*Revue des Energies Renouvelables Vol. 9 N°2 (2006) 113 – 120*
**DOI:** https://doi.org/10.54966/jreen.v9i2.821

# A rule based fuzzy model for the prediction of solar radiation

Radouane Iqdour† and Abdelouhab Zeroual‡

Department of Physics, Cadi Ayyad University, Faculty of Sciences
Semlalia, B.P. 2390, Marrakesh 40000, Morocco.

(reçu le 19 Décembre 2004 - accepté le 21 Juin 2006)

**Abstract** - *The main goal of this investigation is to use the fuzzy systems of Takagi Sugeno (TS) for modelling the daily solar radiation data. The Takagi-Sugeno models are a non-linear techniques, defined by a set of If- Then rules, each of which establishes a local linear input-output relationship between the variables of the model. The TS fuzzy model is trained using data of daily solar radiation recorded on a horizontal surface in Dakhla in Morocco. The predicting results indicate that the Takagi-Sugeno fuzzy model gives a good accuracy of approximately 96 % and a root mean square error lower than 6 %. In addition, the performances of the identified TS fuzzy model are then compared to a linear model using the SOS techniques. The results show the effectiveness of the non linear model.*

**Résumé** – *Le but principal de ce travail est l'utilisation des systèmes flous de Takagi Sugeno (TS) pour modéliser les données de rayonnement solaire journalier. Les modèles de Takagi-Sugeno sont des techniques non linéaires, définies par un ensemble de règles 'Si – Puis', dont chacune établit une relation linéaire d'entrée-sortie entre les variables du modèle. Le modèle flou (TS) est utilisé pour le traitement des données du rayonnement solaire journalier enregistré sur une surface horizontale à Dakhla, au Maroc. Les résultats obtenus indiquent que le modèle flou de Takagi-Sugeno donne une bonne précision approximative de 96 % et une erreur moyenne inférieure à 6 %. En effet, les performances identifiées du modèle flou (TS) sont alors comparées à un autre modèle linéaire en utilisant les techniques de SOS. Les résultats montrent l'efficacité du modèle non linéaire.*

**Keywords**: Daily solar radiation - Modelling - Stochastic process - Takagi-Sugeno systems.

## 1. INTRODUCTION

When an engineer considers the installation in a given place of one or several solar collectors, he needs to know the best possible solar radiation in this place. The statistical knowledge of the characteristics of the sequences for example of the daily solar radiation is necessary for the design and the computation of the performances of the solar systems. These sequences of radiation can be obtained starting from existing measurements, but unfortunately the networks of measurements of the solar radiation are still currently little developed. To overcome these difficulties, several models are developed in the literature for generating sequences of values with the same statistical characteristics as those of sequences observed in nature: average, variance, and probability density function. The mathematical models of regression AutoRegressive (AR), Moving Average (MA) and AutoRegressive Moving Average (ARMA) are among the more used. These models are in general developed using the Second Order Statistics (SOS) or the High Order Statistics (HOS) techniques mainly applicable for stationary processes [1-6].

Since the daily global solar radiation presents a non stationary character and a non gaussian frequency distribution, if we want to apply the models developed using SOS or HOS a transformation of the original data is recommended. Unfortunately, this transformation may influence the prediction precision because the optimal prediction is built on the transformed time series [7-11].

† r.iqdour@ucam.ac.ma
‡ zeroual@ucam.ac.ma

Other models have been developed in the last two decades. Fuzzy identification is among the tools which are effective in the approximation of non-linear systems on the basis of measured data. The fuzzy systems are models where the relationships between the different variables are represented by means of the fuzzy if-then rules with an antecedent and consequent parts. Depending on the form of the consequent part two main types of rule-based fuzzy models are distinguished. The first one is the linguistic fuzzy model where the both parts (antecedent and the consequent) are fuzzy propositions. The second one is the Takagi-Sugeno (TS) fuzzy model of which the antecedent part is a fuzzy proposition and a mathematical function in the consequent part [12-14].

Among these two fuzzy techniques, the Takagi-Sugeno (TS) model [14] which has attracted most attention because the conclusion of each rule is linear and its parameters can be estimated from the numerical data using optimisation methods such as least square algorithms. The fuzzy theory finds its application in several fields such as the recognition problems, the reasoning modelling problems, or the optimisation problems.

In this present work we applied the Takagi-Sugeno fuzzy systems for modelling the daily solar radiation data [7, 8]. A comparison is performed in order to have an idea about the contribution of the fuzzy model with regard to the linear model to improve the precision of the prediction.

## 2. TAKAGI-SUGENO FUZZY SYSTEMS

Lets $Z = \{z_t\}$ to be the database representing the set of the available observations $z_t = (x_t, y_t)$ $(t = 1, \cdots, N)$. The Takagi-Sugeno fuzzy model (TS) consists of aggregating of c fuzzy rules with the following structure:

$$R_k : \text{If } x_t \text{ is } A_k \text{ then } \hat{y}_{t,k} = \beta_0 + x_t' \beta_k \qquad k = 1,2,\ldots,c \quad \text{and} \quad t = 1,2,\ldots,N \tag{1}$$

$R_k : (k = 1,2,\ldots,c)$ indicates $k^{th}$ fuzzy rule, $x_t$ is the input variable $\left(x_t \in R^n\right)$, $\hat{y}_{t,k}$ is the output of the rule k relative to the input $x_t$ and $A_k$ is a fuzzy set and $\beta_k = (\beta_1, \beta_2, \ldots, \beta_n)'$.

The output $\hat{y}_t$ relative to the input $x_t$ obtained after aggregating of $c$ TS fuzzy rules, can be written as a weighted sum of the individual conclusions [16]:

$$\hat{y}_t = \sum_{k=1}^{c} \pi_k(x_t)\, \hat{y}_{t,k} \tag{2}$$

with
$$\pi_k = \frac{\gamma_k \mu_{A_k}(x_t)}{\sum_{j=1}^{c} \gamma_j \mu_{A_j}(x_t)} \tag{3}$$

where $\mu_{A_k}$ is the membership function related to the fuzzy set $A_k$, $y_k = g(\rho_k)$ and $g(x) = \dfrac{1}{1 + e^{-x}}$.

Function $g$ enables the weights $\rho_k$ to be normalized in the sense that $\gamma_k$ will always satisfy the following condition:

$$0 \le \gamma_k \le 1, \qquad k = 1,\ldots,c \tag{4}$$

The latter constraint enables to interpret $\gamma_k$ as an intensity parameter. The normalized intensities are noted $\gamma_k^*$ :

$$\gamma_k^* = \frac{\gamma_k}{\sum_{j=1}^{c} \gamma_j} \tag{5}$$

The rule's influence is given according to the value of $\gamma_k^*$, when the intensity $\gamma_k^*$ is close to 0, we can say that the $c-1$ rules are still sufficient in the approximation, whereas when the intensity $\gamma_k^*$ is close to 1 this rule is significant in computing of the GTS system output [17].

The membership functions are selected Gaussian types:

$$\mu_{A_k}\left( x_t \right) = \exp\left( \frac{-1}{2} \left\| x_t - m_k \right\|_{s_k}^2 \right) \tag{6}$$

$$\left\| x_t - m_k \right\|_{s_k}^2 = \left( x_t - m_k \right)' S_k' \, S_k \left( x_t - m_k \right) \tag{7}$$

The centres $m_k$ and matrix $S_k$ are initialised by projection of the partition obtained from GK algorithm [18]:

$$S_k = \left( F_k^{(x)} \right)^{-1/2} \qquad \text{and} \qquad m_k = V_k^{(x)} \tag{8}$$

$F_k^{(x)}$ and $V_k^{(x)}$ are the projections of the variance covariance matrix, and cluster centres k respectively on the input space.

The identification of a TS model usually requires two types of tuning:

<u>Structural tuning</u>: concerns the determination of the number of rules $c$ and the fuzzy sets to be used in the fuzzy system. Many techniques are available in the literature for c computation [16]. In this study, we used an automatic structural tuning method which exploits the decrement algorithm (DEC) based on a study of the importance, represented by the parameter $\gamma_k$, of the rules on the compute of the output of the TS fuzzy model [17].

<u>Parametric tuning</u>: concerns the estimation of the linear and non linear parameters. The linear parameters $\beta_k$ are identified using the Weighted Least Square (WLS) algorithm. This method encourages the linear experts, parameterised by $\beta_k$, to compete and therefore we can discriminate the action of each expert to get a system more legible. The membership parameters $S_k$ and $m_k$ (non linear parameters) are estimated using the Levenberg-Marquardt method (LM) [19].

## 3. MODELLING DAILY SOLAR RADIATION

In this section, we use the TS fuzzy model for modelling the daily solar radiation. The fig.1 shows the variation of the daily global solar radiation $\left( U_t, \quad t = 1,\ldots,N=365 \right)$ over one year received on a horizontal surface Dakhla (coastal city) in Morocco.

The database $Z = \left\{ \left( x_t, y_t \right) \right\} \ \left( t = 1\ldots1095 \right)$ is constructed from the daily global solar radiation measurements $\left\{ U_t \right\}$, recorded over three years. Therefore, database for fuzzy identification is defined as follows:

$$x_t = \left( U_{t-1}, U_{t-2}, \ldots, U_{t-p} \right) \tag{9}$$

$$y_t = U_t$$

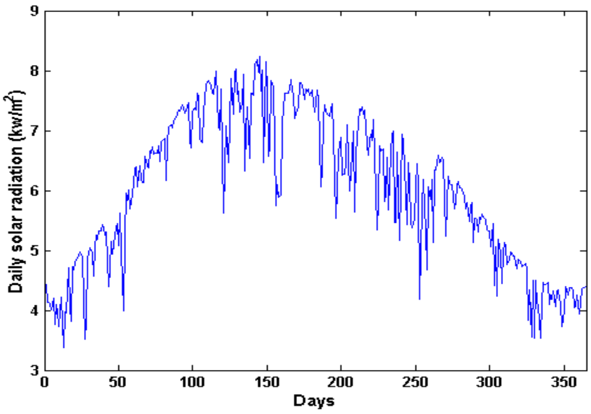


Fig. 1: Annual variation of daily global solar radiation

Where the input of the TS fuzzy model is $x_t$, and the output of the TS fuzzy model is $y_t$. The parameter p is the window size. In the remainder of this work we took $p=1$. In the last section we have tested the influence of this parameter on the precision of the prediction.

The points $(x_t, y_t)$ forming the training database are constructed from the solar measurements over two years. The measurements of the third year are left for the test phase.

The application of the DEC algorithm to determine the optimal structure permitted to obtain a structure with two rules $(c = 2)$. It is worth to not that the DEC algorithm are starting from a structure with ten rules $(c = 10)$. So, we can conclude that eight rules are considered as redundant by the DEC algorithm.

The identified TS fuzzy model is:

$$R_1 : \text{If } x_t \text{ is } A_1 \text{ then } \hat{y}_{t,1} = 1.1581 + 0.767\, x_t$$

$$R_2 : \text{If } x_t \text{ is } A_2 \text{ then } \hat{y}_{t,2} = 0.6751 + 0.8740\, x_t$$

Where $A_1$ and $A_2$ are respectively the fuzzy sets of which the membership functions are presented in the figure 2.

According to the fig. 2 we can remark that the first rule represents the low and medium values of the daily solar radiation. Whereas, the second rule is related to the high values.

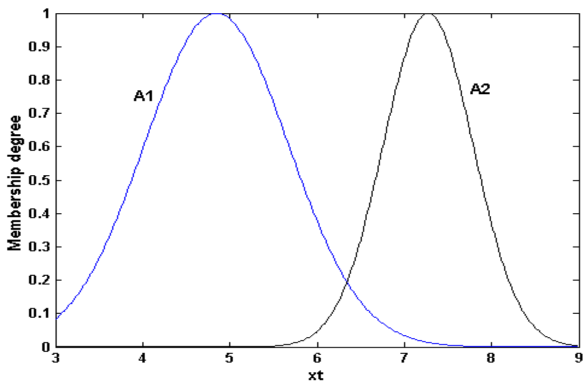


Fig. 2: The obtained Gaussian membership functions

## 4. PREDICTION OF THE DAILY SOLAR RADIATION

The prediction is very useful in solar energy applications because it permits to generate solar data for locations where measurements are not available. In this section, we used the identified TS fuzzy model to predict the daily solar radiation. For that we use the test set of which the data are not using during the training phase.

In figure 3 we present the evolution of the measured and predicted time series of the daily solar radiation using the identified TS fuzzy model. We remark that the predicted and the measured data have similar behaviours, and then we can conclude that the developed TS fuzzy model is adequate to fit the daily solar radiation data.

Another manner to test the validation of the identified model, is to plot the scattering diagram $\left(y_t, \hat{y}_t\right)$ which allows the matching of the measured and predicted values. The distribution of the points $\left(y_t, \hat{y}_t\right)$ around the first bisectrix informs us on the quality of the model. The strong scattering of the points in the superior half or in the inferior half indicates respectively an overestimation or an underestimation.

For our identified TS fuzzy model, in the scatter diagram plotted in fig. 4, we observed a strong scattering of the points $\left(y_t, \hat{y}_t\right)$ around the bisectrix. Thus, the identified fuzzy model provides a reasonable precision.

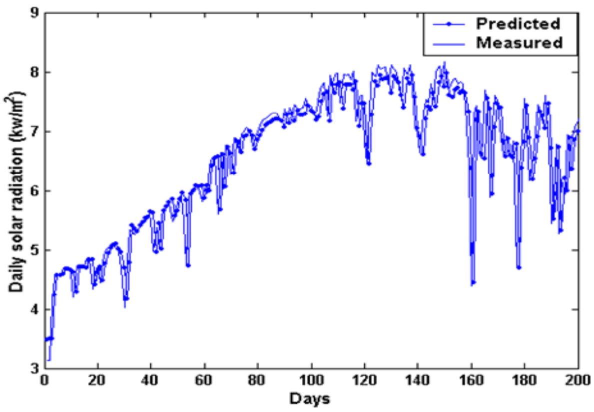


Fig. 3: Comparison between predicted and measured data

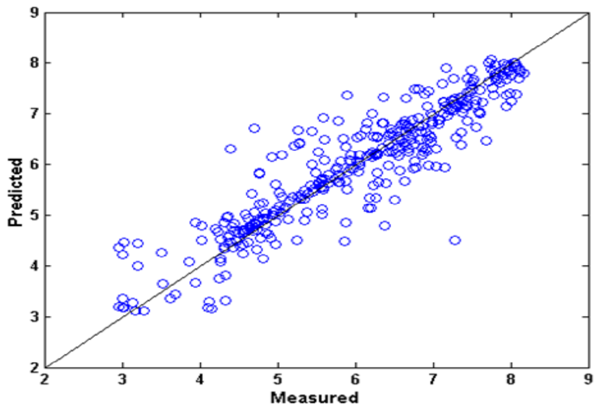


Fig. 4: Scatter diagram of measured and predicted daily solar radiation

Another test can be performed to appreciate the quality of a model to generate a data which have the same frequency distribution as the measurements. The frequency spectres of the real and the generate time series are plotted in the fig. 5. We remark that the spectra of the measured and predicted are comparable.

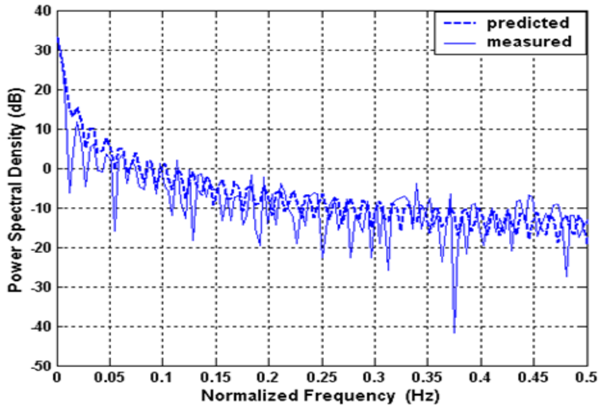


Fig. 5: Spectra of the measured and predicted data

In order to confirm these statistical results we have calculated the criteria RMSE and d (index of agreement) defined as follows:

$$\mathrm{RMSE} = \sqrt{\frac{1}{N}\sum_{i=1}^{N}\left(y_t - \hat{y}_t\right)^2} \tag{10}$$

Where the original time series is $y_t$, the predicted time series is $\hat{y}_t$ and $N$ is the number of patterns.

$$d = 1 - \frac{\sum_{i=1}^{N}\left(\hat{y}'_t - y'_t\right)^2}{\sum_{i=1}^{N}\left(\left|\hat{y}'_t\right| + \left|y'_t\right|\right)^2} \tag{11}$$

Where $\hat{y}'_t = \hat{y}_t - \bar{y}_t$ and $y'_t = y_t - \bar{y}_t$ .

The index d can be any value between 0 and 1; the nearer d is to 1 the better the agreement between observed and predicted values.

## 5. RESULTS AND DISCUSSION

In this section, the prediction results of the identified TS fuzzy model are compared with those obtained using a linear model. Contrary to the TS fuzzy models which suppose that the observed time series result from a determinist process. The linear model supposes that the time series is a signal generated by a linear stochastic process Auto-Regressive Moving Average stochastic process (ARMA).

To identify the linear model ARMA able to describe the daily solar radiation we have employed the Box-Jenkins procedure [1, 2] and the parameters are estimated using the Hannan-Rissanen method [20, 21]. The equation (12)) shows the derived linear regression model (Auto-Regressive (AR) of order two AR(2)):

$$y_t = 0.352\, y_{t-1} - 0.046\, y_{t-2} + \varepsilon_t \tag{12}$$

Where $\varepsilon_t$ is a white noise with mean $0$ and variance $\sigma^2 = 0.824$.

The table 1 represents the prediction performances of each of the identified model (TS fuzzy model and linear model AR (2)).

**Table 1**: Comparison between the two identified models

| Statistical indicators | RMSE | D |
|---|---|---|
| TS Fuzzy model | 0.505 | 96 % |
| Linear model | 0.612 | 89 % |

It can be observed, according to the table 1, that RMSE found is 0.52 kW/m$^2$ and 0.61 kW/m$^2$ respectively for the TS fuzzy model and the linear model which represents less than 8 % and 11 % of the mean of the daily solar radiation data.

The adequacy of fit is also assessed using the agreement index, which returns the percentage of similarity between the measured and the estimated data of the daily solar radiation. We obtain 96 % in the case of the TS fuzzy model and 89 % for the linear model. These results are very satisfactory and the model is considered adequate for such predictions, which strengthens the robustness of the TS fuzzy model.

In the previous study, we have chosen p (window size parameter) equal to 1. In order to evaluate its impact on the precision of the prediction of the daily solar radiation data we have performed a modification of this parameter (p). The following table (Table 2) shows the obtained results.

**Table 2**: Influence of the window size (p) on the quality of the prediction

| Parameter | p = 1 | p = 2 | p = 3 | p = 4 | p = 5 |
|---|---|---|---|---|---|
| TS fuzzy model | 0.505 | 0.497 | 0.488 | 0.481 | 0.473 |

We notice a weak influence of this parameter on the quality of the prediction. We note that the prediction capacity of the identified TS fuzzy model increases with the size of the window. This can be explained by the fact that there is more information about the process in the input of the models what enables it to better predict the daily solar radiation. Whereas, it is worth to not that the difference in performance between the TS fuzzy models of p = 1 and p = 5 do not exceed 0.6 % of t mean of the daily solar radiation data.

## 6. CONCLUSION

In this paper, we have used the TS fuzzy systems for predicting the daily solar radiation. The structure of the TS fuzzy model is identified using a method which permits to determine the optimal structure on automatic manner while being based on the calculation of the importance of the rules in the calculation of the model output. However, the linear parameters are estimated by the WLS algorithm and the non-linear parameters are estimated using the LM method.

Additionally, a comparison is performed between the identified TS fuzzy model and a linear model of which the parameters are estimated using Hannan-Rissanen algorithm.

The obtained results show that the identified TS fuzzy model provides satisfactory performances. So, we conclude that the fuzzy systems can be used as an alternative method to generate solar data for locations where measurements are not available.

## REFERENCES


[1] P.J. Brokwell and R.A. Davis, '*Times Series Theory and Methods*', 2nd Ed, Springer-Verlag, 1991.

[2] G.E.P. Box and G. Jenkins, '*Times Series Analysis, Forecasting and Control*', San Fransisco: Holden-Day, 1970.

[3] A. Zeroual, M. Ankrim and A.J. Wilkinson, '*Stochastic Modelling of Daily Global Solar Radiation Measured in Marrakesh, Morocco*', Renewable Energy, Vol. 6, N°7, pp. 787 - 793, 1995.

[4] S. Safi, A. Zeroual and M.M. Hassani, '*Prediction of Global Daily Solar Radiation Using Higher Order Statistics*', Renewable Energy, Vol. 27, pp. 647 - 666, 2002.

[5] S. Safi and A. Zeroual, '*Modelling Solar Data Using High Order Statistics*', A.M.S.E., Advances in Modelling & Analysis. Vol. 6, N°1, 2, pp. 1 - 16. Advances D-2001.

[6] V. Bahel, R. Srinivsan and Bakhash, '*Statistical Comparison of Correlations for Estimation Of Global Solar Radiation*', Energy 12: pp. 1309 - 1316, 1987.

[7] R. Iqdour and A. Zeroual, '*Modelling Daily Global Solar Radiation Using Fuzzy Systems*', International Conference on Modelling & Simulation - ICMS'04, Valladolid, Spain, 2004.

[8] R. Iqdour and A. Zeroual, '*Modelling Solar Data Using the Takagi-Sugeno Fuzzy Systems*', International Conference on Modelling & Simulation - MS'04, Lyon, France, 2004.

[9] R.J. Stone, '*Improved Statistical Procedure for the Evaluation of Solar Radiation Models*', Solar Energy, Vol. 51, pp. 289 - 291, 1993.

[10] M. Iqbal, '*An Introduction to Solar Radiation*', Academic Press, 1983.

[11] H. Bouhadou, M.M. Hassani, A. Zeroual and A.J. Wilkinson, '*Stochastic Simulation of Weather Data Using Higher Order Statistics, Morocco*', Renewable Energy, Vol. 12, N°1, pp. 21 - 37, 1997.

[12] L. Zadeh, '*Fuzzy Sets - Information and Control*', Vol. 8, pp. 338 - 353, 1965.

[13] W. Pedycz, '*An Identification Algorithm in Fuzzy Relational Systems*', Fuzzy Sets and Systems 13, pp. 153 - 167, 1984.

[14] R. Babuska, '*Fuzzy Modelling tnd Identification*', Thesis, Delft University of Technology, 1996.

[15] T. Takagi and M. Sugeno, '*Fuzzy Identification of Systems and its Application to Modelling and Control*', IEEE Trans Systems Man Cybernet, Vol. 15, pp. 116 - 132, 1985.

[16] A. Fiordaliso, '*Systèmes Flous et Prévision de Séries Temporelles*', Hermes Science, 1999.

[17] A. Fiordaliso, '*Auto-Structuration of Fuzzy Systems y Rules Sensitivity Analysis*', Fuzzy Sets and Systems, Vol. 118, pp. 581 - 586, 2001.

[18] E.E. Gustafson and W.C. Kessel, '*Fuzzy Clustering with a Fuzzy Covariance Matrix*', Proc. IEEE CDC, pp. 761 - 766, 1979.

[19] D.W. Marquardt, '*An Algorithm for Least Squares Estimation of Non-linear Parameters*', J. Soc. Ind. Appl. Math, Vol. 11, pp. 431 - 441, 1963.

[20] E.J Hannan, '*The Estimation of the Order of an ARMA Process*', Annals of Statistics, Vol. 8, pp. 1071 - 1081, 1980.

[21] E.J. Hannan and J. Rissanen, '*Recursive Estimation of Mixed Autoregressive-Moving Average Order*', Biometrika, Vol. 69, pp. 81 - 94, 1982.