a rawinsonde, although the term radiosonde is commonly applied to both. The height profile of these meteorological variables constitutes an upper-air sounding that is known as a radiosonde observation or RAOB. In some cases, a balloon without a radiosonde is tracked by either optical or radar techniques in order to measure only winds. This type of balloon is known as a pilot balloon or simply a pinball, but it is not a radiosonde.

In tropical area the atmosphere have different system of circulation, many parameter can built up the severe weather. And every severe weather accident always recorded the Cumulonimbus Cloud event, what weather do cumulonimbus cloud bring; we can find thunderstorms, gusty, small tornado and maybe a cyclone when over water and warm air. In tropical country Cumulonimbus as impact of convectifiy process, can happens in any time during the day specialy in Rainy season but can happen at any time of the year, but more common in rainy season.

Cumulonimbus clouds are thunderstorm clouds that form if cumulus congestus clouds continue to grow vertically. Their dark bases may be no more than 300 m (1000 ft) above the Earth's surface. Their tops may extend upward to over 12,000 m (39,000 ft). Tremendous amounts of energy are released by the condensation of water vapor within a cumulonimbus. Lightning, thunder, and even violent tornadoes are associated with the cumulonimbus.

The most important hazard impact from cumulonimbus events beside in the industry and other humanity impact, that is highly dangerous for flight. The only sensible defence against the hazards associated with a Cb is therefore to avoid flying into one in the first place. Predicting an individual Cb cell is difficult but it is possible to predict the conditions which will trigger formation of a Cb. Forecasters are therefore able to advise flight crews and controllers of the likely timing, location, direction of movement, and height of cells and whether or not they may be embedded. Airport authorities can plan aircraft movements to take into account the disruption to operations caused by storms, and approach controllers can consider how they will manage en-route, departing, and arriving traffic when storms are in the vicinity. Flight crews can alter their routings to avoid forecast Cb activity or decide to carry extra contingency fuel in case they have to re-route in flight to avoid the storms or burn additional fuel because of the potential use of aircraft de/anti icing systems. Awareness of the conditions which lead to the formation of a Cb, recognition of a developing and mature Cb, and awareness of the signs which indicate the proximity of a Cb will help controllers and flight crews to plan operations to avoid the associated hazards.

In addition to visual recognition, Weather Radar is a particularly valuable aid to avoiding Cb clouds. Airborne weather radar enables the flight crew to identify the areas of the storm cloud which hold the largest water droplets, which indicate the areas with strongest updrafts. The area of the cloud with the most severe turbulence is where the updrafts adjoin the downdrafts; therefore the pilot must avoid flying through the edge of the areas of cloud with the largest water droplets. It should be remembered that a large cloud will absorb a great deal of the radar pulse which may therefore not penetrate all of the way through the storm. This can give a false impression that there are no Cb cells beyond the cell immediately ahead of the aircraft. So as supporting to forecaster adjust the condition of atsmosphere at that time, they must understand the instablity condition from the atmoshpere represent from instablity indeces of atmosphere, and we can get the picture from radiosonde observation analysis.

From reasons above, this paper will concentrate to study about how to identify the Cumulonimbus event using the artificial neural network model to get more accurate result. Use radiosonde observation data which very expensive operation, also present the real dynamics condition in the atmosphere that we know from the simulation and the indeces. With build the new model system, we hope can get better accuracy on Cumulonimbus prediction and study the impact. Future a new reserch can intergrate with all information to the forecaster, public and other numerical model system that has been developed.

2. Data

To predict the occurrence of Cumulonimbus cloud (Cb) event, the radiosonde is chosen as the input data. For classification of predictan, the cumulonimbus it self. Multispectral trasnport satellit (MTSAT) used to provide the area of Cb recorded.

A. Atmosphere Indeces

First of all, develop an application to identify the data of radiosonde indeces with thermodynamic and wind-derived part. Form all algorithm then convert to the numerical calculation program to compute the indeces value and every variable represent different meaning for the atmoshperic condition can be explained. The values obtained for some indices are very different, and we need an objective algorithm to evaluate which method is more related to the observed convective activity.

The program will run using 320 sounding observation during November 2013-April 2014. The observation made by the Indonesian Meteorological Climatological Station (BMKG) (WMO code 96749), the operation launched every 12 hours exclude the operational and analytics problem not used. All the indices computed by the program and saved in the database are here collected in two groups. For each index, in the larger data we will study the climatology , finding a threshold value to compute the skill score for forecasting thunderstorm presence in the associated 12 h.

Table 2-1. List Atmospheric Instability Index

No	**Index**	**No**	**Index**
1	Showalter index	27	850 mb Wind Speed (kt)
2	Lifted index	28	850 mb Dewpoint (C)
3	LIFT computed using virtual temperature	29	700 - 500 mb lapse rate (C/km)
4	SWEAT index	30	Boyden Index
5	K index	31	BRN Shear (m^2/s^2)
6	Cross totals index	32	CAP Strength
7	Vertical totals index	33	CT - Cross Totals
8	Totals totals index	34	DCAPE 0-6 km, MSL
9	Convective Available Potential Energy	35	Delta Theta-e (ePT)
10	CAPE using virtual temperature	36	EHI - Energy Helicity Index
11	Convective Inhibition	37	GOES HMI (Hybrid Microburst Index)
12	CINS using virtual temperature	38	Hail (cm)
13	Equilibrum Level	39	JI - Jefferson Index
14	Equilibrum Level using virtual temperature	40	KO Index
15	Level of Free Convection	41	LFC-LCL height (m)

16	LFCT using virtual temperature	42	LFC - Level of Free Convection (mb)
17	Bulk Richardson Number	43	MDPI - Microburst Day Potential Index
18	Bulk Richardson Number using CAPV	44	NCAPE (Normalized CAPE)
19	Temp [K] of the Lifted Condensation Level	45	S Index
20	Pres [hPa] of the Lifted Condensation Level	46	srH - storm-relative Helicity (0-3 km)
21	Mean mixed layer potential temperature	47	Supercell Composite Parameter (SCP)
22	Mean mixed layer mixing ratio	48	Surface Dewpoint (C)
23	1000 hPa to 500 hPa thickness	49	T1/T2 Gust (kt)
24	Precipitable water [mm] for entire sounding	50	TI - Thompson Index
25	200 mb Wind Speed (kt)	51	TQ Index
26	500 mb Wind Speed (kt)	52	VGP - Vorticity Generation Parameter

B. Cumulonimbus Classification

The studied periode was the month from November 2012 – April 2013 in the Station Meteorological Station (WMO code 9679), managed by Indonesian Meteorological Climatological and Geophisics Agency (BMIKG). Launched twice a day for sounding (0000 and 1200). The total number of sounding collected N= 342 but just 320 that can be used, after filtering with the error and unusable paramater because some problem like balon burst, GPS failed, Cloud > 7 oktas, etc.

The 12-h period was associated with the sounding derrived indices as predictor variable. The predicted variables are built using the weather observation (METAR and Synoptic report) and Cloud Type identification using Remote Sensing from Multi-Functional Transport Sattelites (MTSAT).

Manual observation in Cengkareng Meteorological Station was operated 24 hours per day, with some equipment and legal report exchanges to all Meteorological station in the world under mantained by World Meteorological Organitation (WMO). A METAR weather report is predominantly used by pilots in fulfillment of a part of a pre-flight weather briefing, and by meteorologists, who use aggregated METAR information to assist in weather forecasting. Raw METAR is the most popular format in the world for the transmission of observational weather data. It is highly standardized through the International Civil Aviation Organization (ICAO), which allows it to be understood throughout most of the world.

The second application program was developed by using python colaborate with gsmap and Sataid program. To identify the Cumulonimbus event using data observation from canal MTSAT, the algorithm was found by Tokuno, Masami (1993). The application develop to identify cumulonimbus development in tropical area which is with the high convectivty, and calculate to database the base area around target, expect the research can expand the area of target from 5km.

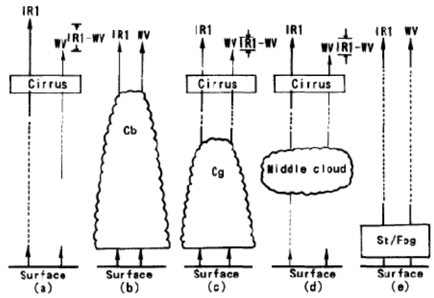

Figure2-1. Scematic representation of cloud types

A scematic representation of cloud types to be classified is shown below, since the top of well developed Cb reaches tropopause, the amount of water vapor in the air column between the cloud top an a sattelite is neglibilibly small, thus it is considered that the radiance observed with a thermal window cannel (11µm band, denoted as IR1) and water vapor channel (6.7 µm band, denoted as WV) is nearly equal. To confirm theorically. Equivalent blackbody temperature to be observed from satellite in each channel is computed for the model atmosphere of mid latitute summer, computed by placing a black body cloyd staring from 7 km at 500 meter increment up to 13 km (tropopause). At the level 1 km below the tropopause the difference becomes 1.5K thus 1.5K is adopted as threshold for Cb. A high level cloud often exists above middle or low level clouds, from radiative transfer point of view this high cloud is semi transparent. Under an assumtion tat the thick cloud below is blackbody. The radiance of thermal window and water vapor channels to be observed from satellite level is expressed as follows.

$$R(WV) = \big(1 - e(WV)\big) * R(1, WV) + e(WV) * R(c, WV)$$

$$R(IR) = \big(1 - e(IR)\big) * R(1.IR) + e(WV) * R(c.WV)$$

$$R(WV) = A * R(IR) + B$$

$$A = \big(R(1.WV) - R(c.WV)/(R(1.IR) - R(c.WV)\big)$$

$$B = R(1.WV) - A * R(1.IR)$$

If cirrus is constant at a certain level then WV and IR be constant. The height of blackbody cloud under cirrus then is estimated from the radiance reaching the base of cirrus.

3. Feature Selection

A. Data normalizaiton

Nonlinier activation function such as the logistic function typically have the squashing role in restricting or squashing the possible ouput from a node to typically, (0,1) or (-1,1). Data normalization is often performed before the training process begins. When nonliniar transfer functions are used at the output nodes, the desired output values must be transformed to the range of the actual output of the network. Even if a linear output transfer function is used, it may still be advantageous to stdardize the outputs as well as the input to avoid computational problems, to meet alqorithm requirement and to facilitate network learning.

The method for input normalization from sounding indeces data, a channel is defined as a set of elements in the same position over all input vectors in the training or test set. That is, each channel can be thought og as an independent input variable. The along channel normalization is performed column by column if the input vectors are put into a matrix. In other words, it normalizes each input variable individually.

The following is the normalization used in this work:

$$x_n = (x_0 - x_{min})/(x_{max} - x_{min})$$

It is unclear weather ther is a need to normalize the inputs because the arc weights could undo the scaling. Form some researh they investigated the effectiveness of liniear and statistical normalization method for classification problems. The data normalization is beneficial in terms of the classification rate and the mean squared error, but the benefit diminishes as network and sample size increase.

B. Principle Component Analysis

Dimensionality reduction of a feature set is a common preprocessing step used for pattern recognition and classification applications and in compression schemes. Principal component analysis (PCA) is one of the popular methods used, and can be shown to be optimal using different optimality criteria. However, it has the disadvantage that measurements from all of the original features are used in the projection to the lower dimensional space. This paper proposes a novel method for dimensionality reduction of a feature set by choosing a subset of the original features that contains most of the essential information, us ing the same criteria as the PCA. We call this method Principal Feature Analysis (PFA). The proposed method is successfully applied for choosing the principal fea tures in selection the best Radiosonde Instability indeces that can work at this area.

Given a random vector $\overline{x}$ of dimension N and its correlation matrix $\overline{\overline{R}}$ we can reduce its dimension to M (with M<N) by Principal Components Analysis in six steps:

1. Find the eigenvectors $\overline{\overline{Q}}$ and eigenvalues λ_i of correlation matrix $\overline{\overline{R}}$:

$$\overline{\overline{R}}\overline{q}_i = \lambda_i \overline{q}_i$$

2. Arrange the eigenvalues in decreasing order:

$$\lambda_1 > \lambda_2 > .. > \lambda_M > .. > \lambda_N$$

3. Pick up the eigenvectors which belong to the first M largest eigenvalues.
4. Calculate compressed vector $\overline{c}$ by $c_i = \overline{x}^T \overline{q}_i$ for i = 1, .., M
5. Use vector $\overline{c}$ for storage, transmission, process, etc.
6. Decode the resulting vector $\overline{c}'$ into N-dimensional vector $\tilde{\overline{x}}'$ using the eigenvector matrix $\overline{\overline{Q}}$.

$$\tilde{\bar{x}}' = \sum_{i=1}^{M} c_i \bar{q}_i$$

To use Principal Components Analysis we need to have a correlation matrix, which defines the similarity between different input vectors. To obtain a correlation matrix $\bar{\bar{R}}$, we construct one by means of observations of different input vectors. We examine for example K different images for constructing matrix $\bar{\bar{R}}$ for a PCA of images. We note $\bar{x}^{(k)}$ as being the k-th observed image.

We use the following empirical approximation of $\bar{\bar{R}}$:

$$\tilde{R}_{ij} = \frac{1}{K}\sum_{k=1}^{K} x_i^{(k)} x_j^{(k)}$$

The more observations are made, the better the approximation $\tilde{\bar{\bar{R}}}$ of $\bar{\bar{R}}$ gets. Instead of matrix $\bar{\bar{R}}$ we use matrix $\tilde{\bar{\bar{R}}}$ in the PCA calculations.

To determine the eigenvectors of correlation matrix $\bar{\bar{R}}$, we have to construct the matrix $\bar{\bar{R}}$ by calculating the outer product of vector $\bar{x}$. In most applications of PCA, this vector $\bar{x}$ is very large, as it represents the data which is to be compressed. The complexity of the calculations are high, namely $O(N^3)$. There is a way in which we can reduce this complexity to $O(K^3)$, where K is much smaller than N, when we use the limited number of observations of vectors $\bar{x}$ to construct the needed eigenvectors for PCA.

Since $\tilde{\bar{\bar{R}}}$ is constructed from different vectors $\bar{x}^{(k)}$, the eigenvalues of $\tilde{\bar{\bar{R}}}$ are in the space which is spanned by $\bar{x}^{(k)}$:

$$\tilde{\bar{q}}_i \in span\{\bar{x}^{(k)}, k = 1,.., K\}$$

Therefore:

$$\tilde{\bar{q}}_i = \sum_{k=1}^{K} \alpha_i^{(k)} \bar{x}^{(k)}$$

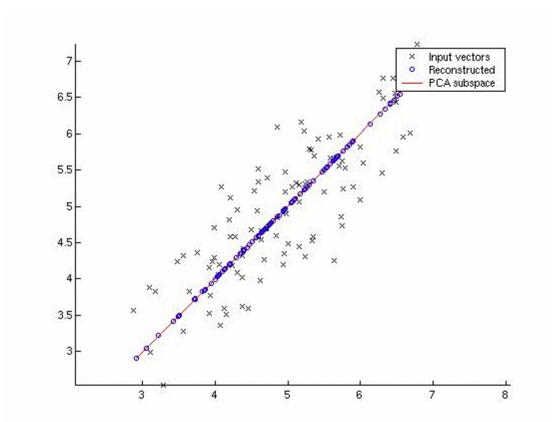

Figure 2-1. Principal Component Analysis used to find the best input variable with the minimal reconstruction error.

We here provide the reduction in complexity. This the so called PCA Kernel method, because it uses the Kernel matrix U, which elements are given by:

$$U_{ij} = \overline{x}^{(i)T}\overline{x}^{(j)}$$

$$\begin{aligned}
&\tilde{\overline{\overline{R}}}\tilde{\overline{q}}_i = \tilde{\lambda}_i\tilde{\overline{q}}_i \rightarrow \\
&\tilde{\overline{\overline{R}}}\sum_{k=1}^{K}\alpha_i^{(k)}\overline{x}^{(k)} = \sum_{k=1}^{K}\tilde{\lambda}_i\alpha_i^{(k)}\overline{x}^{(k)} \rightarrow \\
&\frac{1}{K}\sum_{j=1}^{K}\overline{x}_j^{(k)}\overline{x}_j^{(k)T}\sum_{k=1}^{K}\alpha_i^{(k)}\overline{x}^{(k)} = \sum_{k=1}^{K}\tilde{\lambda}_i\alpha_i^{(k)}\overline{x}^{(k)} \rightarrow \\
&\sum_{k=1}^{K}\alpha_i^{(k)}\sum_{j=1}^{K}\overline{x}_j^{(k)}\overline{x}_j^{(k)T}\overline{x}^{(k)} = K\sum_{k=1}^{K}\tilde{\lambda}_i\alpha_i^{(k)}\overline{x}^{(k)} \rightarrow
\end{aligned}$$

$$\begin{aligned}
&\sum_{k=1}^{K}\alpha_i^{(k)}\sum_{j=1}^{K}\overline{x}_i^{(k)T}\overline{x}_j^{(k)}\overline{x}_j^{(k)T}\overline{x}^{(k)} = K\sum_{k=1}^{K}\tilde{\lambda}_i\alpha_i^{(k)}\overline{x}_i^{(k)T}\overline{x}^{(k)} \rightarrow \\
&\sum_{k=1}^{K}\alpha_i^{(k)}\sum_{j=1}^{K}U_{ij}U_{jk} = K\tilde{\lambda}_i\sum_{k=1}^{K}\alpha_i^{(k)}U_{ik} \rightarrow \\
&\overline{\overline{U}}^2\overline{\alpha}_i = K\tilde{\lambda}_i\overline{\overline{U}}\overline{\alpha}_i \rightarrow \\
&\overline{\overline{U}}\overline{\alpha}_i = K\tilde{\lambda}_i\overline{\alpha}_i
\end{aligned}$$

This last equation can be recognised as the eigenvalue problem. The $\overline{\alpha}_i$ vectors are in fact the eigenvectors of this problem. We can find them by solving the equation.

Now given that $\tilde{\overline{q}}_i = \sum_{k=1}^{K}\alpha_i^{(k)}\overline{x}^{(k)}$ we can determine the principal components:

$$\begin{aligned}
c_i &= \overline{x}^T\tilde{\overline{q}}_i \\
&= \overline{x}^T\sum_{k=1}^{K}\alpha_i^{(k)}\overline{x}^{(k)} \\
&= \sum_{k=1}^{K}\alpha_i^{(k)}\overline{x}^T\overline{x}^{(k)}
\end{aligned}$$

The construction of these principal components in this manner, can be thought of as a neural network. The network has K input neurons, each with N input weights. The output of the network is vector c, which is the compressed version of input vector x.

To decode vector c back into x, we perform the following operation:

$$\tilde{\overline{x}} = \sum_{i=1}^{M}c_i\tilde{\overline{q}}_i$$

The advantage of the Kernel method is that a great reduction in computational complexity is achieved.

4. Artificial Neural Network – Back Propagation

The Backpropagation algorithm is the most popular supervised learning algorithm for feed-forward neural networks.In this algorithm the minimization of the error function is carried out using a gradient-descent technique. The necessary corrections to the weights of the network for each moment *t* are obtained by calculating the partial derivative of the error function in relation to each weight w_{ij}. A gradient vector representing the steepest increasing direction in the weight space is thus obtained. The next step is to compute the resulting weight update. In it simplest form, the weight update is a scaled step in the opposite direction of the gradient. Hence, the weight update rule is

$$\Delta_p w_{ij}(t) = -\varepsilon \cdot \frac{\partial E_p}{\partial w_{ij}}(t),$$

where $\varepsilon \in (0,1)$ is a parameter determining the step size and is called the *learning rate*. The partial derivative of the error for the pattern *p* is given by

$$\frac{\partial E_p}{\partial w_{ij}}(t) = -\delta_{pj} \cdot a_{pi},$$

where δ_{pj} is the error signal of unit *j* and is obtained as follows:

-if unit *j* is an output unit, then

$$\delta_{pj} = f'(net_{pj})(d_{pj} - a_{pj})$$

-if unit *j* is a hidden unit, then

$$\delta_{pj} = f'(net_{pj})\sum_k \delta_{pk} w_{jk}.$$

Hence, the error signals δ_{pj} for the output units can be calculated using directly available values, since the error measure is based on the difference between the desired d_{pj} and actual a_{pj} values. However, that measure is not available for the hidden units. The solution is to back-propagate the δ_{pj} values layer by layer through the network.

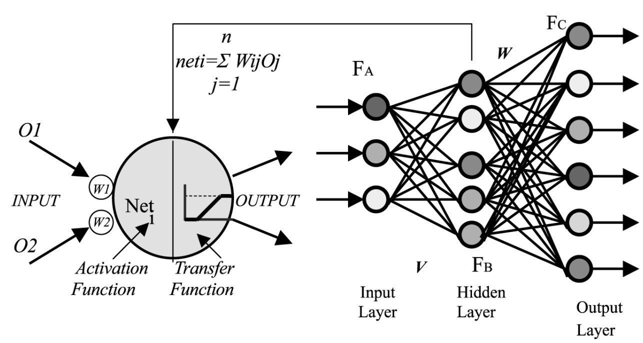

Figure 3-2. back propagation

A *momentum* term was introduced in the Backpropagation algorithm. The idea consists in incorporating in the present weight update some influence of the past iteration. The weight update rule becomes

$$\Delta_p w_{ij}(t) = -\varepsilon \cdot \delta_{pj} \cdot a_{pi} + \alpha \cdot \Delta_p w_{ij}(t-1),$$

where α is the momentum term and determines the amount of influence from the previous iteration to the present one.

The momentum introduces a “damping” effect on the search procedure, thus avoiding oscillation in irregular areas of the error surface and accelerating the convergence in long flat areas. In some situation it possibly avoids the search procedure from being stopped in a local minimum, helping it to skip over those regions without performing any minimization there. In summary, it has been shown to improve the convergence of the Backpropagation algorithm, in general.

5. Result and Discussion

The constrain in building a classification ANN is limiting the number of inputs and of hidden neuron overfitting. It is dificcult to optimize both quantities at the same time. What was done in this work was to try both the raw and preprocessed data after PCA and compare the result.

From input data preprocessed found the best 15 variable as input, a relatively sophisticated way to do this would be to perform a Principal Components Analysis (PCA) on the correlation matrix, and to choose the environmental variable which is most strongly associated with each of the first several principal axes. Thus variable are Lifted Index, Showalter Index, Convective Available Potential Energy, Convective Inhibition, CINS using virtual temperature, Eqilibrium level, Level of Free Convection, LFC-LCL height, Level of Free Convection, Microburst Day Potential Index, Surface Dewpoint, Gust, Thompson Index, Windex, Wet Microburst Severity Index

Table 5-2. Result

Input	**Prediction Accuracy (%)**	
	BPN	**BPN-PCA**
Cb	81.0	95.1
No Cb	63.7	66.7
Overall Percentage	74.9	82.5

The result of the testing prediction as shown above, the accuray of prediction was derived occurences of Cb was predicted correctly, and no occurences of Cb which was predicted correctly also known the accuracy general accuracy result from the model developed. From model with input without PCA process found the result improve in Cb occurrence in two model after PCA. Also accuracy increase significant with dimentionality reduction process, with the limit of input, the best paramater related with cumulonimbus actifity can explained.

Although this research was the first model develop using sounding and ANN to detect Cb, we can compare another typical research with thunderstorm as predictand. Like the result from Ali,A.F et al tried to detect lighting occurrence, just found less accuracy. Manzato in 2005 derive the sounding to ANN model without compare preprocessing input. This research need to expand with larger data and try in other season like condition in tropical area with special characteristi. Unfortunetaly the 2012-2013 rainy season was very paculiar, so the the test sample has been an active case climatology quite different, Cb occurrence higher than other years. Litta, A.J find the thunderstorm occurrence prediction in monsoon season and with the characteristic can be useful in decision making for meteorologist.

6. Conclusion

In this study the performance of combination preprocessing using Principle component analysist applied and compare with raw data pattern. Then the result show neural network after PCA gives better accurary. Back propagation algorithm without PCA is too time consuming and the performance is heavily dependent on the network parameters. But Back propagation give the best result on detecting yes event after PCA process. So using feature selection process much reliable and faster for the cumulonimbus detection. And this method can applied in another data meteorology. Through the implementation of this system, it shown that an intelegent system can be efficently integrated with neural network prediction to predict the Cumulonimbus event.

With this information result, best performance from the experiment become good news for operational used of radiosonde in tropical area. The data of observation can be implemented to make assesment for cumulonimbus prediction speciacllu for airport and early warning center of weather hazard.

In the future studies on Cb prediction using radiosonde indeces, this same scheme used for another season and year will be applied. A possible extention of this kind of work is to include data not only from radiosonde but from another resource like synoptic surface or ensamble with another paramater form numerical weather prediction.

7. References

Ali,A.F.,Johari,D.,Ismail,N.F.,Musirin,I.,Hashim,N., (2011). "Thunderstorm forecasting by using artificial neural network". Proceeding of The 5th International Power Engineering and Optimization Conference, Selangor,Malaysia, pp 369-374.

Ben-Hur,A. and Guyon, I.(2003). " Detecting stable clusters using principal component analysis. In Functional Genomics: Methods and Protocols." M.J. Brownstein and A. Kohodursky (eds.) Humana press, 2003 pp. 159-182

Chakrabarty, H., Murthy, C.A, Bhatacharya,S., Gupta, A.D., (2013). "Application of artificial neural network to predict squall thunderstorm using rawind data." International journal of scienctific & engineering research, Vol 4, Issues 5, pp 1313-1818

Devi, C.J, Reddy, B.S., Kumar, K.V., Nayak, N.R., (2012) "ANN approach for weather prediction using back propagation". International Journal of engineering trends and technology, Volume 3, pp 19-23

Galway, J.G., (1956). The lifted index as a predictor of latent instability. Bull. Am. Meteorol. Soc. 37, 528–529.

Guyon, I., Elissef,A.,(2003) " An introduction to variable and feature selection". Journal of machine learning research 3, pp 1157-1182

Haykin, Simon. (1999)." Neural Networks – a Comprehensive Foundation". Prentice Hall, New Jersey, 2ndedition, 1999. ISBN 0-13-273350-1.

Inoue, T., Satoh, M., Miura,H, Mapes, B,. (2008). " characteristics of cloud size of deep convection simulated by a global cloud resolving model over the western tropical pacific", Journal of the meteorological society of Japan, Vol.86A, pp.1-15

Kuligowski, R.J, Barros,A.P.,(1998). "Experiments in short-term precipitation forecasting using artificial neural networks" . Mon. Weather Rev. 126, 470–482.

Lee, J., Wegner, R.C., Sengputa, S.K., Welch, R.M., (1990). "A neural network approach to cloud classification". IEEE Trans. Geosci. Remote Sens. 28, 846–855.

Litta, A.J, Idicula,S.M., Francis, C.N., (2012),"Artificial neural network model for the rediction of thunderstorms over kolkata", International journal of cumputer application, Vlo 50-No 11, pp 51-55

Ludwiq, O., Nunes, U., Araujo, R.(2014), " Eigenvalue decay: a new method for neural network regularization", Journal neurocomputing,124, pp 33-42

Manzato, A., (2005). "The use of sounding derived indices for a neural network short term thunderstorm forecast". Weather Forecast. 20, 896–917.

Tokuno, M., Tsuchiya.K, (1994), " Classficiation of cloud types based on data of multiple satellite sensors." Adv.Space Res, Vol.14,No.3, pp 199-206

Weng, L.Y, Omar, J., Siah, Y.K., Ahmed, S.K, Abidin, I.Z, (2010), "Lightning forecasting using ANN-BP & Radiosonde", Proceeding of International conference on intellegent computing and cognitive informatics, Kulala Lumpur, Malaysia, pp 153-155

Wilks, D.S., 1995. Statistical Methods in the Atmospheric Sciences.Academic Press. 467 pp.

Wolf, L., Bileschi, S., (2005), " Combining variable selection with dimensionality reduction", Al Memo at massachosetts intitute of technology, cambridge, CBCL Memo 247