# ANALYSIS OF ENTERPRISES OPERATING IN THE CONSTRUCTION INDUSTRY BASED ON KOHONEN NETWORKS

**Marek Vochozka**

**Abstract**

The construction industry is considered very specific. This is due not only to its product, but also to the fact that it indicates the first future development of the entire national economy. That is why many economists are watching the development of the construction industry in particular. It uses a number of tools for year-to-year comparison and an estimate of future developments across the industry. Interestingly, cluster analysis is performed using Kohonen Networks. These are artificial neural networks whose task is to create clusters of businesses according to certain significant features.

The aim of this paper is to analyze the construction industry using Kohonen networks and to estimate its development. Based on the analysis, significantly numerous and defined clusters of businesses were formed. The most important clusters were analyzed. Consequently, we can generalize the result and determine the estimate of the number of businesses that will succeed, stagnate or even fail in the next period. Thanks to this, we can estimate not only the overall growth or decrease in the construction industry in the Czech Republic in the following period, but we can also estimate the structure of the companies in terms of their size, turnover and volume of sales.

**Key words:** construction industry, Kohonen networks, cluster analysis, prediction

**JEL Codes:** C45, G32, M21

## Introduction

The construction industry in the European Union is one of the most important industries that makes a significant contribution to economic growth (Jaafar et al., 2008). Construction occupies a leading position in both direct and indirect job creation and responds very quickly and responsively to a change in investor behavior (Busina and Sikyr, 2014). Thus, the construction industry is the main developer of the constructions, which are an important part

of the investment or, more precisely, the creation of gross fixed capital, in the whole economy (Vochozka et al., 2015).

According to Adinoya and Nor'ainaf (2015), the construction industry plays a vital role in the national economy and its long-term socio-economic development. Its key role lies primarily in the development of housing, facilities and infrastructure necessary for national social and economic development (Mitkus and Trinkuniene, 2008). According to Naderpajouh (2016), the construction industry has its own specificities in particular in the demands on capital, individuality, manual difficulty and the extraordinary requirements.

The only way to increase the competitiveness of an enterprise under very strong competition on the construction market is to control production and trade costs (Siskina et al., 2009). In a free market economy, the company's expenditure management is the starting point for success, so it is very important for building management managers not only to control but also to predict costs caused by managing the process of creating competitive advantages (Kliestik et al., 2018).

The paper deals with the Czech construction industry, so now there are several figures for this industry. Construction output in the Czech Republic increased by 2.1% in 2017. Construction output increased by 5.0% and civil engineering posted a decrease in construction output by 5.7%. The average registered number of employees in enterprises with 50+ employees in the construction sector decreased by 2.0% in 2017. The average nominal wage of these employees increased by 4.6% to CZK 34,478. Construction enterprises with 50+ employees in the Czech Republic closed 61,515 building orders in the year 2017 and this number increased by 0.3%. The total value of these orders increased by 9.3% to CZK 202.6 billion, in the construction industry by CZK 93.3 billion (growth by 8.4%) and in civil engineering by CZK 109.2 billion (growth by 9.9%). The number of building permits issued in 2017 rose by 1.0%, the building authorities issued 84,164 of them (Construction, 2018).

The analysis of the construction industry in the Czech Republic and the estimation of its future development can also utilize modern, innovative methods, such as artificial neural networks, particularly Kohonen's networks.

Kohonen networks are types of neural networks that can be used to group sets of data into different groups. The data is grouped so that the records within the group tend to be the same and the records in the different groups are different. Many experimental results show that Kohonen's networks are very effective for assessing companies (Han and Wang, 2008).

The Kohonen network consists of an input layer that is completely interconnected with the output layer and learns by self-organization, without a teacher. This network has a very

wide range of uses, as it is an alternative network that is applicable to most neural network calculations. It is used mainly for sound editing, speech processing, photos, videos, security applications, and allows the projection of high-dimensional data into lower-dimensional data (Vochozka et al., 2016).

The aim of the paper is to analyze the construction industry in the Czech Republic using the Kohonen networks and then to estimate its development.

# 1 Data and Methods

For the purpose of this paper, a data set will be created, which will include the complete data of the financial statements of 12,584 companies that were active in the Czech Republic in the building industry in 2015. They are therefore entities whose predominant activity is classified in Section F of CZ-NACE Branch Classification of Economic Activities. The set of businesses will be generated from the Albertina database of Bisnode.

The data will be written into a table in Excel. Each line will contain a piece of data from the financial statement of a particular company that will be identified by name and identification number. Businesses that did not perform their core function throughout the reporting period will be removed from the data file. These are companies that were closed during this period (and did not have a significant impact on the direction of the national economy) and companies that started their activities (as well as having no significant influence on the direction of the construction industry in the Czech Republic). Of course, we could still discuss the importance of companies that started their operations on January 1, 2015, and companies that ceased operations on December 31, 2015. However, this does not have a significant impact on meeting the objective of this paper.

Columns of data that do not show any dispersion were also excluded.

The file will then be subjected to cluster analysis using Kohonen networks. For cluster analysis, Dell's Statistica software will be used in version 12. Data Mining will be used as a specific Neural Network tool. Here we select neural networks without a teacher – Kohonen's network. We specify data for analysis – select table with the data set from Excel. In all cases, these are continuous predictors. The file will be divided into three parts: 1. Training Data Set: Represents 70% of enterprises in the set. This Kohonen network will be created on this dataset. 2. Testing Data Set: This is 15% of the enterprise's original set. Using this data set, we verify the parameters of the Kohonen network. 3. Validation dataset: 15% of enterprises in

the file will also be involved. Using this dataset, we will also test the Kohonen network to see whether it is or not usable.

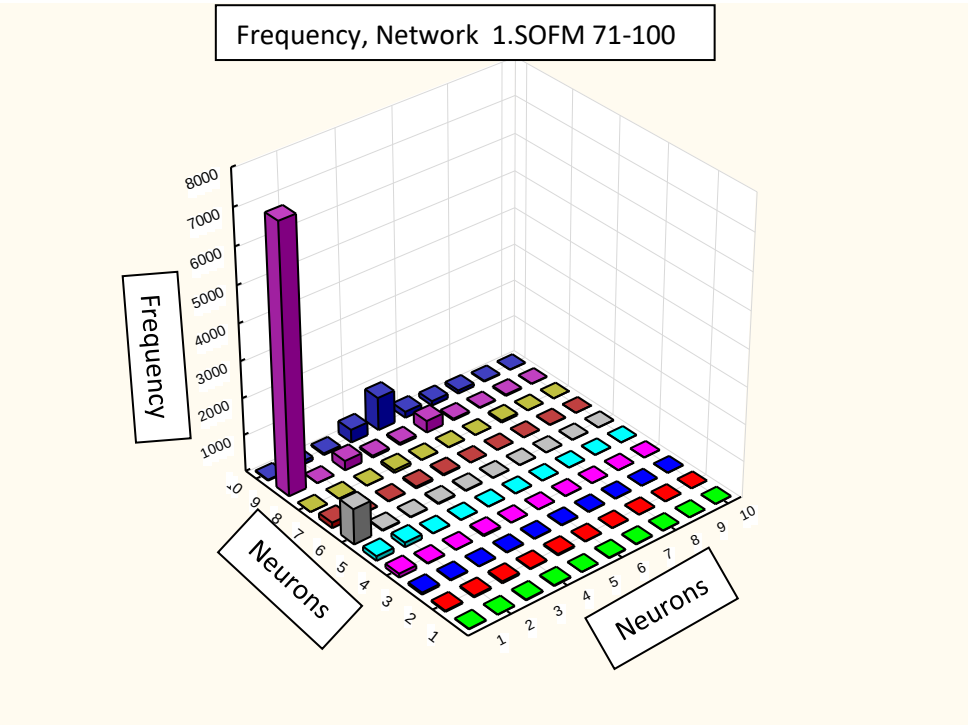
We determine the topological length and the topological width of the Kohonen map at 10. The number of iterations of the calculation will be set at 10,000. However, I should remind that the level of error is decisive. If there is no further iteration to improve the Kohonen network parameters, the training will be terminated before the 10,000th iteration is performed. If the network parameters are improved even at 10,000 iterations, we need to repeat the process and set the higher value of the iterations needed to make sure the result is the best possible. The learning speed will initially be set to 0.1, at the end to 0.02.

The results, ie dividing individual enterprises into clusters of 100, will be imported into the Excel spreadsheet again. Subsequently, the individual clusters will be subjected to an analysis of absolute and selected ratios (or their mean values – arithmetic averages) and the results will be interpreted.

## 2 Results

Based on the applied methodology, clusters were created. The breakdown of enterprise frequencies in each cluster of Kohonen maps is shown in Figure 1.

**Fig. 1: Breakdown of enterprise frequencies in individual clusters of Kohonen maps**




Source: Author.

Figure 1 provides a three-dimensional view of Kohonen maps and the number of companies represented in each cluster. The figure shows that the largest number of enterprises is in the cluster number (9, 1). Cluster (10, 5) follows. On the imaginary third position there is cluster (6, 1). Furthermore, a slightly higher representation of companies in clusters (9, 3), (10, 4) and (9, 6) can be observed. In other clusters, the representation of enterprises is significantly lower. At the same time, it is worth noting that none of the clusters of the Kohonen map are vacant. In order to be able to make a detailed view of the representation of companies in individual clusters, we add the number one figure to the specific values listed in table 1.

**Tab. 1: Business frequency in individual clusters of Kohonen maps**

| | Frequency table (FINAL DATA ID 91)<br>Network: 1.SOFM 71-100<br>Samples: Training, Testing, Validation | | | | | | | | | |
|---|---|---|---|---|---|---|---|---|---|---|
| | 1 | 2 | 3 | 4 | 5 | 6 | 7 | 8 | 9 | 10 |
| 1 | 15 | 8 | 23 | 27 | 3 | 19 | 17 | 12 | 3 | 8 |
| 2 | 32 | 41 | 52 | 24 | 39 | 16 | 8 | 7 | 9 | 4 |
| 3 | 54 | 24 | 10 | 20 | 19 | 20 | 18 | 18 | 5 | 7 |
| 4 | 107 | 23 | 9 | 44 | 7 | 4 | 13 | 4 | 4 | 4 |
| 5 | 127 | 122 | 12 | 10 | 9 | 10 | 3 | 9 | 9 | 2 |
| 6 | 959 | 29 | 19 | 34 | 25 | 15 | 10 | 6 | 11 | 6 |
| 7 | 149 | 13 | 17 | 50 | 36 | 7 | 7 | 6 | 15 | 1 |
| 8 | 5 | 5 | 10 | 76 | 13 | 12 | 2 | 57 | 5 | 1 |
| 9 | 7,245 | 9 | 259 | 52 | 57 | 323 | 42 | 21 | 37 | 14 |
| 10 | 24 | 97 | 48 | 351 | 898 | 167 | 124 | 86 | 25 | 10 |

Source: Author.

It is clear from the table that, in addition to the above clusters (in the table highlighted by yellow cell backgrounds), only six more clusters involve more than one hundred enterprises. However, we will only include clusters highlighted in the table in the next analysis. They contain data on 10,035 businesses, representing four fifths of the entire population. In addition, one of the clusters, namely (9, 1), comprises 7,245 enterprises. It can therefore be assumed that the analysis carried out will have sufficient informative value.

The average values of the file as a whole and the selected aggregates in the assets section are shown in Table 2.

**Tab. 2: Mean values of the assets of the examined file and selected clusters**

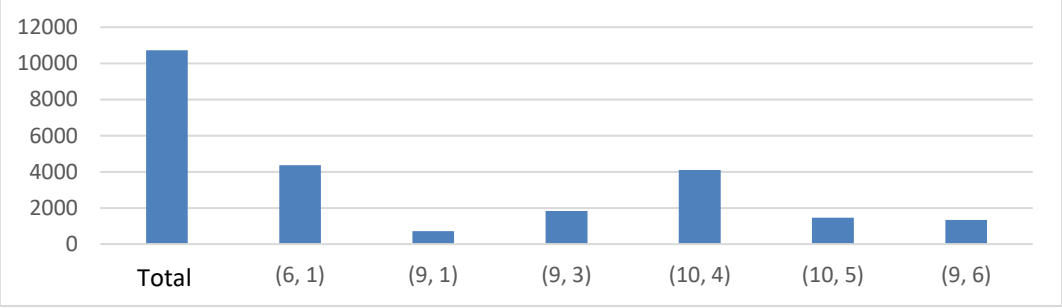
| Item | Total | (6, 1) | (9, 1) | (9, 3) | (10, 4) | (10, 5) | (9, 6) |
|---|---|---|---|---|---|---|---|
| Total assets | 10,719 | 4,369 | 725 | 1,835 | 4,098 | 1,457 | 1,330 |
| Fixed assets | 172 | 44 | 7 | 38 | 17 | 4 | 13 |
| Intangible fixed assets | 7,536 | 3,983 | 651 | 1,719 | 3,795 | 1,375 | 1,254 |
| Tangible fixed assets | 1,387 | 143 | 30 | 82 | 468 | 198 | 128 |
| Grounds | 2,012 | 398 | 49 | 283 | 1,295 | 479 | 483 |
| Buildings | 1,128 | 273 | 49 | 248 | 1,684 | 505 | 571 |
| Separate movable assets and sets of movable assets | 2,494 | 341 | 43 | 68 | 285 | 78 | 63 |
| Long-term financial assets | 22,332 | 9,471 | 1,732 | 5,521 | 13,304 | 3,177 | 4,266 |
| Current assets | 3,218 | 1,280 | 262 | 1,915 | 2,250 | 525 | 711 |
| Inventory | 374 | 13 | 11 | 129 | 614 | 136 | 136 |
| Material | 1,764 | 142 | 29 | 157 | 1,532 | 304 | 495 |
| Unfinished production and semi-finished products | 1,445 | 389 | 69 | 129 | 467 | 136 | 100 |
| Long-term receivables | 11,022 | 4,581 | 830 | 2,097 | 7,839 | 1,731 | 1,988 |
| Short-term receivables | 4,845 | 410 | 101 | 410 | 6,575 | 1,428 | 1,541 |
| Trade receivables (short-term) | 267 | 29 | 8 | 51 | 400 | 81 | 91 |
| State – tax receivables | 232 | 19 | 10 | 100 | 404 | 75 | 67 |
| Other advances provided | 6,043 | 3,222 | 543 | 1,380 | 2,722 | 784 | 1,467 |
| Short-term financial assets | 186 | 8 | 5 | 42 | 108 | 162 | 594 |
| Money | 3,745 | 465 | 78 | 288 | 2,602 | 623 | 878 |
| Bank accounts | 293 | 179 | 31 | 87 | 180 | 70 | 59 |
| Accruals | 33,345 | 14,028 | 2,487 | 7,443 | 17,605 | 4,705 | 5,658 |

Source: Author.

The table clearly shows the difference between the values for the entire construction industry and for each cluster. The average values of the entire construction industry are significantly higher than the average values of clusters with the highest number of enterprises. This may mean that the selected clusters reflected very significantly as one of the parameters for clustering and enterprise size (it did not necessarily have to be the primary parameter, which could be seen in extraordinary way). The most important asset values are shown in separate graphs below.

Figure 1 provides a comparison of average asset values for the construction industry as a whole and for each cluster.

**Fig. 2: Average asset values for the construction industry in total and selected clusters**




Source: Author.

The average balance, ie the value of the assets reaches an average of more than CZK 10 million per business for the construction industry. The other two aggregates reach the average value of assets per enterprise over CZK 4 million. In particular, it is cluster (6, 1) and (10, 4). The lowest average value is reached by the cluster (9, 1). The companies in this cluster accumulate assets on average at 725,000 CZK.

No less interesting indicator that distinguishes companies in individual clusters is certainly the amount of fixed assets in the construction industry. The construction industry is inherently very specific. Above all, it is characterized by a small number of realized products. Production is realized over a long period of time. At the same time, irregular cash flows are typical for enterprises in the construction industry. At the same time, in the construction industry, there is a striving for a great involvement of machine production. Long-term property thus announces a possible marginal rate of capital replacement for work. The more machines an enterprise has, the less workers it has to employ. The construction industry as a whole is in the order of 170,000 CZK of fixed assets per enterprise. The cluster enterprises surveyed use significantly less fixed assets. In particular, only a cluster (6, 1) is reported to be over CZK 40,000 on average per enterprise. And only a cluster (9, 3) with a value of 40,000 CZK per enterprise is approaching. If we assume that the size of the property is growing with the size of the enterprise, it is easy to conclude that the surveyed clusters do not include large or extremely large enterprises operating in the building industry in the Czech Republic.

Another subject being studied is circulating assets. Circulating assets (although not entirely identical to working capital) are a very important part of each enterprise, but the construction company, in particular. It includes inventory, receivables and financial assets. Stocks of building materials are always relatively high in construction sites. Similarly, payment morale in the construction industry is not entirely optimal, so receivables are relatively high. Due to payment discipline, construction companies hold relatively high money. The construction industry as a whole is the highest average value of current assets. Over 2 mil. CZK average show businesses in cluster (10, 4). Just below this boundary are companies in cluster (9, 3).

In this context, we also present monetary values. The results show that the average construction company holds funds in excess of CZK 3.5 million. Enterprises of cluster (10, 4) also hold very high value. Equally interesting are (albeit not surprisingly) the values of sources of financing – capital. Liabilities were selected and their average values were compared. Certainly, the structure of the financing of construction enterprises is worth a

closer examination. Therefore, for each cluster, we have a total average value of own and foreign capital for the construction industry.

Results show that in the construction industry the share of foreign and equity capital is on average 50%. In the case of the total set, both own and foreign capital is significantly above CZK 3.5 million. The best-capitalized cluster is the cluster (10, 4). In this case, foreign capital slightly prevails. Both components of funding amount to approximately 1 million CZK. In the case of two clusters, (6, 1) and (9, 3), foreign capital prevails. In the other three clusters, the values of both components of capital are approximately the same. However, in one case, they are below CZK 0.5 million.

The construction industry indicates potential growth or decline in the economy and a possible recession or even a crisis. This means that we always monitor not only the volume of sales, but also the performance and added value of construction companies. The higher the revenues, the better the national economy. Even if there is even a high added value, it indicates a significant growth in the whole of the economy.

However, due to the scope of the article, we can only focus on selected items of the profit and loss statement. First we compare the average performance values in the construction industry. In this case, the highest value per enterprise has a cluster (10, 4). The performance of the average enterprise of this cluster is close to CZK 35 million. On average, construction companies achieve slightly more than CZK 30 million. However, the third highest value offers cluster (9, 6), only less than 10 million CZK. The achieved economic results were listed. Specifically, it deals with the operating profit or loss, the economic result before tax and the economic result for the accounting period.

For the enterprise, the most important profit category is the operating profit or loss. This is because it represents the results of the company's activity, respectively the results of the activity for which the enterprise was founded. The highest value not only of the operating profit or loss but also of the profit or loss before tax is achieved by the construction companies as a whole. The average value of these two profit categories is close to CZK 1.5 million, the value of the profit for the accounting period is CZK 1 million. The most successful cluster is the cluster (6, 1), the third cluster (10, 4).

So far, we can assume that we have compared businesses of different sizes, which has increased their clustering. The results can be quite distorted. Therefore, I have also dealt with compare them relatively, ie by using the ratio indicators, which to a certain extent will erase differences in the sizes of the enterprises in individual clusters.

## Conclusion

The aim of the paper was to analyze the construction industry in the Czech Republic using the Kohonen networks and then to estimate its development.

The aim of the paper was met. A cluster analysis was performed. From the clusters, the ones with the most represented businesses were selected. Subsequently, an analysis of the absolute values of selected items of the financial statements was performed. A typical cluster representative was determined as the arithmetic average of the cluster examined.

Overall, we can conclude that the construction industry has very good value - the optimal amount of assets, an acceptable financing structure and an adequate economic result. We can even state that ratio indicators show above-average values compared to other investment alternatives. Regarding the results of individual clusters, we can summarize as follows:

1. The most frequently represented companies represented in the clusters represent more than 80% of enterprises active in the construction industry.
2. The examined clusters showed lower, sometimes significantly lower values of the selected items of the financial statements (balance sheet, long-term assets, profit or loss and others).
3. This means that 80% of the enterprises surveyed do not generate 80% of the output of the construction industry.
4. Furthermore it means that the remaining 20% of enterprises are significantly more important for the construction industry than which corresponds with 20%.

We can therefore summarize that we should thoroughly investigate the remaining 20% of businesses and identify which enterprises (or rather clusters) have a significant positive impact on the results of the entire construction industry. If we do, it will allow us to predict the future development of the construction industry in the Czech Republic or rather the entire national economy.

## References

Adinoyi, Y. M., & Nor´aini, Y. (2015). Understanding and Formalization of Innovation: A Panacea to Inefficiency in Construction Industry Project Delivery. *26th International-Business-Information-Management-Association Conference*. Madrid, Spain, 3366-3375.

Construction. (2018). *Kurzy.cz* [online]. [2018-02-14]. Available at: https://www.kurzy.cz/makroekonomika/stavebnictvi/

Han, X., & Wang, L. (2008). Stock Company Comprehensive Assessment Model Based on Kohonen Network. *2008 Second International Conference on Genetic and Evolutionary Computing*. doi:10.1109/wgec.2008.97

Jaafar, M., Aziz, A. R., & Wai, A. L. (2008). Marketing practices of professional engineering consulting firms: Implement or not to implement? *Journal of Civil Engineering and Management, 14*(3), 199-206.

Kliestik, T., Misankova, M., Valaskova, K., & Svabova, L. (2017). Bankruptcy Prevention: New Effort to Reflect on Legal and Social Changes. *Science and Engineering Ethics*, 24(2), 791-803.

Mitkus, S., & Trinkūniene, E. (2008). Reasoned decisions in construction contracts evaluation. *Technological and Economic Development of Economy, 14*(3), 402-416.

Naderpajouh, N., Choi, J., & Hastak, M. (2016). Exploratory Framework for Application of Analytics in the Construction Industry. *Journal of Management in Engineering, 32*(2), 04015047.

Sikyr, M., & Busina, F. (2014). Human resource management in the Czech and Slovak building industry. *8th International Days of Statistics and Economics*. Prague, Czech Republic, 1477-1486.

Siskina, A., Juodis, A., & Apanavičiene, R. (2009). Evaluation of the competitiveness of construction company overhead costs. *Journal of Civil Engineering and Management, 15*(2), 215-224.

Vochozka, M., Rowland, Z., & Vrbka, J. (2015). Evaluation of credibility of civil construction companies in South bohemia. *Innovative Economic Symposium 2015*. České Budějovice, Czech Republic, 145-161.

Vochozka, M., Rowland, Z., Stehel, V., Suler, P., & Vrbka, J. (2016). *Modelování nákladů podniku pomocí neuronových sítí [Modeling the cost of an enterprise using neural networks]*. České Budějovice: ITB. ISBN 978-80-7468-112-7.

**Contact**

Marek Vochozka
Institute of Technology and Business in České Budějovice
Okružní 517/10, 370 01 České Budějovice, Czech Republic
vochozkam@email.cz