# Determination of Fuzzy Relations for Economic Fuzzy Time Series

**Dušan Marček**

Department of Applied Informatics, Faculty of Economic, VŠB-Technical University of Ostrava

**Ostrava 2019**

# Outlines

- Basic principles of identifying input-output functions of systems and forecasting
- A classification table of statistical and NN methods for time series modelling and forecasting
- Statistical ARIMA time series models, B-J approach
- Fuzzy time series and models
- A practical example
- Determination of fuzzy relations
- Conclusion

# Basic principles of identifying input-output functions of systems and forecasting

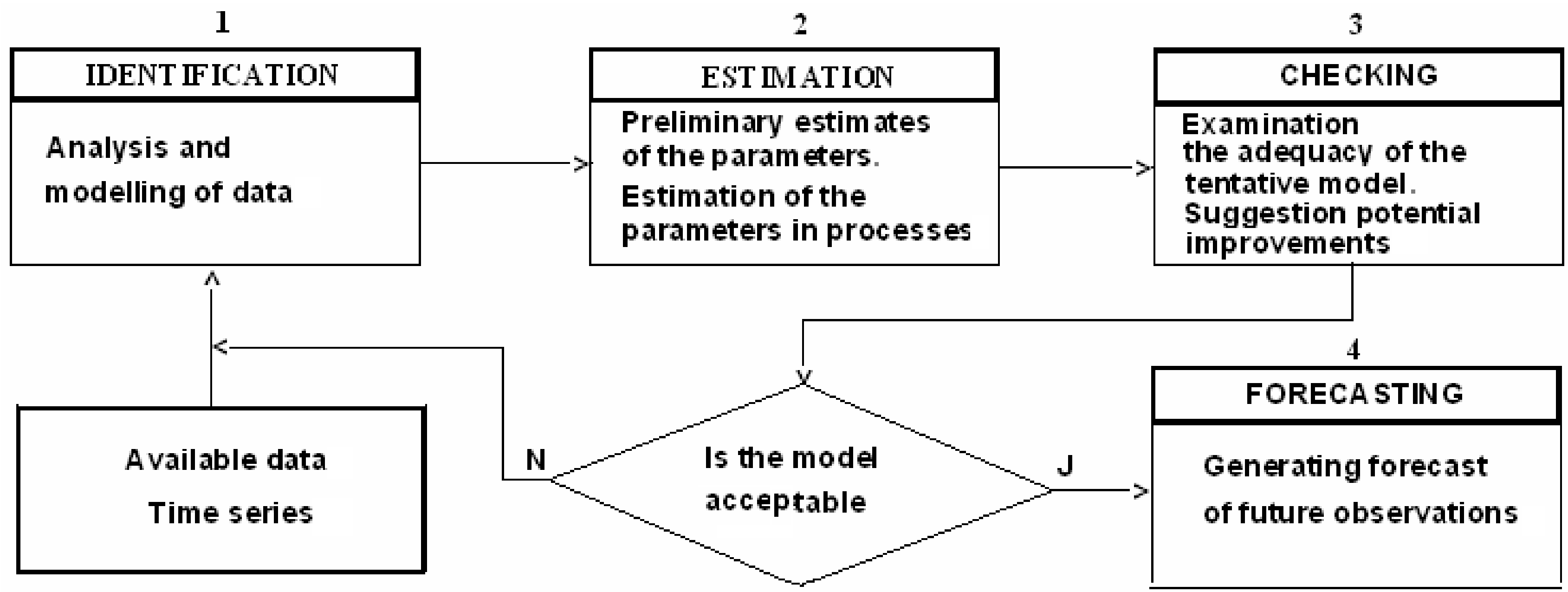
- **There are three major approaches to forecasting – explanatory, time series, machine learning**
- **Explanatory forecasting assumes a cause and effect relationship between the inputs into the system and its output. In this system any change in inputs will affect the output of the system in a predicable way.**
- **Time series**
- **In a special case machine learning such as ANN, fuzzy sets theory offers relatively new ways for improving forecast method.**

# Time series, B-J approach

**Box and Jenkins developed a new modeling approach based on time series analysis and derived from the linear filter known as AR or ARIMA (AutoRegressive Integrated Moving Average) models. The basic ARMA model of orders p, q (ARMA(p,q)) has the form:**

$$y_t = \phi_1 y_{t-1} + \phi_2 y_{t-2} + \ \dots \ + \varepsilon_t + \theta_1 \varepsilon_{t-1} + \theta_2 \varepsilon_{t-2} + \ \dots$$

**The basic steps in Box-Jenkins Procedure are: (1) Identification, (2) Estimation, (3) Diagnostic checking, (4) Forecasting. See next figure**




Flow chart of building an appropriate time series forecast model

# Explanatory approach

**The econometric approach adopted from early days of econometrics is referred to as "AER" or Average Economic Regression and is concerned with the functional form of the multiple regression model in the form**

$$y_t = \beta_0 + \beta_1 x_{1t} + \ldots + \beta_p x_{pt} + u_t$$

**The formulation of an econometric forecasting model requires the following steps:**

- The choice the independent variables
- specification of a functional form of the model
- collection, and analysis of a data set
- model estimation and statistical testing
- evaluation of the model´s forecasting over the ex post period

# Time series models, B-J approach

Quantitative Statistical and Fuzzy Time series Modeling Methods

In practice, there are many processes in which successive observations are dependent, i.e. there exists an observational relation

$$R = \{(y_t, = f(y_t, y_{t-1}), (y_{t-1}, y_{t-2}),\ \ldots\ \} \subseteq Y_t \times Y_{t-1}$$

The most often used model is, however, an explicit function, AR(1) process, (model)

$$f:\ \ Y_{t-1} \to Y_t, \qquad y_t = f(y_{t-1}, \phi_1, \varepsilon_t) = \phi_1 y_{t-1} + \varepsilon_t$$

The AR(1) process is a special case of a stochastic process which is known as the mixed autoregressive-moving average model of the order *(p, q)* which is abbreviated ARMA(*p, q)*

$$y_t = \begin{cases} \phi_1 y_{t-1} + \phi_2 y_{t-2} + \ldots + \phi_p y_{t-p} \\ -\theta_1 \varepsilon_{t-1} - \theta_2 \varepsilon_{t-2} - \ldots - \theta_q \varepsilon_{t-q} + \varepsilon_t \end{cases}$$

All the above time series can be derived from linear combination of independent white noise random variables $\{\varepsilon_t, \varepsilon_{t-1}, \varepsilon_{t-2} \ldots\}$

$$y_t = \mu + \psi_0 \varepsilon_t + \psi_1 \varepsilon_{t=1} + \ldots = \mu + \sum_{j=0}^{\infty} \psi_j \varepsilon_{t-j}$$

# Time series models

## Fuzzy Time series Modeling Methods

In the case of fuzzy time series the fuzzy relational equations can be employed as the models. Analogously to conventional time series models, it is assumed that the observation at the time *t* accumulates the information of the observation at the previous time, i.e. there exists a fuzzy relation such that

$$y_t^j = y_{t-1}^i \circ R_{ij}(t, t-1) \qquad (4)$$

where $y_t^j \in Y_t,\ y_{t-1}^j \in Y_{t-1},\ i \in I,\ j \in J$

Then Yt is said to be caused by Yt-1 only, i.e.

$$y_{t-1}^j \to y_t^i$$

or eqivalemtly

$$Y_{t-1} \to Y_t$$

and

$$Y_t = Y_{t-1} \circ R(t, t-1) \qquad (1)$$

Equation (1) is equivalent to the linguistic condition

$$if\ \ y_{t-1}^i \quad \text{then} \quad y_1^j \qquad (2)$$

The first-order fuzzy time series model can be extended to *p*-order model in the form

$$Y_t = (Y_{t-1} \times Y_{t-2} \ldots \times Y_{t-p}) \circ R_p(t.t-p)$$

or equivalently

$$if\ \ y_{t-1}^{i_1} \text{and}\ \ y_{t-2}^{i_1}\ \ \ldots\ \ \text{and}\ y_{t-p}^{i_1} \quad \text{then} \quad y_1^j \qquad (3)$$

# An practical example (statistical approach)

Let as consider the 514 monthly inflation observations for the forty-three years 1956-1998 (see *Figure 1*). To build a forecast model the sample period for analysis $y_1, \ldots, y_{344}$ was defined, and the ex post period $y_{345}, \ldots, y_{514}$ as validation data set.

http://neatideas.com/data/inflatdata.htm

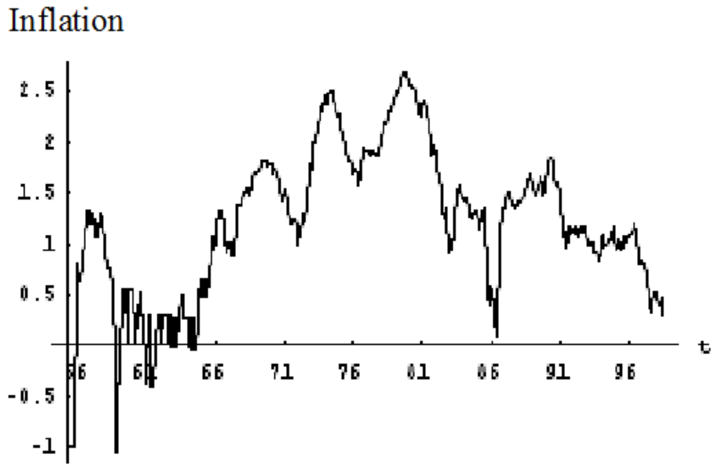



**Figure 1** Natural logarithm of monthly inflation from

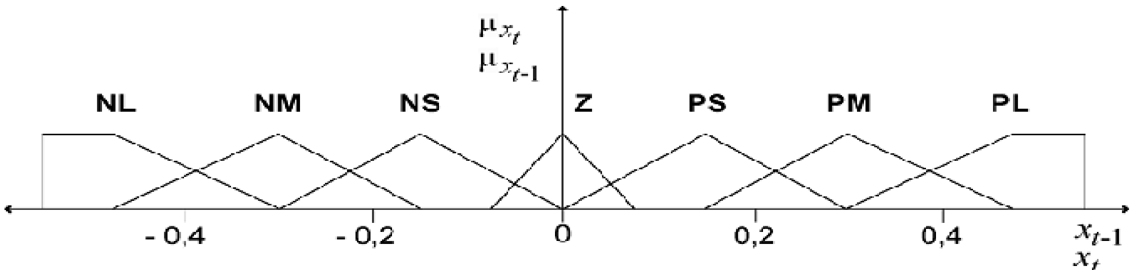
Using time series data and traditional statistical tools as the autocorrelation function (ACF), the partial autocorrelation function (PACF) and the Akaike Information Criterion the model is estimated as

$$\hat{y}_t = -0{,}1248 y_{t-1}$$

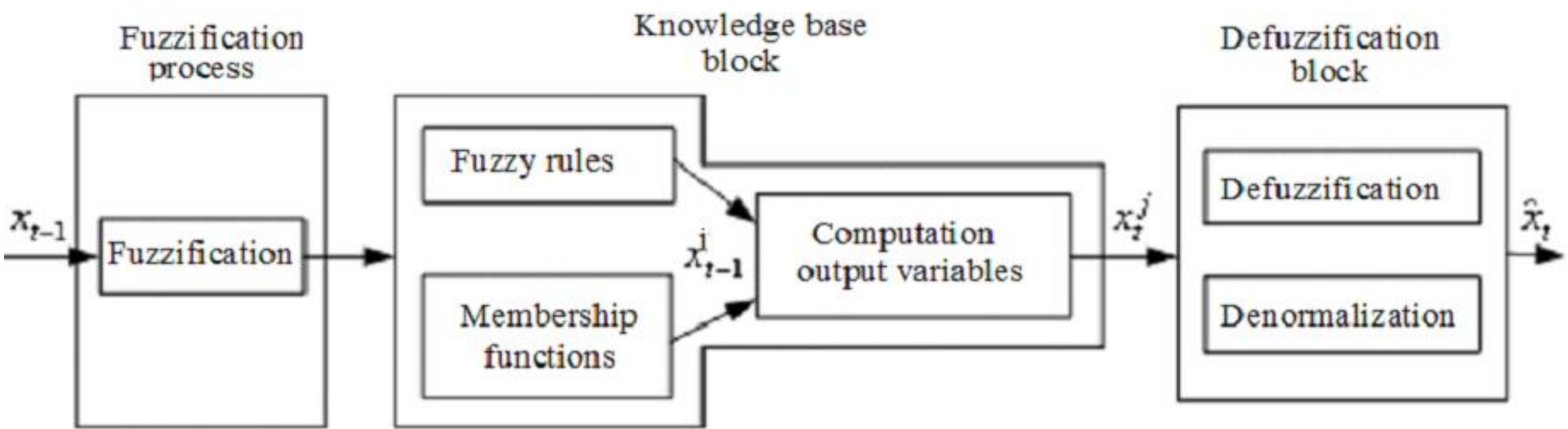
# An practical example (fuzzy time series modelling)

In the fuzzification block, we specified input and output variables. The input variables $x_{t-1}$ as $x_{t-1} = y_{t-1} - y_{t-2.}$, t=3,4,..., and output variable $x_t = y_t - y_{t-1.}$ t=2,3,...
The variablle ranges are as follows: $-0.75 \le x_t, x_{t-1} \le 0.75$

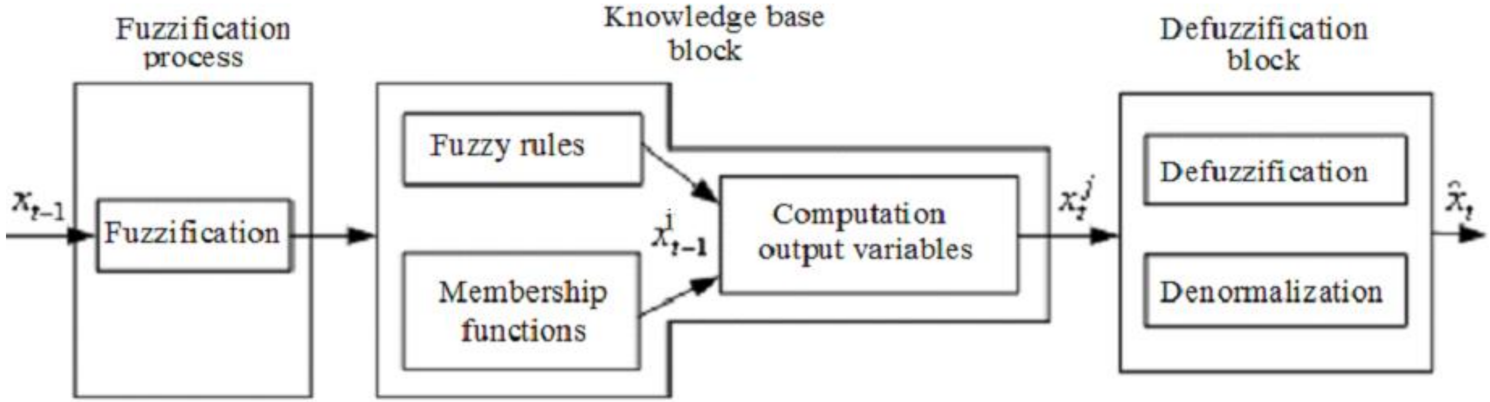
Next, we specified the fuzzy-set values of the input and output fuzzy variables. The fuzzy sets numerically represent linguistic terms. Each fuzzy variable assumed seven fuzzy-set values as follows: NL: Negative Large, NM: Negative Medium, NS: Negative Small, Z: Zero, PS: Positive Small, PM: Positive Medium, PL: Positive Large.
Fuzzy sets contain elements with degrees of membership. Fuzzy membership functions can have different shapes. The triangular membership functions were chosen.



Fuzzy membership functions of fuzzy variables



# An practical example (fuzzy time series modelling)




fuzzy-set values:

NL: Negative Large, NM: Negative Medium, NS: Negative Small, Z: Zero,

PS: Positive Small, PM: Positive Medium, PL: Positive Large.

The input and output spaces we divided into the seven disjoint fuzzy sets. From membership function graphs $\mu_{t-1}, \mu_t$ shows that the seven intervals [-0,75; -0,375], [-0,375; -0,225], [-0,225; -0,075], [-0,075; 0,075], [0,075; 0,225], [0,225; 0,375], [0,375; 0,75]

correspond to NL, NM, NS, Z, PS, PM, PL, respectively.

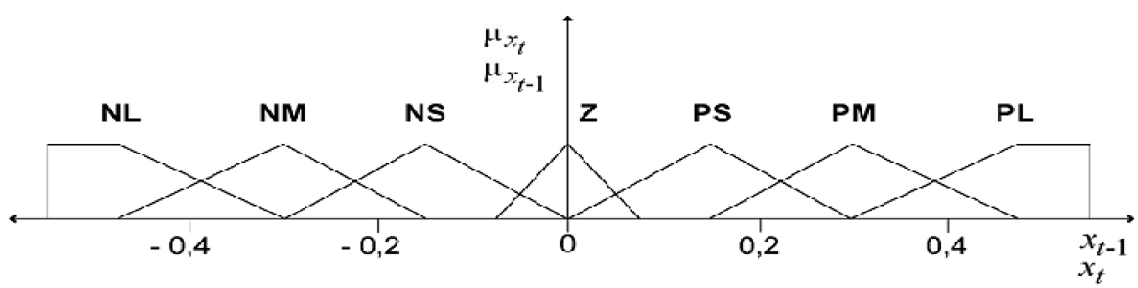



Fig. 2 Fuzzy membership functions of fuzzy variables

# An practical example (fuzzy time series modelling)

×

Next, we specified the fuzzy rule base or the fuzzy relation bank. The above specified interval $-0.75 \le x_t, x_{t-1} \le -0.75$ portioned into seven non-uniform subintervals that represented the seven fuzzy sets values NL, NM, NS, Z, PS, PM, PL assumed by fuzzy variables $x_{t-1}$ $x_t$

The Cartesian product of these subsets defines 7 x 7 = 49 fuzzy cells in the input-output product space $R^2$. These fuzzy cells equal fuzzy rules. Thus, there are total 49 possible rules and thus 49 possible fuzzy relations.

We can represent all possible fuzzy rules as 7-by-7 linguistic matrix (see next Figure ). The idea is to categorize a given set of distribution of input vector $\mathbf{x}_t = (x_{t-1}, x_t)$ into 7 x 7 = 49 klassses, and then represent any vector by the class into which it falls.

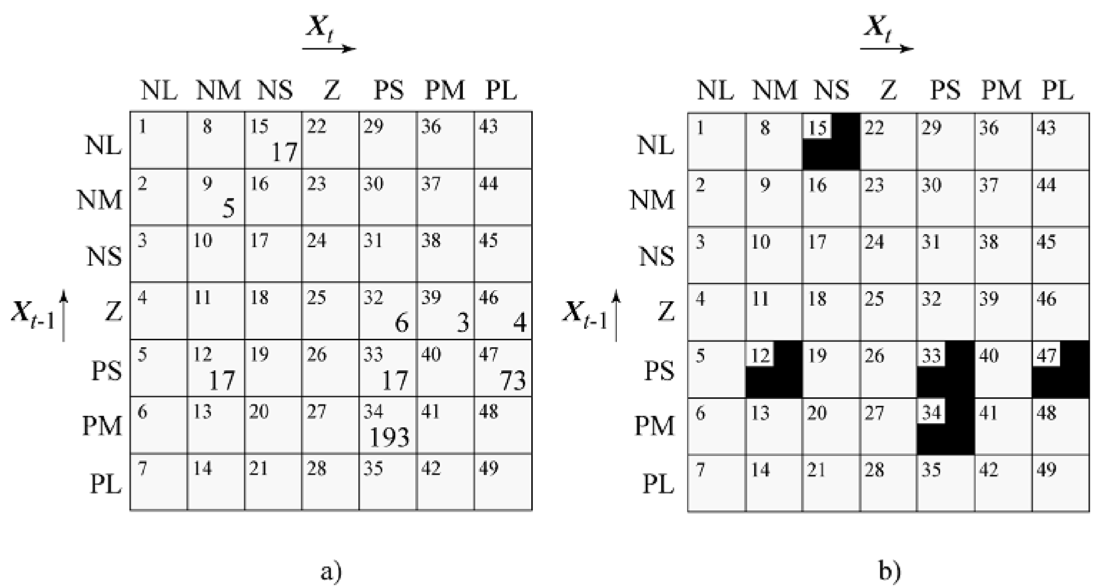
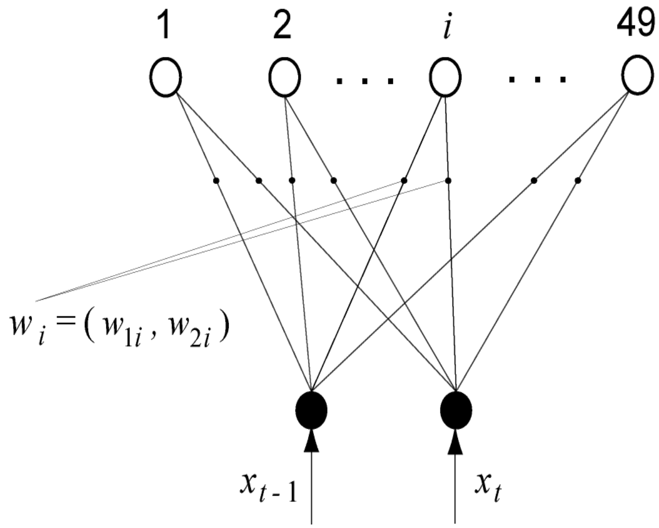



Figure 4: Distribution of input-output data $(x_{t-1}, x_t)$ in the input-output product space $X_{t-1} \times X_t$ (a). Bank of fuzzy rules of the time series modelling system (b)

# An practical example (fuzzy time series modelling)




$$\left.\begin{aligned} \tilde{w}_{1i'} &\leftarrow \tilde{w}_{1i} + \eta\ (\tilde{x}_{1t} - \tilde{w}_{1i}) \\ \tilde{w}_{2i'} &\leftarrow \tilde{w}_{2i} + \eta\ (\tilde{x}_{2t} - \tilde{w}_{2i}) \end{aligned}\right\} \quad \text{if } i = i' \qquad \left.\begin{aligned} \tilde{w}_{1i'} &\leftarrow \tilde{w}_{1i} - \eta\ (\tilde{x}_{1t} - \tilde{w}_{1i}) \\ \tilde{w}_{2i'} &\leftarrow \tilde{w}_{2i} - \eta\ (\tilde{x}_{2t} - \tilde{w}_{2i}) \end{aligned}\right\} \quad \text{if } i \neq i'$$

Where $i'$ is the winning unit defined as

$$\left\|\tilde{\mathbf{w}}_{i'} - \tilde{\mathbf{x}}_t\right\| \leq \left\|\tilde{\mathbf{w}}_i - \tilde{\mathbf{x}}_t\right\|$$

IF $x_{t-1}^{i}$ = PM THEN $x_t^{j}$ = PS

# An practical example (fuzzy time series modelling)

When the input value, say $x_{t-1}^{l} = x_{344}^{l}$ ,is applied to the model (4), the output fuzzy value $x_t^j = x_{345}^j$ can be calculated. It is possible to compute the output fuzzy value $x_t^j$ by the following simple procedure consisting of three steps:

- Compute the membership function values $\mu_{NL}(x_{t-1}), \ldots \mu_{PL}(x_{t-1})$ for the input $x_{t-1}$ using the membership functions in Figure 2
- Substitute the computed membership function values in fuzzy relations (2), (3).
- Apply the max-min composition to obtain the resulting value of fuzzy relations. $x_t^j$

Following the above principles, we have obtained the predicted fuzzy value for the inflation $x_t = x_{345}^j = 0.74933$ .

The inflation values in the output $x_t^j$ , $t = 345, 346, \ldots$ are not very appropriate for a decision support because they are fuzzy sets. To obtain a simple numerical value in the output universe of discourse, a conversion of the fuzzy output is needed. This step is called defuzzification. The simplest defuzzification scheme seeks for the value $\hat{x}_t$ that is of middle Membership in the output fuzzy set. Hence, this defuzzification method is called the Middle of Maxima, abbreviated MOM. Following this method, we have obtained the predicted value for the $\hat{x}_{345}$ = - 0.15. The remaining forecast for ex-post forecast period $t$ =346, 347, … may be generated in a similar way

# An practical example (fuzzy time series modelling)

As a final point, let us examine what has been gained by use of fuzzy time series model over an ordinary AR(1) model for the output $\hat{x}_{345}$. For this purpose, we have computed prediction limits on the one-step-ahead forecast from the AR(1) model, and fuzzy time series model. The 95 percent interval around the actual inflation value based on the statistical theory is

*x345±u1-α/2σε(1+φ12)1/2*=(−0.0442; 0.05043)

where $\hat{x}_{345}$ represents the forecast for period $t$ = 345 made at origin $t$ = 344, $u_{1-\alpha/2}$ is a 100(1-$\alpha$ / 2) percentage of the standard normal distribution, and $\hat{\sigma}_{\varepsilon}$ an estimate of the standard deviation of the noise. An intuitive method for constructing confidence intervals for fuzzy time series model is simply the defuzzification method First of Maxima and First of Minima to obtain the prediction limits on the one-step-ahead forecast. In our example, the "confidence" interval for fuzzy time series value $\hat{x}_{345}$ = 0.00312 is (-0-30256 to 0.3088). The actual value for the AR(1) model does not fall within the forecast interval, and moreover, its sign is opposite to the forecast value sign.

**Thank you**