Adaptive Kalman Filter Forecasting for Road Maintainers

Jānis Pekša
Institute of Information Technology
Faculty of Computer Science and
Information Technology
Riga Technical University
Kalku street 1, Riga, Latvia
Janis.Peksa@rtu.lv

Abstract—**The article considers the road monitoring weather-stations which collects raw observations that are processed to be able to make the necessary forecasting for future decisions. For the road maintainers those predictions are crucial to make decisions daily. When it comes to the winter season when road safety is very important; however, the road condition is also affected by the snow and icing. In order to improve safety on the road network the road maintainers are trying to use every possible way to be able to provide it. A number of methods have been studied and compared to clarify the parameter required by Kalman filter, which can be improved by making forecasting more accurate. Several road monitoring weather-stations are merged into one region because they are relatively close to each other and it is assumed that there are common conditions in one region that may indicate changes in road conditions. The corresponding algorithms are applied for each region and then compared to each other. Adaptive Kalman filter is generalized in the relevant article in order to have a general understanding of how to correctly apply the approach. The main result of this article is a comparison with the different methods, which are finally compiled in a single table.**

Keywords— Kalman filter, noise covariances, time-series, forecasting.

I. Introduction

The road maintenance is a complex decision-making process that needs to be done within a specific time frame. Pastime may affect the essential operations of the road network in total. One of them is driving conditions when traffic speed is rapid. The higher the speed rate the greater the probability that accidents are possible [1]. As know, the road conditions are subject to rapid changes in surface temperature and precipitation sum. The road conditions most fluctuations are observed during the winter period; however, the road condition is also affected by the snow and icing. The road maintenance is performed for specific Road Sections belonging to a Region. The region is the Republic of Latvia road network and the dataset is taken from VAS "Latvijas valsts celi". The Region consists of 52 road monitoring weather-stations that are relatively distant from each other. To be able to respond to changes

in the environment on road sections near the road surface are located road monitoring weather-stations and cameras operated by different entities. The road monitoring weather-stations collects raw observations which are processed to be able to make the necessary forecasting for future decisions. For road maintainers, those predictions are crucial to making decisions daily. The road monitoring weather-stations and cameras operated by different entities can help for the decision-making. The road maintainers controlling smart road signs are available to give warning messages to drivers on a specific stage of the road network [2]. For instance, driving conditions are evaluated according to surface temperature and precipitation sum provided by roadside weather-stations as well as weather service. However, the weather service provides data only at the regional level. In the road maintenance case, one of the adjustments evaluates a need to de-ice the roads [3]. The road conditions context element is computed using surface temperature (°C) and precipitation sum (mm). The warning adjustment is triggered only if the locking conditions are met preventing unnecessary nervousness. The road maintainers receive signals via interfaces from Enterprise Resource Planning Systems (ERPs) that have received a warning that one of the parameters has changed dramatically and reached a critical rule [4]. The critical rule depends on the definition of the road maintainers. To be able to predict future parameters forecasting comes in handy for this task.

As already known forecasting consists of pre-processing data and missing data should be completed to perform predictions successfully [5]. In the previous work, missing data are filled for the data set time-series parameters. To be effective prediction will be applied Kalman filter. The Kalman filtering, also known as linear quadratic estimation, is an algorithm that uses a series of measurements observed over time, containing statistical noise and other inaccuracies [6]. To be able to create Adaptive Kalman filter one of the parameters should be modified dynamically changing the noise parameter also known as the noise covariances. The noise covariances are found by those algorithms:

- Simple moving average – arithmetic moving

average calculated by adding recent closing prices and then dividing that by the number of periods in the calculation average [7];

- Arithmetic mean – the most common type of standard is the arithmetic mean, in other words, an average value [8];
- Bayesian average – a method of estimating the mean of a population using outside information [9];
- k-Means Clustering algorithm – a method of vector quantization, originally from signal processing, that is popular for cluster analysis in data mining [10];
- Expectation-Maximization algorithm – an iterative method to find the maximum likelihood of parameters in statistical models representing the optimal weights [11].

The optimal weight algorithms are compared to each other finding the best algorithm for the Kalman filter.

The objective of this paper is to find the most suitable algorithm for the noise covariances to be applied for Kalman filter to make forecasting for few parameters of the weather-stations to trigger warnings signal for the ERP systems.

The paper is structured as follows. Section 2 background, Section 3 adaptive Kalman filter and Section 4 concludes.

II. BACKGROUND

To meet the objective, the need for mathematical formulas and methods that are available in the literature are represented.

A. Simple moving average

A simple moving average (SMA) is an arithmetic mean of the variable calculated by adding the last closing prices and then dividing them by the number of time periods is calculated on average [7]. Most of the advantages:

- The SMA is a technical indicator for determining if an asset price will continue or reverse a bull or bear trend [7];

The SMA is calculated as the arithmetic average of an asset's price over some period [7];

- The SMA can be enhanced as an exponential moving average (EMA) that more heavily weights recent price action [7].

The SMA can be formulated as follows [7]:

$$SMA = \frac{A_1 + A_2 + ... + A_n}{n} \tag{1}$$

where A refers to the price of an asset at period n and n is the number of total periods. The price parameter is assumed that the price parameter is equal to the observation parameter in this case. A simple variable average is adjustable because it can be calculated for another time period by merely adding the end price of the security for several time periods and then dividing that total by the number of time periods indicating the average price of the

security over the time period. A simple variable average smoothes out volatility and facilitates access to securities price trends. If the simple moving average points up, it means that the price of the security increases. If it points down, it means that the price of security is decreasing. The longer the average period, the simpler the ordinary variable. The shortest moving average is more volatile, but its reading is closer to the source data [7].

B. Arithmetic mean

The arithmetic mean is a sum of the numbers divided by how many numbers are being averaged, in other words, an average value. Formulation of the arithmetic mean can be represented as follows [8]:

$$AM = \frac{1}{n}\sum_{i=1}^{n} a_i = \frac{a_1 + a_2 + ... + a_n}{n} \tag{2}$$

where n numbers are given, each number denoted by (where $i = 1, 2, ..., n$), the arithmetic mean is the sum of the as divided by n.

C. Bayesian average

A Bayesian average is a method of estimating the average number of people who use external information, especially the existing beliefs that have been taken into account in the calculation. A Bayesian average method is described as follows [9]:

$$\overline{x} = \frac{C_m + \sum_{i=1}^{n} x_i}{C+n} \tag{3}$$

where if the previous mean m and constant C are used in the Bayesian average value calculation, a value is assigned that is proportional to the size of the typical data set. The value is higher if the expected difference between the datasets is small. This is less if the data sets are expected to differ significantly. This is equivalent to adding C data points of value m to the data set. It is the weighted average of the previous average m and the sample [9]. The disadvantage is that the Bayesian average can be used if the size of the dataset is small for the data selected by the experiment; they are small enough. This method can be disabled for further investigation.

D. k-Means Clustering algorithm

k-means clustering is a method for quantitative determination of vectors, initially from signal processing, which is accessible for cluster analysis in data mining. The purpose of k-means clustering is to divide n observations into k clusters where each observation belongs to a cluster with the closest mean value serving as a cluster prototype. k-means clustering can be described as a set of observations $(x_1, x_2, ..., x_n)$, where each observation is a d-dimensional real vector, k-means clustering aims to partition the n observations into $k (\leq n)$ sets $S = \{S_1, S_2, ..., S_k\}$ so as to minimize the within-cluster sum of squares [10].

$$\underset{S}{\arg\min} \sum_{i=1}^{k} \sum_{x \in S_i} ||x - \mu_i||^2 = \underset{S}{\arg\min} \sum_{i=1}^{k} |S_i| Var S_i \tag{4}$$

where is the mean of points in S_i.

E. Expectation–Maximization algorithm

In statistics, an expectation–maximization (EM) algorithm is an iterative method to determine the maximum opportunity or maximum posterior (MAP) parameter estimation in statistical models, where the model depends on the latent variables. The EM iteration alternates with step E that creates a function to predict the probability of the log using the current estimation of the parameters, and the maximization M step, which calculates the parameters, maximizing the expected log probability found in step E. These parameter calculations are used to determine the distribution of latent variables in the next E phase [11].

Given the statistical model which generates a set X of observed data, a set of unobserved latent data or missing values Z, and a vector of unknown parameters θ, along with a likelihood function $L(\theta;X,Z)=p(X,Z|\theta)$, the maximum likelihood estimate (MLE) of the unknown parameters is determined by maximizing the marginal likelihood of the observed data [11]:

$$L(\theta; X) = p(X|\theta) = \int p(X, Z|\theta)dZ \tag{5}$$

F. The noise covariances algorithm comparison

To use the Kalman filter to determine the internal state of the process that has been assigned only a series of noisy observations, the process must be modeled according to the Kalman filter system. Kalman filters noise links consisting of covariance of Q process noise covariance and R surveillance noise. These parameters will be determined by algorithms and then compared. The Kalman filter has many materials and a lot of research on noise linking methods. In practice, Q and R are often considered as design variables. A common approach to choosing cravings, where Q is chosen as a diagonal weight matrix [6]. The above algorithms are used for comparison in this paper.

For experiment purposes, observations are taken between the closest three weather-stations that are close to each other and all the road monitoring weather-stations are separated by region $(R_1, R_2, ..., R_n)$. In other words, algorithms will be used and then compared results with each other. In this way, they are getting the the Root Mean Square Error (RMSE) is used for this precise compensation. The RMSE is the standard deviation of the residuals prediction errors. The interval used is 5 minutes, which consists predominantly of data entering between stations at 5 minutes interval. The region has 30 actives of 52 total road monitoring weather-stations near the road surface, only active ones are considered. These weather-stations by region is combined as follows shown in Table 1 as the previous results indicate the nearest weather-stations.

TABLE I. THE COMBINED REGION FROM THE ROAD MONITORING WEATHER-STATIONS

Region	Station name	Location (°)
R1	Adazi	57.0903, 24.3133
	Garkalne	57.0688, 24.4900
R2	Admini	56.5931, 25.5464
	Livani	56.4365, 26.0723
	Nicgale	56.1551, 26.3433
R3	Annenieki	56.6671, 23.0608
	Kalnciems	56.8360, 23.5761
	Laci	56.8945, 23.7118
	Sloka	56.9295, 23.6150
	Talsi	57.1522, 22.6902
R4	Apvedcels	56.8436, 24.0192
	Dalbe	56.7511, 23.8937
	Saulkalne	56.8465, 24.4101
	Kaibala	56.6608, 24.8981
	Vircava	56.6401, 23.7763
R5	Daugavpils	56.7511, 26.6438
	Kraslava	55.8976, 27.2933
	Zilupe	56.3850, 28.0688
	Rezekne	56.5401, 27.2991
R6	Durbe	56.5752, 21.2939
	Rudbarzi	56.6452, 21.8846
	Saldus	56.6710, 22.2898
	Nica	56.2145, 21.1329
R7	Inciems	57.2741, 24.8949
	Sigulda	57.1689, 24.9683
	Melturi	57.2212, 25.2308
R8	Smiltene	57.3881, 25.9739
	Valmiera	57.5035, 25.3271
	Stalbe	57.3781, 25.0512
	Strenci	57.6512, 25.8638

The regions consist from two till five road monitoring weather-stations located relatively near to each other. As a result, eight regions are created for the noise covariances algorithms calculation. Calculating results from each region use the average arithmetic approach. The road monitoring weather-stations observations per weather-station are taken into account surface temperature (°C) and precipitation sum (mm). Observations are taken from winter month 15th of January 2018, therefore, for each region representing surface temperature (°C) and precipitation sum (mm) average in each region are shown in Figure 1. Since the observation is much then cut out from 1:00 AM to 2:00 AM is reflected. As can be seen in two of the regions no precipitation sum has values only 0 mm. In the seventh region, the temperature has apparently fluctuated within one hour.

The corresponding algorithms are used for each region and their results are reflected in Table 2.

TABLE II.   Accuracy for each algorithm

Region	Surface temperature (°C), RMSE				
	Simple moving average	Arithmetic mean	Bayesian average	k-Means Clustering	Expectation–Maximization
R1	0.0288	0.0502	0.1010	0.0382	0.0201
R2	0.3626	0.1052	0.2717	0.1612	0.0972
R3	0.3807	0.3525	0.3738	0.3727	0.3101
R4	0.3633	0.4372	0.3927	0.3693	0.3092
R5	0.5422	0.5577	0.5321	0.5216	0.4948
R6	0.8393	0.7994	0.7369	0.7745	0.6925
R7	0.6273	0.7065	0.6514	0.6426	0.5829
R8	0.4819	0.5016	0.4796	0.5127	0.4420

The Expectation–maximization algorithm has proven in this experiment that there is one of the most accurate compared to other algorithms. The Expectation–Maximization algorithm is used further with Kalman filter and is called "Adaptive Kalman filter", which will be able to take account of the existing algorithm factor and use it to make more accurate forecasts for the raw data observations. Each time a new portion of observations comes in from the road monitoring weather-stations, the corresponding algorithm will be called to calculate an already new parameter required by Adaptive Kalman filter to be more likely to make forecasts every time. The next step is to define the Adaptive Kalman filter, which needs to be tested in a long-term way with data set. The data set already in place to be able to determine the effectiveness of this approach, which is necessary for forecasting because, as the best way to predict, a number of data sources such as Waze, Twitter, and other useful resources are needed to face clarity and well results. Obtaining an algorithm that can be applied Kalman filter case will continue in the next section.
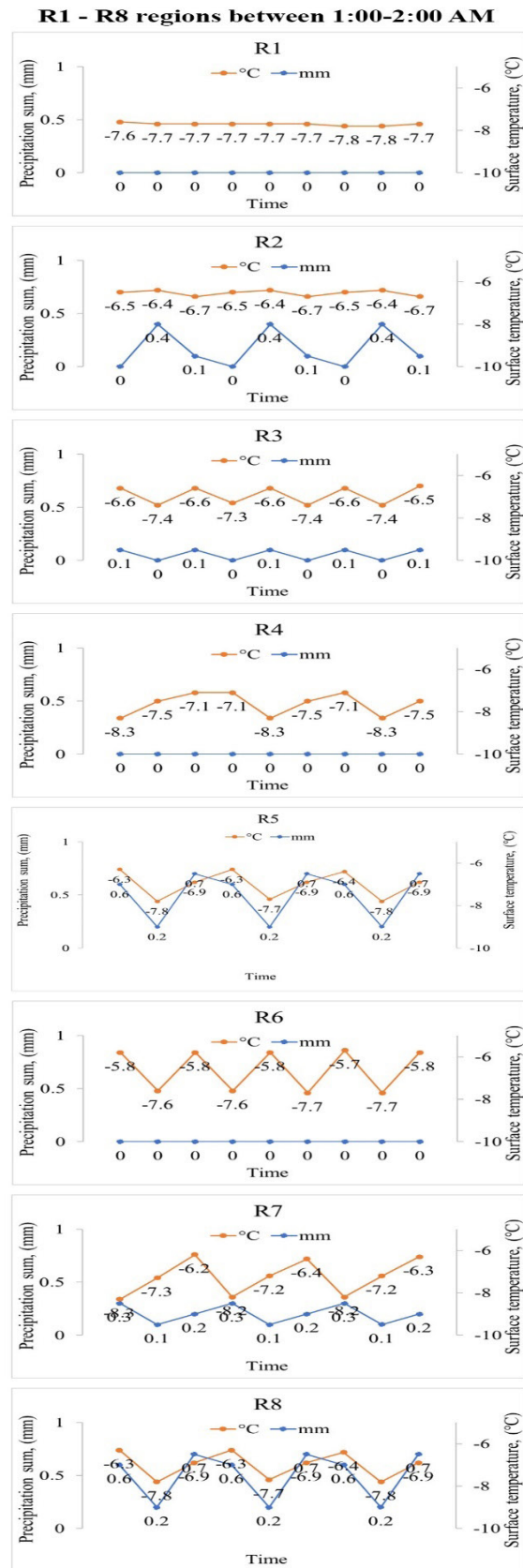

Fig. 1. R1 – R8 regions between 1:00 AM and 2:00 AM

III. ADAPTIVE KALMAN FILTER

The Expectation–maximization algorithm has proven in the experiment that it is best able to be more accurate for noise covariances algorithm comparison. The algorithm will be used for the Adaptive Kalman filter. It is initially necessary to define the Kalman filter to be able to work with it. Mathematical formulation of Kalman Filter as follows [6]:

$$x_k = Ax_{k-1} + Bu_k + w_{k-1} \tag{6}$$

the Kalman filter addresses the general problem of trying to estimate the state $x \in R^n$ of a discrete-time controlled process that is governed by the linear stochastic difference equation. With a measurement $y \in R^m$ that is [6]:

$$y_k = Hx_k + v_k \tag{7}$$

The random variables w_k and v_k represent the process and measurement noise respectively. They are assumed to be independent of each other, white, and with normal probability distributions: $p(w) \approx N(0,Q)$ and $p(v) \approx N(0,R)$. In practice process covariance Q and covariance of measurement noise R matrices can change with every step or measurement, but here they are constant n * n matrix. A refers to the state of the previous time move to the current stage if there is no driving function or process noise n * 1. The matrix B refers to optional control input state m * n of matrix H in the measurement equation refers to the state measurement y_k [6].

The Kalman filtering process has two stages: the forecasting step, where the next state. The system is predicted by previous measurements and the update phase where the current state of the system is assessed taking into account the measurement over a given time period. Actions are translated into equations as follows [6]:

- Prediction

$$X_k^- = A_{k-1} + X_{k-1} + B_k U_k \tag{8}$$
$$P_k^- = A_{k-1} P_{k-1} A_{k-1}^T + Q_{k-1} \tag{9}$$

- Update step

$$V_k = Y_k \quad H_k X_k^- \tag{10}$$
$$S_k = H_k P_k^- H_k^T + R_k \tag{11}$$
$$K_k = P_k^- H_k^T S_k^{-1} \tag{12}$$
$$P_k^- = P_k^- \quad K_k S_k K_k^T \tag{13}$$

where X_k^- and P_k^- are the predicted mean and covariance of the state, respectively, on the time step k before seeing the measurement. X_k and P_k are the estimated mean and covariance of the state, respectively, on time step k after seeing the measurement. Y_k is mean of the measurement on time step k. V_k is the innovation or the measurement residual on time step k. S_k is the measurement prediction covariance on the time step k. K_k is the filter gain, which tells how much the predictions should be corrected on time step k [6]. Adaptive Kalman filter has not been tested with experiments in this article.

IV. CONCLUSIONS

The road maintainers can analyze many sources to make decisions during the winter season to be able to serve and maintain the current road surface state more effectively. The road monitoring weather-stations allows collecting raw observations close to the road surface; allowing raw observations to be processed. To be able to predict these observations is needed. After processing the raw observations concerned, they may be used for forecasting purposes. Forecasting techniques are different types of algorithms and approaches. Studies and experiments that demonstrate the accuracy of the algorithm or approach are necessary to assess the best.

The following methods and algorithms are summarized in this article: Simple moving average, Arithmetic mean, Bayesian average, k-Means Clustering algorithm, and Expectation-Maximization algorithm. As well as an experiment with the VAS "Latvijas valsts celi" data set. The best of these algorithms is identified accordingly in this case the Expectation-Maximization algorithm shows the best RMSE. As well as the approach of Adaptive Kalman filter to make forecasting and apply the corresponding algorithm to improve forecasting results.

Further work requires an experiment with the corresponding Adaptive Kalman filter using the Expectation-Maximization algorithm with real-time data. Connecting the corresponding algorithm to raw observations and comparing it to the holding critical rules that have been defined by the road maintainers.

REFERENCES

[1] [1] J. B. Edwards, "Speed adjustment of motorway commuter traffic to inclement weather," Transp. Res. Part F Traffic Psychol. Behav., vol. 2, no. 1, pp. 1–14, Mar. 1999, https://doi.org/10.1016/S1369-8478(99)00003-0

[2] [2] Y.-Y. Nguwi and A. Z. Kouzani, "Detection and classification of road signs in natural environments," Neural Comput. Appl., vol. 17, no. 3, pp. 265–289, Jun. 2008, https://doi.org/10.1007/s00521-007-0120-z

[3] [3] J. P. Gibbs and W. G. Shriver, "Can road mortality limit populations of pool-breeding amphibians?," Wetl. Ecol. Manag., vol. 13, no. 3, pp. 281–289, Jun. 2005.

[4] [4] C. Taylor, S. Pollard, S. Rocks, and A. Angus, "Selecting Policy Instruments for Better Environmental Regulation: a Critique and Future Research Agenda," Environ. Policy Gov., vol. 22, no. 4, pp. 268–292, Jul. 2012.

[5] [5] J. Honaker and G. King, "What to Do about Missing Values in Time-Series Cross-Section Data," Am. J. Pol. Sci., vol. 54, no. 2, pp. 561–581, Apr. 2010.

[6] [6] G. Bishop, G. W.-P. of SIGGRAPH, undefined Course, and undefined 2001, "An introduction to the kalman filter," byron.soe.ucsc.edu.

[7] [7] F. R. Johnston, J. E. Boyland, M. Meadows, and E. Shale, "Some properties of a simple moving average when applied to forecasting a time series," J. Oper. Res. Soc., vol. 50, no. 12, pp. 1267–1271, Dec. 1999.

[8] [8] E. Jacquier, A. Kane, and A. J. Marcus, "Geometric or Arithmetic Mean: A Reconsideration," Financ. Anal. J., vol. 59, no. 6, pp. 46–53, Nov. 2003.

[9] [9] J. Huber, K. T.-M. Letters, and undefined 2001, "On the similarity of classical and Bayesian estimates of individual mean partworths," Springer.

[10] [10] A. Likas, N. Vlassis, J. V.-P. recognition, and undefined 2003, "The global k-means clustering algorithm," Elsevier.

[11] [11] T. M.-I. S. processing magazine and undefined 1996, "The expectation-maximization algorithm," ieeexplore.ieee.org.