Time Series Analysis by Fuzzy Linear Regression

Richard POSPÍŠIL (richard.pospisil@upol.cz)
Faculty of Arts, Palacký University in Olomouc, Czech Republic

Bogdan OANCEA[1] (bogdan.oancea@faa.unibuc.ro)
Faculty of Business and Administration, University of Bucharest,

Miroslav POKORNÝ (miroslav.pokorny@vsb.cz)
Faculty of Electrical Engineering and Computer Science, VSB - Technical University of Ostrava, Czech Republic

Gabriela ŠTEFANOVÁ (gabriela.stefanova@zssenicenh.cz)
Elementary school Senice na Hané, Czech Republic

ABSTRACT

Fuzzy set theory constitutes the theoretical background for abstractly formalizing the vague phenomenon of complex systems. Vague data are defined herein as specialized fuzzy sets, i.e., fuzzy numbers, and a fuzzy linear regression model is described as a fuzzy function with such numbers as vague parameters. We applied a generic algorithm to identify the associated coefficients of the model, and provide both analytically and graphically, a linear approximation of the vague function, together with description of its potential application. We also provide an example of the fuzzy linear regression model being employed in a time series with economic indicators, namely the evolution of the unemployment, agricultural production, and construction between 2009 and 2011 in the Czech Republic. We selected this period since it represents the period when the financial and economic crisis started, and a certain degree of uncertainty existed in the evolution of economic indicators. Results take the form of fuzzy regression models in relation to variables of the time-specific series. For the period 2009-2011, analysis confirmed assumptions held by the authors on the seasonal behaviour of such variables and connections between them. In 2010, the system behaved in a fuzzier manner; hence, relationships between variables were vaguer than otherwise, brought about by factors such as difference in the elasticity of demand, state interventions, globalization, and transnational impacts.

Keywords: *fuzzy set, fuzzy linear regression, genetic algorithms, time series*
JEL Classification: *C22, C51, C49, C65*

1. Coressponding author

1. INTRODUCTION

Regression models are used in engineering practice wherever there is a need to reflect independent variables together with the effects of other unmeasured disturbances and influences. In classical statistical regression, it is assumed that the relationships between dependent variables and independent variables of the model are well-defined and sharp. Although statistical regression has many applications, problems can occur under the following circumstances: the number of observations is inadequate (a small data set); there are difficulties verifying distributional assumptions; vagueness affects connections between input and output variables; ambiguity exists surrounding events or the degree to which they occur; inaccuracy and distortion are introduced by linearization (Shapiro, 2005). In real-world applications, these conditions are non-specific and vague. This is particularly true when modelling complex systems that are difficult to define or measure or where a human element is incorporated into the model.

Fuzzy set theory constitutes the theoretical background for abstract formalization of the vague phenomenon of complex systems. In this paper, we define vague data as specialized fuzzy sets, i.e., fuzzy numbers, and the fuzzy linear regression model is defined as a fuzzy function with such numbers as vague parameters. Estimating the uncertainty of a regression model by applying a fuzzy approach does not require adherence with the presumptions of the classical statistical regression.

Fuzzy regression analysis was performed based on the analysis of the time-specific series (hereinafter referred to as “time series”) containing selected macroeconomic variables, which could be of a seasonal character in relation to the national economy. These include indicators of construction production (CPT), agricultural production (APT) and the rate of unemployment (UNT) in the Czech Republic from 2009 to 2011. We selected this specific period since it represents the moment when the financial and economic crisis started, and a certain degree of uncertainty was recorded in the evolution of economic indicators. The data sets used in this study were provided by the Czech Statistical Office (2002).

The choice of macroeconomic variables was based on their seasonal character and interrelationships. While construction and agricultural yield increase during the period of spring to autumn, the level of unemployment in this period generally decreases, although they may evolve in the same period differently; numerous factors affect this, e.g., variation in elasticity of demand for construction and agricultural yield, various levels and forms of state intervention in these segments of the national economy, and the influence of foreign trade and globalization. Along with these forms of the trajectories of variables pertaining to construction and agricultural production, the unemployment rate may not behave completely normally. The cause of this phenomenon can be seen *inter alia* in the limited elasticity of labour supply, the strong influence of trade unions and the entire labour-related system of social security, which in aggregate distort the labour market. The rest of the paper is organized as follows. Section 2 briefly presents the main concepts used by the fuzzy regression analysis, and section 3 is dedicated to the identification of the fuzzy regression model. Next, we presented the results of the fuzzy regression analysis for the selected time series together with some comments on them in section 4 and 5. The paper ends with a section of conclusions.

2. FUZZY REGRESSION ANALYSIS

The ordinary linear regression model of the investigated system (Seber and Lee, 2003) comprises a linear combination of values of its input variables, as in equation (1) below:

$$Y = A_0 x_0 + A_1 x_1 + \cdots + A_n x_n = \sum_{i=0}^{n} A_i x_i \qquad i = 0,1,\ldots,n \qquad [1]$$

where $(x_0,\ldots,x_n)$ are input (independent) variables ($x_0 = 1$), $(A_0, A_1,\ldots,A_n)$ are the regression coefficients and Y is the output (dependent) variable.

The conventional regression model assumes that system characteristics are defined as crisp and precise, and deviations between the observed and estimated values of the dependent variable stem from observational errors. The origin of a deviation between the observed and estimated value for the dependent variable may not be significantly caused by poor local variables of the system structure. The causes of these variations do not align with the very sharp nature of system parameters. Such fuzzy phenomenon must also be reflected in the fuzziness of the corresponding parameters of the model.

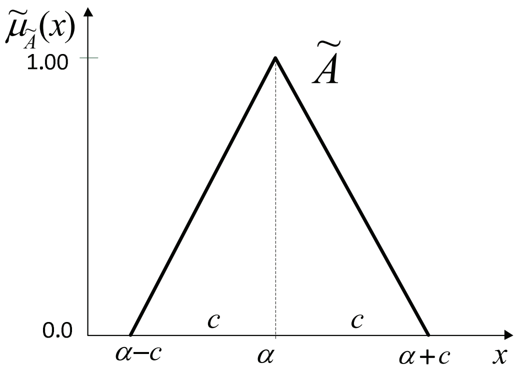
The evolution of the indeterminate regression model occurs through development of the model of vagueness, using the formalization of uncertainty rather than numerical intervals (Ishibushi and Tanaka, 1990; Poleshchuk and Komarov, 2012). Regression models reflecting the vagueness of the modelled systems are called fuzzy regression models (Buckley et al., 2008), (Heshmaty and Kandel, 1985), (Kacprzyk and Fedrizzi, 1992), (Pokorný, 1993), (Shapiro, 2005), (Tansu, 2012). The indeterminate nature of the fuzzy regression model is represented by the fuzzy output values $\tilde{Y}$ and the fuzzy regression coefficients $\tilde{A}$ in the form of specialized fuzzy sets, i.e., fuzzy numbers (Negoita, 2000), (Novák et al., 1999). The form of the fuzzy linear regression model is given by equation (2) below:

$$\tilde{Y} = \tilde{A}_0 x_0 + \tilde{A}_1 x_1 + \dots + \tilde{A}_n x_n = \sum_{i=0}^{n} \tilde{A}_i x_i \qquad i = 0,1,\dots,n \qquad [2]$$

where $(\tilde{A}_0, \tilde{A}_1, \dots, \tilde{A}_n)$ are fuzzy regression coefficients (fuzzy numbers); the fuzzy number $\tilde{A}$ is defined by its triangular shape membership function $\mu_{\tilde{A}}(x)$ (see figure 1);

Triangular membership function of fuzzy number $\tilde{A}$

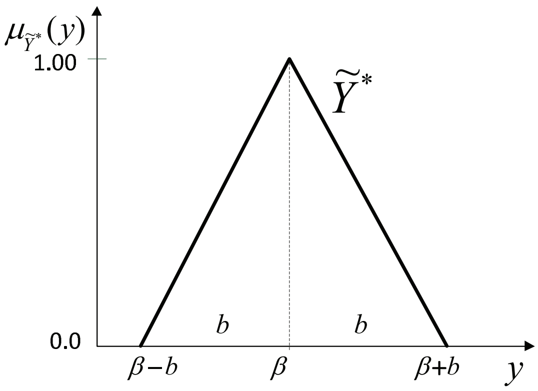
Figure 1



where α is the mean value (core) of fuzzy number $\tilde{A}$ and c is half of the width of the carrier bearing $\tilde{A}\{\alpha, c\}$. The output variable $\tilde{Y}$ of fuzzy regression model 2 is a fuzzy number defined by the triangular membership function (see figure 2). The estimated value $\tilde{Y}^*$ is defined in the form $\tilde{Y}^*\{\beta, b\}$, respectively. The observed value $\tilde{Y}^0$ is denoted in the form $\tilde{Y}^0\{y^0, d\}$ (see Figure 3);

Triangular membership function of fuzzy numbers $\tilde{Y}^*$

Figure 2



where the estimated value β is the mean value (core) of estimated output fuzzy number $\tilde{Y}^*$ and b is half of the width of the carrier bearing $\tilde{Y}^*\{\beta, b\}$.

The parameters β, d are computed applying the principles of fuzzy arithmetic (Mordeson and Nair, 2001); the mean value β is given by equation (3):

$$\beta = \alpha_0 x_0 + \alpha_1 x_1 + \ldots + \alpha_n x_n = \sum_{i=0}^{n} \alpha_i x_i \quad [3]$$

and fuzziness *b* is given by equation (4):

$$b = c_0|x_0| + c_1|x_1| + \ldots + c_n|x_n| = \sum_{i=0}^{n} c_i|x_i| \quad [4]$$

where fuzzification of the examined value y^0 is conducted via fuzzy interval d.

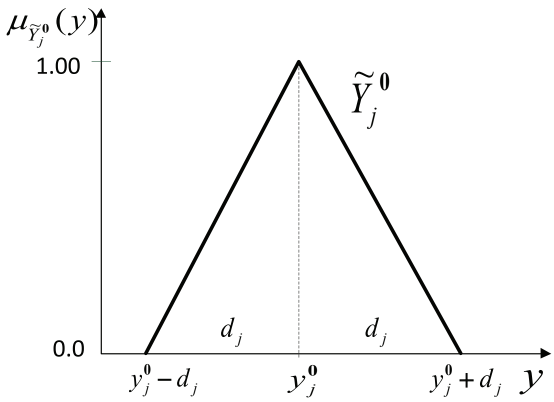
3. IDENTIFICATION OF THE FUZZY REGRESSION MODEL

3.1 Observed output variable y^0 fuzzification

We employed a version where the input variables x comprise crisp numbers, while the observed values $\tilde{Y}^0$ are triangular fuzzy numbers to define the type of the fuzzy regression model. Thus, fuzzy number $\tilde{Y}^*$ is taken as the estimated value and fuzzy number $\tilde{Y}^0$ is the observed value for the model output variable. The fuzziness d_j of observed fuzzy value $\tilde{Y}_j^0$ at step observation *j* is determined via the values gauged at steps (*j+1*) and (*j-1*), respectively (see figure 3).

Triangular membership function of fuzzy numbers $\tilde{Y}^0$

Figure 3



This means that the fuzzy number $\widetilde{Y}_j^0$ is of an unequal triangular type; values for d_j are calculated by equation (5):

$$d_j = \frac{1}{2}\left|y_{j+1}^0 - y_{j-1}^0\right| \qquad [5]$$

3.2 Computation of Fuzzy Regression Coefficients $\widetilde{A}$

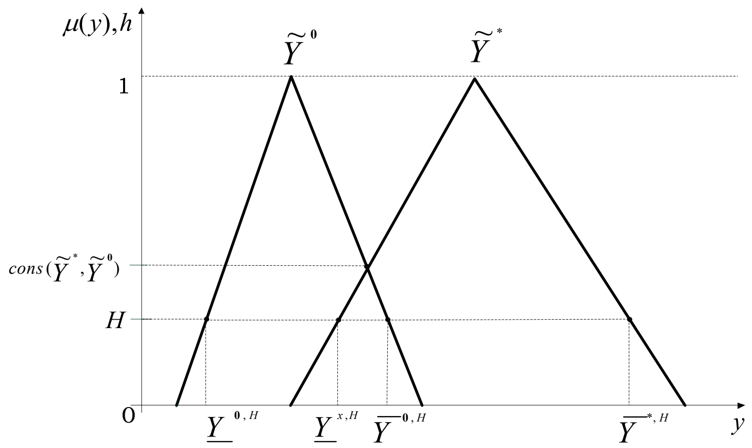
Finding values α and c as search parameters of fuzzy regression coefficients $\tilde{A}_i$ (see Figure 1) is an optimization problem.

The fit of the linear regression fuzzy model to the given data is determined through the Bass-Kwakernaak index *H* (see Figure 4) (Cetintav and Zdemir, 2013), (Kacprzyk and Fedrizzi, 1992). The adequacy of the observed and estimated values is conditioned by equation 6 – the maximum intersection (consistency) of two fuzzy sets; the estimated values for $\widetilde{Y}^*$ and the observed ones for $\widetilde{Y}^0$ must exceed the set value *H* (see Figure 4).

$$\max_y\{\mu_{\widetilde{Y}^0}(y) \wedge \mu_{\widetilde{Y}^*}(y)\} = \text{Cons}\left(\widetilde{Y}^0, \widetilde{Y}\quad\right) \geq H \qquad [6]$$

Adequacy of Linear Regression Model

Figure 4



A good estimation of $\widetilde{Y}^*$of the observed output value $\widetilde{Y}^0$ is only forthcoming if relation (6) is fulfilled.

Relation (6) is satisfied under the following conditions:

$$\underline{Y}^{,H} \leq \overline{Y}^{0,H} \quad [7]$$

$$\underline{Y}^{0,H} \leq \overline{Y}^{,H} \quad [8]$$

Considering the determined level *H,* the boundary of intervals $Y^{,H}$ and relations (3) and (4), the formulae below are derived:

$$\underline{Y}^{,H} = -(1-H)\sum_{i=0}^{n} c_i |x_i| + \sum_{i=0}^{n} \alpha x_i \quad [9]$$

$$\overline{Y}^{,H} = (1-H)\sum_{i=0}^{n} c_i |x_i| + \sum_{i=0}^{n} \alpha x_i \quad [10]$$

$$\underline{Y}^{,0} = y^0 + (1-H)d \quad [11]$$

$$\overline{Y}^{,0} = -y^0 + (1-H)d \quad [12]$$

Taking $j = 1, 2, \ldots$ and m as the number of observations, conditions (7) and (8) are formulated in final form:

$$\sum_{i=0}^{n}\sum_{j=1}^{m} \alpha_{i,j} x_{i,j} + (1-H)\sum_{i=0}^{n}\sum_{j=1}^{m} c_{i,j} |x_{i,j}| \geq y_j^0 + (1-H)\overline{d}^0 \quad [13]$$

$$-\sum_{i=0}^{n}\sum_{j=1}^{m} \alpha_{i,j} x_{i,j} + (1-H)\sum_{i=0}^{n}\sum_{j=1}^{m} c_{i,j} |x_{i,j}| \geq -y_j^0 + (1-H)\underline{d}^0 \quad [14]$$

$$c_{ij} \geq 0 \quad [15]$$

The requirement for adequacy of the estimated and observed values (6) is complemented by the need for minimal total uncertainty of the identified fuzzy regression function:

$$\sum_{i=0}^{n}\sum_{j=1}^{m} c_{i,j} \rightarrow \min, \qquad i = 0, 1, \ldots, n, \;\; j = 1, 2, \ldots, m \quad [16]$$

where $i = 1, 2, \ldots$, and n is the number of input values of the regression function, $j = 1, 2, \ldots$ and m is the number of observations.

Then, the optimization problem can be set:
a) minimization of fuzzy model vagueness (16)
b) under condition (6)

To solve the minimization problem (16) under condition (6), numerous authors have employed the linear programming method (Arabpour and Tata, 2008), (Kacprzyk and Fedrizzi, 1992). Nevertheless, a genetic algorithm method is used herein to solve this problem (16). The main reason for this is that the authors are oriented towards unconventional methods of artificial intelligence to substantiate their quality and efficiency in solving complex tasks (Oancea et al, 2021). Genetic algorithms are a representative of evolutionary methods, and their higher computational complexity is tackled nowadays by the availability of cheap high-performance computers. Thus, they are widely used to solve optimization problems and are commonly employed to identify fuzzy regression models geared towards discerning optimal fuzzy regression coefficients as triangular fuzzy numbers.

Identification of fuzzy regression coefficients - the fuzzy numbers $\tilde{A}_0, \tilde{A}_1, \ldots, \tilde{A}_n$ - was split into two tasks:

a) discerning the mean value (core) α_i of fuzzy number $\tilde{A}_i$ and

b) discerning c_i as a half of the width of the carrier bearing $\tilde{A}_i = \{\alpha_i, c_i\}$.

These tasks were solved via a genetic algorithm. Identification of α_i and c_i was carried out first. Thus, the optimization of the fuzzy linear regression model comprised a two-step process where two genetic algorithms, designated GA1 and GA2, were applied.

For identification of the mean value (core) α_i of fuzzy number $\widetilde{\mathrm{A}}_i$, minimization of the fit function J_1 is defined in the form below:

$$\min J_1 = \min \frac{1}{m} \sum_{j=1}^{m} \left(y_j^0 - \beta_j\right)^2 \qquad [17]$$

where the genetic algorithm GA1 is used. For identification of c_i as a half of the width of the carrier bearing $\tilde{A}_i$, minimization of the fit function J_2 is defined as:

$$\min J_2 = \min \sum_{j=1}^{m} \sum_{i=0}^{n} \left|c_{j,i}\right| \qquad [18]$$

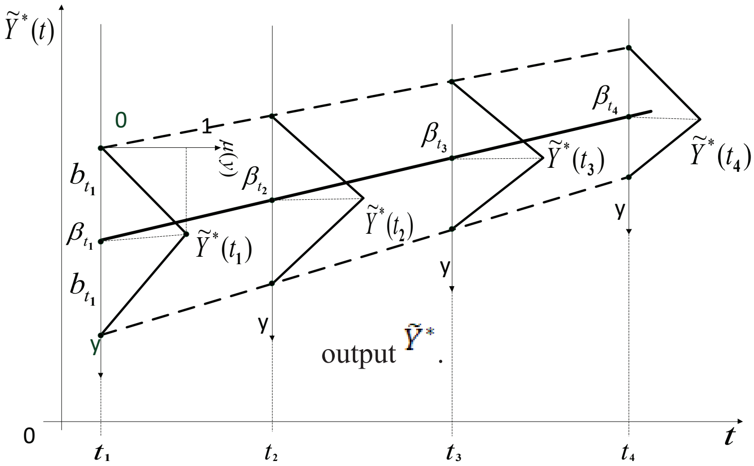
where the genetic algorithm GA2 with three constraints (13), (14), (15) is used. Minimization of the fit function J_2 is based on prior identification of the role of the mean value (core) α_i and utilizes identified values of α_i to determine the width of carrier bearing α_i.

4. ANALYSIS OF TIME SERIES FUZZY REGRESSION

The fuzzy linear regression model can express not only the analytical linear approximation of multivariate functions, but also the size of its uncertainty (vagueness, fuzziness) in the form of an indeterminate potential. Figure 5 shows the graph of a one-dimensional fuzzy regression function together with the appropriate linear approximation and the potential of the estimated fuzzy

One-Dimensional Fuzzy Linear Regression Function

Figure 5



The one-dimensional fuzzy time series regression model can express trends and seasonal cycles. Both are enhanced by the potential that defines the size of the vagueness of the model and defines the range in which the value of the trend and seasonal cycles may fall.

The one-dimensional fuzzy linear regression model of a time series trend is given by the formula

$$\tilde{Y} = \tilde{A}_0 + \tilde{A}_1 t \qquad t = 1, 2, \ldots \qquad [19]$$

The value of a seasonal deviation in every month MSD (as a fuzzy number) is calculated for each year $r = 1,2,\ldots, L$ and for each month $k = 1, 2, \ldots, 12$ as the difference between the trend value and the actual value to be estimated:

$$\text{MSD} = \left(\tilde{Y}^0_{r,k} - \tilde{Y}^*_{r,k}\right),\ r = 1,2,\ldots, L,\ k = 1, 2, \ldots, 12 \qquad [20]$$

The central value of fuzzy number MSD is calculated as the difference between the central values $\tilde{Y}^0_{r,k}; \tilde{Y}^*_{r,k}$, and the fuzziness is calculated as the sum of fuzziness of fuzzy numbers $\tilde{Y}^0_{r,k}; \tilde{Y}^*_{r,k}$.

The seasonal cycle is then defined as the time series of 12 seasonal deviations for 12 months. A seasonal deviation for a given month $k = 1, 2, ..., 12$ is calculated as the average value for the month of year $r = 1, 2, ...,$ L of the considered time series.

$$\tilde{Y}_k^* = \frac{1}{L}\sum_{r=1}^{L}\left(\tilde{Y}_{r,k}^0 - \tilde{Y}_{r,k}^*\right);\ ,\ r = 1,2,\ldots,L,\ k = 1, 2, \ldots, 12 \qquad [21]$$

For example, the seasonal variation for the first month of January is calculated as the mean of the seasonal variations for January of the considered $r = 3$ years:

$$\tilde{Y}_1^* = \frac{1}{3}\sum_{r=1}^{3}\left(\tilde{Y}_{r,1}^0 - \tilde{Y}_{r,1}^*\right) \qquad [22]$$

The values of monthly deviations are calculated as fuzzy numbers. The core of fuzzy number $\tilde{Y}_k^*$ is calculated as the mean difference between the cores, while the uncertainty is calculated as the mean of the sum of fuzziness. Thus, 12 fuzzy numbers are calculated that pass into the timeline of 12 months as a curve of the cores and their potential.

5. ANALYSIS OF SELECTED ECONOMIC VARIABLES

Modelling economic variables with high degree of uncertainty is very difficult, especially during economic and financial crisis. The variability of such variables is subject to several influences, both exogenous and endogenous, some of which are hardly predictable at all or have a prominent degree of fuzzitivity. The relative effect of non-economic influences upon the evolution of the selected economic variables has risen in importance, as various subjects on the market – households and companies – adapt their level of consumption, investment, and savings based on an uncertain future. Apart from rational evaluation of the relevant economic data, they are also under pressure from several influences from the areas of psychology, politics, demographic development, natural circumstances, foreign affairs, etc., and the so-called transactional motive is replaced with the motive of caution.

The time series used in this analysis consist in twelve measured values of the selected variables from 2009 to 2011. The analysed period was selected based on the beginning of the economic crisis, as 2009 was the first year when the crisis fully proceeded throughout the year. The selection of variables was methodically chosen with regard to the mutual interconnectivity and their relative importance in the economy. This was the reason why two primary variables from GDP were analysed (construction and agricultural production) as well as a secondary variable (unemployment), which is in causal relationship to the two previous variables. Both construction and agricultural production are variables with a highly seasonal cycle, which is mirrored in the evolution of unemployment in both directions following a delay. Simultaneously with this assumption, construction, and agricultural production, however seasonal, may act differently, caused by the obvious differences in the characteristics of these variables. While the elasticity of demand of agricultural production may be very low, it is very high for construction, thus households and companies postpone their consumption and investments for a more favourable period. Therefore, the decrease in the extent of construction works leads to increase in unemployment, yet this is not true for agricultural production. The latter is applicable to the so-called Giffen goods effect, where demand for agricultural production does not rise or fall significantly, as the effect of income fully negates the influence of substitution. Although an increase in unemployment along with a decrease in the level of agricultural production tends to be affected by the seasonal cycle, the increase in unemployment as it relates to the construction sector is usually dictated by decrease in demand for building works.

Construction and agricultural production usually increase during the spring to autumn and undergo decline in the winter. An opposite trend is shown by the unemployment rate, though, which drops in the spring to autumn, and unemployment generally reaches its highest values in the winter (seasonal unemployment). While this is purely conjecture, several certified works and empirical observations support it. Indeed, the economy could have behaved unpredictably in the crisis during the studied years (2009 - 2011). It is interesting to note the positive values for agricultural production when consumers simply cannot significantly reduce demand, and great fluctuations are evident, including negative values in the construction sector. However, for both types of production, despite the diversity, the seasonal behaviour of variables are somewhat apparent, especially if a long-term perspective is taken.

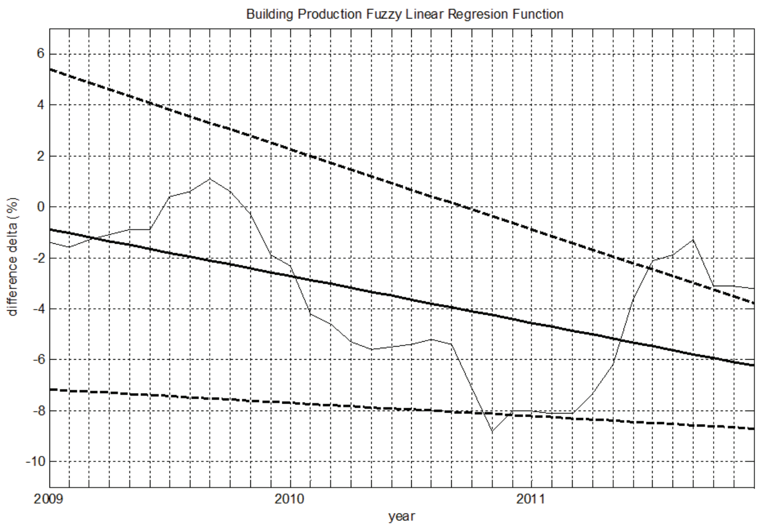
Unemployment behaves in the opposite way, since should production decline (GDP), unemployment goes up and vice versa (due to fluctuation - decrease/increase in demand). Hence, unemployment secondarily shows a seasonal character. It adheres *inter alia* to the so-called Okun law formulated in the 1960s, which states that if there is a decline in GDP by 2%, the unemployment rate grows by 1% or the proportion is approximately 2:1.

Identification of the time series fuzzy regression models was made by applying standard algorithms in Optimtoolbox MATLAB software (Matlab, 2018).

The results are shown as fuzzy regression models of the time series for construction production (CPT; figures 6 and 7), agricultural production (APT; figures 8 and 9) and unemployment (UNT; figures 10 and 11). Figures 6-11 represent their fuzzy trends and fuzzy seasonal cycles. Appropriate regression coefficients of regression functions are presented in the form $A\{\alpha; c\}$

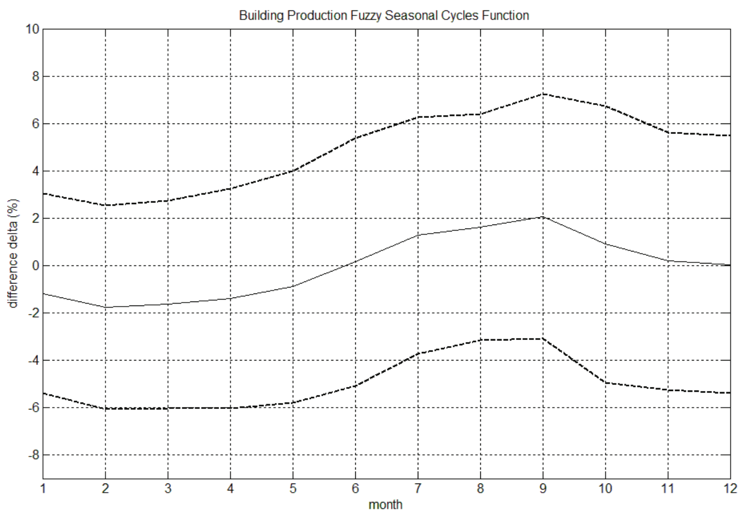
Construction Production - Fuzzy Linear Regression Function

Figure 6



Construction Production - Fuzzy Seasonal Cycles Function

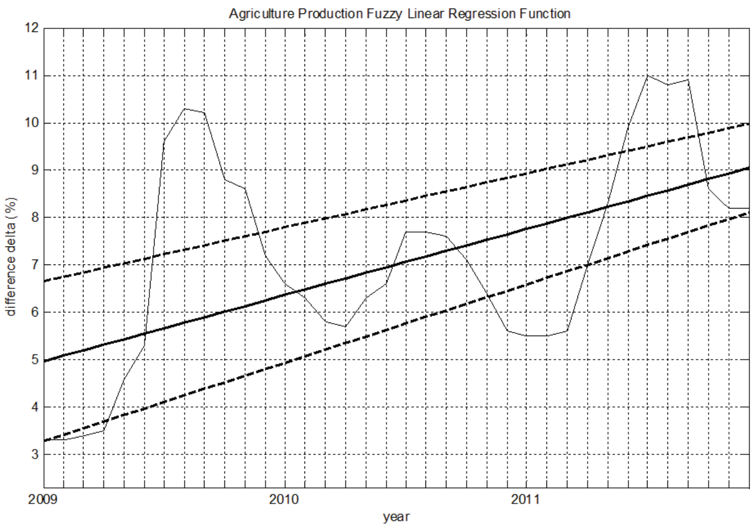
Figure 7



$$A_0\{4.9646; 0.7891\}$$
$$A_1\{0.1157; 0.0243\}$$

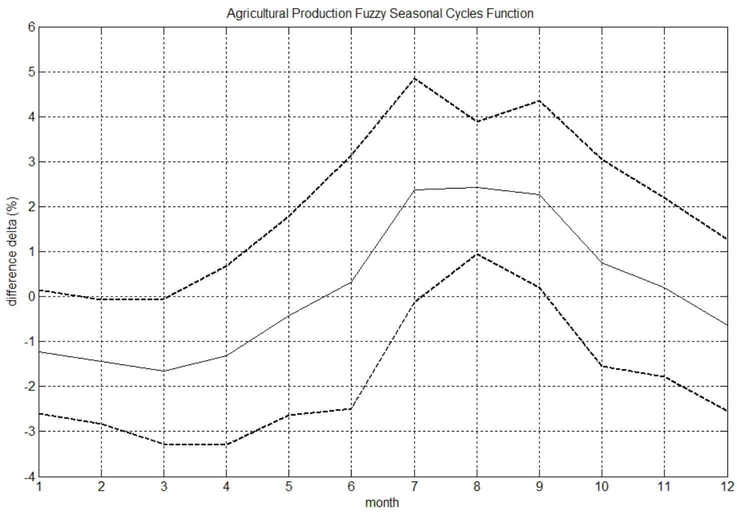
Agricultural Production - Fuzzy Linear Regression Function

Figure 8



Agricultural Production - Fuzzy Seasonal Cycles Function

Figure 9

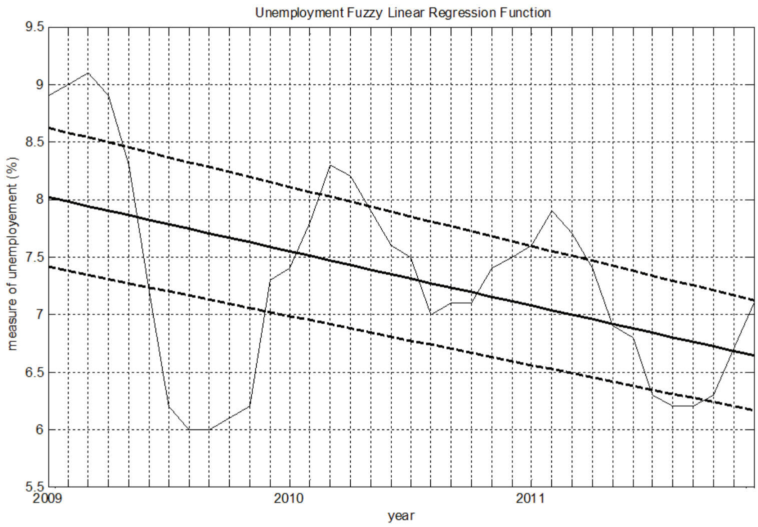


$$A_0\{7.9438; 0.5790\}$$

$$A_1\{-0.0385; 0.0057\}$$

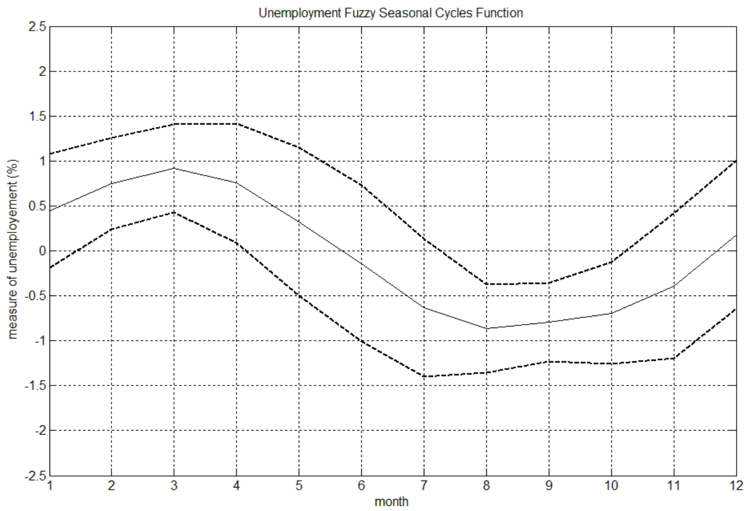
Measure of Unemployment - Fuzzy Linear Regression Function

Figure 10



Measure of Unemployment - Fuzzy Seasonal Cycles Function

Figure 11



The results of the time series analysis of the macroeconomic variables (UNT, CPT and APT) show the interdependence of some of these variables, but in some cases, they also reflect a certain degree of vagueness, i.e., fuzzitivity.

This concerns both the interdependence of CPT and APT variables in relation to UNT, and in some periods the same trend of CPT and APT (2009 and 2011) and their opposite trend in 2010. The dependencies selected above were confirmed by several scientific methods and long-term professional empirical observation (Boeckh, 2010). However, some of the assumptions mentioned in Boeckh (2010) were not confirmed by the presented work for reasons that can be satisfactorily explained.

One of these relates to market failure due to the global economic crisis. This does not refer to a state of stagnation or moderate inflation, but current so-called stagflation, which was once a relatively rare phenomenon. It is a combination of two failures of macroeconomic equilibrium, namely economic stagnation, or rather stagnation of GDP growth, and rising prices (inflation). The existence of this type of failure raises serious national economic problems with an impact on the fiscal and monetary policy of the country and an emphasis on the contradictory nature of these failures, especially the choice of current fiscal expansionary and restrictive monetary instruments of economic policy (Rubin, 2011). What also plays a role here is the global type of economy and thus limited effectiveness of measures at a national level, which is especially true for small and open economies, such as in the Czech Republic.

Another important influence on the variables UNT, CPT and APT over time is referred to as “time lag” in the economy. This is a series of delays resulting from the characteristics of an economic process based on a premise that from the moment when the problem (failure) emerged, while observing the issue with conclusive measurable economic tools (recognition lag), some time (delay) always passes; then there is a particular time interval needed for making a decision and choosing tools for correcting the failure, and a time interval is required to implement the necessary tools, including their positive effect (implementation lag). This fact significantly reduces the efficiency of economic policy and, together with its global character, fundamentally affects economic activity (Friedman, 1994). It also refers to the mutual correlation of all the variables and their existent and proven fuzzitivity.

Another significant circumstance affecting the fuzzitivity of the monitored system is distortion of the market by existing governmental and political interference. In the monitored set of variables, the APT variable is especially affected by agricultural subsidies at national and European levels and deflects the behaviour of particular economic market agents (Mises, 2009). To some extent, this also applies to the UNT variable, which is influenced, for example, by the minimum wage, state employment policy, amount of social benefits and a variety of other interventions that unilaterally deflect the labour market out of the free market. A relatively free market environment exists only through construction sector. Government interventions that tend to grow definitely increase the vagueness of the variable behaviour of the monitored system.

The intermediate effects of the crisis, which gradually changed within the three observed years from a financial to an economic crisis, also have an indubitable effect on the high fuzzitivity of the investigated system. The crisis concurrently spread slowly from individual market subjects to a crisis of public budgets and state debt crisis. This phenomenon, being much stronger in Eurozone countries than the Czech Republic, exerted an imminent influence upon foreign demand, upon which the Czech economy, being small and open, was very dependent to a certain extent. This was especially the case of agricultural production, semi-products and foodstuffs (i.e., generally APT), and to a smaller extent the export of construction materials, construction workers and investment construction units (generally CPT). Foreign influences, however, tended to follow another route. It was mostly the large import of agricultural production into the Czech Republic, where also typical and traditional agricultural products of both cattle and plant-based production were imported into the Czech Republic, as well as technical and construction materials. This import narrowed the operating space for Czech manufacturers and their supply was limited. In this context, foreign influences have the largest affect upon the variable UNT, as the free movement of the labour force is one of the freedoms of the European free market. The analysed economical areas of APT and CPT are rather less demanding in matters of the labour force qualifications; therefore, they are most affected by the tide of a foreign labour force. This feature cannot be influenced at the national level; thus, it has an imminent influence upon the growth of UNT and is one of the reasons for its high fuzzitivity.

Despite the above-mentioned facts, it is still possible to observe dependencies in the monitored variable set described in this work. In 2009 and 2011, there was a similar trend for CPT and APT seasonal cycles in the summer (from June to September), with a clear decrease in UNT, and in 2011 this trend was even stronger than in 2009. It is a well-known phenomenon of production growth (in this case, CPT and APT) with a parallel decrease of unemployment during summer, or more precisely, with the rise of unemployment during winter, which is known as seasonal unemployment. In 2010, the system behaved fuzzily with an unproven dependency of CPT and APT on UNT. The same year the cycle amplitude of APT was significantly lower than that of CPT, caused by elasticity of demand, i.e. the proportion of change in the demanded quantity and price. Elasticity of APT (agricultural production and foodstuffs) is much lower, sometimes almost zero, compared to the elasticity of CPT (private and public construction works combined), which has a high elasticity. Therefore, the trend zones and seasonal cycles of APT were significantly narrower than the ones for CPT and sustainably achieved smaller fluctuations. The same was also true for the fuzzitivity of the relation between the APT and UNT variables, which was significantly lower than between CPT and UNT.

Comparing the trend zones for CPT and APT in relation to UNT during the monitored period, the existence of the Okun law can be demonstrated, this being an empirical relationship between the cyclical movements of GDP (herein the CPT and APT variables) and UNT. The law says that if actual GDP drops towards a potential one by 2%, the unemployment rate (UNT) increases by approximately 1%. This relationship applies to the GDP (not only to the sum of CPT and APT); however, but contradictory movement of these variables is also proved by this study. While the APT trend zone sees growth (CPT relatively stagnant), the UNT trend zone decreases, i.e. when production increases, unemployment decreases. This phenomenon can be observed in the variables during the period of 2009-2011, while in the last year of the period the phenomenon shows itself most strongly.

All the variables investigated above have an immediate effect on the fiscal area of the economic policy of the state. While the level of production of actual GDP (parts of which are also APT and CPT) affects the level of tax allocation for public costs, the level of UNT affects the level and rate of their later redistribution. However, the demand for APT and CPT exerts an immediate effect on monetary matters pertaining to the economic policy of the state. Economic entities then react to an anticipated drop in inflation by trying to obtain interest-bearing assets by selling other assets. Through this, they attempt to reduce losses from holding liquid assets that they had obtained by continuous inflation. Such purchases of new assets, however, lead to rise in prices and drop in real pay-off, meaning that even an expected increase in inflation brings about reduction in the interest rate. In economic literature, this effect is called the Mundell-Tobin effect.

From the results of the time series analysis of UNT, CPT and APT herein, the authors demonstrate the interdependence of these variables, and in some points even their high fuzzitivity. This is mainly due to the global nature of the economy, protracted economic crisis, time delays, especially state interventions, and political measures, which influence the free market and national economy.

6. CONCLUSIONS

In classical statistical regression, it is assumed that the relationship between dependent variables and independent variables of a model is well defined and sharp. Although statistical regression has many applications, problems can occur in situations when the number of observations is inadequate (a small data set), or difficulties arise in verifying distributional assumptions, vagueness affects the relationship between input and output variables, ambiguity exists in events or degree to which they occur, or inaccuracy and distortion are introduced by linearization.

In this study we defined vague data as specialized fuzzy sets - fuzzy numbers and devised a fuzzy linear regression model as a fuzzy function with such numbers as vague parameters. Determining the uncertainty of the regression model via a fuzzy approach does not require that the above presumptions are met.

A genetic algorithm was applied to identify the fuzzy coefficients of the model. The linear approximation of the vague function together with its potential was presented analytically and graphically.

Several assumptions, concerning the evolution of the CPT, APT and UNT variables, their seasonality and relationship between them were proved by performing fuzzy regression analysis of the selected variables. In the first (2009) and third (2011) years of observation there was a common and seasonal growth in CPT and APT, while the increase of CPT was lesser in extent due to a higher elasticity of demand for construction, as well as to the full impact of the economic crisis in this segment of economy. The assumption that UNT dropped during the studied period along with the increase of CPT and APT was also confirmed by fuzzy regression analysis.

The fuzzy regression analysis of the time series of CPT, APT and UNT revealed the non-standard behaviour of the monitored variables in 2010. This constituted the third and most profound year of the crisis, and the full influence of the state was evident alongside the huge impact of globalization on the small and open economy of the Czech Republic. Delay certainly played a role here, emerging in the economy during the second studied year (2010). That year the system of indicators behaved fuzzily and the interdependence of CPT and APT on UNT was not proven by the model; moreover, the model behaved in a much vaguer fashion, i.e., fuzzily, in connection with CPT to UNT rather than APT to UNT. The cause of the phenomenon lay in the limited elasticity of demand for agricultural production, or, for example, in the rising price of agricultural commodities throughout the period. State intervention and transnational influences on the APT and UNT variables were so great that they could be seen as one of the causes of the non-standard and fuzzy behaviour of these variables during the year.

REFERENCES

1. **Arabpour, A.R., Tata, M.**, (2008), *"Estimating the Parameters of a Fuzzy Linear Regression Model"*, Iranian Journal of Fuzzy systems, Vol.5, No. 2, pp.1-19.

2. **Boeckh, J. A.,** (2010), *The Great Reflation, Wiley and Sons*, Inc., USA.

3. **Buckley, J.J., Jowers, L.J.** (2008), *Fuzzy Linear Regression I*. Studies in Fuzziness and Soft Computing, Vol. 22, Springer.

4. **Cetintav, B., Zdemir, F.**, (2013), *LP Methods for Fuzzy Regression and a New Approach* E.Krause, Eds. In Synergies of Soft Computing and Statistics for Intelligent Data Analysis, volume 22, Springer.

5. **Friedman, M.**, (1994), *Money Mischief: Episodes in Monetary History*, First Harvest Edition, USA.

6. **Heshmaty, B., Kandel, A.**, (1985), "*Fuzzy linear regression and its applications to forecasting in uncertain environment*", Fuzzy Sets and Systems no. 15, pp. 159-191.

7. **Ishibuchi, H., Tanaka, H.**, (1990), "*Identification of fuzzy parameters by interval regression models*", Electronics and Communications in Japan (Part III: Fundamental Electronic Science) 73.12, pp. 19-27.

8. **Kacprzyk, J., Fedrizzi, M.**, (Eds.) (1992), *Fuzzy Regression Analysis*. Studies in Fuzziness and Soft Computing, Physica-Verlag HD.

9. **MATLAB,** (2018), 9.7.0.1190202 (R2019b). *Natick, Massachusetts*: The MathWorks Inc.

10. **Mises, L.v.,** (2009), *The Theory of Money and Credit*. New Haven: Yale University Press.

11. **Mordeson, N.J., Nair, P.S.**, (2001), *Fuzzy Mathematics*. Physica-Verlag.

12. **Negoita, C.**, (2000), Fuzzy Sets, New Falcon Publications.

13. **Novák,V., Perfilieva, I., Močkoř, J.**, (1999), *Mathematical Principles of Fuzzy Logic*. Kluwer, Boston.

14. **Oancea B., Pospíšil R., Jula M.N., Imbrișcă C-I**, (2021), "*Experiments with Fuzzy Methods for Forecasting Time Series as Alternatives to Classical Methods*", Mathematics, 9(19):2517. https://doi.org/10.3390/math9192517.

15. **Poleshchuk, O, Komarov, E.**, (2012), *A fuzzy linear regression model for interval type-2 fuzzy sets*, 2012 Annual Meeting of the North American Fuzzy Information Processing Society (NAFIPS), pp. 1-5, DOI: 10.1109/NAFIPS.2012.6290970.

16. **Pokorný, M.** (1993), *Fuzzy Nonlinear Regression Analysis*. Doctoral Thesis (in Czech) VUT Brno, Faculty of Electrical Engineering, Brno.

17. **Reeves, C.R., Rowe, J.E.,** (2002), *Genetic Algorithms: Principles and Perspectives Kluwer Academic Publishers*, New York.

18. **Rubin, I.S.**, (2011), *The Politics of Public Budgeting*. Chatham: Chatham House, USA.

19. **Seber, G.A.F., Lee, A.J**. (2003), *Fuzzy Regression Analysis*. Wiley Series in Probability and Statistics.

20. **Shapiro, A.F.**, (2005), *Fuzzy regression models*. Available at: http://www.soa.org/research/arch/2006/arch06v40n1-ii.pdf.

21. **Czech Statistical Office**, (2022), https://www.czso.cz/csu/czso/home

22. **Tansu, A.,** (2012), *Fuzzy Linear Regression: Fuzzy Regression*. Lambert Academic Publishing.