# Neural Network System to Forecast the Soybean Exportation on Brazilian Port of Santos

Emerson Rodolfo Abraham, João Reis, Adriane Paulieli Colossetti, Aguinaldo
Souza, Rodrigo Carlo Toloi

**HAL Id: hal-01707279**
**https://inria.hal.science/hal-01707279**

Submitted on 12 Feb 2018

# Neural Network System to Forecast the Soybean Exportation on Brazilian Port of Santos

Emerson Rodolfo Abraham [*1,2], João Gilberto Mendes dos Reis [1,6],
Adriane Paulieli Colossetti [3], Aguinaldo Eduardo de Souza [1,4], and
Rodrigo Carlo Toloi [1,5]

[1] Paulista University - UNIP, PPGEP, São Paulo, Brazil
[2] Metodista University, São Bernardo do Campo, Brazil
[3] FIAP, São Paulo, Brazil
[4] UNIBR, São Vicente, Brazil
[5] IFMT, Rondonópolis, Brazil
[6] UFGD - PGA, Dourados, Brazil
* Corresponding author:emerson.abraham@stricto.unip.br

**Abstract.** Agricultural products are an important part of the Brazilian economy. In soybean production, the country is the second largest producer with 114.0 million tons in the 2016/2017 harvest. Mato Grosso state is the largest Brazilian producer with 30.5 million tons and the port of Santos is mainly requested by being the largest port in Latin America. However, the poor infrastructure of the transport road causes bottlenecks when dispatching soybean through the major ports. Artificial Neural Networks (ANN) are used worldwide in logistics; therefore, we propose to design, train and simulate an ANN on MatLab©software to forecast the demand of soybean produced in Mato Grosso and exported through the port of Santos. The value of 9.0 million tons was predicted for 2017 as an increase of about 26.5% compared with the 2016 movement of 7.1 million tons. In addition, it was noticed that 5.9 million tons were moved only in the first five months (Jan-May) of transactions in 2017.

**Keywords:** Artificial intelligence · Intelligent systems · Artificial neural network · Soybean · Logistics.

## 1 Introduction

The expected global production of soybean in the 2016/2017 harvest was 351.3 million tons [1]. Brazil is the second largest producer of soybean with 114.0 million tons. A major Brazilian producer is the Mato Grosso state located in the midwest of the country with 30.5 million tons [2]. Despite these numbers, Brazil's challenge is drained this production that generally use the port of Santos, and it is in average about two thousand kilometers away.

In a competitive world, companies need excellence in logistics management and some factors are essential to their success such as cost cutting, meeting clients' requirements, fast delivery and high service quality [3]. In this context,

among several logistics activities, one of them is distinguished because it corresponds to 60% of the logistics costs, transport, which has a direct influence on client satisfaction [4]. In Brazil, for instance, the main model used for cargo traffic is road transport with 1.7 million kilometers of roads [5].

Research done by a non-governmental Brazilian organization, the National Confederation of Transport - CNT [6] analyzed approximately 103 thousand kilometers of roads and observed that 58.2% have shown some sort of problem in their general status (flooring, signaling and geometry of the ground). This problem has been increased in the past few years with the increment of transportation volume and the cargo capacity of vehicles. Additionally, the CNT study [6] showed that only the damages on flooring generates an average increase of 24.9% in the cost of transport and it is estimated at R$ 2.3 billion (around of 755 million US dollars in March 2017) losses to the transporters because of over 700 million liters of diesel wasted in 2016. Finally, research also indicated that in order to adjust the Brazilian road network, it would be necessary to spend R$ 292.5 billion (around of 94.7 billion US dollars in March 2017). It is worth mentioning that in 2015 only R$ 9.3 billion (around of 3.1 billion US dollars in March 2017) had been authorized.

At the same time, while road transportation raises many issues railroad transportation is neglected in the country. Even though it is considered the largest one in Latin America with an approximate territorial extension of 28 thousand kilometers, it is still very small compared with developed countries [5].

The influence of transport infrastructure on the Brazilian economy is huge because, historically, the country is moved by agricultural production and many products are produced far from ports.

The Brazilian port system has a fundamental importance in commerce. In 2016, the ports in Brazil moved 998 million tons; considering the 37 public ports, the port of Santos is the highlight moving 113.8 million tons in its commercial operations [7]. The structure of docking complex in Santos contributes to the position of the largest port in Latin America [8]. With a building area of 7.8 million $m^2$, the port of Santos was the one that received the most ships in 2016, about 13 ships per day [7]. This port has an area of primary influence that includes the states of Mato Grosso, Mato Grosso do Sul, Goias, Minas Gerais and São Paulo. Together, they represent 67% of the Gross National Product and 56% of the Brazilian Commerce Balance [7].

Port of Santos has been in great demand and consequently it can suffer bottlenecks in its operations. Besides that, deficiencies related to the Brazilian road transport infrastructure and the lack of investments in the sector stimulates the research for solutions that do not depend on government's initiative.

Thus, the use of intelligent systems appear as an important tool to optimize the use of transportation networks. Intelligent Systems are used in optimization solutions and logistics processes worldwide and the most used techniques are Genetic Algorithms, Fuzzy Logic and ANN. These systems are part of the

research in Artificial Intelligence (A.I) and have been mainly used in pattern recognition and demand prediction especially in non-linear problems [3].

The objective of this work is to build, train and simulate an ANN able to forecast the soybean exportation demand (metric tons) from Mato Grosso in the port of Santos. This system can help policy-making improve forecasts and consequently, establish a priority for the investments in this soybean traffic route.

## 2 Methodology

This paper uses ANN knowledge to predict Brazilian Soybean exportation by the Santos Port.

### 2.1 ANN

Intelligent systems such as ANN are able to solve problems uniquely and "creatively". They have been intended to emulate human behavior in making decisions through the learning processes. In addition, they are fault-tolerant because they can extract useful results from an incomplete data set. ANNs should be designed to solve dynamic problems and are not suitable for classical problems. They are trained from situations that have already happened (historical data) and the main approaches are to solve problems related to pattern recognition and prediction of future values according to past occurrences [9,10,11]. The computational logic of an ANN is similar to human neural networks, in other words, it consists of several computational units known as connected neurons. Each neuron (Fig. 1) is associated with specific weights (intensity and signal of the connection) in several input values. The neurons are activated in function of inferences and propagate the signal to others until the last units. [9,10,11].
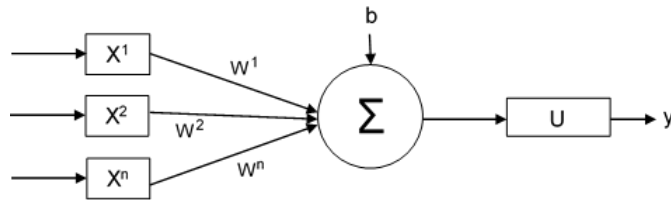


Fig. 1: Artificial Neuron. Adapted from: [12]

### 2.2 Sample

The data were collected on Information Systems of Foreign Trade (ALICE WEB). This system is based on the information extracted from the Foreign Trade Integrated System (SISCOMEX) and is maintained by the Ministry of Industry,

Foreign Trade and Services. It divulges the Brazilian Statistics of export - import transactions [13].

A database query was performed using the following filter:

1. Posição (position) - SH 4 dígitos: 1201 - Soja, mesmo triturada (Soybeans, whether or not broken); 2. UF: 52 - Mato Grosso (state); 3. Porto: 4117 - Santos - SP (port); 4. Via (way): 1 Marítima; 5. Período (period): 1997 – 2016 (yearly production).

The system generated a spreadsheet in MS Excel, which was easily imported into an array in the Matlab software. In this study, the soybean exportation data analysis is presented in metric tons and links the Mato Grosso state production with the port of Santos in the last 16 years (2001-2016). The first years of the series (1997, 1998, 1999 and 2000) were used for delays and were not plotted.

### 2.3   Computacional Tool

To build, train and simulate an ANN we used the software Matlab©R2016b and its neural network toolbox. This toolbox provides a graphic interface and allows the creation of a time series application. A Non-linear Autoregressive (NAR) solution was adopted because it is an excellent solution for a unique data set, without external interferences [14]. The ANN considered 80% of data for training, 5% for validation, and 15% for testing the network. In each section, data were selected randomly. According to Mirabdolazimi and Shafabakhsh [15] "Determination of proper structure, education algorithm, transmission functions and number of neurons in hidden layers are among the most important factors in the designing process of neural networks". The network was generated and trained in an open loop form with 50 neurons and four delays. The training algorithm used was the Bayesian regularization that provides the best output for a small data set and noises (Fig. 2) [14].
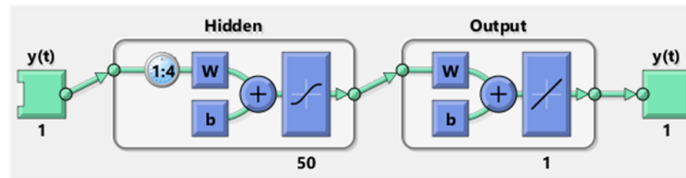


Fig. 2: Neural Network. MatLab©R2016b

After training, a closed loop form simulation was done; according to [14] "this function replaces the feedback input with a direct connection from the output layer creating a multi step prediction. A Step Ahead form simulation was done too; according to [14] the new network returns the same outputs as the original network plus one step ahead. These algorithms were provided by the Matlab software.

## 3  Results and Discussion

The training reached the best value for minimum errors after 152 interactions and 6 seconds of processing in order to minimize the mistakes between the data (target) and the feedback. Also an estimated linear regression with the correlation between the variables was performed (Fig. 3).
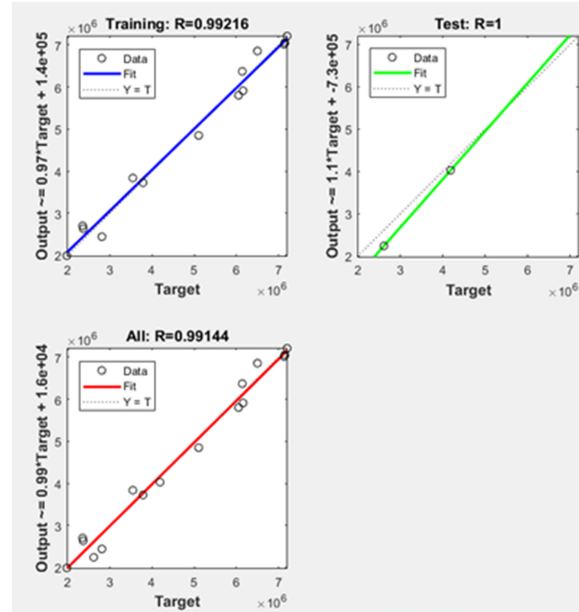


Fig. 3: Regression and Correlation. MatLab©R2016b

The results shows that both training and test surpass the minimum value of correlation (0.9), which means that the model is positively correlated with soybean data analysis. The next step was to plotted a time series using the data according to Fig. 4.

This is the original series that represents the cargo traffic of the soybean exportation from Mato Grosso to the Port of Santos in a period of 16 years (2001-2016). According to Laboissiere et al. [12] the analysis of time series is referred to as a sequence of data specified at regular time intervals during a period. Consequently, the time series analysis is used to determine structures and patterns in historical data and develop a model that predicts their behavior. They are normally treated by means of regression models. The error represents the difference between the data (target) and the feedback generated by the neural network. The time is represented by the x- axis and the y- axis represents the amount in tons. It is noticeable that the movement in tons has increased over the years. Another interesting observation is that there was a significant increase in demand after 2007. Although these tendencies exist, the
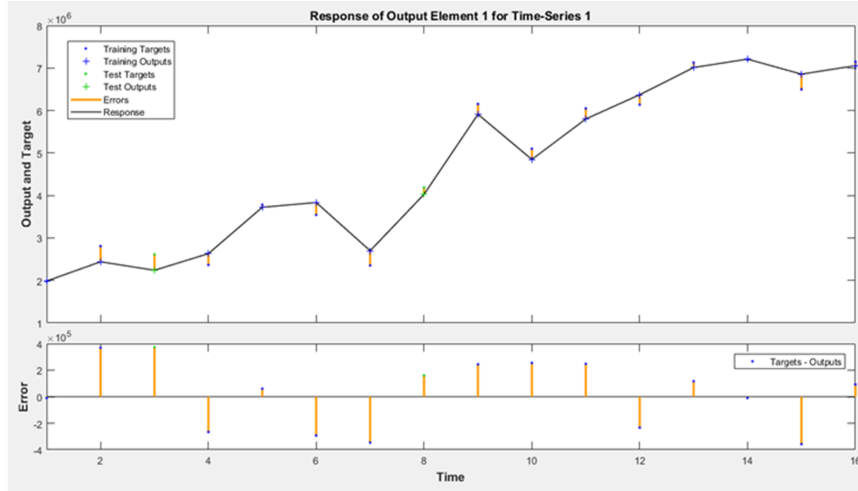
Fig. 4: Original Time Series. MatLab©R2016b

series showed some aleatory data which is interesting to analyze with ANN that deal with non-linear problems. Afterwards we used the data to plot a time series in closed loop form (Multi Step Prediction) as shown in Fig. 5.
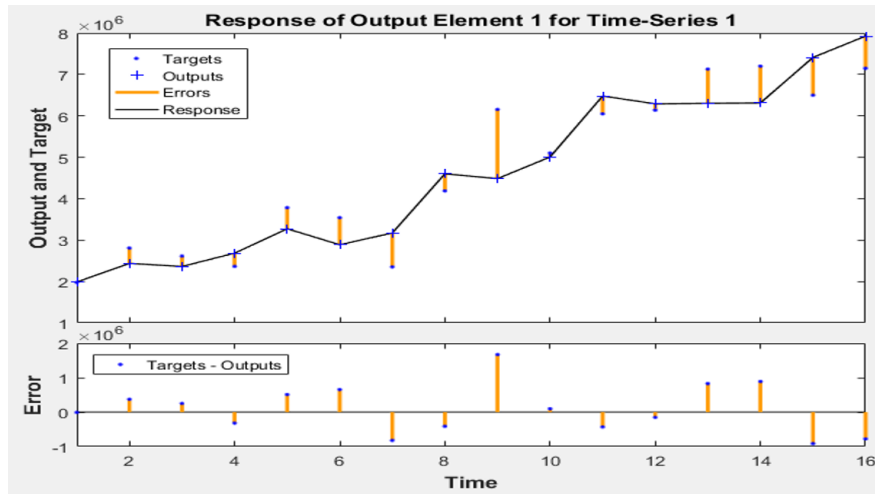


Fig. 5: Multi Step Prediction. MatLab©R2016b

The multi step prediction was performed each year of the series based on previous years. The first four years that were not plotted served as a delay for the 2001 forecast. The other years were predicted on the basis of all the years

prior to the analysis. This type of simulation is interesting to determine the real capacity of the network to make predictions, since it is possible to perceive the actual data (target) and the predicted data (output). The difference is the error in the forecast. The calculated mean error was 12.57%. Finally, we used the algorithm provided by Matlab software to generate a time series step ahead as shown in Fig. 6
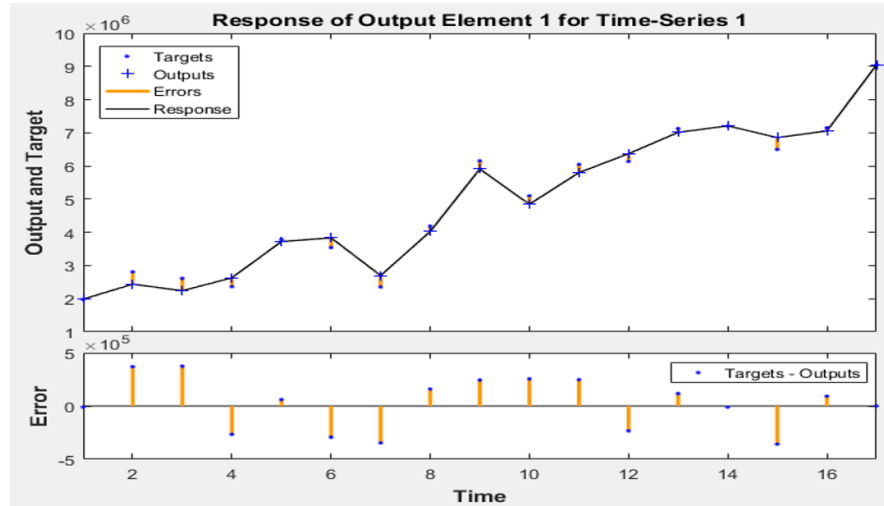


Fig. 6: Step Ahead Prediction. MatLab©R2016b

This time series represents the feedback of the data which allows to inferred demand prediction in 2017 for the port of Santos. The value of 9.0 million tons was predicted for 2017 as an increase of about 26.5% compared with the 2016 movement of 7.1 million tons. Considering the mean error of 12.57% in making predictions, demand is expected to vary between 7.9 and 10.1 million tons. A database query was performed on ALICEWEB on Jun / 2017 and it was noticed that 5.9 million tons were moved only in the first five months (Jan-May) of transactions in 2017, meaning an increasing demand and the alignment of the model proposed to make predictions.

## 4   Conclusions and Outlook

Brazil is the second largest producer of soybean in the world and the state of Mato Grosso is responsible for over a third of the country's production. The largest part of soybean from this state is dispatched to the main port of Latin America, the port of Santos. However, the trade in Brazil faces issues in dispatching the production to the port of Santos by road transport. This system is the most widely used in Brazil and presents poor logistics infrastructure. Be-

sides that, bottlenecks can occur in the port due to delays in operations. Therefore, this work quested to use an ANN algorithm that simulates human behavior to generate a time series that represents the movement of soybean coming from Mato Grosso to the port of Santos in the last 16 years. It was verified that the cargo traffic has been increasing year after year practically uninterrupted. The errors between the data inserted in the system and the feedback of the network is related to the technology that simulates the human knowledge for making decisions, therefore being dynamic and subject to learning. Furthermore, in future research we intend to use a larger data set in the time series and compare with the performing training results.

## Acknowledgments

## References

1. USDA United States Department of Agriculture: World Agricultural Supply and Demand Estimates  566 (2017)
2. EMBRAPA: https://www.embrapa.br/soja/cultivos/soja1/dados-economicos
3. Ko, M., Tiwari, A., Mehnen, J.: A Review of Soft Computing Applications in Supply Chain Management. Applied Soft Computing 10(3), 661–674 (2010)
4. Ballou, R.H.: Logística Empresarial, Transportes, Administração de Materiais e Distribuição Física. Ed.Atlas, São Paulo (2007)
5. Ministério dos Transportes: http://www.transportes.gov.br
6. CNT Confederação Nacional dos Transportes and SEST Serviço Social do Transporte and SENAT Serviço Nacional de Aprendizagem do Transporte : Pesquisa CNT de Rodovias 2016 : relatório gerencial. – 20.ed. (2016)
7. CODESP Companhia Docas do Estado de São Paulo: Análise do movimento físico do porto de santos dezembro de 2016 (2016)
8. CEPAL: http://www.cepal.org
9. Artero, A.O.: Inteligência Artificial: Teórica e Prática. São Paulo: Livraria da Física (2009)
10. Russell, S., Norvig, P.: Artificial Intelligence: A Modern Approach. Prentice Hall, 3 edition edn. (2011)
11. Wang, J., Tsapakis, I., Zhong, C.: A space–time delay neural network model for travel time prediction. Engineering Applications of Artificial Intelligence 52, 145–160 (2016)
12. Laboissiere, L.A., Fernandes, R.A., Lage, G.G.: Maximum and minimum stock price forecasting of brazilian power distribution companies based on artificial neural networks. Applied Soft Computing 35, 66 – 74 (2015)
13. ALICE WEB: http://aliceweb.mdic.gov.br/
14. Beale, M.H., Hagan, M.T., Demuth, H.B.: Neural Network Toolbox. Getting Started Guide (2017)
15. Mirabdolazimi, S., Shafabakhsh, G.: Rutting depth prediction of hot mix asphalts modified with forta fiber using artificial neural networks and genetic programming technique. Construction and Building Materials 148, 666 – 674 (2017)