NEW MODEL OF CPI PREDICTIONS USING TIME SERIES OF GDP WITH SOLVING GDP-CPI DELAY HYPOTHESIS

David Dudáš[1]*, Petr Dlask[1]

[1] Faculty of Civil Engineering; CTU in Prague, Thákurova 7, 166 29 Praha 6, Czech Republic

Abstract

The existence of correlation between economical situation of a state and its effects on construction production is generally known. Thanks to economical and legal specifications of construction field, it is possible to qualitatively forecast future development based on macroeconomic indicators. Macroeconomic indicator of GDP (gross domestic product) is used towards mentioned goal. This paper strives to quantify general relations between these values: GDP and CPI (construction production index). For this purpose, calculation will be used of correlation coefficient of time series based on statistical data of every state in the European Union after solving delay hypothesis. These action create an analytical map of EU with levels of GDP – CPI dependencies which are used in prediction algorithm. Prediction algorithm is created in relation to levels of GDP - CPI dependences and data from their times series. Algorithm is tested on creating CPI predictions for year 2012 for EU states. Output values are compared with real values from Eurostat. Differences are compared with predictions based on neural network and on random predictions with significantly better results.

Key words

Construction production index; correlation CPI and GDP**;** forecasting construction production; predictions

To cite this paper: *Dudas, D., Dlask, P. (2014). New model of CPI predictions using time series of GDP with solving GDP-CPI delay hypothesis, In conference proceedings of People, Buildings and Environment 2014, an international scientific conference, Kroměříž, Czech Republic, pp. 152-163, ISSN: 1805-6784.*

*Corresponding author: Tel.: +420-603-295-212
E-mail address: david.dudas@fsv.cvut.cz

1 INTRODUCTION

People constantly try to predict future development of various events. That is more easily done in some areas than in others. For example, a lot of effort is devoted to short-term or medium-term economic development forecasting [1]. Of course, there is always a degree of uncertainty and degree of subjectivity. Research goal of this paper is to objectively investigate a relationship between change of macroeconomic indicator GDP (gross domestic product) [2] representing economical development and change of CPI (construction production index) [3] in countries of the EU (European Union) and to predict future development based on this research. With all necessary data gathered, research focuses on analysis and logical connections between time series of these indicators. We also react to a discussion of impossibility of future development prediction based on independent relationships between GDP and construction production. The primary goal of this paper is to create an instrument to predict future values of CPI based on known past values and values of GDP. Such instrument would create predictions based on impartial technical analysis without human interference. Secondary goal is to create an analytical map of Europe with representation of uncertainty of future construction development in each individual state based on its GDP indicator.

2 LITERATURE REVIEW

The topic of predictive modelling is nowadays the area of high interest [4] as it can be widely used both in academic and industrial research. In the field of economy, there is a lot of focus aimed at predictions of GDP [5, 6, 7]. However, there has been reported a limited number of using predictive modelling and forecasting at the intersection of economy and construction industry. Especially estimating construction production on macroeconomic level and related technical analysis can be extremely useful for national and enterprise resources planning. To our best knowledge only two groups so far have focused on analysis of construction production time series [8, 9]. Abdelhamid and Everett [8] in their work explored the use of time series analysis in evaluating the results of experiments to improve construction productivity. Hwang [9] in more recent paper presented the concept of forecasting construction costs using time series indexes. We believe that merging the idea of GDP forecasting and using construction production time series analysis to investigate the relationship between construction production and economic performance of the region can be advantageous and can bring even more insight to this emerging interdisciplinary research field.

3 DATA SOURCE

We obtain necessary data - time series of GDP and CPI from free database Eurostat [10] which gathers data from individual states of the European Union or from the EU as a single unit [11]. It should be noted that this paper focuses only on states of the European Union and not on European states in general. European states outside the EU (e.g. Norway, Switzerland) share specific aspects, followed naturally from their not being in the EU, which could distort final results. Methodology of gathering data for the database is under Eurostat management, and it is available for inspection at their web pages, see references.

Two time series were chosen for further analysis: *Construction production index - annual data - percentage change* [12] and *Real GDP per capita, growth rate and totals* [13]. Both tables are modified to represent percentage change in index compared to previous year, see table 1.

Tab. 1: Example of data from Eurostat database representing whole EU – change in index (%) of GDP and CPI. (source: Eurostat)

	1996	1997	1998	1999	2000	2001	2002	2003	2004	2005	2006	2007	2008	2009	2010	2011	2012
GDP	1,7	2,6	2,8	2,8	3,6	1,9	1	1	2,1	1,6	2,9	2,8	-0,1	-4,6	1,8	1,4	-0,5
CPI	-2,6	-0,6	0,5	3,1	2,8	0,9	0,6	2,0	0,9	2,5	3,3	2,6	-2,9	-7,8	-3,5	1,3	

See figure 1, there is an example of two time series from the database presented in line chart (data from the whole EU). We can see that there is an obvious similarity in tendency of both series. However, the goal is to analyse this similarity using statistical methods [14].

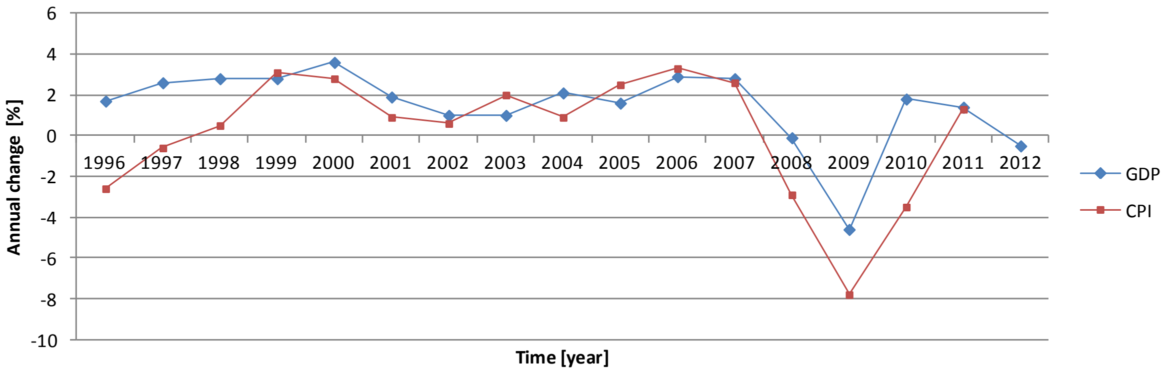


Fig. 1: Line chart of time series of change in GDP and change in CPI (data from whole EU) (source: Eurostat)

4 CORRELATION OF TIME SERIES

This section focuses on methods of analysis of relationship between development of GDP and CPI. The dependency rate between these values is represented by correlation coefficient. Correlation coefficient is a dimensionless number. Its values are expected to range from zero to one, numbers near value one represent strong dependence and numbers near zero represent low dependence or independence. Values from minus one to zero, representing negative dependence, are possible but not expected. For calculation we use formula for correlation (1) and covariance (2) [14].

$$R(i,j) = \frac{C(i,j)}{\sqrt{C(i,i)C(j,j)}} \tag{1}$$

$$C(i,j) = E\{[i - E(i)][j - E(j)]\} \tag{2}$$

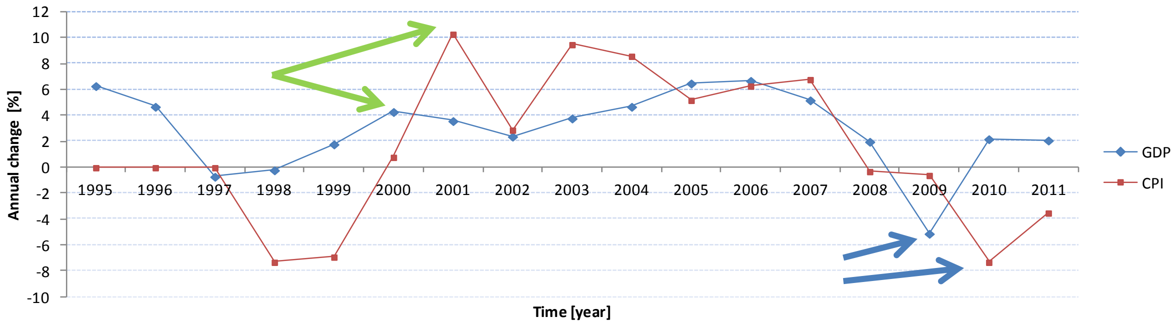
Where i and j are time series vectors, R is correlation coefficient and C is covariance coefficient.

According to general principles, we can interpret values of correlation coefficient R as follows [15]:

$R \leq 0.36$ – low or weak correlations (dependences)
$R \in (0{,}36;\ 0{,}67)$ – modest or moderate correlations (dependences)
$R \geq 0{,}67$ – strong or high correlations (dependences)

4.1 Time shift (lag) of time series – delay hypothesis

For purpose of analysis using correlation coefficient, there is a need to analyze time series without any time shift (lag) [16]. Original hypothesis was that there would be a certain time shift in all GDP-CPI time series. Assumption was that there would be at least a year's delay of construction production when compared to GDP. This conclusion is based on logical reasoning of known characteristics of construction industry: long term investment, long term planning, and backlog of work for construction company, etc. [17] therefore long reaction time of industry development to economical events (e.g. economical growth or economical crisis). This idea is also supported by visualized data from some countries (e.g. Czech Republic), figure 2. This hypothesis needs to be tested.



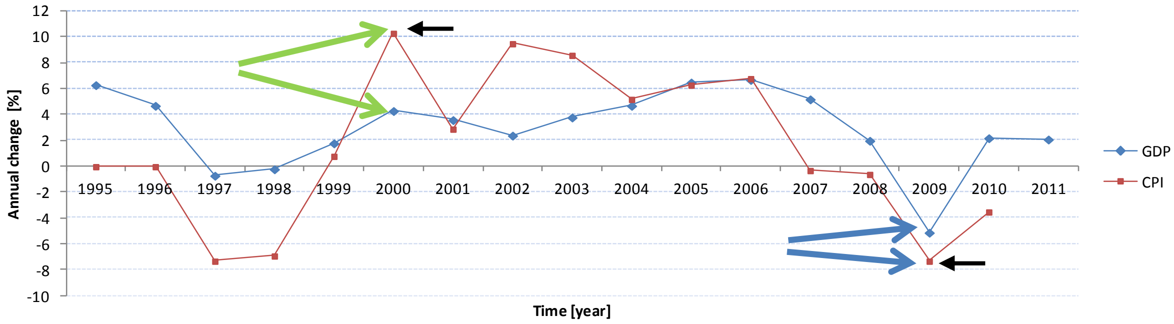


Fig. 2: Time series with time shift (top) and compensation of time shift (bottom). Example of data Czech Republic. (source: author)

We calculate correlation coefficients for time series of GDP and CPI with and without artificial time shift (which was meant to compensate for the originally present time lag). This is done in case of ten randomly chosen EU states. Time shift was set to one year, see figure 2. Results are shown in table 2.

Tab. 2: Correlation coefficient of time series with and without compensation of a time shift (source: author)

	Correlation coefficient	
	Without time shift - y	With 1 year shift - z
Belgium	0,1176	-0,0007
Germany	0,3546	0,2696
United Kingdom	0,7207	0,0686
Czech Republic	0,5449	0,7105
Poland	0,7309	0,5554
Spain	0,7460	0,7131
Sweden	0,6159	0,3319
Finland	0,8743	-0,0049
Slovakia	0,6915	0,3550
Portugal	0,3760	0,3579
Average	**0,5772**	**0,3356**

Results indicate that the original hypothesis is wrong and that time series have more significant correlation without time shift compensations. Question is whether this result can be applied for every state of interest (the remaining EU states). We confirm that by using a pair t-test [18]. Result of this test proves that our original assumption is not correct at a significance level of 5% and that our time series have higher correlations without any time shift. This result does not fit data from the Czech Republic, but t-test shows this discrepancy as insignificant in the big picture. Therefore, in further calculations we use the correlation without time shift compensations.

4.2 Correlation coefficient of change in GDP and CPI time series

This paper uses time series of GDP and CPI change of every state of EU, from the starting point of circa twenty years ago. Data for calculations is represented by two vectors of time series GDP and CPI from one country of the EU. These vectors share the same length, as appropriate when the same stretch of time is considered. Calculation is done according to formula (1) and (2). Table 3 shows the calculated correlation coefficients.

Tab. 3: Correlation coefficients of change in GDP and in CPI in EU countries (source: author)

Belgium	0,1176	France	0,5618	Austria	0,2550
Bulgaria	0,7427	Italy	0,7223	Poland	0,7309
Czech Republic	0,5449	Cyprus	0,9084	Portugal	0,3760
Denmark	0,3501	Latvia	0,8779	Romania	0,7029
Germany	0,3546	Lithuania	0,9104	Slovenia	0,7200
Estonia	0,8244	Luxembourg	0,2504	Slovakia	0,6915
Ireland	0,8102	Hungary	0,5897	Finland	0,8743
Greece	0,5083	Malta	0,5346	Sweden	0,6159
Spain	0,7460	Netherlands	0,5236	United Kingdom	0,7207

We can see the dependency rate between GDP and CPI in individual states using the correlation coefficient in table 3. By the definition of correlation coefficient [15], we can assume that there is a possibility of prediction of one time series based on a development of the other one in countries with high correlations. The question is which one of them is the cause and which one is the effect. This question was answered in other paper with the following conclusion: „GDP tends to lead the construction flow, not vice versa“[19]. Combined with this conclusion, we can assume that there is a possibility of prediction of CPI development based on data from GDP. This idea is discussed further in the next section.

For better interpretation of results, an analytical map was created to illustrate the rate of correlation in individual states of European Union. The rate of correlation is related to color scale in the map, for example see figure 3. The whole map is featured in appendix 1.

Where correlation is low, the development of construction production is believed to be run by different causes than GDP. Those may be: political interference, impact of international trade or atypical features of construction industry in the country [20]. These ideas should be subjected to further research, as they reach beyond the scope of this paper.

4.3 Prediction of CPI change based on GDP development

For prediction of development in construction production we use data which state the development of GDP. For calculations of coefficients in previous chapter we used time series of GDP up to year 2011. Now we are using data of GDP including also years 2012, 2013 (prediction) and 2014 (prediction).

Prediction algorithm uses Monte Carlo method [21]. After simulation, using pseudorandom number generator, of future development of construction production (separately for years 2012, 2013 and 2014), correlation coefficient between GDP and construction production is calculated with new simulated values. This new coefficient is compared to the original coefficient in table 3. If it fits confidence interval [22] (formula (3)), predicted values are noted. This algorithm gives us a set of possible future developments. This set is transformed to three values: maximum value, minimum value and mean value. Diagram of algorithm is presented in figure 4.

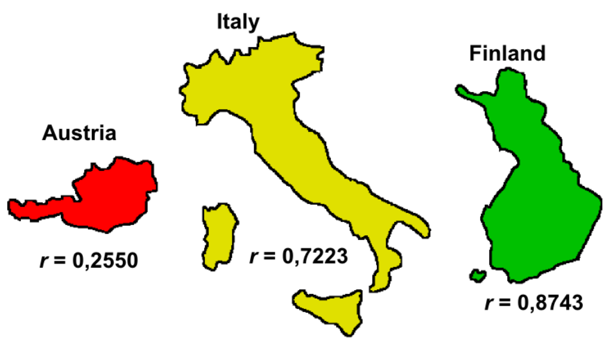


Fig. 3: Example of an analytical map with correlation coefficients of individual countries. More in appendix 1 (source: author)

$$R_2 \in (R_{1min}\,; R_{1max}) \tag{3}$$

Confidence interval (R_{1min};R_{1max}) of correlation coefficient R_1 is calculated using Fishers transformation [23]; therefore, formula (4) is implied.

$$R_{1max} - R_1 \neq R_1 - R_{1min} \tag{4}$$

Confidence interval has a significance level α = 0,45; therefore, predicted result are distributed as it is suggested in figure 5.

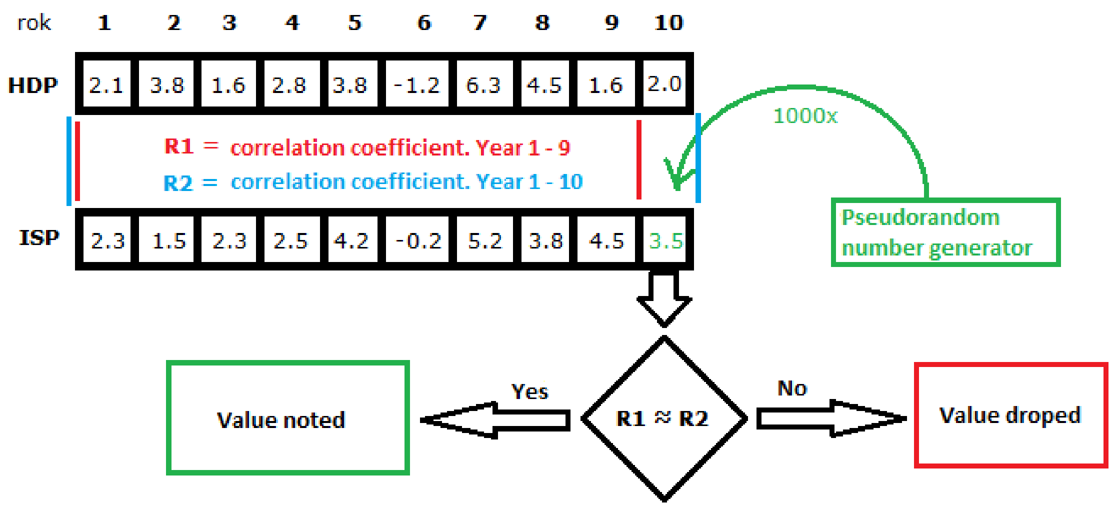


Fig. 4: Algorithm of predictions based on correlation coefficient (timeline is approximate) (source: author)

Finland was chosen as an example for prediction of future development in CPI. Reason of choosing Finland is its high correlation coefficient (R = 0,8743). Though the correlation coefficient is high, there is an inaccuracy. That is the reason why results are distributed according to the diagram in figure 5. See predictions in figure 6.

There are different scenarios of development (maximum, minimum and mean values) of CPI for 2012 at figure 6. Prediction for the following period, year 2013, must be based on a specific scenario from previous periods.

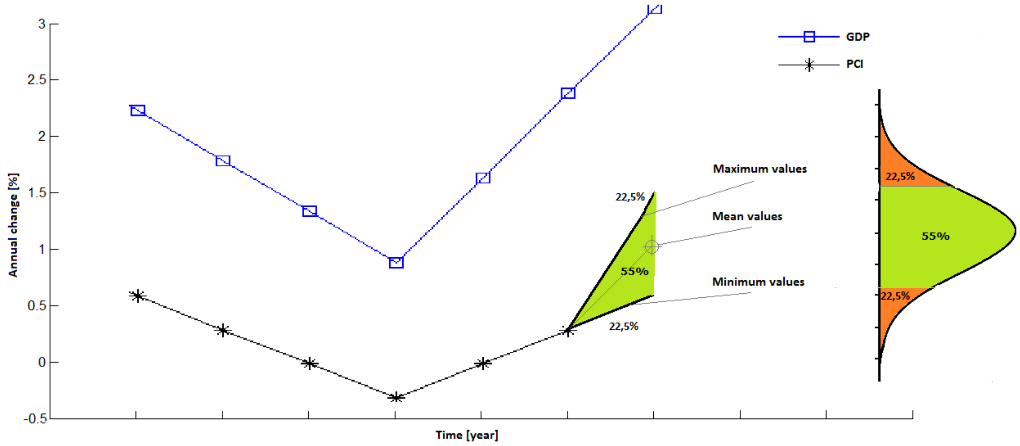


Fig. 5: Diagram of results distributions (source: author)

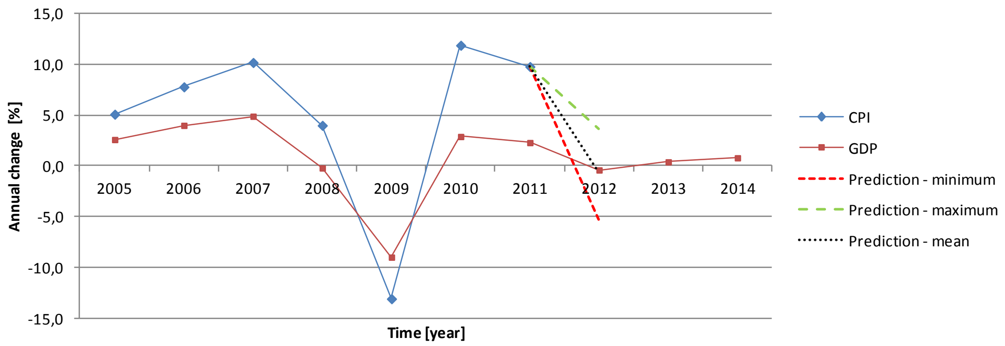


Fig. 6: Predictions of change in CPI for Finland year 2012 (source: author)

It is important to say about these predictions that in case of using presented algorithm it is necessary to consider scenarios from previous prediction period. Otherwise, it may lead to misinterpretation of results. See figure 7. Future predictions based on correlation coefficient have tendencies to compensate for past extreme values. There are predictions for Finland in 2013 as an example in appendix 2.

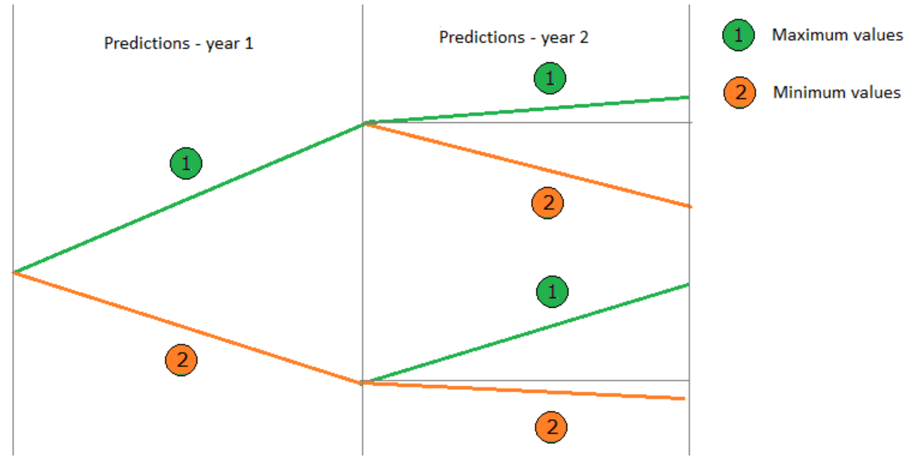


Fig. 7: Presentation of possible compensations in future predictions two years ahead (source: author)

Correlation coefficient has a significant role in predicting result. A similar prediction was calculated for Austria, which is a country with low correlation coefficient (R = 0,2550). Prediction for 2012 is illustrated in figure 8. It is obvious that the trend of construction production does not match development of GDP, and future tendencies are quite uncertain.

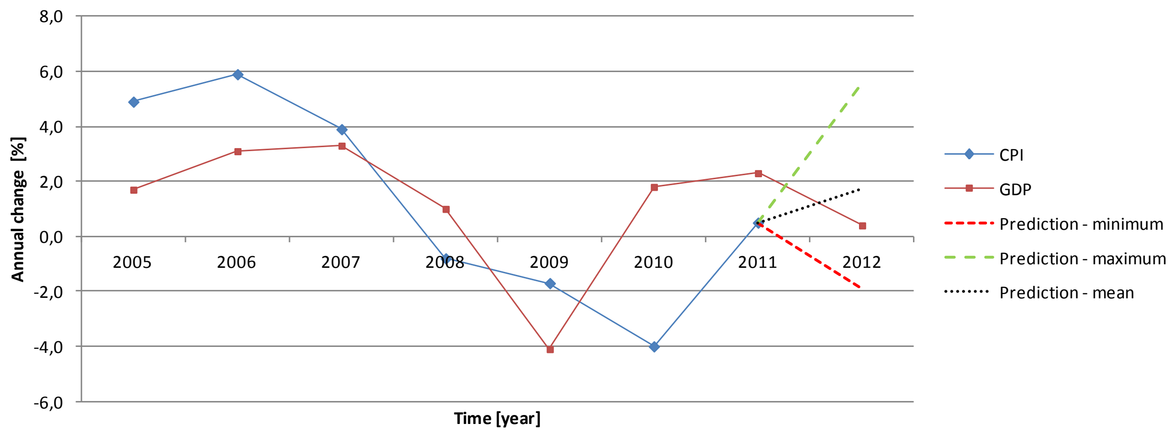


Fig. 8: Prediction of CPI for Austria 2012 based on low correlation coefficient (source: author)

5 DISCUSSION

To discuss quality of prediction algorithm and predicted values, predictions are created for year 2012 (mean values). Predictions are created only for EU states with strong CPI – GDP correlation ($R > 0{,}670$). These values are compared to real CPI values from Eurostat for 2012. See table 4.

Mean and standard deviations from absolute value of difference between prediction and real values are calculated with results of 7,71 for mean deviation and 5,47 for standard deviation.

These values are compared with completely random predictions and predictions based on neural network. Results are shown in table 5.

Predictions based on neural network were created using Feedforward[24] neural network structure with Back–propagation learning algorithm [25] specifically *Sequential order incremental training w/learning function* with 2000 learning cycles. Neural network processed the same data input information as presented algorithm.

Tab. 4: Prediction of CPI for Austria 2012 based on low correlation coefficient (source: author)

State	Eurostat	Presented algorithm	State	Eurostat	Presented algorithm
Bulgaria	-0,50%	5,15%	Lithuania	-7,20%	-0,64%
Czech Republic	-7,40%	-2,95%	Hungary	-6,70%	-5,10%
Estonia	18,40%	0,61%	Netherlands	-8,10%	-4,56%
Ireland	-5,40%	-9,61%	Poland	-4,70%	-0,37%
Spain	-5,40%	-8,23%	Slovenia	-16,90%	-9,75%
Italy	-13,90%	-0,87%	Slovakia	-12,20%	-5,15%
Cyprus	-22,10%	-5,66%	Finland	-1,50%	-0,47%
Latvia	14,10%	5,07%	United Kingdom	-7,70%	-5,95%

Tab. 5: Comparison of prediction results (source: author)

	Presented algorithm	Neural network	Random predictions
Mean error	7,71	8,68	14,13
Standard deviation of error	5,47	6,75	7,37

Results show that presented algorithm reaches more accurate results than other two alternatives.

6 CONCLUSION

This paper investigates correlation between development of GDP and construction production over a period of the last 20 years. It finds out that correlation between these time series must be calculated without time shift, in spite of original assumptions. It proves that this principle works for the whole Europe with statistically insignificant exception of a few states.

The paper demonstrates a table of correlation coefficients of all states of European Union. These results are presented in an analytical map with colour scale. This map helps to decide if there is a possibility of predicting future construction production based on development of economy represented by GDP. On the map, we can see large areas with low correlation e.g. Benelux, Germany, Austria etc. They represent quite unpredictable development in these areas and it should be a subject of further research.

An algorithm was created for forecasting future development of construction production based on development of GDP. This idea is projected onto the example of Finland. Calculated results from other countries are compared with real values from Eurostat and compared with predictions made by using neural network and random forecasting. Presented algorithm shows significantly better results.

ACKNOWLEDGEMENT

This work was supported by the Grant Agency of the Czech Technical University in Prague, grant No. SGS14/016/OHK1/1T/11.

REFERENCES

[1] Young, A. H. Reliability and Accuracy of the Quarterly Estimates of GDP (1993). *Survey of Current Business*. October, 29-43.

[2] Yamarone, R. *The trader's guide to key economic indicators* (2012). Hoboken: John Wiley & Sons.

[3] Eurostat (2013) *Eurostat Home*. Available at: http://epp.eurostat.ec.europa.eu/statistics_explained/index.php/Construction_production_(volume)_index_overview (Accessed: 9 Jan 2013).

[4] Maddalli, G a Kajal Lahiri (2009). *Introduction to econometrics*. 4thed. Chichester, U.K.: Wiley.

[5] Baffigi, A., Golinelli, R. and Parigi, G., Bridge (2004). Models to forecast the euro area GDP. *International Journal of Forecasting*, **20**, pp. 447-460.

[6] Haubrich, J. and Dombrosky, A. (1996). Predicting Real Growth Using the Yield Curve. *Economic Review*. Q1: 26-35.

[7] Rünstler, G., Barhoumi, K., Benk, S., Cristadoro, R., Den Reijer, A., Jakaitiene, A., Jelonek, P., Rua, A., Ruth, K. and Van Nieuwenhuyze, C. (2009). Short-term forecasting of GDP using large datasets: a pseudo real-time forecast evaluation exercise. *Journal of Forecast*ing. **28**, pp. 595–611.

[8] Abdelhamid, T. and Everett, J. (1999). Time Series Analysis for Construction Productivity Experiments. *Journal of Construction Engineering and Management*, **2**, pp. 87–95.

[9] Hwang, S. (2011). Time Series Models for Forecasting Construction Costs Using Time Series Indexes. *Journal of Construction Engineering and Management*, **9**, pp. 656–662.

[10] Eurostat (2013) Eurostat Home. Available at: http://epp.eurostat.ec.europa.eu /Accessed 9 Januar 2013.

[11] Eurostat (2013) Introduction. Available at: http://epp.eurostat.ec.europa.eu/portal/page/portal/about_eurostat/introduction (Accessed: 9 Januar 2013).

[12] Eurostat (2013) Construction production index - Data Explorer. Available at: http://appsso.eurostat.ec.europa.eu/nui/show.do?dataset=sts_coprgr_a&lang=en (Accessed: 9 Januar 2013).

[13] Eurostat (2013) GDP - Data Explorer. Available at: http://epp.eurostat.ec.europa.eu/tgm/table.do?tab=table&plugin=1&language=en&pcode=tsdec100 (Accessed 9 Januar 2013).

[14] Riley, K., Hobson, M., and Bence, S (2006). *Mathematical methods for physics and engineering*. 3rded. Cambridge: Cambridge University Press.

[15] Taylor, R. (2013). Interpretation of the Correlation Coefficient: A Basic Review. *Journal of Diagnostic Medical Sonography*. **6**, pp. 35-39.

[16] Arlt, J. and Radkovský, Š. (2011). The lag analysis in modelling of relationship of economic time series. *Politická ekonomie*. **1**, pp. 58-73.

[17] Ofori, G. (1990). *The Construction Industry: Aspects of Its Economics and Management*. Singapore: Singapore University Press.

[18] Douglas, C.M., Runger, G.C., Hubele N.F. (2010). Engineering Statistics.Wiley, 5th ed..

[19] Tse, R. and Ganesan, S. (2013). Causal relationship between construction fows and GDP: evidence from Hong Kong. *Construction Management and Economics*. **15**, pp. 371-376.

[20] Hillebrandt, Patricia M. (2000). *Economic theory and the construction industry*. 3rded. London: MacMillan.

[21] Fishman, G. S. (1995). Monte Carlo: Concepts, Algorithms, and Applications. *New York: Springer*.

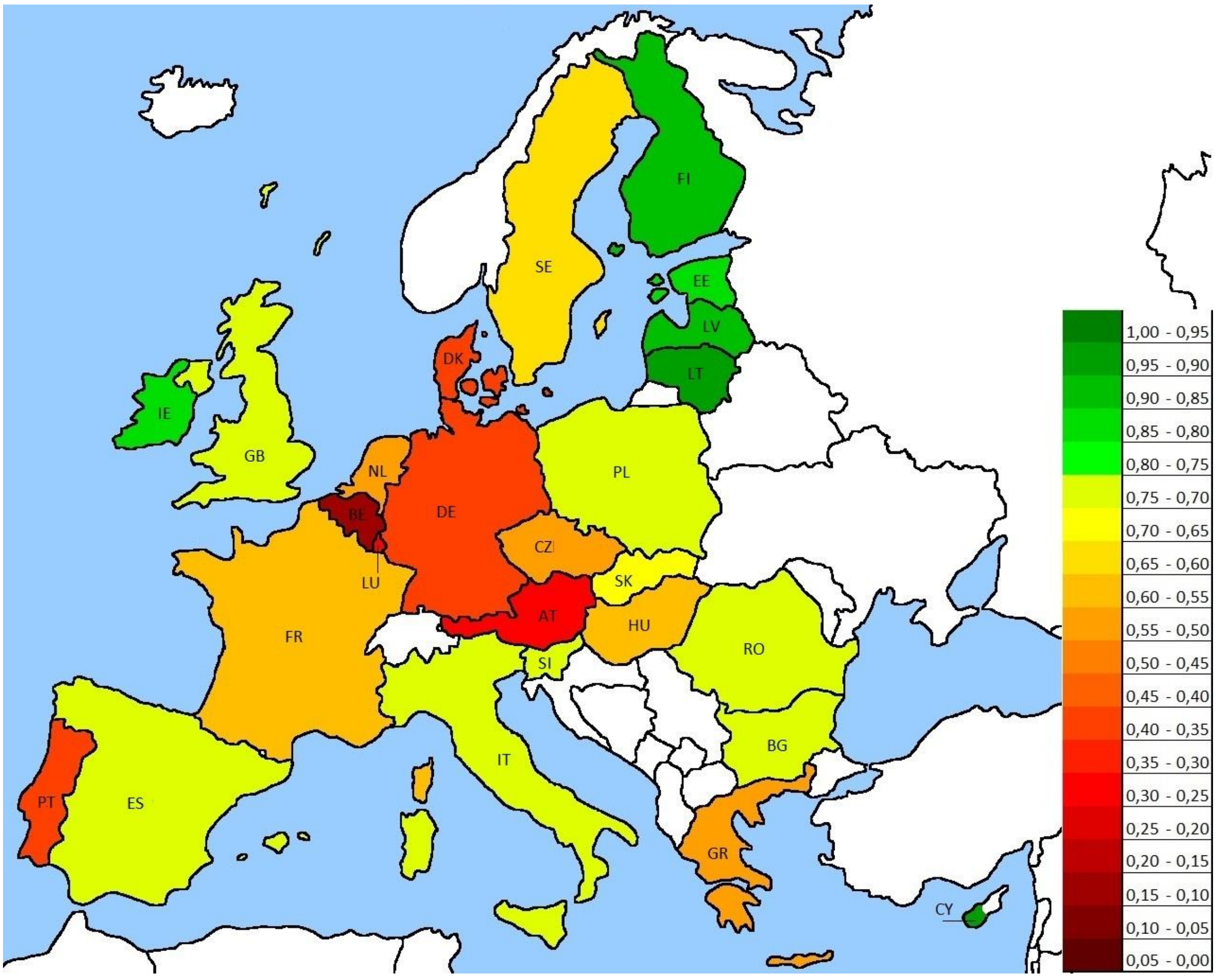
[22] Cox D.R., Hinkley D.V. (1974). *Theoretical Statistics*, Chapman & Hall.

[23] Hoteling, H. (1953). New light on the correlation coefficient and its transforms. *Journal of the Royal Statistical Society: Series B.* **2**, pp. 193-232.

[24] Sandberg, I.W., Lo, J.T., Fancourt, C.L., Principe, J.C., Katagiri, S., Haykin, S. (2001). Nonlinear *Dynamical Systems:Feedforward Neural Network Perspectives.* Wiley.

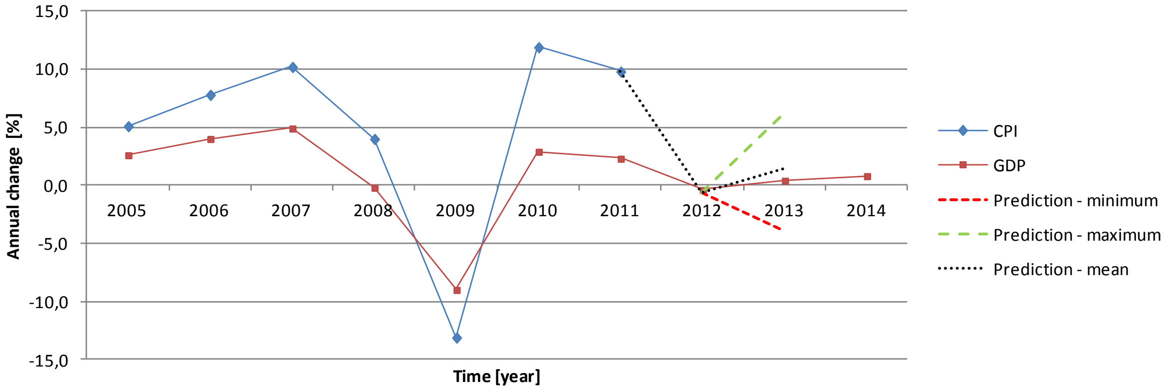
[25] Iwasa, K., Kugler, M., Kuroyanagi, S., Iwata, A. (2009). A Back Propagation Training Method for Multilayer Pulsed Neural Networks Using Principle of Duality. *Advances in Neuro-Information Processing: Lecture Notes in Computer Science.* **5507**, pp. 300-307.

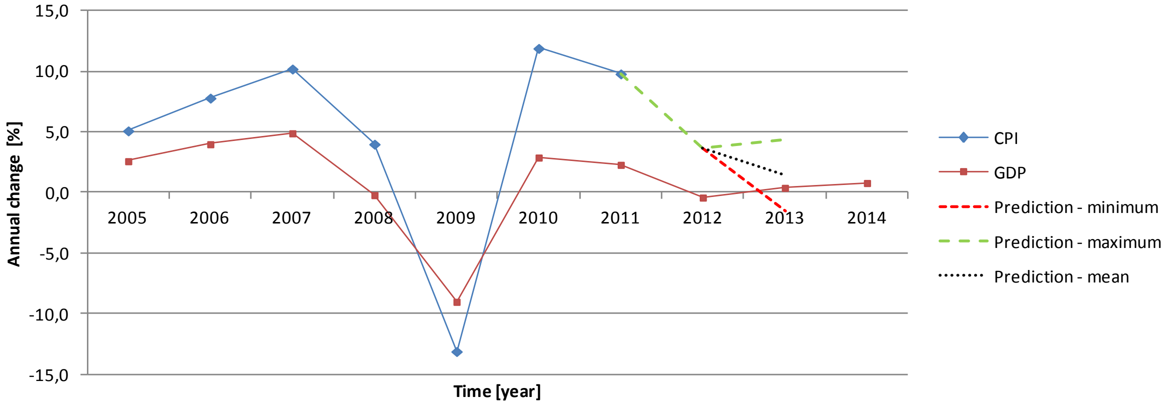
APPENDIX 1 – CORREALATION COEFFICENTS OF GDP-CPI TIME SERIES

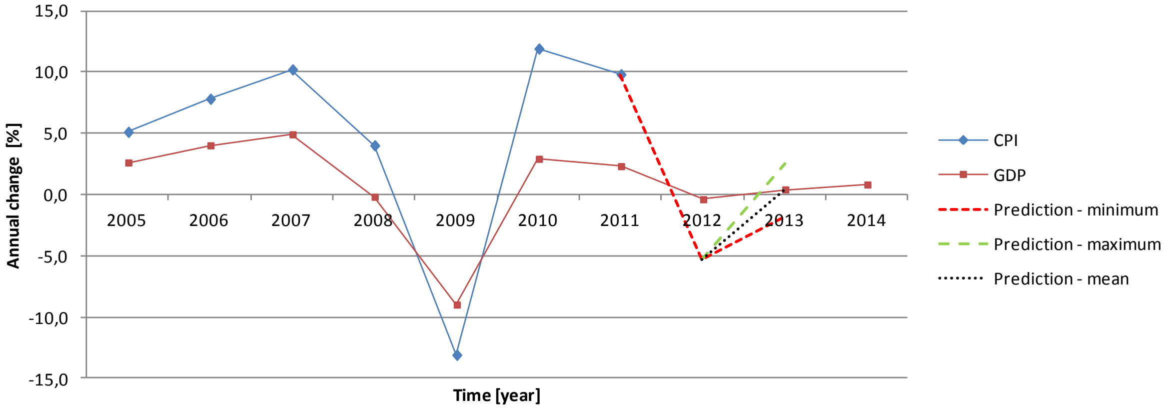


(source: autor)

APPENDIX 2 – PREDICTIONS OF CPI CHANGE FOR FINLAND IN 2013







(source: author)