Short-term forecasting of intermodal freight using ANNs and SVR: Case of the Port of Algeciras Bay.

J.A. Moscoso-López

Intelligent Modelling of Systems Research Group. University of Cádiz. Polytechnic School of Engineering (Algeciras) – Avda. Ramón Puyol s/n, Algeciras, 11202 Spain

I.J. Turias[1], M. J. Jiménez-Come[1], J. J. Ruiz-Aguilar[1]

[1]Intelligent Modelling of Systems Research Group. University of Cádiz. Polytechnic School of Engineering (Algeciras) – Avda. Ramón Puyol s/n, Algeciras, 11202 Spain

M. Cerbán

Research Group Transport and Innovation Economic. University of Cádiz. Faculty of Economics – Avda. Duque de Nájera s/n, 11002 Cádiz, Spain

ABSTRACT

Forecasting of future intermodal traffic demand is very important for decision making in ports operations management. The use of accurate prediction tools is an issue that awakens a lot of interest among transport researchers. Intermodal freight forecasting plays an important role in ports management and in the planning of the principal port activities. Hence, the study is carried out under the motivation of knowing that modeling the freight transport flows could facilitate the management of the infrastructure and optimize the resources of the ports facilities. The use of advanced models for freight forecasting is essential to improve the port level-service and competitiveness. In this paper, two forecasting-models are presented and compared to predict the freight volume. The models developed and tested are based on Artificial Neural Networks and Support Vector Machines. Both techniques are based in a historical data and these methods forecast the daily weight of the freight with one week in advance. The performance of the models is evaluated on real data from Ro-Ro freight transport in the Port of Algeciras Bay. This work proposes and compares different approaches to determine the best prediction. In order to select the best model a multicomparison procedure is developed using several statistical test. The results of the assessed models show a promising tool to predict Ro-Ro transport flows with accuracy.

1. INTRODUCTION

Nowadays, port management demands accurate making decision tools as traffic flow forecasting that could improve level of service. Ports play an important role in the supply chain and the knowledge of the flows with several days in advance could be an important competitive advantage. In European Union, all the import freight from third countries must be checked in ports facilities by border inspection post. The agility of the import processes in ports largely depends on the traffic flow behavior in short terms. Therefore, the anticipation of the future values of traffic is an important making decision tool in

planification of the transport system (Vlahogianni et al., 2004). Import process is very complex in the port phase due to the multiple delays and variations of the scheduled arrivals. These problems imply continuous variations and modifications in the port operations planning (Bilegan et al., 2008; Romero et al., 2013). The prediction in ports is gaining importance in the case of perishable freight because if delays or bottlenecks exist the values of the cargo may lose value. During last decades, the solution of many traffic problems has been addressed by several forecasting approaches as statistical methods and artificial neural networks (ANNs). ANNs are commonly used to solve non-linear functions as freight flows (Amin et al., 1998; Moscoso-López et al., 2014; Ruiz-Aguilar et al., 2014, 2015, 2016; Sun et al., 2012; Vlahogianni et al., 2005, 2004). Recently, Support Vector Machines (SVM) have been applied in solution of forecasting problems in transport field obtained good performance (Bhattacharya et al., 2014; Castro-Neto et al., 2009; Marković et al., 2015).. The remainder of the work is organized as follows. Section 2 discusses the methodology of ANNs and SVM, Section 3 gives a brief description of the database and experimental procedure, Section 4 discusses the results and finally conclusions are drawn in Section 5.

2. METHODOLOGY

2.1. Artificial Neural Networks

The ANNs are universal approximators, this is one of the advantages of the ANN models in relation of other nonlinear models. (Hornik et al., 1989). In addition to this, ANNs have a good performance in complex problems due to their learning features based on historical data.
In this paper, a multilayer perceptron (MLP) with feedforward connections has been used due to MLP is the most used configuration model in literature. This configuration consists in an input layer, an output layer and one or several hidden layer. The most widely used model for time series modelling and forecasting is a single hidden layer network.

The configuration of the hidden layer is composed by the combination of several numbers of neurons in order to obtain the best generalization error (using different samples that it used in the training stage). Backpropagation learning procedure (Rumelhart et al., 1986) has been used in this work. In Backpropagation Neural Networks (BPNN), the weights are adjusted to minimize the mean squared error between the target output and the network output.

To calculate the error function instead to obtain it analytically, an iterative process of minimizing a measure of the error made by the model, based on the available samples can be used. The Levenberg-Marquardt (LM) algorithm has been used in order to optimize the training procedure. The Gauss-Newton approximation used in LM algorithm provide sufficient robustness and velocity (Hagan and Menhaj, 1994). A regularization approach has been applied in order to obtain the configuration with best generalization. Early Stopping is

the regularization approach used in this work.

This neural approach models the relationship between *X* and *Y* in the form (through a weighted structure of layers, usually input-hidden(s)-output),

$$Y = g(\sum_{j=0}^{M} w_{kj} \cdot f(\sum_{i=0}^{D} w_{ij} \cdot X_i)) \tag{1}$$

Where W_{ij} is the matrix of the connection weight input layer with *D* being the total number of the inputs units and W_{kj} the matrix of the connection weight hidden layer with *M* number of units.

2.2. Support Vector Machines for Regression (SVR)

Support Vector Machines (SVM) are a powerful machine learning method based on statistical learning theory (Vapnik, 1998). SVM implements the structural risk minimization principle instead of the usual empirical risk by minimizing the fitting error of ANNs, thus increasing the generalization of the model. Originally, SVM were initially developed to solve classification problems, nevertheless with the introduction of Vapnik´s ε-intensive loss function SVM have been undertaken for regression problems. SVM for regression are called SVR.

The problem to be solved is shown in the following equation:

$$y(x) = w^T \phi(x) + b \tag{2}$$

Where *w* is the weight vector, *b* is the bias term and $\phi(x)$ represents the kernel function applied on the input vectors, *x*. The vector *w* and the constant *b* are estimated by minimizing the following regularized risk function.

$$R = \frac{1}{2}\|w\|^2 + \frac{1}{N} C \sum_{i=1}^{N} |y_i - f(\mathrm{x}_i)|_{\varepsilon} \tag{3}$$

Where *C* is the regularization parameter and $y_i\text{-}f(x_i)$ represents the loss function. The ε-intensive loss function as defined as:

$$|y - f(x)|_{\varepsilon} = \begin{cases} y - f(x) - \varepsilon, & \text{if } |y - f(x)| \geq \varepsilon \\ 0, & \text{otherwise} \end{cases} \tag{4}$$

Where ε is related to the tolerance error. Using Lagrange multiplier techniques, the minimization of equation (3) leads to the following dual optimization problem. Minimize,

$$\min_{w,b,\xi,\xi^*} \frac{1}{2}\|w\|^2 + C\sum_{i=1}^{N}(\xi_i + \xi_i^*) \tag{5}$$

Subject to the constraints

$$\begin{cases} y_i - (w^T\phi(x_i) + b) \le \varepsilon + \xi_i \\ (w^T\phi(x_i) + b) - y_i \le \varepsilon + \xi_i^* \\ \xi_i, \xi_i^* \ge 0 \end{cases} \tag{6}$$

For $i=1,2,\ldots,N$

Training errors above ε are denoted as ξ_i^* whereas training bellow $-\varepsilon$ are denoted as ξ_i. Once the quadratic optimization problem with the constraints is solved, the parameter vector w in (2) is obtained:

$$w = \sum_{i=1}^{N}(\beta_i^* - \beta_i)\phi(x_i) \tag{7}$$

Where β_i^*, β_i are obtained by solving a quadratic program and are the Lagrangian multipliers.

Finally, the SVR regression function is obtained as:

$$f(x) = \sum_{i=1}^{N}(\beta_i^* - \beta_i)K(x_i, x_j) + b \tag{8}$$

Where $K(x_i, x_j)$ is called the kernel function, and the value of the equals of the kernel equals the inner product of two vectors, x_i, and x_j, in the feature space $\varphi(x_i)$ o $\varphi(x_j)$. This kernel function maps the input data into a higher dimensional feature space where a linear regression can be performed.

In this work the Gaussian radial basis functions (RBF) has been used as kernel function As follow,

$$K(x_i, x_j) = \exp\left(\frac{-(x_i - x_j)^2}{2 \cdot \sigma^2}\right) \tag{9}$$

Where σ is the width of Kernel function.

3. EXPERIMENTAL PROCEDURE

The Strait of Gibraltar is one of the main maritime routes in the world, is located between the Mediterranean Sea and the Atlantic Ocean in the narrowest connection point between Europe and Africa. At the same time, the Port of Algeciras Bay is the main access to the European Union from North Africa ports (Tanger-Med mainly) in Ro-Ro freight transport. The maritime line Tanger-Med–Algeciras are covered in less of two hours of sailing. In 2015, more than 239,000 trucks crossed the Strait of Gibraltar between Port of Algeciras Bay and Tanger-Med by Ro-Ro vessels. Transportation of fresh vegetables is the largest Ro-Ro freight in the Strait of Gibraltar and therefore have an important impact in ports operations. This is the principal reason to choose fresh vegetables in this study, due to, the

transit time in port should be as short as possible. It´s very important that ports operations are carried out efficiently, otherwise the goods will lose their value as hours or days go by. The flows in this transportation are very unstable with a great variations during the week and during the year. The dataset used in this work was provided by the Port Authority of Algeciras Bay and is composed of the daily records during eight years of the Ro-Ro freight (in Kg) of fresh vegetables. The period of the dataset is from 2000 to 2007 both included.

The main challenge of this paper is to establish a reliable forecasting system which it could be used as making decision tool in ports environments. The number of input data defines the input layer features. In this work, the fresh vegetable cargo weight (in Kg) per day characterizes the input and the number of input nodes depends on the number of the past samples selected to be considered as inputs. Moreover, the output was the forecasted value of the vegetable freight (in Kg) for the prediction horizon of 7 days.

Two forecasting-models, based on the daily historical data, are presented and compared in order to predict the daily fresh vegetable weight with seven days of prediction horizon. The forecasting-models are Artificial Neural Network and Support Vector Machines for Regression. Each model is composed using different values of their parameters. The different configurations of the forecasting models are collected in Table 1, where are represented all the parameters of their architecture configurations.

	ANN	SVR
Time horizon	7	
Lags (*s*)	1;7	
Autoregressive window (*n*)	1,2,3,4,7,14,21,28,52	
Hidden units (*nhhiden*)	1,2,3,4,5,10,15,20	-
Kernel function	-	Gaussian
C	-	0.05;0.5;1:5;10:10:50
Epsilon (ε)	-	$2^{-8}, 2^{-7}, \ldots, 2^{-2}$
Gamma (γ)	-	$2^{-8}, 2^{-7}, \ldots, 2^{-2}$

Table 1: Parameter values of the forecasting models.

A total of 81 and 756 models have resulted from the combination of the variables of ANNs and SVR, respectively. The training in the case of ANNs is usually a stochastic and unstable process. As the weights of the network are initialized at random and training patterns are presented in random order, ANNs training will typically be different in value and performance. In addition, small changes in the training set can lead to completely different trained networks with different performance even if the nets had the same initial weights. For the models based on SVR technique, the performance depends on the combination of the parameters described in Table 1. In order to asses and compare the prediction performance of the proposed models, three performance indexes: Agreement Index (*d*), Correlation Coefficient (*R*) and Mean Absolute Percentage Error (*MAPE*).

$$d = 1 - \frac{\sum_{i=1}^{n}(F_t - O_t)^2}{\sum_{i=1}^{n}(|F_t - \bar{O}| - |O_t - \bar{O}|)^2} \tag{10}$$

$$R = \frac{\sum_{i=1}^{n}(O_i - \bar{O})(F_i - \bar{F})}{\sqrt{\sum_{i=1}^{n}(O_i - \bar{O})^2 \cdot \sum_{i=1}^{n}(F_i - \bar{F})^2}} \tag{11}$$

$$MAPE = \frac{100}{n} \sum_{i=1}^{n} \frac{|O_t - F_t|}{O_t} \tag{12}$$

Where O is the observed value, F the forecasted value and n the number of observations. Furthermore, in order to obtain the optimal configuration for each model two-fold cross-validation (2-CV) was applied. In this procedure the data is randomly partitioned into two equal groups. One group is applied to test the model, whereas the other group is used to training the model. This procedure was repeated twice, until each group has been used to test the model. This iterative procedure were performed 20 times for the estimation of the performance indexes.

4.RESULTS

Models based on ANNs and SVR were presented in this work to analyse the prediction for seven-day prediction of the time series for vegetable freight by Ro-Ro traffic in the Strait of Gibraltar. To construct the input of the time series several configurations were considered as; several size of autoregressive window (n=1,2,3,4,7,14,21,28,52) and two lags in the past (s=1 and s=7). The autoregressive window is n previous values taken in the past. The steps (lags) of the autoregressive in this work are 1 and 7 days. In order to obtain the best generalization ability, a 2-CV has been applied for all models. The number of the repetitions of each configuration was 30. Table 2 collects the performance index of the best models obtained for each technique (ANNs and SVR) and each lags (s=1 and s=7). In order to simplify the results only the models with the highest values have been collected. In turn, Table 2 shows the values of the parameters for each best model. According to Table 2, ANNs have obtained better values of d and R indexes than SVR in both cases of s. In particular, the best values for d and R are 0.9380 and 0.8898, respectively and they have been obtained with an ANN model using with a configuration of 2 neurons in the hidden layer, 52 samples of autoregressive size window with lag of 7 days. On the other hand, in a general way, MAPE index performed better with SVR models than ANNs. The best value of *MAPE* is 72.1644. It has been obtained by a SVR model with 1 sample of autoregressive size window, lag of 1 day and the values of the SVR parameters C, ε and γ are 40, 2^{-6}, 2^{-3} respectively.

Technique	Steps (lags)	Performance Index		Parameters				
		Index	Value	*nhiddens*	n	C	ε	γ
ANN	1	d	0.9275	2	28	-	-	-
		R	0.8750					
		MAPE	80.8648					
SVR (RBF)		d	0.9192	-	1	40	2^{-6}	2^{-3}
		R	0.8512					
		MAPE	**72.1644**					
ANN	7	d	**0.9380**	2	52	-	-	-
		R	**0.8898**					
		MAPE	88.1064					
SVR (RBF)		d	0.9188	-	1	30	2^{-8}	2^{-5}
		R	0.8504					
		MAPE	72.8593					

Table 2: Forecasting performance and parameters of the ANNs and SVR best models for 7 days of prediction horizon.

Figure 1 shows the comparison of the best models with the observed values for each technique and each lag. As Figure 1 shows, the values obtained with the SVR models are fitted better with the observed values than the values obtained with ANNs.
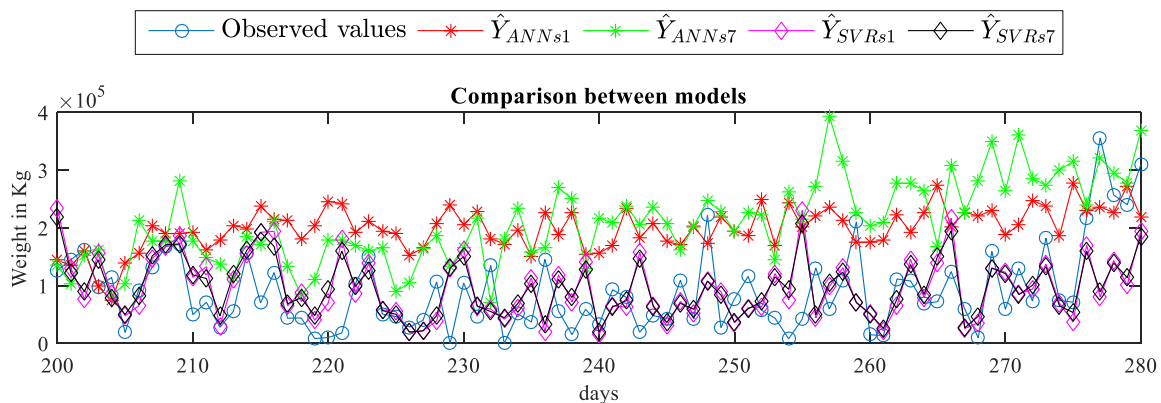

Figure 1: Comparison between prediction models and observed values.

5.CONCLUSION

The knowledge of the future demand in port operations is a challenge in port planning. Freight flow forecasting is an important making decision tool in the ports management and could improve the quality service in the logistic node of the Strait of Gibraltar. This work analyses and compares the prediction (in kg) of the import vegetable freight in Port of Algeciras Bay. The prediction has been carried out by Artificial Neural Networks and Support Vector Machines for Regression. Both methods have obtained well performance in the seven day ahead prediction. Furthermore, it has been proven that SVR could be a promising tool in freight flow forecasting problems and in particular to predict vegetable Ro-Ro flows in the Strait of Gibraltar slighly better than ANN models did.

6.REFERENCES

AMIN, S.M., RODIN, E.Y., LIU, A.-P., RINK, K., GARCÍA-ORTIZ, A. (1998). Traffic Prediction and Management via RBF Neural Nets and Semantic Control. *Comput. Civ. Infrastruct. Eng.* 13. pp. 315–327.

BHATTACHARYA, A., KUMAR, S.A., TIWARI, M., TALLURI, S. (2014). An intermodal freight transport system for optimal supply chain logistics. *Transp. Res. Part C Emerg. Technol.* 38. pp. 73–84.

BILEGAN, C.I., CRAINIC, T.G., GENDREAU, M., (2008). Forecasting freight demand at intermodal terminals using neural networks–an integrated framework.

CASTRO-NETO, M., JEONG, Y.S., JEONG, M.K., HAN, L.D., (2009). Online-SVR for short-term traffic flow prediction under typical and atypical traffic conditions. *Expert Syst. Appl.* 36. pp. 6164–6173.

HAGAN, M.T., MENHAJ, M.B., (1994). Training feedforward networks with the Marquardt algorithm. *IEEE Trans. Neural Networks* 5. pp. 989–993.

HORNIK, K., STINCHCOMBE, M., WHITE, H., 1989. Multilayer feedforward networks are universal approximators. *Neural Networks* 2. pp. 359–366.

MARKOVIĆ, N., MILINKOVIĆ, S., TIKHONOV, K.S., SCHONFELD, P. (2015. Analyzing passenger train arrival delays with support vector regression. *Transp. Res. Part C Emerg. Technol*. 56. pp. 251–262.

MOSCOSO-LÓPEZ, J.A., RUIZ-AGUILAR, J.J., TURIAS, I., CERBÁN, M., JIMÉNEZ-COME, M.J. (2014). A Comparison of Forecasting Methods for Ro-Ro Traffic: A Case Study in the Strait of Gibraltar, in: *Proceedings of the Ninth International Conference on Dependability and Complex Systems DepCoS-RELCOMEX*. June 30–July 4, 2014, Brunów, Poland. Springer. pp. 345–353.

ROMERO, G., DURÁN, G., MARENCO, J., WEINTRAUB, A. (2013). An approach for efficient ship routing. *Int. Trans. Oper. Res*. 20. pp. 767–794.

RUIZ-AGUILAR, J.J., TURIAS, I.J., JIMÉNEZ-COME, M.J. (2015). A two-stage procedure for forecasting freight inspections at Border Inspection Posts using SOMs and support vector regression. *Int. J. Prod. Res.* 53. pp. 2119–2130.

RUIZ-AGUILAR, J.J., TURIAS, I.J., JIMÉNEZ-COME, M.J. (2014). Hybrid approaches based on SARIMA and artificial neural networks for inspection time series forecasting. *Transp. Res. Part E Logist. Transp.* Rev. 67. pp. 1–13.

RUIZ-AGUILAR, J.J., TURIAS, MOSCOSO-LOPEZ, J.A., I.J., JIMÉNEZ-COME, M.J., CERBAN, M. (2016). Forecasting of Short-Term Flow Freight Congestion: A Study Case of Algeciras Bay Port (Spain). *DYNA* 83 (195).pp.163-172.

RUMELHART, D.E., HINTON, G.E., WILLIAMS, R.J. (1986). Learning internal representations by error propagation. *Parallel Distrib. Process. Explor. Microstruct. Cogn*. pp. 318–362.

SUN, S., HUANG, R., GAO, Y. (2012). Network-Scale Traffic Modeling and Forecasting with Graphical Lasso and Neural Networks. *J. Transp. Eng.* 138. pp. 1358–1367.

VAPNIK, V. (1998). *Statistical learning theory*. 1998. Wiley, New York.

VLAHOGIANNI, E.I., GOLIAS, J.C., KARLAFTIS, M.G. (2004). Short-term traffic forecasting: Overview of objectives and methods. Transp. Rev. 24. pp. 533–557.

VLAHOGIANNI, E.I., KARLAFTIS, M.G., GOLIAS, J.C. (2005). Optimized and meta-optimized neural networks for short-term traffic flow prediction: A genetic approach. Transp. Res. Part C Emerg. Technol. 13. pp. 211–234.