ESANN 2014 proceedings, European Symposium on Artificial Neural Networks, Computational Intelligence and Machine Learning. Bruges (Belgium), 23-25 April 2014, i6doc.com publ., ISBN 978-287419095-7. Available from http://www.i6doc.com/fr/livre/?GCOI=28001100432440.

Multi-Step Ahead Forecasting of Road Condition Using Least Squares Support Vector Regression

Sirvio, Konsta[1] and Hollmén, Jaakko[2]

Department of Information and Computer Science, Aalto University
School of Science, P.O.Box 15400, FI-00076 Aalto, Finland

Abstract. Network-level multi-step road condition forecasting is an important step in accurate road maintenance planning, where correct maintenance activities are defined in place and time of road networks. Forecasting methods have developed from engineering models to non-linear machine learning methods that make use of the collected condition and traffic data of the road network. Least Squares Support Vector Regression gives significantly the best results compared to Radial Basis Function networks or multiple linear regression.

1 Introduction

Road infrastructure is an import public asset of most of the countries and for example in case of Finland there are approximately 80 000 kilometres of public roads whose estimated balance value was approximately 15 billion euros in 2003 although higher estimates of 37 billion euros have been proposed.[11] The infrastructure deteriorates over time and the value depreciates unless funds are allocated on maintenance. In case of Finland annual net maintenance costs during 2012 were less than 500 million euros. A common problem is that the maintenance budget is too restricted in order to maintain the roads in adequate condition.[3] The poorer the road condition is the more road users pay in terms of increased Vehicle Operating Costs (VOC) as fuel consumption and vehicle maintenance increase. Overall, annual vehicle-related VOC is estimated to be 20 billion euros in Finland.[7] Percentually small decreases in VOC has thus high cost impact, but also environmental impact.

Maintenance planning is a procedure to practice predictive maintenance, where future maintenance operations are allocated on the road network. Ideally, the total costs to the whole society are minimised by finding correct maintenance treatments on the right road segments on the right time. This requires forecasting of road condition, cost modelling of the maintenance operations and spatio-temporal optimisation of the operations.

Although that road condition and traffic data is collected annually from the road networks by road authorities the methods are not using the full potential of the collected data as engineering models are used in practice. These models are often based on just a few variables and parameters that are calibrated to variable conditions.[6] Our approach is based on nonlinear machine learning methods making use of the data more thoroughly both spatially as temporally.

The paper introduces the utilised data in Chapter 2, after which the problem formulation is described in Chapter 3 and computational methods are presented

ESANN 2014 proceedings, European Symposium on Artificial Neural Networks, Computational Intelligence and Machine Learning. Bruges (Belgium), 23-25 April 2014, i6doc.com publ., ISBN 978-287419095-7.
Available from http://www.i6doc.com/fr/livre/?GCOI=28001100432440.

in Chapter 4. The forecasting results are given in Chapter 5 that is followed by the conclusions and discussion in Chapter 6.

2 Description of the data

Road Administration in Finnish Transport Agency collects annually road condition data. However, the collection interval on the same road segment varies according to the importance of the road and only the most important segments goes through data collection procedures at least once a year. The spatial distance of measurement points and the lenght of the aggregated variables vary depending on the variable, but when the condition data is stored in the road condition database the variables describing the surface condition used in network-level planning is averaged to 100 metre long road segments.[10]

When the surface condition of paved roads is concerned, the measured variables are typically different types of cracking, potholes, ravelling edge break, roughness and rutting. With the current survey technology the last two variables have been automatically collected in Finland by profilometres that have laser beams measuring the distance and thus determining the depth of rutting and so-called International Roughness Index (IRI). The other variables are collected by visual inspection although there is new technology emerging.[12]

In this study, the data from the road condition database is taken for the whole Finnish road network administered by Finnish Transport Agency between the years 2003 and 2010. IRI and rutting are the dependent variables to be forecasted. Explanatory variables are: road number, road section, carriageway number, survey direction, lane, start distance of the segment from the road start, region, survey day, survey month, survey year, survey season, previously measured rutting, previously measured IRI, previously mesured left IRI, previously measured right IRI, previously measured left water rutting, previously measured right water rutting, ridge, start inclination and end inclination.

IRI describes the road user experience of road condition and it is traditionally measured by a string in the survey equipment moving vertically due to road roughness. The variable is measured by how many millimetres the string moves on 1 metre driven on the road.[8] Rutting is defined as longitudinal depressions in the wheel paths of asphalt concrete pavements and it is a distance measure.[2] Currently, the measurements can be done by a laser beam.

The data set covers 8 years of data. The survey time interval varies between different segments as there can be one or two survey annually or more infrequently. Also, maintenance treatments had been taken place on some of the road segments. Therefore, the data was pre-processed so that 12 separate data sets were formed so that each set included road segments with equal number of condition surveys without any maintenance treatments between the first and last survey. It was noticed that some of the data contained measurement errors and therefore those segments were deleted that had an absolute difference of more than 5 mm/m for IRI in each data set. The criterion could be stricter, but as the study concentrates on comparison of different methods data problems would

ESANN 2014 proceedings, European Symposium on Artificial Neural Networks, Computational Intelligence and Machine Learning. Bruges (Belgium), 23-25 April 2014, i6doc.com publ., ISBN 978-287419095-7.
Available from http://www.i6doc.com/fr/livre/?GCOI=28001100432440.

not matter as all the methods face the same issue and therefore comparisons can be made. Immediate following measurements were excluded from the forecasts since the results would not have much practical value. The data was normalised before applying the methods.

3 Problem formulation

Two road condition surveys have been conducted and IRI and rutting have been measured. The date for both surveys is known, but the values for IRI and rutting of the latter survey is unknown. The idea is to forecast the condition survey values of IRI and rutting, y at time t+k. The n explanatory variables include time-invariable variables, past condition values and the temporal information that are all denoted by x^i at time t. The forecasted values are obtained by some function f and normal error term, ϵ is assumed as:

$$y_{t+k} = f(\sum_{i=1}^{n} x_t^i) + \epsilon_{t+k}, \tag{1}$$

4 Methods

4.1 Overview

Both linear and non-linear methods are used for comparison purposes. Previous value of the forecasted variables are used as the very basic baseline. Linear regression model is used as another baseline method. The main focus is at non-linear hybrid methods. Since the amount of data is over 1 200 000 rows in the biggest set the data sets are first clustered by k-means++ so that each cluster contains at maximum 12 000 data rows to make it computationally feasible.[1] For each cluster 10-fold cross-validation was applied so that each time 9/10 part of the data was used for training Radial Basis Function (RBF) and Least Squares Support Vector Regression (LS-SVR) and 1/10 part of the data was used for validation. Once this was done for all the 10 data chunks the process was repeated altogether 5 times. Mean and standard deviation of the Mean Squared Errors (MSE) were calculated and reported. Overview of the utilised methods is depicted in Figure 1. Traditionally, mathematical engineering models have been used for road condition forecasting, where calibration factors had been adjusted with the help of some explanatory variables so that the model represents a significant part of the network. That approach does not take into account the differences between each 100 metre long road segment such as the geometry. Also much potentially affecting data is excluded.

4.2 Least Squares Support Vector Regression

Instead of a quadratic programming problem of Support Vector Regression (SVR) a linear system is solved in Least Squares Support Vector Regression and the regression function takes the form:

ESANN 2014 proceedings, European Symposium on Artificial Neural Networks, Computational Intelligence and Machine Learning. Bruges (Belgium), 23-25 April 2014, i6doc.com publ., ISBN 978-287419095-7.
Available from http://www.i6doc.com/fr/livre/?GCOI=28001100432440.

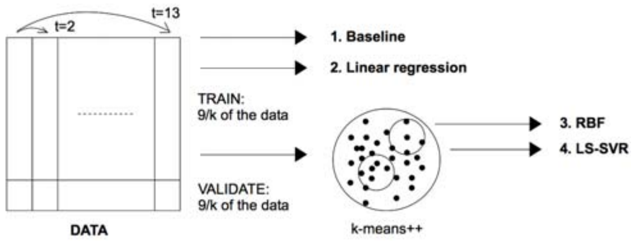

Fig. 1: Overview of the methods.

$$y(x) = \sum_{i=1}^{N} \alpha_i K(x, x_i) + b \tag{2}$$

The kernel function $K(x, x_i)$ can have several forms, but radial basis function $K(x, x_i) = e^{-\frac{\|x-x_i\|_2^2}{\sigma^2}}$ is commonly used. Parameters α_i and b are solution to:

$$\begin{bmatrix} 0 & \overrightarrow{1}^T \\ \overrightarrow{1} & \Omega + \gamma^{-1} I \end{bmatrix} \begin{bmatrix} b \\ \alpha \end{bmatrix} = \begin{bmatrix} 0 \\ y \end{bmatrix}$$

where $y = [y1, ..., y_N]$, $\overrightarrow{1} = [1, ..., 1]^T, \alpha = [\alpha_1, ..., \alpha_N]^T$ and the Mercer condition $\Omega_{kj} = \varphi(x_k)^T \varphi(x_j)$ and $k, j = 1, ..., N$.[5]

5 Results

Table 1 summarises forecasting errors of IRI and rutting for the selected methods. The values are mean values of Mean Square Errors (MSE) of all the cross-validation runs. Average tandard deviations of the MSE are presented in parenthesis. Standard deviation is calculated from all the individual forecasts in case of the Baseline. Forecasts producing minimum errors are shown as bold.

In case of RBF the radius and number of iterations were selected heuristically to be 5 and 10 respectively. In case of LS-SVR two parameters were optimised in one data set (t=12) and used with all the sets. The regularisation parameter,γ for IRI was 4.71 and 5.02 for rutting. In case of RBF kernel function parameter,σ^2 the optimised value was 14.07 for IRI and 4.25 for rutting. Compared methods were the following: M1: No change from the previous measurement, M2: Linear regression, M3: K-means++ -clustering and Gaussian RBF for each cluster, M4: K-means++ -clustering and Least Squares Support Vector Regression for each cluster

According to the results it can be said that linear regression defeats the baseline and LS-SVR defeats the RBF with the reported problem. Comparison of the results of the linear regression and RBF it can be seen that RBF outweighs linear regression in case of rutting with all the other data sets except with t=9. In case of IRI RBF gives better results than linear regression with t=7, 10, 11,

ESANN 2014 proceedings, European Symposium on Artificial Neural Networks, Computational Intelligence and Machine Learning. Bruges (Belgium), 23-25 April 2014, i6doc.com publ., ISBN 978-287419095-7. Available from http://www.i6doc.com/fr/livre/?GCOI=28001100432440.

Data set	M1	M2	M3	M4
2 (IRI)	0.129 (0.70)	0.117 (0.66)	0.186 (0.69)	**0.037 (0.17)**
2 (RUT)	3.55 (14.1)	2.74 (12.85)	2.59 (8.08)	**0.79 (2.74)**
3 (IRI)	0.122 (0.67)	0.111 (0.64)	0.256 (0.84)	**0.031 (0.15)**
3 (RUT)	5.97 (16.12)	3.66 (14.43)	3.26 (8.97)	**0.92 (2.81)**
4 (IRI)	0.143 (0.68)	0.124 (0.63)	0.155 (0.91)	**0.044 (0.21)**
4 (RUT)	9.04 (17.91)	4.55 (14.65)	3.89 (9.71)	**2.08 (4.86)**
5 (IRI)	0.138 0.67)	0.155 (0.91)	0.150 (0.76)	**0.034 (0.13)**
5 (RUT)	12.07 (19.89)	5.30 (14.83)	4.62 (10.73)	**3.14 (8.86)**
6 (IRI)	0.161 (0.67)	0.044 (0.21)	0.140 (0.58)	**0.046 (0.19)**
6 (RUT)	15.26 (22.90)	5.92 (25.60)	5.18 (11.73)	**1.42 (3.79)**
7 (IRI)	0.150 (0.65)	0.117 (0.49)	0.112 (0.40)	**0.040 (0.16)**
7 (RUT)	18.04 (25.20)	6.24 (14.57)	5.40 (12.19)	**1.45 (4.00)**
8 (IRI)	0.161 (0.69)	0.122 (0.50)	0.135 (0.43)	**0.048 (0.17)**
8 (RUT)	19.77 (25.96)	6.35 (13.54)	5.47 (11.65)	**1.55 (4.15)**
9 (IRI)	0.168 (0.73)	0.128 (0.54)	0.141 (0.47)	**0.051 (0.18)**
9 (RUT)	22.20 (28.63)	6.67 (13.46)	6.84 (13.46)	**1.67 (4.31)**
10 (IRI)	0.187 (0.78)	0.137 (0.57)	0.129 (0.49)	**0.059 (0.20)**
10 (RUT)	25.39 (32.04)	6.66 (13.26)	5.91 (12.28)	**2.06 (4.71)**
11 (IRI)	0.218 (0.86)	0.148 (0.54)	0.126 (0.48)	**0.073 (0.26)**
11 (RUT)	26.39 (34.58)	6.90 (12.93)	6.23 (11.42)	**2.09 (5.11)**
12 (IRI)	0.270 (1.18)	0.162 (0.43)	0.152 (0.39)	**0.099 (0.25)**
12 (RUT)	27.31 (38.17)	6.90 (12.91)	6.35 (12.01)	**2.08 (4.86)**
13 (IRI)	0.411 (1.87)	0.159 (0.36)	0.127 (0.28)	**0.070 (0.17)**
13 (RUT)	44.07 (51.95)	7.70 (12.49)	5.53 (8.99)	**2.88 (5.73)**

Table 1: MSE of forecasted IRI values.

12 and 13. Standard deviation of the MSE decrease similarly to the mean MSE. Increasing number of the data set indicates longer forecasting time horizon in average and it would be logical to assume decreasing accuracy, but this is not always the case. One explanation is problems with the surveyed condition data as the pre-processing criterion was arbitrary as it cannot be surely known, which values are erroneous.

6 Conclusions

The non-linear methods performed significantly better than the linear ones. In case of IRI, mean MSE of the best method was between 17 and 37 percent of the mean MSE for the baseline and between 27 and 66 percent of the mean MSE for the second best method. In case of rutting the best method performed also well as the corresponding values were between 6 and 26 percent compared to the baseline and between 25 and 68 percent compared to the second best method depending on the data set i.e. the forecasting time step. When compared to previous study the methods seem to give significantly better results even with

ESANN 2014 proceedings, European Symposium on Artificial Neural Networks, Computational Intelligence and Machine Learning. Bruges (Belgium), 23-25 April 2014, i6doc.com publ., ISBN 978-287419095-7.
Available from http://www.i6doc.com/fr/livre/?GCOI=28001100432440.

longer time horizons.[9] Road administrations with historical condition data sets should switch to non-linear forecasting models since network-level road condition forecasting is done for planning purposes infrequently and therefore the increased computation time is not an issue.

Support Vector Regression -based methods are a luring option for forecasting since domain-specific knowledge is not needed in model creation. Better results could probably be achieved by using different clustering and optimising the parameters for each data sets separately.

Besides Finland the methods are applicable to road networks of other countries with collected condition, location and other data. Besides road networks the methods can be applied to predictive maintenance of other infrastructure assets, where condition is monitored systematically.

7 Acknowledgements

We would like to express our gratitude to J. Meriläinen, V. Männistö and T. Toivonen from Finnish Transport Agency for providing data and information.

References

[1] Arthur, D. and Vassilvitskii, S. (2007). k-means++ The Advantages of Careful Seeding. Symposium on Discrete Algorithms (SODA).

[2] Archilla, Adrian R.; Madanat, Samer. Development of a Pavement Rutting Model From Experimental Data. University of California Transportation Center. UCTCReprint No. 452, 2000.

[3] Finnish Transport Agency. Liikenneviraston Toiminta- ja Taloussuunnitelma Vuosille 2012-2015. Helsinki, 2011. (In Finnish)

[4] Haykin, Simon. Neural Networks. A comprehensive foundation, Prentice-Hall Inc. London, United Kingdom, 1999.

[5] Jiang, Jingqing et al. Adaptive and Iterative Least Squares Support Vector Regression Based on Quadratic Renyi Entropy. In proceeding of: The 2008 IEEE International Conference on Granular Computing, GrC 2008, Hangzhou, China, 26-28 August 2008.

[6] Kerali, Henry G.R.; Odoki, J.B.; Stannard, Eric E.Volume one. Overview of HDM-4. The Highway Development and Management Series. The World Road Association 2006.

[7] Linturi, Risto; Kuusi, Osmo; Ahlqvist, Toni. Suomen Sata Uutta Mahdollisuutta: Radikaalit Teknologiset Ratkaisut. Eduskunnan tulevaisuusvaliokunnan julkaisu 6/2013. Helsinki, 2013. (In Finnish)

[8] Sayers, M.W., and Karamihas, S.M. Little Book of Profiling. University of Michigan Transportation Research Institute, 1998.

[9] Sirvio, Konsta; HollmÃ©n, Jaakko. Spatio-temporal road condition forecasting with Markov chains and artificial neural networks. In Emilio Corchado, Ajith Abraham, and Witold Pedrycz, editors, Third International Workshop in Hybrid Artificial Intelligent Systems (HAIS'08), volume 5271 of Lecture Notes in Artificial Intelligence, pages 204-211. Springer-Verlag, September 2008. Burgos, Spain.

[10] Törnqvist, Jouko et al. Teiden käyttöikäsuunnittelu - lähtökohtia tavoitetilan märittelyyn. VTT Technical Research Centre of Finland, 2008. (In Finnish)

[11] Uimonen, Sakari. Suomen infrastruktuuripääoma: Tiet. Helsinki, VATT, Government Institute for Economic Research, 2007. (In Finnish)

[12] Äijö, Juha. Automated Crack Measurement Test in Finland 2004. Finnra Reports 5/2005. ISSN 1459-1553, ISBN 951-803-443-5, Finnish Road Administration, 2005.