IEEE Second International Conference on Data Stream Mining & Processing
August 21-25, 2018, Lviv, Ukraine

On the Equivalence between AR Family Time Series Models and Fuzzy Models in Signal Processing

Anna Walaszek-Babiszewska
Department of Computer Science
Opole University of Technology
Opole, Poland
a.walaszekbabiszewska@gmail.com

Marek Rydel
Department of Computer Science
Opole University of Technology
Opole, Poland
m.rydel@po.opole.pl

Nataliia Kashpruk
Department of Computer Science
Opole University of Technology
Opole, Poland
n.kashpruk@gmail.com

***Abstract*— In the paper an advanced analysis of the relationships between statistical *Autoregressive* (AR) type models and fuzzy models have been presented. The examined family of AR type models includes *Autoregressive models of order p*, AR(p), *Threshold AR* (TAR) as well as *Smooth Transition Autoregressive* (STAR) models. On the other hand, fuzzy models representing different approach, characteristic for *Computational Intelligence* technics, have been tested for time series analysis and forecasting. The data have been taken from financial market. The research can enrich knowledge which is useful for experts using both approaches to modelling.**

I. Introduction

Statistical models worked out in the area of mathematical statistics played a breakthrough role in time-discrete signal processing and were widely employed in a number of fields of science. These models were mainly developed for application in econometrics and control theory. Some clarification in the theory of time series, primarily on account of its application in control theory, was introduced by the work written by Box and Jenkins concerning linear models of time series: *Autoregressive* (AR), *Moving Average (*MA), their combinations *Autoregressive Moving Average* (ARMA) and *Autoregressive Integrated Moving Average* (ARIMA) [1]. Works in the field of econometrics resulted in another types of stochastic models including modelling of nonlinear time series, e.g. *Threshold Autoregressive* (TAR), *Smooth Transition Autoregressive* (STAR), *Self-Exciting Threshold Auto-Regressive* (SETAR), *Auto-Regressive Conditional Heteroscedasticity* (ARCH), as well as a number of other models. Together with the development of methods and techniques of artificial intelligence, fuzzy models and neuro-fuzzy models started to be applied for analysis and forecast of time series. During a few decades of existence of C*omputational Intelligence,* the literature on the subject presents still not enough comparisons of both approaches to building models, identification and effects in terms of convenience of application, accuracy, computational volume, etc. The works of J.L. Aznarte and J.M. Benitez constitute an exception. In their paper we find the following proposition:

„The STAR (Smooth Transition Autoregressive) model is functionally equivalent to an Additive TSK Fuzzy Rule-Based (FRB) model with only one term in the rule antecedents." [2].

The aim of the article is theoretical analysis of building models from autoregressive family (AR, TAR and STAR models) as well as fuzzy models for indicating opportunities of gaining and using knowledge useful for constructing these models. The main criteria of the comparison analysis include:

- type of the state-space domain granulation,
- mathematical form of the models from AR family and fuzzy Mamdani's as well as Takagi-Sugeno-Kang's (TSK) models,
- statistical metrics, as Mean Squared Error (MSE) and autocorrelation function of residues.

Theoretical analysis will be supported by computational examples.

II. Autoregressive Family of Models

A. Autoregressive Models of Time Series

Autoregressive model constitute a scheme of a time-discrete stochastic process $\{X_n\}$, n=1,2,… which assumes that future values of the process stand for a linear combination of its p past values

$$x_n = a_0 + a_1 x_{n-1} + ... + a_p x_{n-p} + \varepsilon_n \qquad (1)$$

where p≥1, $\{\varepsilon_n\}$ is a *white noise process* of the finite variance, $\sigma_\varepsilon^2 < \infty$, and a covariance function $\mathrm{cov}(\varepsilon_i, \varepsilon_j) = 0, i \neq j$. Such model defined by (1) is known as the autoregressive model of order p, $AR(p)$.

When applied this model to forecast future values of the process, $\hat{x}_n$, parameters $a_0, a_1, ..., a_p$ can be determined by the least squares method and past values of the process $\{x_1, x_2, ..., x_K\}$:

$$\sum_{n.} [x_n - \hat{x}_n]^2 = \min_{\overline{a}} \sum_{n=p+1,...,K} [x_n - (a_0 + ... + a_p x_{n-p})]^2 \qquad (2)$$

The assumption of stationarity of the process $\{X_n\}$ secures that $AR(p)$ model with parameters $\hat{a}_0, \hat{a}_1, ..., \hat{a}_p$ estimated on the base of large sample $\{x_{p+1}, x_{p+2}, ..., x_K\}$,

$K >> p$ suites sufficiently for whole data population.

B. Threshold Autoregressive and Smooth Transition Autoregressive Models

In a practice the realizations of stochastic processes rather do not meet assumptions of stationarity, especially the assumption concerning the constant value of the mean process value. This is the reason that autoregressive models are not very good approximation for whole series.

The piecewise approach to modelling, by using e.g. the *threshold* autoregressive *model* (TAR) usually improves such approximation. In different subspaces of χ, process $\{X_n\}$ is described by local autoregressive models x_n^r. Transition from a local model to another one is described by the so-called transition function, λ_r, which assumes the value of 1 for each *r-th* subspace, whereas in the remaining ones it is equal to zero. The global model may be written as :

$$x_n = \sum_r \lambda_r x_n^r \ , \qquad (3)$$

where x_n^r is an autoregressive model of the form (1) defined in the *r-th* range of the selected variable. The boundaries of the intervals stand for the thresholds of the model.

Smooth transition autoregressive (STAR) model contains continuous transition functions $\lambda_r(b^r, x_{n-1}^r)$ which define the location and shape of the transition between local autoregressive models x_n^r. Transition function $\lambda_r(b^r, x_{n-1}^r)$ takes its values from unit interval [0,1] and usually is a nonlinear function of independent variables.

III. Fuzzy Models in Time series Analysis

A. Fuzzy Rule-Based Linguistic Model

The basis of constructing fuzzy models of systems is the input-output space division, X^p, into sub-areas where behaviour of the modelled system can be described by one conditional statement. It is the so-called information granulation process. The operation is analogous to space division in constructing TAR and STAR models.

Let us analyse the dependence $x_n = \varphi(x_{n-1})$, defined in space χ^2, which is modelled as the set $\{R_i\}$, i=1,2,…,I of conditional linguistic rules of the form:

$$R_i\text{: } If \ \ (x_{n-1} \ is \ A_i) \ Then \ (x_n \ is \ A_i) \ . \qquad (4)$$

The input-output space, $\chi \times \chi$, is divided by fuzzy sets, $A_i \times A_i$, i=1,2,…,I. The antecedent of the rule defines fuzzy condition and the consequent part of the rule defines fuzzy conclusion. Fuzzy sets are most often defined by piece-wise linear membership functions (5), (6) or Gaussian membership function (7):

$$\mu_{A_i}(x; a_i, m_i, c_i) = \begin{cases} \dfrac{x - a_i}{m_i - a_i}, & a_i \le x \le m_i \\ \dfrac{c_i - x}{c_i - m_i}, & m_i < x \le c_i \\ 0, & (x < a_i) \cup (x > c_i) \end{cases} \qquad (5)$$

The form of membership fuzzy sets function is responsible for transformation of input information, that is to say for fuzzyfication process.

$$\mu_{A_i}(x; a_i, b_i, c_i, d_i) = \begin{cases} \dfrac{x - a_i}{b_i - a_i}, & a_i \le x \le b_i \\ 1, & b_i < x \le c_i \\ \dfrac{d_i - x}{d_i - c_i}, & c_i < x \le d_i \\ 0, & (x < a_i) \cup (x > d_i) \end{cases} \qquad (6)$$

$$\mu_{A_i}(x) = \exp\left(-\frac{(x - m_i)^2}{2\sigma_i^2}\right) . \qquad (7)$$

In the fuzzyfication process the numerical value of input $x_{n-1}^{\bullet}$ is transformed into the activation level of a rule

$$\tau_i = \mu_{A_i}(x_{n-1}^{\bullet}) \ , \qquad (8)$$

according to linear relation, (5) and (6), accordingly to non-linear relation (7) or in relation changing numerical values $x_{n-1}^{\bullet}$ to a constant value equal to 1 if $x_{n-1}^{\bullet} \in (b_i, c_i]$, according to (6).

When the fuzzy reasoning procedure runs in compliance with Mamdani-Assilan formula, the fuzzy conclusion membership function is determined as follows [3]:

$$\mu_{A'}(x_n) = \max[\min[\tau_i, \mu_{A_i}(x_n)] . \qquad (9)$$

The *maximum* and *minimum* operations correspond to the logical union and intersection of fuzzy sets, whereas $\tau_i = \mu_{A_i}(x_{n-1}^*)$ is the level of activation of *i-th* formula for numerical value of input x_{n-1}^*.

Another operation influencing the transformed signal is defuzzyfication. In continuous space, $\chi \subset R$, non-fuzzy value of output , x_n^*, constitutes the following weighted value, dependent on the area below the function $\mu_{A'}(x_n)$ of fuzzy conclusion *A'*:

$$x_n^* = \int_\chi x_n \mu_{A'}(x_n) dx_n \Big/ \int_\chi \mu_{A'}(x_n) dx_n \qquad (10)$$

To sum up, while applying Mamdani's linguistic model for time series modelling the result of reasoning in the form of numerical output value $x_n^* \in \chi$, for a given value of premise (input variable), $x_{n-1}^* \in \chi$, generally constitutes a non-linear dependence $x_n^* = \varphi(x_{n-1}^*)$ mainly due to reasoning and deffuzyfication procedures. As it appears in fuzzyfication process it is possible to preserve linearity of transformation.

B. Simplified Method of Fuzzy Reasoning

Applying the Simplified Method of Fuzzy Reasoning , we obtain non-fuzzy output x_n^* as the weighted average of centroides, m_i, of the output variable fuzzy sets [4]:

$$x_n^* = \frac{\sum_i \tau_i \cdot m_i}{\sum_i \tau_i}. \qquad (11)$$

Generally, as regards model with one input variable, the levels of activation of rules meet the condition $\sum_i \tau_i = 1$ and then the output value

$$x_n^* = \sum_i \tau_i \cdot m_i \qquad (12)$$

may be the linear dependence if $\tau_i(x_{n-1}^\bullet)$ is the linear function, that is to say when fuzzy sets of membership function (5) have been chosen. Moreover, for $\tau_i = 1$, when only one formula is active, the output value in the local model is a constant value, $x_{n,i} = m_i$.

C. Takagi-Sugeno-Kang's Fuzy Model

Using Takagi-Sugeno-Kang's (TSK) fuzzy model [5], for modelling dependencies $x_n = \varphi(x_{n-1})$ we build set $\{R_i\}$, i=1,2…,I of rules of the form :

$$R_i\text{: } If \quad (x_{n-1} \text{ is } A_i) \text{ Then } x_n^i = \alpha_i x_{n-1}. \qquad (13)$$

These rules differ from the form (4) in that, there is a non-fuzzy function of the input numerical values in formula successor (13), in this case this being a linear function. Usually, parameters α_i, i=1,2…,I are known. A single rule provides a local linear model. The global model is obtained as the weighted sum of active rule outputs

$$x_n^* = \frac{\sum_i \tau_i \cdot x_n^i}{\sum_i \tau_i}, \qquad (14)$$

where τ_i is the activation level of *i-th* rule, (8). Assuming that $\sum_i \tau_i = 1$, for the input x_{n-1}^*, relationship (14) leads to the form :

$$x_n^* = \sum_i \tau_i \cdot \alpha_i x_{n-1}^* \qquad (15)$$

Taking into account that $\tau_i(x_{n-1}^*)$ is a linear or exponential function, formula (15) does not provide a linear dependence $x_n = \varphi(x_{n-1})$ any more . While for $\tau_i = 1$ with only one rule being active, the output value is identical with the equation in the *i-th* rule consequent's part

$$x_n^* = x_n^i = \alpha_i x_{n-1}^*. \qquad (16)$$

Hence, the fuzzy sets in a part of premise rule (13) of the TSK model cannot be entirely arbitrary. They are usually sets of trapezoid membership functions where the linear part (increasing or decreasing) corresponds to that part of space χ which belongs to two fuzzy sets simultaneously. Reasoning provides smoothing, according to (14), of two linear models described by two rules.

IV. Exemplary Calculations

In the research a real time series of WIG20, Polish market indicator, $\{x_n\}$, n=1,2,…,150, $x \in \chi = [1630, 2080]$, was used to demonstrate features of tested models. Preliminary test of the series was composed of: calculations of a mean value, variance and autocorrelation function of the series. According to that, space χ has been divided into subspaces by the threshold x_t=1855 for building TAR model (see Fig.1 and Fig. 2). Autocorrelation function of the series proved, that $\{x_n\}$ is not a realization of the white noise process but constitutes the realization of the long memory stochastic process. Therefore searching for time series models is justified.

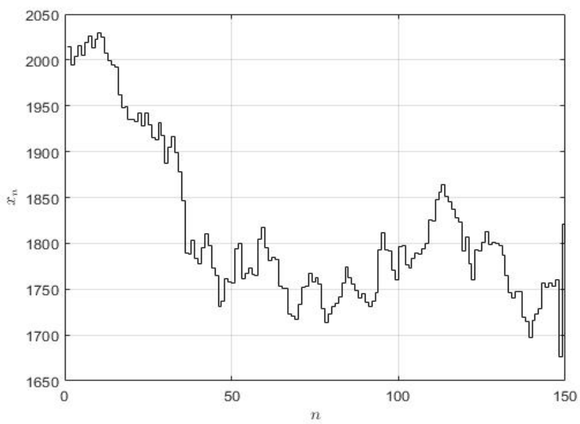

Fig. 1. Time series x_n of WIG 20

It is assumed to search for models containing only one lagged value of the series, $\hat{x}_n = \hat{\varphi}(x_{n-1})$, which means a one-step prognosis. The models AR(1), TAR and TSK of rules numbers 3 and 5 have been established. The mean square errors MSE of models and MSE related to variance of row data D_x^2, have been computing. Moreover, autocorrelation functions of residues for all models have been calculated and presented in Fig. 3.

The models have the following form:

AR(1) $\hat{x}_n = 0.9991 x_{n-1}$

TAR(1) $\hat{x}_n = \begin{cases} 0.9979 x_{n-1}, & x_{n-1} \geq 1855 \\ 0.9997 x_{n-1} & x_{n-1} < 1855 \end{cases}$

TSK (3)

R_1: *If* $(x_{n-1}$ *is* $A_1)$ *Then* $x_n^1 = 0.6601 \cdot x_{n-1} + 621.6$

R_2: *If* $(x_{n-1}$ *is* $A_2)$ *Then* $x_n^2 = 1.296 \cdot x_{n-1} - 550.1$

R_3: *If* $(x_{n-1}$ *is* $A_3)$ *Then* $x_n^3 = 1.389 \cdot x_{n-1} - 794.4$

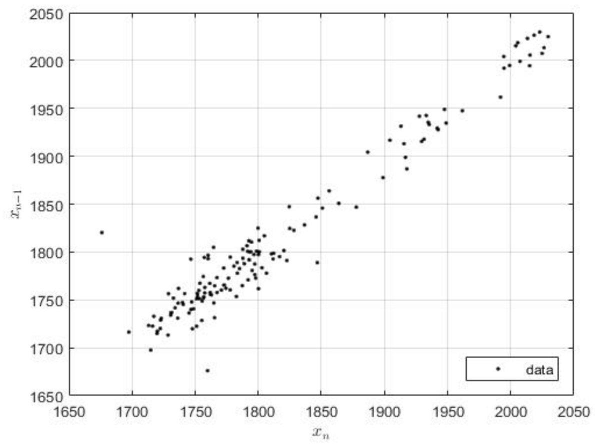

Fig. 2. Diagram $x_n = \varphi(x_{n-1})$ of data

where membership functions of particular fuzzy sets A_1, A_2, A_3 are triangular with following parameters:

$$\mu_{A_1}(x; a_1, m_1, c_1) = \mu_{A_1}(x; 1630, 1675, 1855),$$
$$\mu_{A_2}(x; a_2, m_2, c_2) = \mu_{A_2}(x; 1675, 1855, 2035),$$
$$\mu_{A_3}(x; a_3, m_3, c_3) = \mu_{A_1}(x; 1855, 2035, 2080).$$

TABLE I. PARAMETERS OF RESIDUES OF PARTICULAR MODELS

Model	**MSE**	MSE / D_x^2, [%]
AR(1)	451.49	5.78
TAR(1)	288.18	3.67
TSK(3)	371.43	4.73
TSK(5)	320.51	4.08

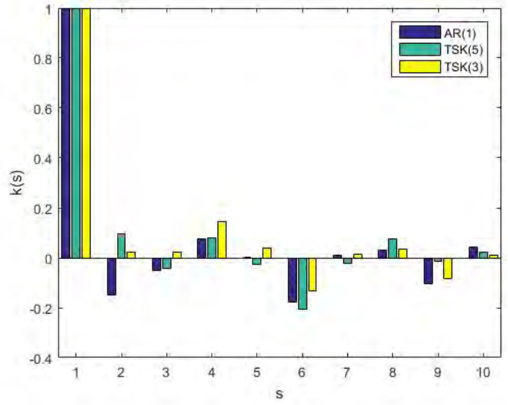

Fig. 3. The autocorrelation function of residues of tested models

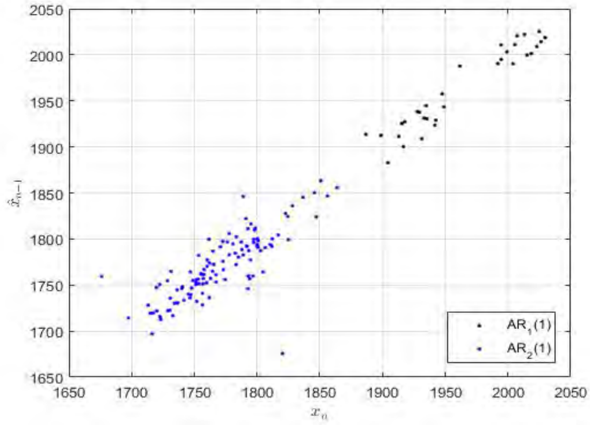

Fig. 4. Diagram $\hat{x}_n = \varphi(x_n)$ of TAR model and data

Models AR(1) and TAR contain one parameters each, close to 1, however in the consequent parts of TSK models two-parameters linear relationships, $\hat{x}_n = a x_{n-1} + b$, are included. The values of MSE shown in Table 1 and diagram presented in Fig.4 prove that TAR model consisting of two AR(1) models gives the best mapping of input time series. The autocorrelation function of residues point out that there is a correlation on a level lower than 0.2. Hence it can be assumed that all the models describe changeability of input series satisfactorily.

V. CONCLUSIONS

Taking into account theoretical analysis of the models conducted in paragraphs I. – III. the following differences in the procedures of creating AR linear models and fuzzy rule-based models may be pointed out :

- AR linear models are created as *Least Squares* approximation of the entire data set.
- In fuzzy rule-based modelling each rule constitutes a local model built on a part of data set and aggregation procedure plays the role of fuzzy transition.
- Linguistic fuzzy models with Mamdani-type reasoning and Simplified Method of Fuzzy Reasoning

give mainly nonlinear dependencies $x_n = \varphi(x_{n-1})$ due to the specificity of reasoning and deffuzyfication procedures.

- Fuzzy rule-based TSK models are closest to TAR and STAR models; local linear models represented by particular rules are aggregated as weighted sum where weight coefficients are not constant but depend on an input variable.
- In order to obtain a linear mapping of local TSK, $x_n^i = \varphi(x^i{}_{n-1})$, it is advisable to apply trapezoid fuzzy sets in the input variable space.

The conducted exemplary calculations for a given time series show more comparability of accuracy of the tested models although they differ in structure and even the equation form. The choice of model and method of its obtaining is a matter of the user's choice.

REFERENCES

[1] G. Box and G. Jenkins, Time Series Analysis: Forecasting and Control, San Francisco: Holden-Day, 1970.

[2] J. L. Aznarte and J. M. Benitez, “The links between statistical and fuzzy models for time series analysis and forecasting,” in: Time Series Analysis, Modeling and Applications; A Computational Intelligence Perspective, W. Pedrycz and Shyi-Ming Chen, Eds. Intelligence Systems Reference Library, vol. 47, Springer, pp. 1-30, 2013.

[3] E. H. Mamdani and S. Assilan, “An experiment in linguistic synthesis with a fuzzy logic controller,” International Journal of Man-Machine Studies, 20(2), pp. 1-13, 1970.

[4] R. R. Yager and D. P. Filev, Essentials of Fuzzy Modeling and Control, John Wiley & Sons, Inc. 1994.

[5] T. Takagi and M. Sugeno, “Fuzzy identification of systems and its applications to modeling and control,” IEEE Transactions on Systems, Man and Cybernetics, 15, pp.116-132, 1985.