

# Comparison of neural networks and regression time series in estimating the development of the EU and the PRC trade balance

Jaromír Vrbka<sup>1,\*</sup>, Zuzana Rowland<sup>1</sup> and Petr Šuleř<sup>1</sup>

<sup>1</sup>Institute of Technology and Business in České Budějovice, School of Expertness and Valuation, Okružní 517/10, 37001 České Budějovice, Czech Republic

**Abstract.** China, by GDP, is the second largest economic power, and hence also a key player in the field of international relations. As far as the EU is concerned, it is China's largest trading partner. From this point of view, it is clear that monitoring export and import development between these partners is essential. This paper therefore aims to compare two useful methods, namely the accuracy of time series alignment through regression analysis and artificial neural networks, to assess the evolution of the EU and the People's Republic of China trade balance. Data on the export and import trends of these two partners since 2000 have been used, and it is clear that the trade balance was completely different that year than it is now. The development over time is interesting. The most appropriate curve is selected from the linear regression, and from the neural networks three useful neural structures are selected. We also look at the prediction of future developments while taking into account seasonal fluctuations.

**Key words:** trade balance, export and import, linear regression, neural networks

## 1 Introduction

Balance of payments is the payment value of all economic transactions between the country and other countries over a period, usually a calendar year. The balance of payments is made up of export and import transactions of goods and services, the movement of investment resources, gifts and transfers of money or gold [1]. The balance of payments is in the form of an account where credit and debit items are recorded. In the case of surplus exports, the trade balance is an asset, otherwise it is a liability. This is related to the balance of payments balance, which tracks the difference between export and import on the current account [2].

China is with its GDP the second largest economic power and the key player in the field of international relations [3]. On the development of the Chinese economy we can see progress in transforming the low-cost economy into innovation and high-tech products [4]. The first trade relations of the European Union with China were established in 1975. The Trade and Economic Cooperation Agreement, which includes three basic principles of

---

\* Corresponding author: [vrбка@mail.vstecb.cz](mailto:vrбка@mail.vstecb.cz)

cooperation – business, economic and the tasks of the Joint Committee became the fundamental document of trade relations between the EU and China in 1985 [5]. International trade is also subject to a political dialogue that seeks to correct the Long-Term Policy for China-European Relations [6]. Until 2020, the EU-China Strategic Partnership Agenda applies, covering the area of security, exchange and sustainable development [7].

Successful business ties between China and the EU can be expressed by a figure from 2015 when trade exchanges amounted to 520.897 billion EUR. Looking at the European Union's trade balance, we must say that it is unfavorable [8]. The difference between export and import is almost doubled in favor of imports from China [9]. In the years 2014-2015, exports grew by almost 16 percent. The European Union is China's biggest trading partner, which was confirmed by the Chinese National Statistics Office, which stated that imports from China in 2017 amounted to 400 billion USD. The table below shows the development of exports and imports over a five-year period [10].

**Table 1.** The development of exports and imports over a five-year period

Export (USD trillions)	Import (USD trillions)	Return (USD trillions)	Balance (USD billions)	
2012	2.05	1.82	3.87	230.31
2013	2.21	1.95	4.16	259.01
2014	2.34	1.96	4.30	382.46
2015	2.27	1.68	3.95	593
2016	2.09	1.58	3.68	510.7
2017	2.21	1.73	3.94	476

Source: [10].

Taking into account the diversity of the sector in which the European Union and China are linked, we can expect a positive development of cooperation, but we should not forget the dangerous factors [11]. These factors can be included, for example, in cybernetics, and in particular their security. Cyber security is today a concept that employs all three of the largest trading partners, China, the European Union and the United States. International trade in the future will also affect relations between China and the United States, whose relations are still not politically stable. Another risk, which could be negatively affected, is the problem of dumping. The European Union is now addressing sixteen cases of likely imports of goods below the market price. China is facing a problem of large distribution at the expense of efficiency of one industry. China's economic growth carries various risks. In addition to the internal ones, it is mainly the fact that the economies of the EU countries that have failed to adapt to the effects of the financial crisis can only keep exports, of which a significant proportion goes to China [12]. Given that Chinese quality and technical level of production are rising, commitments on the EU's current account balance of payments are deepened by mutual commercial relationships, which negatively affects the unemployment rate and government debt. Thanks to acquisitions, China is buying the know-how of EU companies, thus gradually taking over its own export markets. In this way, businesses are losing their competitiveness, and economic risk is rising [13]. The European-Chinese market can certainly offer us much more, the question is whether China will realize its priorities and effectively transform them into satisfaction for both parties.

## 2 Data and methods

Data for analysis is available on the website of the World Bank etc. For the analysis, trade balance data between the EU and the People's Republic of China (hereinafter referred to as the PRC) will be used. This will be the difference between total exports and imports between the two trading partners from the EU perspective. The time period for which the data will be available is the monthly balance starting from January 2000 and ending in July 2018. This is about 223 input data. The unit is the Euro.

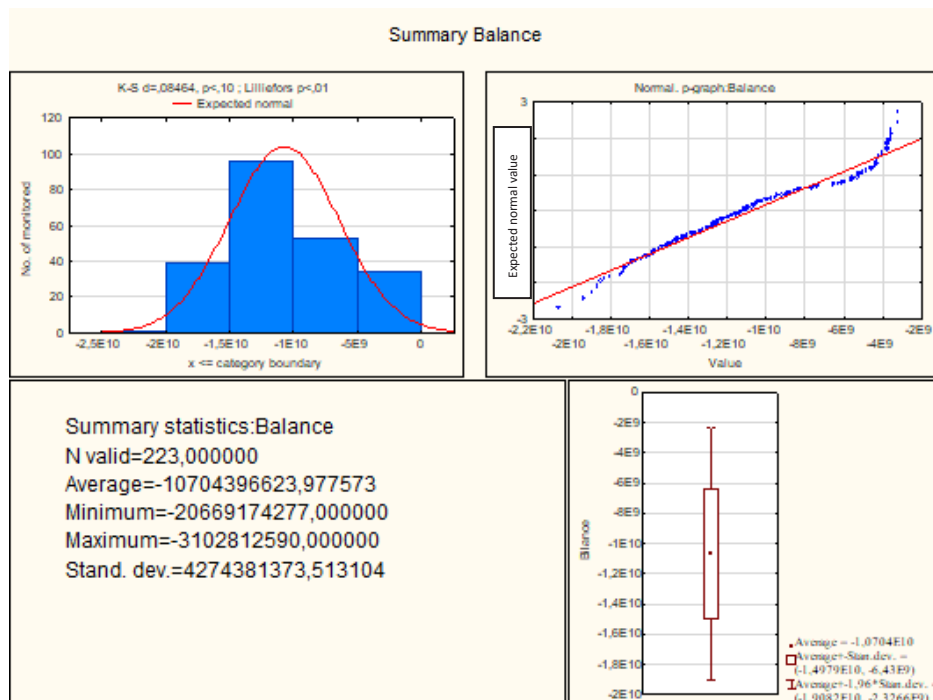
The descriptive characteristics of the data are given in Table 2.

**Table 2.** Characteristics of the data file

Samples	Month (Input variable)	Balance (Output (goal))
Minimum (Training)	36526.00	-2.066917E+10
Maximum (Training)	43252.00	-3.102813E+09
Average (Training)	39916.52	-1.066424E+10
Standard deviation (Training)	1949.82	4.128087E+09
Minimum (Testing)	36586.00	-1.888250E+10
Maximum (Testing)	43282.00	-3.118169E+09
Average (Testing)	39702.30	-1.022066E+10
Standard deviation (Testing)	2174.13	4.848833E+09
Minimum (Validation)	36951.00	-1.937546E+10
Maximum (Validation)	43040.00	-3.497008E+09
Average (Validation)	40047.88	-1.137918E+10
Standard deviation (Validation)	3096.16	7.279221E+09
Minimum (Overall)	36526.00	-2.066917E+10
Maximum (Overall)	43282.00	-3.102813E+09
Average (Overall)	39904.26	-1.070440E+10
Standard deviation (Overall)	1963.77	4.274381E+09

Source: Authors.

Interesting, of course, is price development over time. Figure 1 therefore shows the selected statistical characteristics in graphical form, including the histogram of the input data.



**Fig. 1.** Graph of basic statistical characteristics  
 Source: Authors.

For data processing, DELL's Statistica version 12 will be used. First, a linear regression will be performed. Afterwards, neural networks will be used for regression.

Linear regression will be performed on the monitored data sample for the following functions:

- Linear,
- Polynomial,
- Logarithmic,
- Exponential,
- Multiparameter of weighed distances,
- Multiparameter negative-exponential smoothing.

First, the correlation coefficient, i.e. the dependence of trade balance on time, will be calculated. Next, we will work with a level of significance of 0.95.

Then regression will be performed using neural structures. We will generate multilayer perceptron networks and neural networks of basic radial functions. Time will be the independent variable. We will determine the trade balance between the EU and the PRC as the dependent variable. We divide the time series into three sets – training, testing, and validation. The first group will have 70% of input data. Based on the training set of data, we generate neural structures. In the remaining two sets of data, we always leave 15% of the input information. Both groups will serve us to verify the reliability of the found neural structure, or the found model. The delay of the time series will be 1. We will generate 10,000 neural networks. We will preserve 5 of them with the best characteristics<sup>†</sup>. In the

<sup>†</sup> We will be orientated using the smallest square method. We will terminate network generation if there is no improvement, i.e. to reduce the sum of squares. Thus, we will preserve those neural

hidden layer, we will have at least two neurons, at most 50. In the case of the radial basic function, there will be at least 21 neurons in the hidden layer, at most 30. For the multiple perceptron network we will consider these distribution functions in the hidden layer and in the output layer:

- Linear,
- Logistic,
- Atanh,
- Exponential,
- Sinus.

Other settings are left by default (using the ANS tool – automated neural network).

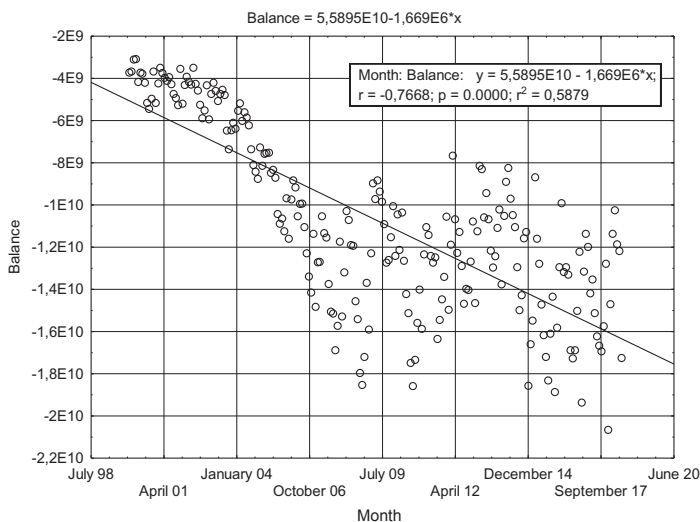
In conclusion, we compare the results of linear regression and regression using neural networks. Comparison will not take place in the form of residue analysis (minimum, maximum values, spread of residues, etc.), but at expert level and experience of the assessor, economist.

### 3 Results

#### Linear regression

The correlation coefficient is -0.7668, which represents a significant statistical indirect dependence of the trade balance on the development over time. The determination coefficient reaches 0.9001.

A point graph was constructed (see Figure 2), where the points were intersected by a regression curve, in this case linear. The line parameters are shown in the figure.



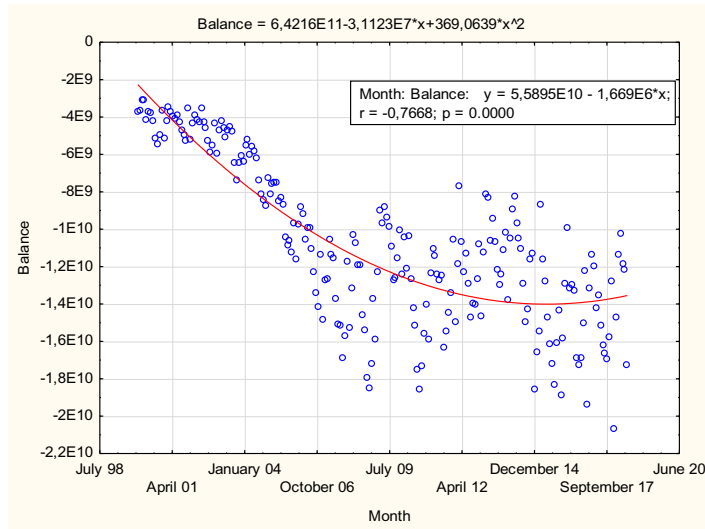
**Fig. 2.** The Eu and the PRC point graph intersected by the regression curve – linear curve  
 Source: Authors.

The solid line represents the regression function

Figure 3 represents the intersection of the point graph with a polynomial function.

---

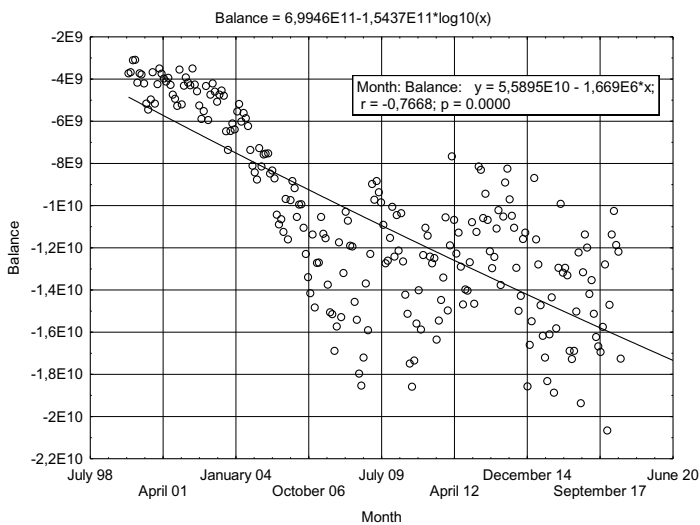
structures whose sum of squares of residue to the actual EU and PRC trade balance will be as low as possible (ideally zero).



**Fig. 3.** The EU and the PRC trade balance point graph intersected by a regression curve – polynomial function  
Source: Authors.

The polynomial function better outlines the trade balance developments between the EU and the PRC than is the case with the linear function.

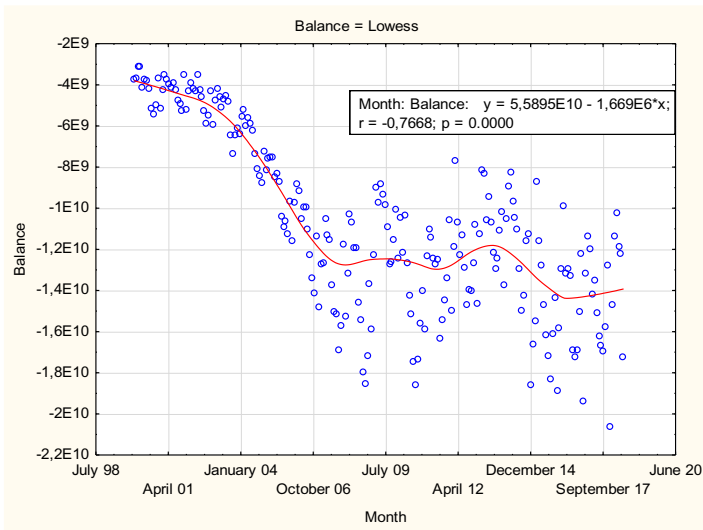
Figure 4 shows a point graph intersected by a logarithmic function.



**Fig. 4.** EU and PRC trade balance point graph intersected by a regression curve – logarithmic function  
Source: Authors.

Due to the shape of the curve and the location of individual points in the graph, it is obvious that the logarithmic function rather copies the linear function and is not suitable for the regression and will not estimate the evolution of the EU and the PRC trade balance optimally.

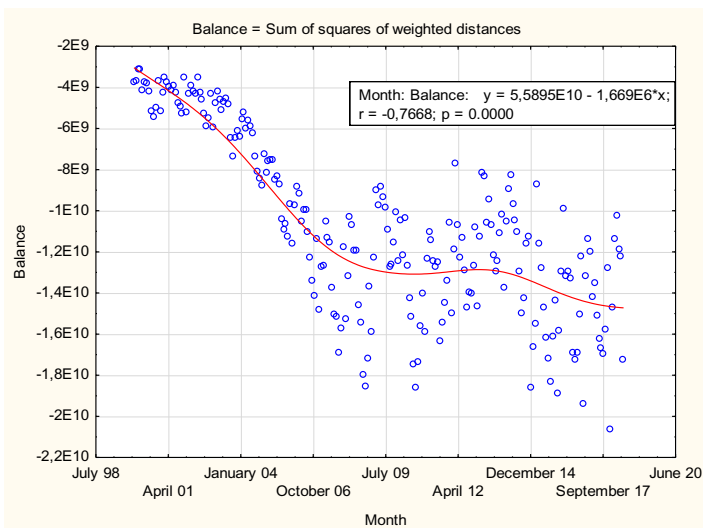
Figure 5 represents the LOWESS (locally weighted scatterplot smoothing) or LOES (locally estimated scatterplot smoothing) function.



**Fig. 5.** The EU and the PRC trade balance point graph intersected by a regression curve – LOWESS  
Source: Authors.

Regression is calculated for such a curve at partial intervals. The LOWESS curve aligns the time series and monitors its trend appropriately. LOWESS curve intersection looks very interesting. It can fairly reveal the trend of the trade balance of the partner countries.

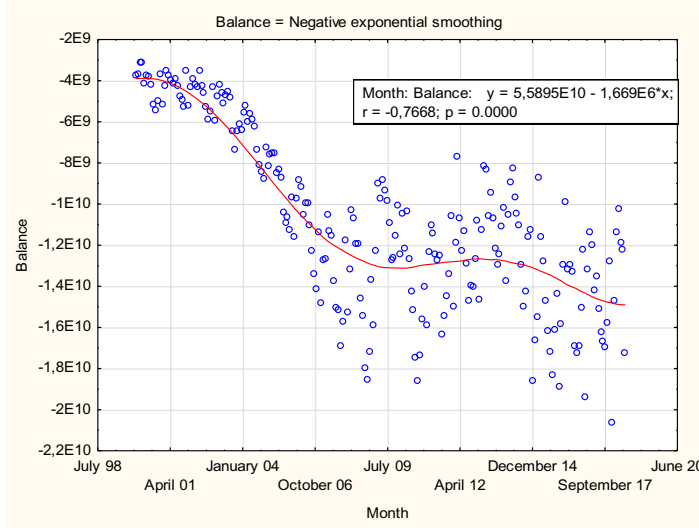
Figure 6 provides a point graph of the EU and PRC trade balance, intersected by smallest squares of weighed distances function.



**Fig. 6.** The EU and PRC trade balance point graph intersected by the regression curve – Sm. sq. of weighed distances  
Source: Authors.

The curve fairly well copies the development of the trade balance over its entire interval. However, we must state that in detail, the curve is less accurate than the previous LOWESS curve.

Figure 7 provides the intersection of the function obtained by smallest squares of negative-exponential smoothing.



**Figure 7.** EU and PRC trade balance point graph intersected by a regression curve – Sm. sq. negative exponential smoothing  
Source: Authors.

Even this curve appears to be interesting and appropriate for a possible prediction.

As noted above, the correlation coefficient indicates a significant statistical indirect dependence of the target variable on development over time. Nevertheless, it is very distant from the value of -1. If we only evaluated the results by comparing the development of the trade balance between the EU and the PRC and the shape of the regression curve while taking into account a simple linear regression, we could certainly say that the LOWESS curve is closest to the development. This follows curves obtained by the smallest squares method, namely negative-exponential smoothing and weighed distances. All three copy the basic development of trade balance between the EU and the PRC.

### Neural structures

Based on the established procedure, 10,000 neural networks were generated. Five networks have been preserved, the ones showing the best parameters. Their overview is given in Table 3.



**Table 3.** Overview of preserved neural networks

Network name	Training perf.	Testing perf.	Valid. perf.	Training error	Testing error	Validation error	Training algorithm	Error function	Activ. of hidd. lyr.	Output activ. funct.
RBF 1-24-1	0.905183	0.905623	0.926697	1.500796E+18	2.368994E+18	1.609908E+18	RBFT	Sum of sq.	Gauss	Identity
RBF 1-29-1	0.908078	0.881347	0.926506	1.459978E+18	2.571730E+18	1.559369E+18	RBFT	Sum of sq.	Gauss	Identity
RBF 1-26-1	0.908992	0.894534	0.933502	1.443607E+18	2.334177E+18	1.522580E+18	RBFT	Sum of sq.	Gauss	Identity
RBF 1-29-1	0.887327	0.887465	0.928457	1.767475E+18	2.534595E+18	1.365003E+18	RBFT	Sum of sq.	Gauss	Identity
RBF 1-29-1	0.912620	0.881659	0.927402	1.388482E+18	2.542374E+18	1.485960E+18	RBFT	Sum of sq.	Gauss	Identity

Source: Authors.

These are only the neural networks of the basic radial function. The input layer has only one variable - time. In the hidden layer, neural networks contain from 24 to 29 neurons. In the output layer, we have a single neuron and the only output variable is the export balance of the EU and the PRC. For all networks, the RBFT training algorithm was applied. In addition, all neural structures used the same function to activate the hidden layer of neurons, namely the Gaussian curve. They also use the same function to activate the outer layer of neurons, and this function is identity (see Table 3).

Training, testing and validation performance is also interesting. In general, we are looking for a network that has the same performance in all sets of data (we remind that the data was randomly distributed). The error should be as small as possible.

The performance of individual sets of data is expressed as a correlation coefficient. The values of the individual data sets according to specific neural networks are presented in Table 4.

**Table 4.** Correlation coefficients of individual data sets

	Balance (Training)	Balance (Testing)	Balance (Validation)
1.RBF 1-24-1	0.905183	0.905623	0.926697
2.RBF 1-29-1	0.908078	0.881347	0.926506
3.RBF 1-26-1	0.908992	0.894534	0.933502
4.RBF 1-29-1	0.887327	0.887465	0.928457
5.RBF 1-29-1	0.912620	0.881659	0.927402

Source: Authors.

The table shows that the performance of all preserved neural structures is approximately identical. Slight differences do not affect the performance of individual networks. The correlation coefficient of all training data sets ranges from almost 0.89 to more than 0.91. The value of the correlation coefficient of test data sets reaches more than 0.887. The correlation coefficient of the validation set of data for all neural networks is above 0.926. In order to select the most suitable neural structure, we need to analyse the results obtained. Table 5 provides the basic statistical characteristics of each set of data for all neural structures.

**Table 5.** Statistics of individual data sets according to preserved neural structures

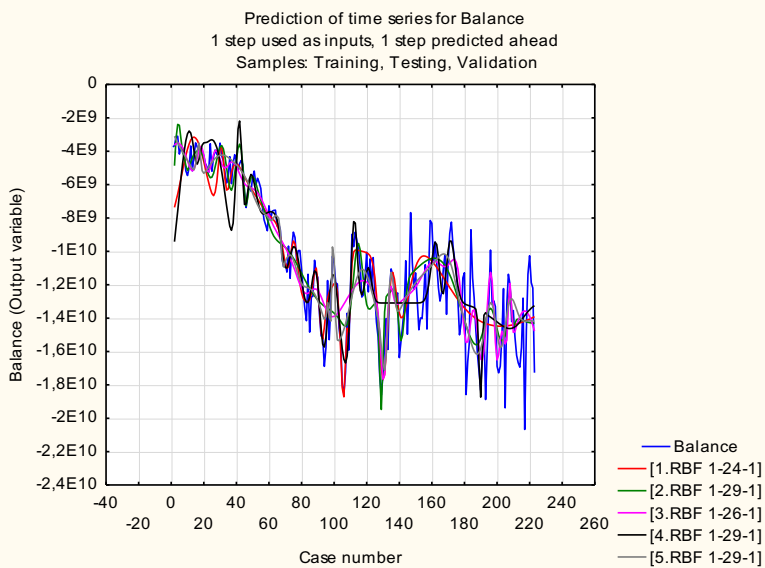
Statistics	1.RBF 1-24-1	2.RBF 1-29-1	3.RBF 1-26-1	4.RBF 1-29-1	5.RBF 1-29-1
Minimum prediction (Training)	-1,871511E+10	-1,947388E+10	-1,766263E+10	-1,873705E+10	-1,711803E+10
Maximum prediction (Training)	-3,194390E+09	-2,393646E+09	-3,512380E+09	-2,187089E+09	-3,099840E+09
Minimum prediction (Testing)	-1,817011E+10	-1,535797E+10	-1,487201E+10	-1,584944E+10	-1,566790E+10
Maximum prediction (Testing)	-3,152334E+09	-2,445794E+09	-3,543154E+09	-3,310599E+09	-3,302471E+09
Minimum prediction (Validation)	-1,594921E+10	-1,545983E+10	-1,583557E+10	-1,677209E+10	-1,606970E+10
Maximum prediction (Validation)	-3,210110E+09	-3,745279E+09	-4,477256E+09	-2,682479E+09	-4,243404E+09
Minimum residues (Training)	-6,487865E+09	-6,484048E+09	-7,099190E+09	-6,816668E+09	-6,486602E+09
Maximum residues (Training)	4,486943E+09	3,800359E+09	4,809439E+09	5,740114E+09	4,622375E+09
Minimum residues (Testing)	-4,556967E+09	-4,745440E+09	-4,659938E+09	-5,208299E+09	-4,529743E+09
Maximum residues (Testing)	5,079201E+09	6,510952E+09	5,428001E+09	5,388652E+09	6,731699E+09
Minimum residues (Validation)	-4,903684E+09	-4,318523E+09	-3,827757E+09	-4,820044E+09	-4,821168E+09
Maximum residues (Validation)	3,076065E+09	3,131742E+09	2,597524E+09	3,247646E+09	2,224139E+09
Minimum standard residues (Training)	-5,295914E+00	-5,366276E+00	-5,908597E+00	-5,127380E+00	-5,504870E+00
Maximum standard residues (Training)	3,662602E+00	3,145223E+00	4,002856E+00	4,317614E+00	3,922790E+00
Minimum standard residues (Testing)	-2,960697E+00	-2,959129E+00	-3,050094E+00	-3,271460E+00	-2,840887E+00
Maximum standard residues (Testing)	3,299997E+00	4,060054E+00	3,552819E+00	3,384744E+00	4,221871E+00
Minimum standard residues (Valid.)	-3,864755E+00	-3,458285E+00	-3,102089E+00	-4,125572E+00	-3,955020E+00
Maximum standard residues (Valid.)	2,424349E+00	2,507907E+00	2,105085E+00	2,779725E+00	1,824561E+00

Source: Authors.

Ideally, the individual statistics of the neural network are cross-sectional in all sets (minimum, maximum, residue, etc.). However, in the case of aligned time series, the differences are minimal. Therefore, we are not able to clearly identify which of the preserved neural networks has the most appropriate results.

Figure 8 is a graph showing the real trade balance between the EU and the PRC, as well as the development of predictions using individual generated and preserved networks.

The graph shows that all neural networks predict the trade balance at different intervals slightly differently. However, the similarity between predictions of individual networks is not important, but the similarity (or degree of consistency) with actual trade balance developments is. Based on the optical comparison of the curves, we can say that the RBF 1-24-1 and RBF 1-29-1 networks show at certain intervals weaker results than the remaining three preserved neural networks. The networks 2. RBF 1-29-1, 3. RBF 1-29-1, and 5. RBF 1-29-1 respects the curve guidelines that assess the development of the trade balance and at the same time tend to perceive the extremes of this curve.



**Fig. 8.** Line graph – The EU and the PRC trade balance predicted by neural networks compared to the real trade balance in the monitored period  
Source: Authors.

## 4 Conclusion

The aim of the paper was to compare the accuracy of time series alignment using regression analysis and neural networks on the example of the trade balance between the EU and the People's Republic of China.

In general, each prediction is given by a certain degree of probability with which it is to be fulfilled. As we predict the future development of any variable, we try to estimate the future development of this variable on the basis of previous years' data. Although we can include most of the factors influencing the target variable in the model, we always simplify reality, and we always work with a certain degree of probability that some of the predicted scenarios will be fulfilled. It is only a mathematical variable that evaluates the relationship of two or more variables that were generated in the past. And we are trying to predict future developments in accordance with past developments. As with linear regression, as well as regression using neural networks, there is simplification – and quite substantial. We only work with two variables – input (time) and output (trade balance EU X PRC). So, we completely ignore other input variables, which often have a significant impact on the trade balance between the EU and the PRC (international political situation, taxation in both countries, production factor prices, state export support and many others). However it is still true that we estimate aggregated variables better than the partial ones. For some analyses, we can use regression results using neural networks that show a higher degree of accuracy. This may be trend-setting for estimating industry development for the need for individual business valuation, etc.

At the same time, we can state that due to the great simplification of reality, it is not possible to predict the emergence of extraordinary situations and their impact on the EU and PRC trade balance (perhaps in the short term, yes, but certainly not long term). Ideal would be a prediction for the next few days, but for such a short prediction it is currently not possible to obtain data.

The trade balance of both countries can be determined on the basis of statistical methods, causal methods and intuitive methods. In this case, we have been comparing statistical methods. However, they only gave us a possible framework for the development of the monitored variable. It is important to work with information on the possible future development of the economic, political or legal environment. If we can predict their development, we can then project it into the monitored trade balance. At the same time, however, the personality of the evaluator comes into play - an economist who, on the basis of his knowledge and experience, corrects the price determined by the statistical methods and specified on the basis of causal links.

Optically, the LOWESS curve was the best out of linear regression, followed by the smallest squares curve by negative-exponential smoothing, and the curve obtained by the smallest square-weighted distance. Of the neuronal networks, three of the five preserved neural structures proved to be usable in practice: 2. RBF 1-29-1, 3. RBF 1-29-1 and 5. RBF 1-29-1. If we look at the performance from a correlation coefficient point of view, only three of the neural networks listed here are left to be used, of which there is practically no difference.

The aim of the paper was fulfilled.

Given the real trade balance between the EU and the PRC, it is advisable to predict future developments while taking into account seasonal fluctuations.

## References

1. C. Z. Liu, F. Wang, Empirical Study on elasticity approach to balance of payments in China – Synchronously discussing the influence of RMB appreciation to payments balance of current account. *International Conference on Management Science and Engineering*, (2004)
2. J. Stefaniak, J. Kuczevska., Export Specialization in Services of the Visegrad Countries. *Equilibrium: Quarterly Journal of Economics and Economic Policy*, **11**(2), 265-284, (2016)
3. J. L. Zamora Ortiz, Balance of Payments Constrained Growth in China: An Application of the Autoregressive Distributed-Lag Modelling Approach. *Revista Ecorfan*, **6**(15), (2015)
4. L. Yan, Does foreign direct investment provide desirable development finance? The case of China. *China and World Economy*, **15**(2), (2007)
5. I. L. Rocha, Manufacturing as driver of economic growth. *PSL Quarterly Review*, **71**(285), (2018)
6. P. Balaz, *Globalization of the World Economy and the Impact of Chinas Expansion on the Competitiveness of EU*, 14-25, (2014)
7. E. Soukiazis, Are the Transition Economies – Balance-of-Payments Constrained? An Aggregate and Multisector Approach Applied to Central and Eastern Europe. *Eastern European Economics*, **55**(5) (2017)
8. C. Angelache, P. Anghel, Main Developments and Perspectives of the European Union. *Romanian Statistical Review*, **2**, 57-79, (2017)
9. J. Mayer, Towards a new normal growth strategy: China in comparative perspective. *Journal of Chinese Economic and Business Studies*, **14**(2), (2016)
10. Embassy of the Czech Republic in Beijing, *China: Foreign Trade and Investment* [online]. Available at: <https://www.businessinfo.cz/cs/clanky/cina-zahranicni-obchod-a-investice-19052.html> (2018)

11. P. Balaz, J. Pasztorova, *Changes in Development of Interantional Trade in the Period of Financial Crisis – Consequences for Economic Interests of the Slovak Republic*, (2014)
12. R. Higgs, Against the Whole Concept and Construction of the Balance of International Payments. *Independent Review*, **23**(2), (2018)
13. J. Schymik, Globalization and the evolution of corporate governance. *European Economic Review*, **102**(48), (2018)