

# Estimation of Precipitable Water Vapor Using an Adaptive Neuro-fuzzy Inference System Technique

Wayan Suparta, Kemal Alhasa

► **To cite this version:**

Wayan Suparta, Kemal Alhasa. Estimation of Precipitable Water Vapor Using an Adaptive Neuro-fuzzy Inference System Technique. David Hutchison; Takeo Kanade; Madhu Sudan; Demetri Terzopoulos; Doug Tygar; Moshe Y. Vardi; Gerhard Weikum; Khabib Mustofa; Erich J. Neuhold; A Min Tjoa; Edgar Weippl; Ilsun You; Josef Kittler; Jon M. Kleinberg; Friedemann Mattern; John C. Mitchell; Moni Naor; Oscar Nierstrasz; C. Pandu Rangan; Bernhard Steffen. 1st International Conference on Information and Communication Technology (ICT-EurAsia), Mar 2013, Yogyakarta, Indonesia. Springer, Lecture Notes in Computer Science, LNCS-7804, pp.214-222, 2013, Information and Communicatiaon Technology. <10.1007/978-3-642-36818-9\_22>. <hal-01480177>

**HAL Id: hal-01480177**

**<https://hal.inria.fr/hal-01480177>**

Submitted on 1 Mar 2017

**HAL** is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L'archive ouverte pluridisciplinaire **HAL**, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.



# Estimation of precipitable water vapor using an adaptive neuro-fuzzy inference system technique

Wayan Suparta\*, Kemal Maulana Alhasa

Institute of Space Science (ANGKASA), Universiti Kebangsaan Malaysia, 43600 Bangi,  
Selangor Darul Ehsan, Malaysia  
wayan@ukm.my, kemalalhasa@gmail.com

**Abstract.** Water vapor has an important role in the global climate change development. Because it is essential to human life, many researchers proposed the estimation of atmospheric water vapor values such as for meteorological applications. Lacking of water vapor data in a certain area will be a problem in the prediction of current climate change. Here, we reported a novel precipitable water vapor (PWV) estimation using an adaptive neuro-fuzzy inference system (ANFIS) model that has powerful accuracy and higher level. Observation of the surface temperature, barometric pressure and relative humidity from 4 to 10 April 2011 has been used as training and the PWV derived from GPS as a testing of these models. The results showed that the model has demonstrated its ability to learn well in events that are trained to recognize. It has been found a good skill in estimating the PWV value, where strongest correlation was observed for UMSK station ( $r = 0.95$ ) and the modest correlation was for NTUS station ( $r = 0.73$ ). In general, the resulting error is very small (less than 5%). Thus, this model approach can be proposed as an alternative method in estimating the value of PWV for the location where the GPS data is inaccessible.

**Keywords:** PWV, Adaptive neuro-fuzzy inference system, Estimation, Meteorological applications

## 1 Introduction

One of the most important factors in meteorology is determining the rate of water vapor in the atmosphere. Water vapor content and its variability have an important role in human activities and environments. It has been balancing the energy in the preservation atmosphere, the weather process and ultimately important for operational weather forecasting. Water vapor comes from the evaporation of sea water, ground water, rivers, swamps, glaciers, snow, and water in the atmosphere in the form of clouds. Therefore, accurate estimation of water vapor circulation is of primary importance for monitoring and prediction of precipitation rates to indicate climate change takes place at small scales as well as at regional scales. Many efforts have been invested to study the water vapor changes, especially from the science communities such as meteorology, hydrology and climatology [1]. Atmospheric water vapor

content in terms of precipitable water vapor (PWV) was derived from GPS and the surface meteorological data with superior in temporal and spatial resolution [2],[3]. This technique is still had good capability in comparison with the ground-based techniques that are traditionally made through balloon-borne radiosondes and water vapor radiometers (WVRs). Although the GPS networks currently distributed thousands around the world, there is a still lacking of GPS data, especially in the remote locations that could potentially give the extreme effects on the local weather conditions.

One of the tools of soft computing is the adaptive neuro-fuzzy inference system (ANFIS). ANFIS is a class of adaptive networks that are functionally equivalent to fuzzy inference systems. ANFIS represents both the Sugeno and the Tsukamoto fuzzy models. It has emerged as a powerful tool in solving problems in engineering and non engineering. With uses a hybrid-learning algorithm, it has an ability to do adjustments of rules by using learning set of data and allow the rules to adapt. Thus, it is profoundly suitable for control, pattern recognition and forecasting task. Several researchers have been employed the ANFIS techniques due to its computational speed, robustness, and also its ability to handle complex problems of non-linear functions. Among them is estimation of evaporation using Artificial Neural Networks (ANN) and ANFIS techniques [4], which found that the ANFIS models are better than ANN although the difference is small. Estimation of time series on earthquake events with mapping function [5], and estimation of subsurface strata of Earth [6] reported that the interpretation using ANFIS technique will give the promising results with much fewer percentage error.

The main objective of this study is to investigate the potential use of ANFIS model to predict PWV that influenced by meteorological factors. The second is to evaluate the performance of ANFIS model in estimating the water vapor value by comparing the PWV data obtained from the Global Positioning System (GPS).

## **2 Methodology**

### **2.1 Data Collection**

Two types of data were used in proposed ANFIS. There are the surface meteorological data (pressure, temperature and relative humidity) as the original input and PWV data as the target output. The tested data were taken from the Universiti Malaysia Sabah Kota Kinabalu, Malaysia (**UMSK:** 6.03°N, 116.12°E and height of 63.49 m) and Nanyang Technological University, Singapore (**NTUS:** 1.350°N, 103.680°E and height of 75.38 m). Sabah as a main base of this study has an equatorial climate and naturally more affected by two types of surges associated with East Asian Winter Monsoon that much of the rainfall received during this period. For this work, data gathered from 4 to 10 April 2011 is processed. The meteorological data were taken data of air temperature, barometric pressure, and relative humidity. The PWV data were collected using the ground-based GPS receiver, and the surface meteorological data were collected using the Paroscientific MET4A Broadband meteorological sensors. All the data collected were taken at one-minute interval.

## 2.2 Adaptive Neuro-Fuzzy Inference System

Adaptive neuro-fuzzy inference system is a method that combines neural networks and fuzzy inference system. This method using Sugeno inference model or Takagi-Sugeno-Kang (TSK) fuzzy structures interface, which adapt the rules that are used to fix the parameters so that resulting in a minimum error. ANFIS consists of five components, namely input and output database, fuzzy generator, fuzzy inference system and adaptive neuro. To study the estimation of PWV mode, combination between Sugeno model and TSK inference as a fuzzy inference system (FIS) and adaptive neuron were employed. The optimization method is done by using a hybrid learning algorithm.

Suppose there are two inputs  $x_1, x_2$  to the node and one output is  $y$ , then there are two rules in the base Sugeno models [7]:

$$\text{Rule 1: if } x_1 = A_1 \text{ and } x_2 = B_1 \text{ then } f_1 = p_1x + q_1y + r_1$$

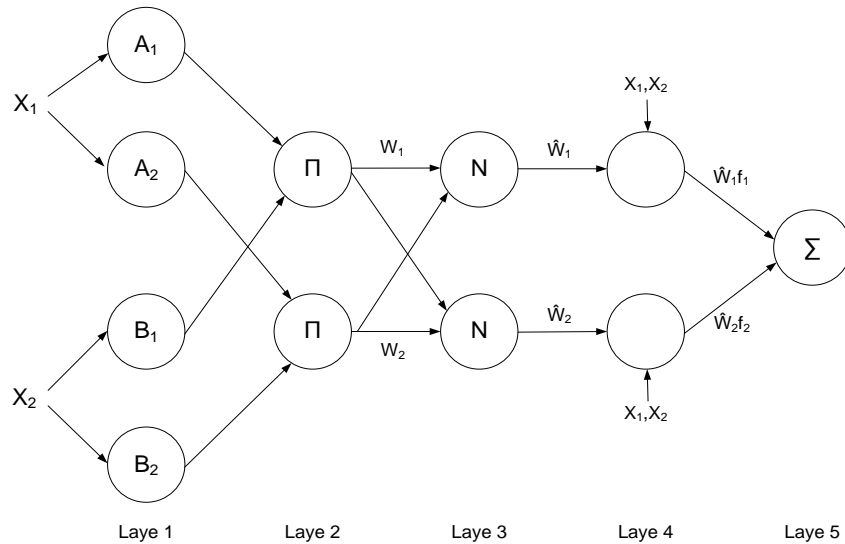
$$\text{Rule 2: if } x_2 = A_2 \text{ and } x_2 = B_2 \text{ then } f_2 = p_2x + q_2y + r_2$$

where  $x_1$  and  $x_2$  are the crisp inputs to the node, and  $A_1, B_1, A_2, B_2$  are fuzzy sets, the reasoning mechanism on this Sugeno model is

$$f = \frac{w_1f_1 + w_2f_2}{w_1 + w_2} = \bar{w}_1f_1 + \bar{w}_2f_2 \quad (1)$$

where  $w_1$  and  $w_2$  are the degree of membership of a fuzzy set, while  $f_1$  and  $f_2$  are linear equations for the output from a first-order of Sugeno inference model.

ANFIS network consist of five layers (see Figure 1) as follows:



**Fig. 1.** Architecture of Adaptive Neuro Fuzzy Inference system

**Layer 1:** This layer serves as fuzzyfication process. The output of node  $i$  in layer 1 is marked as  $O_{1,i}$ . Every node  $i$  in this layer is adaptive to the parameters of an activation function or the output:

$$O_{1,i} = \mu_{A_i}(x), \quad i = 1,2 \quad (2)$$

$$O_{1,i} = \mu_{B_{i-2}}(y), \quad i = 3,4 \quad (3)$$

where  $x$  or  $y$  is the input value to node  $i$  and  $A_i$  or  $B_{i-2}$  is a linguistic label (fuzzy set) associated to the node. While  $O_{1,i}$  is membership degree of fuzzy set. For example, suppose the membership function of fuzzy set  $A$  is given as follows:

$$\mu_A(x) = \frac{1}{1 + \left| \frac{x-c}{a} \right|^{2b}} \quad (4)$$

where  $\{a,b,c\}$  are the parameter set, generally value of  $b$  is equal to 1. As the value of these parameters change, the bell curve shape will change as well. The parameter in this layer is usually called the name of the premise parameters.

**Layer 2:** Every node from this layer is fixed node with labeled  $\pi$ , which output of the result is the incoming signal to entire node. Every output node in this layer present the firing strength of each rule, generally used the AND operator (t-norm operator).

$$O_{2i} = w_i = \mu_{A_i}(x)\mu_{B_i}(y), i = 1,2 \quad (5)$$

For convenience, outputs from this layer generally are called normalized firing strengths.

**Layer 3:** Every node in this layer is non-adaptive or fixed with denoted by  $N$ . Every node serves only to calculate the ratio between the firing strength of  $i^{\text{th}}$  rule towards total firing strength of all rules.

$$O_{3i} = \bar{w}_i = \frac{w_i}{w_1 + w_2} \quad (6)$$

where  $\bar{w}_i$  is normalized firing strengths output of layer 3 and  $(p_i x + q_i y + r_i)$  is parameter set on first order Sugeno fuzzy inference system model. The parameters in this layer are referred to as consequent parameters.

**Layer 5:** A single node in the fifth layer is a fixed node which is the sum of all incoming signals.

$$O_{5i} = \sum_i \bar{w}_i f_i = \frac{\sum_i w_i f_i}{\sum_i w_i} \quad (7)$$

Thus, the fifth layer will build an adaptive network that is functionally equivalent to the first order Sugeno fuzzy inference Model.

### 3 Result and Discussion

In this study, ANFIS was set with the epoch number of 700, step size 0:01 with a membership function 'gaussmf'. The learning method (adaptive network) used the hybrid algorithm. The data used in this study is the surface meteorological data and the GPS PWV data as mentioned in section 2.1. The data are constructed into two parts, training data and checking data. The data used for training as much 2/3 data and 1/3 the data used for checking.

Before the data input is processed by the ANFIS, the first work is to construct initial FIS. In forming the initial FIS, needed a plan. Among them is the establishment of membership functions and rule formation. The curve used in the formation of the membership function is 'gaussmf'. This is due to the Gaussian curve represent the changing of continuous data. PWV is a continuous, which means its value can change over time depending on the season changes, particularly temperature and air pressure conditions. A Gauss curve using two parameters: the central curve domain and the standard deviation that shows the width of the curve. The two values are determined by using Fuzzy C-Means clustering. The value of the central domain curve is taken from the average value of the data cluster and the standard deviation values were obtained from the standard deviation of a data cluster. The formation rules in Sugeno-Takagi type FIS using a linear equation. There are four linear equations in this FIS that has been optimized to estimate PWV. The reasoning mechanisms on this Sugeno model after training by ANFIS are:

$$f_1 = 0.1309P + 2.567T + 0.6413H - 196.9 \quad (8)$$

$$f_2 = 0.06896P + 2.991T + 0.6319H - 157.3 \quad (9)$$

$$f_3 = -0.1251P + 2.851T + 0.6539H + 39.91 \quad (10)$$

$$f_4 = 0.3626P + 3.031T + 0.8667H - 421.5 \quad (11)$$

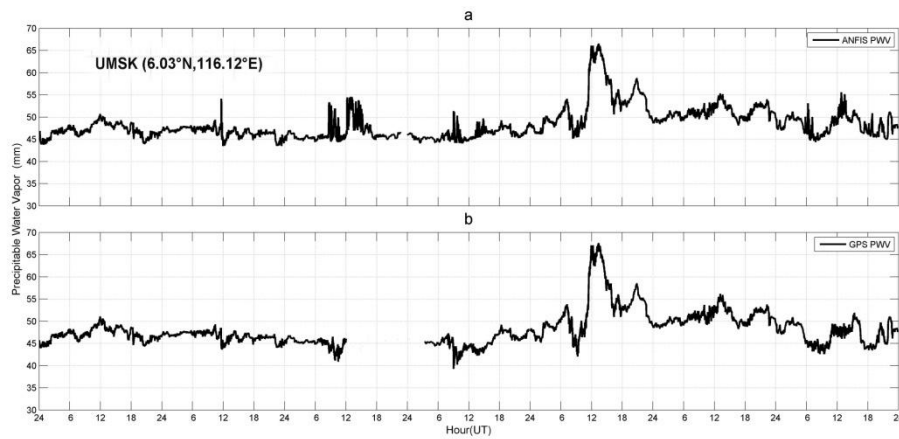
where  $P$  is the surface pressure (mbar),  $T$  is the surface temperature (in degree Celsius), and  $H$  is the relative humidity (in percent). If the predicate  $\alpha$  for the four rules are  $W_1$ ,  $W_2$ ,  $W_3$ , and  $W_4$ , then the weighted average can be calculated as

$$PWV = \frac{W_1f_1+W_2f_2+W_3f_3+W_4f_4}{W_1+W_2+W_3+W_4} = \overline{W_1}f_1 + \overline{W_2}f_2 + \overline{W_3}f_3 + \overline{W_4}f_4 \quad (12)$$

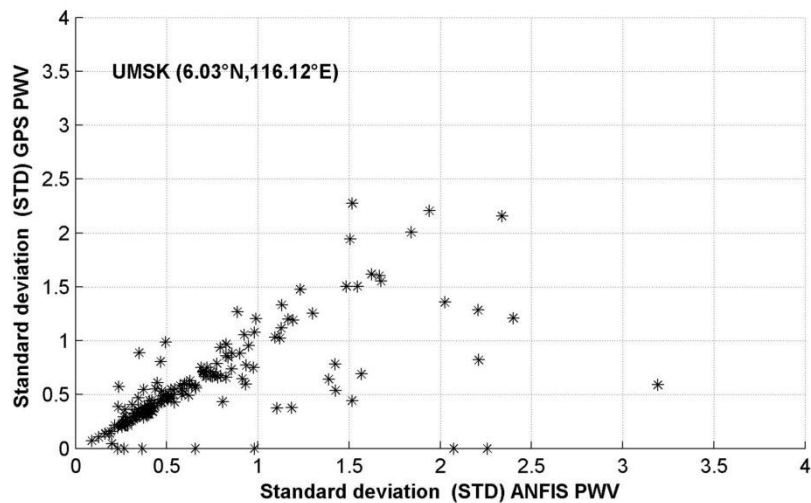
#### 3.1 PWV Result at UMSK

Figures 2a and 2b show the PWV result between ANFIS technique and observed by GPS at UMSK station. In general, the pattern of PWV predicted by ANFIS follows the pattern of PWV observations. Comparing to both figures, data are found blank in Figure 2b due to lack of GPS data at that time and only the surface meteorological

data are available and therefore, the time series is contrast to Figure 2a. This is the advantage of the ANFIS model, which use the surface meteorological data as input. ANFIS model was successfully estimating the PWV value and clearly follow the pattern of PWV observation. Figure 3 shows the scatterplot standard deviation (STD) of PWV between ANFIS and GPS at the UMSK station. The STD calculation was done with one-hour interval of data. The STD pattern between ANFIS and GPS showed a binomial trend where the maximum curve reached to 2.4 mm.



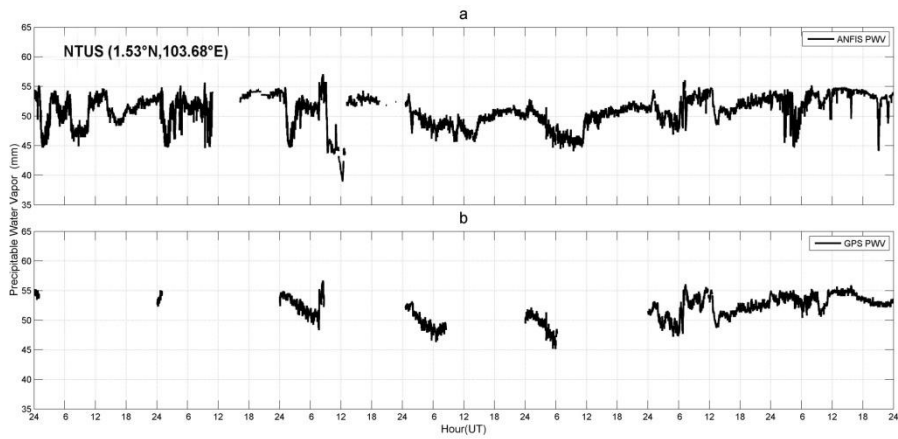
**Fig. 2.** PWV results between prediction using (a) ANFIS technique and (b) observation using GPS at UMSK station, Malaysia



**Fig. 3.** The scatterplot of STD PWV between ANFIS and GPS for UMSK station

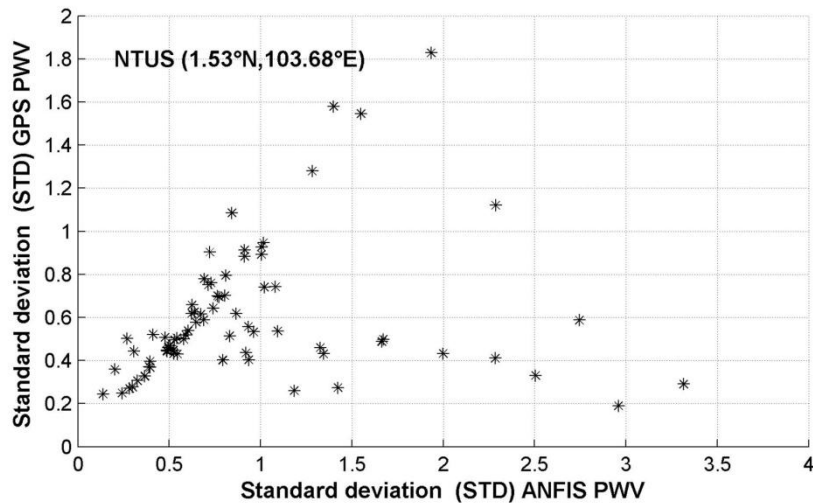
### 3.2 PWV Result at NTUS

A similar PWV result between ANFIS and GPS was also found for NTUS station as depicted in Figures 4a and 4b. From Figure 4a, the final shape of PWV ANFIS is clearly similar to the shape of PWV GPS. Looking at the Figure 4b, some data blank for GPS PWV was recorded at NTUS station. This is same case happened like UMSK station, where there is no GPS data provided at this time. From this lacking, ANFIS model successfully solved the problem as demonstrated in Figure 4a. Figure 5 shows the scatterplot of standard deviation (STD) of PWV between ANFIS and GPS. The STD trend at NTUS was found a similar pattern with the UMSK station with a maximum curve reached to 1.82 mm.



**Fig. 4.** PWV results between prediction using (a) ANFIS technique and (b) observation using GPS at NTUS station, Singapore.





**Fig. 5.** The scatterplot of STD PWV between ANFIS and GPS for NTUS station

From two cases of examination of ANFIS technique at UMSK and NTUS stations shows that the ANFIS prediction results clearly follow the pattern of PWV observed by GPS in the Southeast Asia region, especially, Malaysia and Singapore. The results of PWV comparison at two example stations are given in Table 1.

**Table 1.** Statistical parameters of the PWV comparison

Quantity	UMSK	NTUS
$R$	0.95	0.73
RMSE	1.29 mm	1.68 mm
Mean	-0.06 mm	-1.31 mm

The correlations between the predicted results and observational data show that the ANFIS technique can be used to predict the PWV value, using meteorological data as the input parameter. The relationship with the highest correlation value is observed at UMSK station ( $r = 0.95$ ), while the modest value is observed at NTUS station ( $r = 0.73$ ). From calculation of RMSE value, the resulting error is very small which less than 5%. The PWV difference between the prediction error and the observation are obtained -0.06 mm for UMSK station and -1.31 mm for NTUS station. The negative value indicates that the PWV value from ANFIS is lower than those of PWV obtained from GPS. On the other hands, this method is more convenient, cost-effective and practical when compared to the GPS observations. The promising method only needs the meteorological data (pressure, temperature and relative humidity) as input parameters and very helpful for meteorological station with absent of GPS data.

## 4 Conclusion

In this study, a new method for estimating the PWV value was developed by using one of the tools soft computing, ANFIS techniques. The correlation coefficient between the predicted results and the observational data was found strongest relationship. Looking at UMSK and NTUS stations as an example of examination, the highest correlation coefficient was found at UMSK and the modest correlation coefficient was at NTUS with a resulting error was less than 5%.

A model approaches adaptive neuro fuzzy system interfaces with the input meteorological data (temperature, pressure and relative humidity) can be used as an alternative method in the estimation of PWV value when absent of GPS data in a particular station. In the future, it is recommended to use more extensive data to improve the estimation of PWV value. It is either by adding a new parameter or increases the amount of interval data used to further clarify the accuracy of PWV model developed.

## Acknowledgment

This work was partially supported by the Ministry of Higher Education Malaysia (MOHE) under grants UKM-LL-07-FRGS0211-2010 and PKT 1/2003.

## References

1. Zhang, S., Xu, L., Ding, J., Liu, H., Deng, X.: Advance in Neural Network Research & Application: A Neural Network Based Algorithm for the Retrieval of Precipitable Water Vapor from MODIS Data. LNEE, vol 67, pp. 909-916. Springer, Heidelberg (2010)
2. Bevis, M., Businger, S., Herring, T.A., Rocken, C., Anthes, R.A., Ware, R.H.: GPS Meteorology Remote Sensing of Atmospheric Water Vapor Using the Global Positioning System. J. Geophys. Res. 97, 15787-15801 (1992)
3. Suparta W., Abdul Rashid, Z.A., Mohd Ali, M.A., Yatim, B., Fraser G.J.: Observation of Antarctic Precipitable Water Vapor and Its Response to The Solar Activity Based on GPS sensing. J. Atmos. Sol.-Terr. Phys. 70, 1419-1447 (2008)
4. Kumar, P., Kumar, D., Jaipaul, Tiwari, A.K.: Evaporation Estimation Using Artificial Neural Networks and Adaptive Neuro-Fuzzy Inference System Techniques. Pakistan Journal of Meteorology. 8, 81-88(2012)
5. Joelianto, E., Widiyantoro, S., Ichsan, M. Time Series Estimation on Earthquake Events Using ANFIS with Mapping Function. International Journal of AI. 3, 37-63(2009)
6. Srinivas, Y., Stanley, Raj.A., Hudson O.D., Muthuraj, D., Chandrasekar, N.: Estimation of Subsurface Strata of Earth Using Adaptive Neuro-Fuzzy Inference System (ANFIS). Acta Geod. Geoph. Hung. 47(1), 78-89(2011)
7. Jang, J.S.R.: ANFIS: Adaptive Network-Based Fuzzy Inference Systems. In: IEEE Trans. Syst., Man Cybern. 23. pp.665-685(1993)