Tomi Seppälä – Antti Ruotoistenmäki – Fridtjof Thomas

OPTIMAL SELECTION AND ROUTING OF ROAD SURFACE MEASUREMENTS

HELSINKI SCHOOL OF ECONOMICS
WORKING PAPERS
W-424

Tomi Seppälä – Antti Ruotoistenmäki – Fridtjof Thomas

OPTIMAL SELECTION AND ROUTING OF ROAD SURFACE MEASUREMENTS

Quantitative Methods of Economics and
Management Science

October
2007

HELSINGIN KAUPPAKORKEAKOULU
HELSINKI SCHOOL OF ECONOMICS
WORKING PAPERS
W-424

HELSINGIN KAUPPAKORKEAKOULU
HELSINKI SCHOOL OF ECONOMICS
PL 1210
FI-00101 HELSINKI
FINLAND

ISSN 1235-5674
(Electronic working paper)
ISBN 978-952-488-137-1

Helsinki School of Economics -
HSE Print 2007

Abstract

The condition of the road network, collected using high-speed devices as part of normal traffic flow, is an essential input to the maintenance process at all decision-making levels. Data collection is relatively inexpensive compared to maintenance needs; yet its benefits should be evaluated and the data collection process made as effective as possible. Our objective in this paper is to evaluate the benefits of road surface measurements, using a decision theoretical approach combined with optimisation of measurement route. We develop an integer linear programming model with route constraints and an objective function that maximises the expected length of road to be reclassified using new measurements for updating the belief of road network condition. The elements of an access matrix are used for evaluating the connectivity of the optimised measurement route.

A simplified network model is used for illustrating the calculation method, which is then transferred on to the network of main and regional roads of Uusimaa Road District in Finland. The results validate the proposed method and also reveal the need for further development. For example, one-carriageway roads are normally measured in one direction only. In our example, we use the same benefits for both directions. Based on the results of this work, it can be concluded that the emphasis in the measurement policy should be shifted from measuring some roads every year to measuring all roads in both directions every other year.

Keywords: route optimisation, access matrix, measurement policy

1 Introduction and objectives

The condition of the road network is an essential input to the maintenance process at all decision-making levels. The condition is measured by using high-speed devices participating in the normal traffic flow. Relatively inexpensive, the measurements, compared to maintenance budgets, are easily taken for a large part of the road network. Yet the total expenditure on road surface measurements may be considerable. The road manager should therefore evaluate the benefits of measurements and utilise the collected information in decision-making as effectively as possible.

What are then the benefits of taking new measurements? Maintenance needs are assessed based on the condition information. This may be done by comparing the measured condition values to trigger values for classifying road sections into categories. In the absence of recent measurements, the current condition may be projected from previous measurements using statistical models. However, uncertainty is connected with both measured and modelled values, and this may result in inaccuracies in the estimated maintenance needs. The benefits may then be evaluated by assessing how much this uncertainty can be reduced by taking new measurements, resulting in more accurate assessment of maintenance needs (Ruotoistenmäki et al. 2006).

Our objective in this paper is to evaluate the benefits of road surface measurements, using a decision theoretical approach combined with optimisation of measurement route. In Section 2, the problem setting is described together with the principles of our methodical solution. In Section 3, a stylised example is presented to illustrate our calculation method. Full analysis using a test network constructed from the condition data bank of Finnish Road Administration, is presented in Section 4. Finally, the conclusions are presented in Section 5.

2 Methods

In road maintenance works programming, the decisions to be made concern which road sections are maintained and what maintenance work types are selected. In the decision-making situation, several decision criteria prevail. These are, for instance, budget constraints, scheduling the maintenance works and the condition of the road sections. In this paper, we take a closer look at the last one of these criteria - condition.

Our current belief of the network condition is based on the distribution of condition variables, such as roughness and rut depth. The distribution is composed of the condition variables for a finite number of road sections, whose condition values include uncertainty. For each individual section, the registered condition is the result of some form of averaging over a large number of measurements. For example, the rut depth attributed to a single road section is the mean value of a number of maximum rut depth measurements obtained from a larger number of transversal road surface profiles. Such an average is therefore approximately normally distributed by the central limit theorem and we therefore assume here that the condition variables for individual sections can be modelled as normally distributed random variables. This normality is conditional on the (unknown) true value for the respective road section. The discrete distribution of all conditions of the finite number of road sections is a mixture over these normal distributions. Empirically, the resulting distribution for the condition of the individual road sections in a road network under continued operation is often close to log-normal.

According to our current belief of the condition of a road section, each section is classified into one of four categories in accordance with whether that section should be included into the next maintenance program:

1. No Action.
2. Warning.
3. Action.
4. Must Do.

Normally, we have previous information on the current condition of roads. This information is, for example, based on previous measurements and knowledge of the degradation of condition from performance models and engineering experience. The value of road surface measurements lies in the fact that they cast light over the current true condition of the measured road sections.

In our particular context one is essentially interested in the resulting classification of a road section into one of the four categories from No Action to Must Do. Additional road surface measurements are therefore motivated if we would expect that the measured road sections might be reclassified based on the additional measurements. In other words, to measure a road section when we have no indication that we might change our perception of that section is of little value for the problem at hand. The value of the measurements arises from the fact that we might reclassify sections and therefore might compile a different and more appropriate maintenance program. A desirable route for a measurement vehicle is therefore a route that contains a large number of road sections which we can expect to reclassify, compared to the current classification. We formalize this approach in the following.

First, let us consider the situation where additional measurements are actually collected. Formally, we view our current belief about a road section's condition as prior information which is updated with a new observation in order to obtain the posterior distribution of the actual condition. This is illustrated as: $\text{posterior} \propto \text{likelihood} \times \text{prior}$. The likelihood follows from our knowledge of the measurement process. This posterior distribution would allow a classification of the road section that may or may not be different from the one based on the prior information (before or without additional measurements).

Now, before deciding whether to collect an additional measurement, we can evaluate the probability that the values to be measured will lead to a reclassification of that road section. We do that with the help of the *predicted* posterior density, i.e. a distribution that describes our likely posterior knowledge *before* the additional measurements are actually taken. This predictive distribution is the basis for our assessment whether it is likely that we will reclassify road sections and, consequently, whether it is worth it to actually collect these measurements.

The posterior distribution of the road condition $\pi(\theta \mid x)$ is proportional to the product of the likelihood $f(x \mid \theta)$ of the utilised statistical model and the prior distribution of the parameter, $\pi(\theta)$, i.e.

$$\pi(\theta \mid x) = \frac{f(x \mid \theta)\pi(\theta)}{\int_\theta f(x \mid \theta)\pi(\theta)} \propto f(x \mid \theta)\pi(\theta)\,, \tag{1}$$

where θ indicates the actual condition for a specific road section. In our model, we utilize normal distributions for the prior and the likelihood. In that case even the resulting posterior distribution is a normal distribution that is completely specified by mean (expected) value μ and standard deviation σ.

Our current belief of the road condition is represented by the prior distribution $\pi(\theta)$, where $\theta \sim \mathrm{N}(\mu, \tau^2)$, based e.g. on a previous measurement x_{t-1}. The measurement process is assumed to produce a normally distributed measurement value so that the measured values x are correct on average, i.e. $X \sim \mathrm{N}(\theta, \sigma^2)$. The posterior distribution for θ is then (Berger 1985, p. 128)

$$\theta | x_t, x_{t-1} \sim N\left(\mu(x_t), \frac{\tau^2\sigma^2}{\tau^2+\sigma^2} \right). \tag{2}$$

Now, the measurement x_t is not yet taken, but it has the predictive density (Berger 1985, p. 95)

$$x_t | x_{t-1} \sim N\left(\mu, \tau^2+\sigma^2\right). \tag{3}$$

Thus, replacing σ^2 in Equation (2) with $\tau^2+\sigma^2$ from Equation (3) yields the following predictive posterior distribution for θ:

$$N\left(\mu, \frac{\tau^2(\tau^2+\sigma^2)}{2\tau^2+\sigma^2} \right).[1] \tag{4}$$

This principle is illustrated in Figure 1, where the increase of θ indicates deterioration. The prior distribution of an individual road section, $\pi(\theta), \theta \sim N\left(\mu, \tau^2\right)$, is shown in the dotted curve. The predictive posterior distribution for that road section based on Equation (4) is illustrated by the solid curve. Maintenance threshold values for No Action, Warning, Action and Must Do levels are shown as vertical lines. According to the example shown in Figure 1, we classify a particular road section into Warning category based on previous condition information, because most of the probability mass of $\pi(\theta)$ falls into the associated interval for values of θ. We then calculate the probability that the road section will be classified differently by evaluating the predictive posterior distribution using the same criteria of, in this case, predictive posterior density mass for the various intervals associated with the classification. Different classification means a classification into any other class and corresponds in this particular example to the predictive posterior density mass given to all values of θ not belonging to the Warning level. This probability corresponds to the combined areas under the solid curve to the left of the leftmost vertical line

[1] Of course, once measurements are actually taken, the information should actually be updated with help of Equation (2), resulting in a true posterior distribution for θ.

between categories No Action and Warning and to the right of the vertical threshold line between categories Warning and Action.

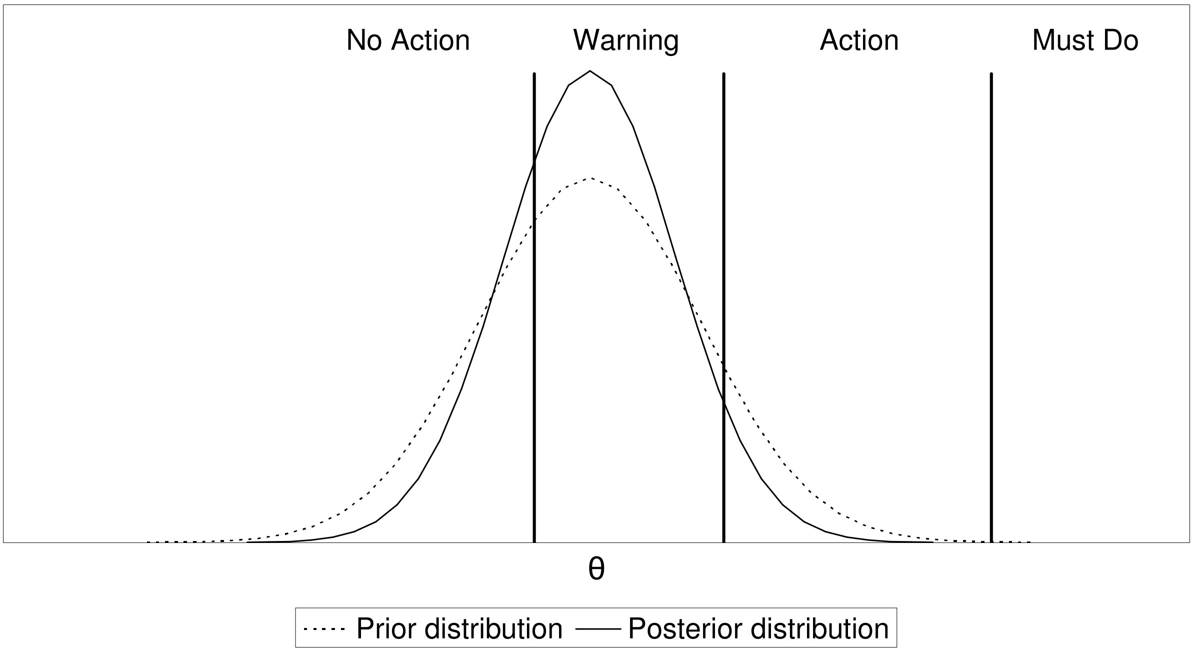

Figure 1. Evaluating whether additional measurements are likely to result in reclassification of a road section.

We determine this probability of reclassification of a section in case of additional measurements for each single section in the road network. Our aim is conditional on given resources (funding, maximum kilometres to measure, etc.), to find a route that maximises the number of sections we expect to be reclassified. This route enhances our decision-making most. Equivalently, we aim at maximising the expected length of road to be reclassified by a given measurement effort.

3 Illustration of the optimisation method

3.1 Stylised example

To give the reader a clear picture of the calculation method used, in Figure 2 we present a stylised example of a network, which is then applied to a network of in-service roads.

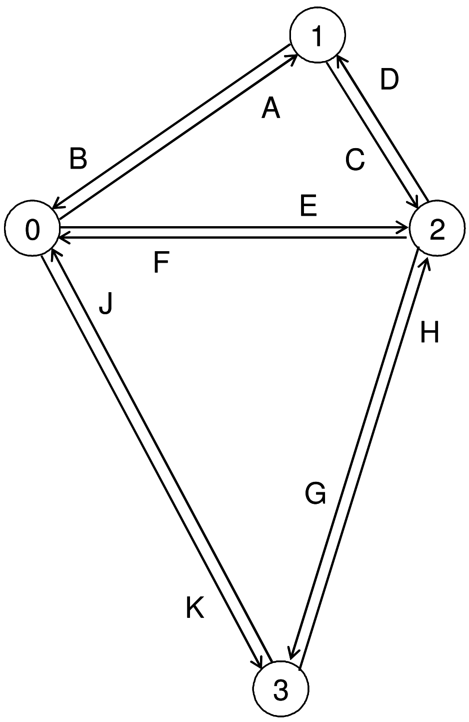

Figure 2.Network for the stylised example.

The stylised example consists of five physical sections, each one of which can be measured in two directions. This results in a total of ten sections. In our example, we use roughness in terms of the logarithm of *IRI*[2] to represent the condition θ. For each section, we have the current belief of logarithmic roughness with the associated standard deviation of 0.168 (=16.8 %) at each section. The standard deviation of the

[2] International Roughness Index (*IRI*) represents the vertical movement of passenger and vehicle per distance travelled (unit mm/m). For further explanation, the reader is referred to Sayers et al. (1986a, 1986b) and UMTRI (2007).

measured logarithmic *IRI* is assumed 0.118 (=11.8 %)[3]. The standard deviation of the conditional posterior density for the new measurement x_t from Equation (3) is then

$$\sqrt{\tau^2+\sigma^2} = \sqrt{0.168^2+0.118^2} = 0.2053, \tag{5}$$

and the standard deviation of the predictive posterior distribution of condition θ, according to Equation (4) is

$$\sqrt{\frac{\tau^2\left(\tau^2+\sigma^2\right)}{2\tau^2+\sigma^2}} = \sqrt{\frac{0.168^2\left(0.168^2+0.118^2\right)}{2*0.168^2+0.118^2}} = 0.130. \tag{6}$$

The classification of the sections according to *IRI*, and the associated threshold values are presented in Table 1. The threshold values of the original untransformed *IRI* values reflect those of a high-class paved road network.

Table 1. Condition classification according to IRI.

Classification	*IRI* [mm/m]	ln(*IRI*)
N = No Action	<1.5	<0.4055
W = Warning	1.5 - 2.0	0.4055 - 0.6930
A = Action	2.0 - 3.0	0.6931 - 1.0986
M = Must Do	>3.0	>1.0986

The length of sections and the expected value of the current belief of logarithmic *IRI* for each section, simulated from a normal distribution $N \sim (0.404, 0.380)$[4] are presented in Table 2, along with the current classification and the probability of reclassification. Gains from taking new measurements are defined as the product of the probability of reclassification and section length, and shown in the rightmost column in Table 2.

[3] These values are based on Ruotoistenmäki et al. (2006), who developed measures of accuracy for measured and modelled condition values. The accuracy 16.8 % of the current belief of the logarithmic *IRI* is based on the logarithmic *IRI* predicted from two year old measurements, whose accuracy is 11.8 %.

[4] These values reflect distribution of log(IRI) on a high-class paved road network, and are from the data set where the accuracy of measured and modelled condition values were verified by Ruotoistenmäki et al. (2006).

Table 2. Data for stylised example.

Section ID	From	To	Length of section [km]	Current belief, Expected value	Current classification	p [1)]	gain (g) = Length * p
A	0	1	4	0.13	N	0.016	0.062
B	1	0	4	0.49	W	0.324	1.297
C	1	2	3	0.71	A	0.459	1.378
D	2	1	3	0.58	W	0.279	0.837
E	0	2	5	0.05	N	0.003	0.017
F	2	0	5	0.62	W	0.335	1.673
G	2	3	7	0.78	A	0.260	1.820
H	3	2	7	0.03	N	0.002	0.015
J	3	0	8	0.86	A	0.130	1.037
K	0	3	8	0.62	W	0.335	2.676

1) p = Probability of reclassification for this section

The probability of reclassification p is calculated as the sum of the probabilities that the new classification is downgraded and that the new classification upgraded. For example for road section C, this probability would be calculated as

$$p = \Phi\left(\frac{0.71 - 0.6931}{0.130}\right) + \left(1 - \Phi\left(\frac{1.0986 - 0.71}{0.130}\right)\right) = 0.459\,, \tag{7}$$

where Φ denotes the cumulative density of the standard normal distribution. The expected gain (expected length of reclassified road section) is the probability of reclassification multiplied by the section length, i.e. for this example 3*0.459=1.378.

In order to complete our aim to find a route that maximises the expected length of road to be reclassified using the new measurements, we need to maximise the sum of the rightmost column in Table 3, the product of section length and the probability of reclassification. Constraint in the optimisation is the funding available for measurements.

3.2 Optimisation

Let x_{ij}=1, if road section from node i to node j is to be measured, and 0 otherwise ($i = 0,\ldots,n; j = 0,\ldots,n$). Let c_{ij} and g_{ij} be the corresponding cost and gain for measuring, respectively, and let B denote the available budget. In addition, we need a balancing equation for each node that will be passed, i.e. the number of arriving and departing arcs to and from each node on the route are equal. Otherwise additional sections would have to be driven to be able to apply the optimal solution in practice. This, of course, would mean additional costs. We will also require that the measuring vehicle starts from node 0 and also finishes the route in node 0. This can be achieved by requiring that there is at least one arc leaving from node 0. The problem then becomes an integer linear programming problem with binary decision variables x_{ij} as follows (the sums are taken over all the possible routes):

$$
\begin{aligned}
&Max \quad G = \sum_{i=0}^{n} \sum_{\substack{j=0 \\ i \neq j}}^{n} g_{ij} x_{ij} \\
&s.t. \quad \sum_{i=0}^{n} \sum_{\substack{j=0 \\ i \neq j}}^{n} c_{ij} x_{ij} \leq B \\
&\sum_{\substack{j=0 \\ i \neq j}}^{n} x_{ij} = \sum_{\substack{j=0 \\ i \neq j}}^{n} x_{ji} \ \ for\ all\ nodes\ i \\
&\sum_{j=1}^{n} x_{0j} \geq 1 \\
&x_{ij} \in \{0,1\}\ for\ all\ i, j; i \neq j
\end{aligned}
\tag{8}
$$

The problem can easily be solved with specialised software, e.g. AMPL[5], which we have used to obtain the optimal solutions for the different cases in this paper. For the input values given in table 2, where cost of measurement € 43.30 / km and a budget constraint of € 1200 are presented, the optimal solution is shown in figure 3. Solving problem (8) yields the optimal solution where sections C, D, F, H and K are measured. The optimal expected length of reclassified road is 6.579 km at a cost of € 1125.8.

[5] www.ampl.com

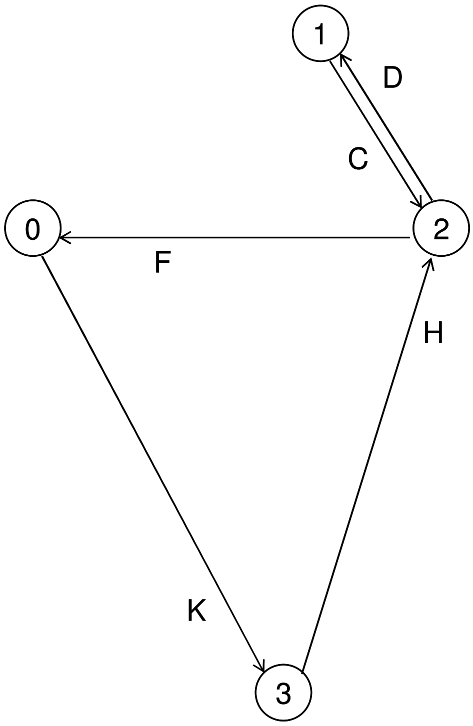

Figure 3. Solution of equation (8) for the input values given in table 2, cost of measurement € 43.30 / km and a budget constraint of € 1200. Sections C, D, F, H and K are measured. The optimal expected length of reclassified road is 6.579 km with a cost of € 1125.8.

Clearly, the solution is consistent, i.e. it is connected so that one vehicle can be used for driving all sections selected for measurement in the optimisation process. The balancing constraint of equation (8) ensures that the number of departing and arriving arches match at each node. However, this property does not guarantee the consistency of the route.

3.3 Access matrix

But how do we ensure the consistency of the route? One obvious solution is to use several vehicles, but then again the route of each individual vehicle has to be connected. Otherwise more sections have to be driven through to unite the route and this would be a violation of the budget constraint used in the optimisation. For practical purposes, slight deviation from the budget constraint may well be a feasible solution, and on small networks, the consistency of the route is easy to check on a map. However, on large networks, an analytical model is desirable for securing the consistency of an optimised route within the budget limits.

We developed a mathematical approach for detecting the consistency of the optimised route, based on the concept of *access matrix*. The elements of an access matrix **A** indicate whether a connection between two nodes exists. The value of an element $a_{ij} = 1$, if a connection between nodes i and j exists, otherwise $a_{ij} = 0$. The diagonal elements of $\mathbf{A} \equiv 0$, indicating that one cannot return to a departing node when driving along one arc only. Furthermore, the upper triangular matrix and the lower triangular matrix are mirror images of each other around the diagonal, i.e. if a connection from node i to node j exists, also a connection from node j to node i exists[6]. The elements of the sum matrix

$$\sum_{i=1}^{n} A^i = A^1 + A^2 + \ldots + A^n \tag{9}$$

indicate the number of alternative routes from node i to node j in the network of n nodes when a maximum of n arcs are driven. In this case, the diagonal elements of $A \neq 0$, which means one can return to departure node after driving a maximum of n arcs.

Likewise, the sum of matrices, according to equation (9), indicates the number of alternative routes from node i to node j in an optimised solution. If an element = 0 on a row i or column i for a node i that is on the route, then that node is not connected to one or more of the other nodes $j \neq i$ on the route. More precisely, a node i for which $a_{ij} = 0$ $(i \neq j)$, is not connected with node j. However, if a node i is not on the route, all elements of that row i and column i equal 0, and it can be ignored.

[6] Actually, this need not be the case. One-way streets exist, for which $a_{ij} \neq a_{ji}$. In this case, our method can be applied as such, even though our examples represent cases for which $a_{ij} = a_{ji}$.

In order to check the consistency of an optimised route, the equation (9) is applied to the access matrix of that solution, and the sum matrix is analysed. An example is shown in Table 3, where the access matrix of the original stylised example is shown in the left pane and the access matrix of the solution in Figure 3 in the right pane.

Table 3. Access matrix of the original stylised example (left pane) and a consistent solution (right pane).

	Original access matrix = A (Figure 2)				A consistent solution (Figure 3)			
	To				To			
From	Node 0:	Node 1:	Node 2:	Node 3:	Node 0:	Node 1:	Node 2:	Node 3:
Node 0:	0	1	1	1	0	0	0	1
Node 1:	1	0	1	0	0	0	1	0
Node 2:	1	1	0	1	1	1	0	0
Node 3:	1	0	1	0	0	0	1	0
	$A^1 + A^2 + A^3 + A^4$				$A^1 + A^2 + A^3 + A^4$			
	To				To			
From	Node 0:	Node 1:	Node 2:	Node 3:	Node 0:	Node 1:	Node 2:	Node 3:
Node 0:	22	16	22	16	1	1	2	2
Node 1:	16	14	16	14	2	2	3	1
Node 2:	22	16	22	16	3	3	3	2
Node 3:	16	14	16	14	2	2	3	1

From Table 3, we can see that both in the original network and in the consistent solution, when driving *n* arcs, a number of routes from a node to any other node (including the departing node), can be made. This is obvious from figure 3, but for large networks, a visual examination of an optimised solution may be cumbersome. Furthermore, the method presented here can be programmed as a part of an optimisation application.

3.4 Driving or measuring?

In practice it is possible to turn the measurement apparatus off so that driving along the sections can be less costly. In this way we can choose the sections most beneficial for measuring. We can then drive but not measure other routes so that the measurement route remains consistent.

We will assume that the cost of driving without measurement on a given section from node i to node j is estimated as d_{ij}. Let x_{ij}=1, if road section from node i to node j is being measured, and 0 otherwise; as before. In addition, we need another binary variable y_{ij}=1, if road section from node i to node j is being used for travelling without measuring, and 0 otherwise. With this modification the model is as follows:

$$
\begin{aligned}
&Max \quad G = \sum_{i=0}^{n}\sum_{\substack{j=0\\ i\neq j}}^{n} g_{ij}x_{ij} \\
&s.t. \quad \sum_{i=0}^{n}\sum_{\substack{j=0\\ i\neq j}}^{n} \left(c_{ij}x_{ij} + d_{ij}y_{ij}\right) \leq B \\
&\sum_{\substack{j=0\\ i\neq j}}^{n} \left(x_{ij} + y_{ij}\right) = \sum_{\substack{j=0\\ i\neq j}}^{n} \left(x_{ji} + y_{ji}\right) \; for\ all\ nodes\ i \\
&\sum_{j=1}^{n} \left(x_{0j} + y_{0j}\right) \geq 1 \\
&x_{ij} + y_{ij} \leq 1 \; for\ all\ i, j; i \neq j \\
&x_{ij} \in \{0,1\} \; for\ all\ i, j; i \neq j
\end{aligned}
\tag{10}
$$

We introduce a new constraint that means that each section can either be measured or travelled on only. A reasonable estimate for the travelling cost is a fixed percentage of the measurement cost per kilometre. Using the same numerical data as before and assuming the travelling cost to be 20 % of the measurement cost, the optimisation problem was solved. The optimal solution is shown in Figure 4. The dashed lines on sections A, D and H indicate travelling without measuring. The measured sections are B, C, F and K. The total expected length of reclassified road is 7.024 km, which is more than in our previous solution (6.579) but the cost is less (€ 987.24 instead of € 1125.8).

The access matrix can similarly be used for checking the consistency of the optimised measurement route. We are only interested in whether the route is connected, not which routes are measured, or which ones are driven on only. Thus, we set all elements a_{ij} = 1 that are either driven or measured. From Figure 4 and Table 4 we see that the route is consistent. This makes this approach a feasible one in a real road measuring problem. It should be noted that if travelling cost equals measurement cost, the equation (10) is reduced back to equation (8).

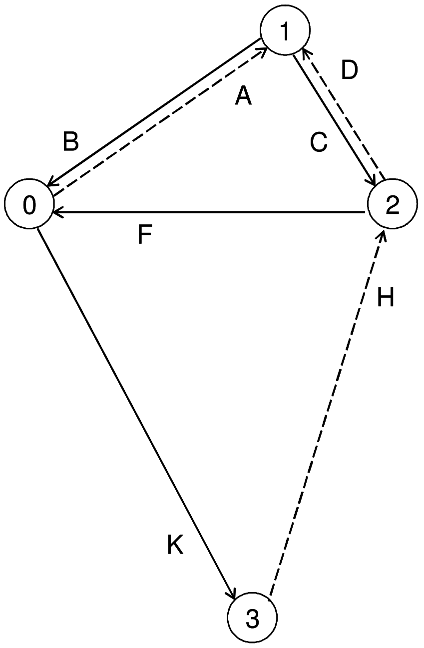

Figure 4. Solution when route constraints are considered and driving without measuring has a lower cost than measuring. Sections B, C, F and K are measured, A, D and H driven on only. The optimal expected length of reclassified road is 7.024 km with a cost of € 987.24.

Table 4. Access matrix of a consistent solution shown in Figure 4.

	A consistent solution (Figure 6)			
	To			
From	Node 0:	Node 1:	Node 2:	Node 3:
Node 0:	0	1	0	1
Node 1:	1	0	1	0
Node 2:	1	1	0	0
Node 3:	0	0	1	0
	$A^1 + A^2 + A^3 + A^4$			
	To			
From	Node 0:	Node 1:	Node 2:	Node 3:
Node 0:	6	6	6	4
Node 1:	8	8	6	4
Node 2:	8	8	6	4
Node 3:	4	4	4	2

4 Validation of the optimisation method

4.1 Test network

In the previous section, we developed a decision theoretical framework combined with a route optimisation method for assessing the benefits of road surface measurements. We also illustrated our method on a stylised network. For its further validation, we now present a calculation example using data from an in-service road network. The network consists of the main and regional roads of Uusimaa Road District in Finland. The road network also embraces Helsinki, the capital city of Finland.

In the condition data bank (CDB) of the Finnish Road Administration (Finnra), the road network is divided into management sections with an average length of approximately 5 kilometres. We divided our network into sections by placing nodes at management section change points in the CDB. The resulting network, shown in Figure 5, consists of a total of 289 nodes and 700 arcs. The total length of the network is 2 952.3 kilometres. Helsinki is located in the middle of the southern coast of the Uusimaa District, and main roads radially start from there or surround the capital as rings.

All routes can be driven in both directions and two routes in opposite directions exist between every two nodes, as in the stylised example shown in Figure 2. This enables us to complete a drivable measurement route. In addition, even though the road condition in the opposite directions is correlated, it is different, and the expected gains should be different for the different directions.

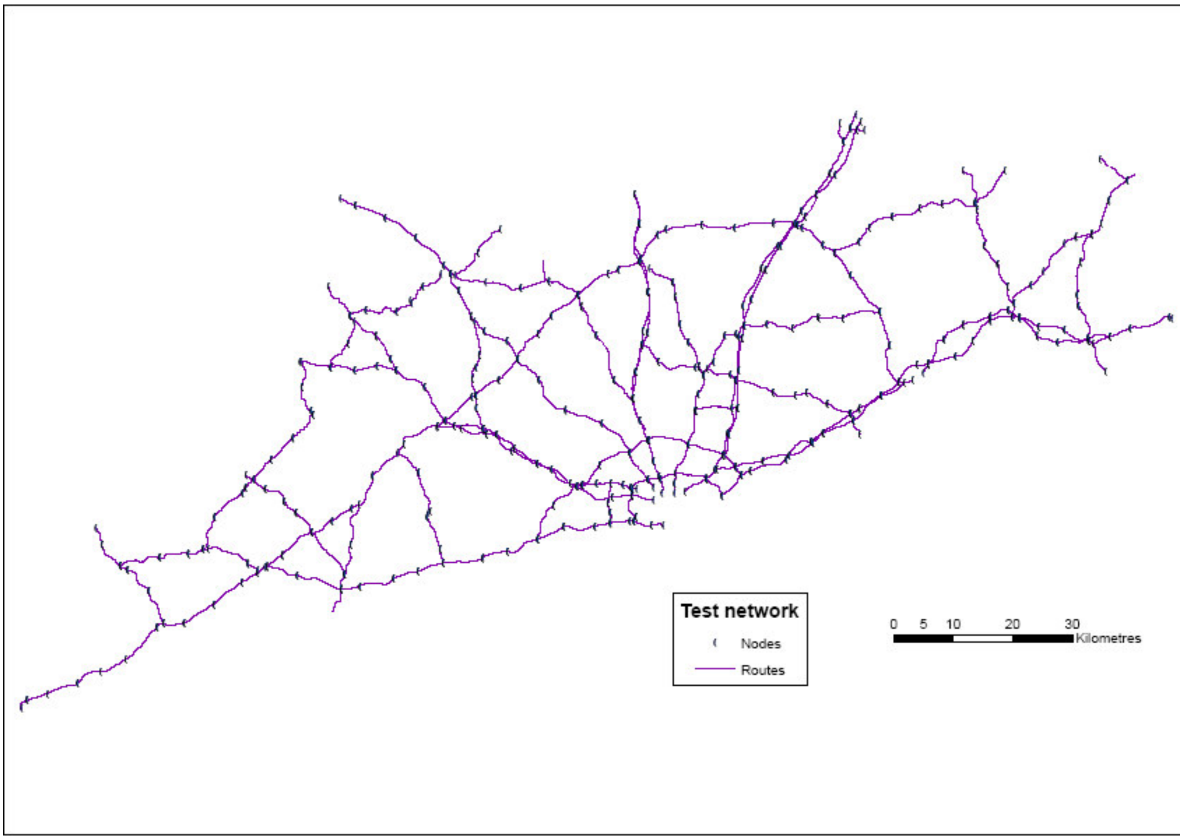

Figure 5. Test network of main and regional roads of Uusimaa road district in Finland.

Data is stored in the CDB in 100-meter sections. We extracted roughness measurements (*IRI*) from 2003 to 2006 for this network and calculated the expected benefits of taking new measurements for each 100-meter section. We then calculated the gain for each route as the average of gains from 100-meter sections, and multiplied it with route length to produce gains for maximisation according to Equation (10). The current measurement policy is to measure one-carriageway roads in one direction only, whereas for two-carriageway roads all lanes in both directions are measured. For two-carriageway roads, we used the actual measurements to derive gains in both directions separately, but for one-carriageway roads, we used the same gain from measurement in one direction for both directions.

4.2 Measures of accuracy

For the stylised example in the previous section, we used the measures of accuracy as defined by Ruotoistenmäki et al. (2006) to calculate the standard deviation of the conditional posterior density for the new measurement and the standard deviation of the predictive posterior distribution of condition according to equations (5) and (6). These measures of accuracy are based on equipment not in use any more. Instead, we developed new measures of accuracy using data from current measurement vehicles and a similar procedure as used by Ruotoistenmäki et al. (2006). These measures are based on data collected from our sample network from 2003 to 2006. These accuracies, presented in Table 5, are considerably better than the previous ones.

Table 5. Measuring and modelling accuracy from data collected between 2003 and 2006 using the method developed by Ruotoistenmäki et al. (2006).

		Accuracy	Eq (11)	
Measured		6.1%	0.0042	is the right hand side
Modelled using data from previous measurements	1+2+3 yrs	7.2%	0.0014	are the left hand sides
	1+2 yrs	7.3%	0.0016	
	1+3 yrs	7.3%	0.0017	
	1 yr	7.8%	0.0024	
	2+3 yrs	9.2%	0.0047	
	2 yr	9.7%	0.0058	
	3 yr	11.7%	0.0100	

Ruotoistenmäki et al. (2006) developed the following inequality to calculate the maximum excess variance for which it is more beneficial to use a model than to take a new measurement in order to assess the current road condition:

$$s_Y^2 - s^2 \leq \frac{2c}{k_1 + k_2} \tag{11}$$

Here s_Y^2 is the variance associated with modelled values, s^2 is the variance associated with measured values, c is the measurement cost and k_1 and k_2 are loss coefficients for agency and user losses due to untimely maintenance. The values from Equation 11 are shown in the rightmost column in Table 5. From these values it can be concluded that if a road is measured in the previous year, it is more beneficial to model the current condition of the road from previous measurements than to measure it again this year. If no data is available from the preceding year, it is more beneficial to measure than to use a model. According to this calculation, the roads should be measured every other year.

4.3 Updating knowledge of road condition

For each 100-meter section, we have the current belief of the logarithmic roughness from the CDB, with the associated standard deviation of 0.078 (=7.8 %), 0.097 and 0.117 for one, two and three year old measurements, respectively. Data from four years enables modelling based on up to three year old measurements. Thus, we estimated the standard deviation based on four year old measurements at 0.130 and based on measurements older than that at 0.140. The standard deviation of the measured logarithmic *IRI* is 0.061 (=6.1 %). The standard deviation of the conditional posterior density for the new measurement $x_t|x_{t-1}$ for one-year old measurements from Equation (3) is now

$$\sqrt{\tau^2 + \sigma^2} = \sqrt{0.078^2 + 0.061^2} = 0.099\,, \tag{12}$$

and the standard deviation of the predictive posterior distribution of the condition θ, according to Equation (4) is

$$\sqrt{\frac{\tau^2(\tau^2+\sigma^2)}{2\tau^2+\sigma^2}} = \sqrt{\frac{0.078^2(0.078^2+0.061^2)}{2*0.078^2+0.061^2}} = 0.061\,. \tag{13}$$

The classification of the sections according to *IRI*, and the associated threshold values are the same as before, and are presented in Table 1. One-carriageway roads are measured in one direction only, and in that case, the same value for gain is used for both sections in different directions between the same nodes. For two-carriageway roads, actual measured values are used for determining gains for the different directions separately.

4.4 Budget limit

We used the following procedure for selecting the budget limit for optimisation: As was concluded from the values of Equation (11) in Table 5, a section should be measured only if it was not measured in the previous year. A total of 705 km out of 2 950 km meet this criteria, and the expected length of reclassified road is 70.9 km, which means the average probability of reclassification of one km of road is 70.9 / 705 = 0.1. These roads are shown in Figure 6 where it can be seen that they are mainly located in the low-volume part of the network. This is due to the current measurement policy, where main roads are measured every year and minor roads every three years. The measurement cost of these roads is 705 km * 43.3 €/km = 30 521 €. This cost does not change if the sections' locations on the network change. Consequently, we set the budget limit at this lump sum. The question then is whether we can gain greater benefit for the same measurement budget by using our optimisation model.

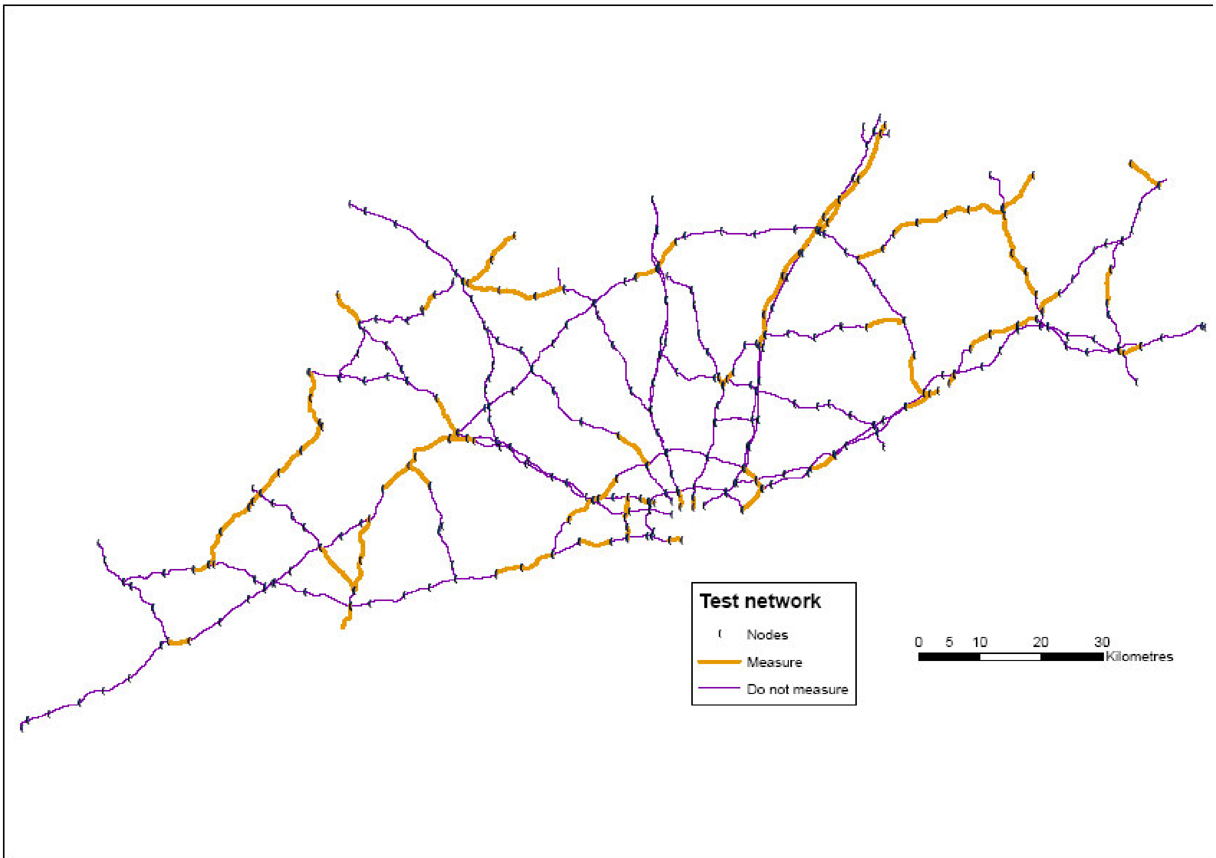

Figure 6. Measurement program of main and regional roads of Uusimaa Road District in Finland, based on accuracy of existing condition information.

4.5 Results

Optimisation was done using the AMPL[7] software. The solution is shown in Figure 7. Interestingly enough, all sections except two that are driven are also measured. The budget limit is € 30 521, and the budget used is € 30 468, which results in a measurement program of 703.7 km. The expected length of the roads in the optimised measurement program, which is the value of objective function in Equation (10), is 102.3 km. This gain is 44 % higher than the gain from measurement route that complies with the current measurement policy shown in Figure 6. The average probability of reclassification of one km of road in the optimised solution is 102.3 / 703.7 = 0.145. From Figure 7 it can also be seen that the optimised solution is concentrated in the low-volume part of the network. This is the part of the network where deterioration of roads in terms of roughness is greatest and where the expected benefits of taking new measurements are highest.

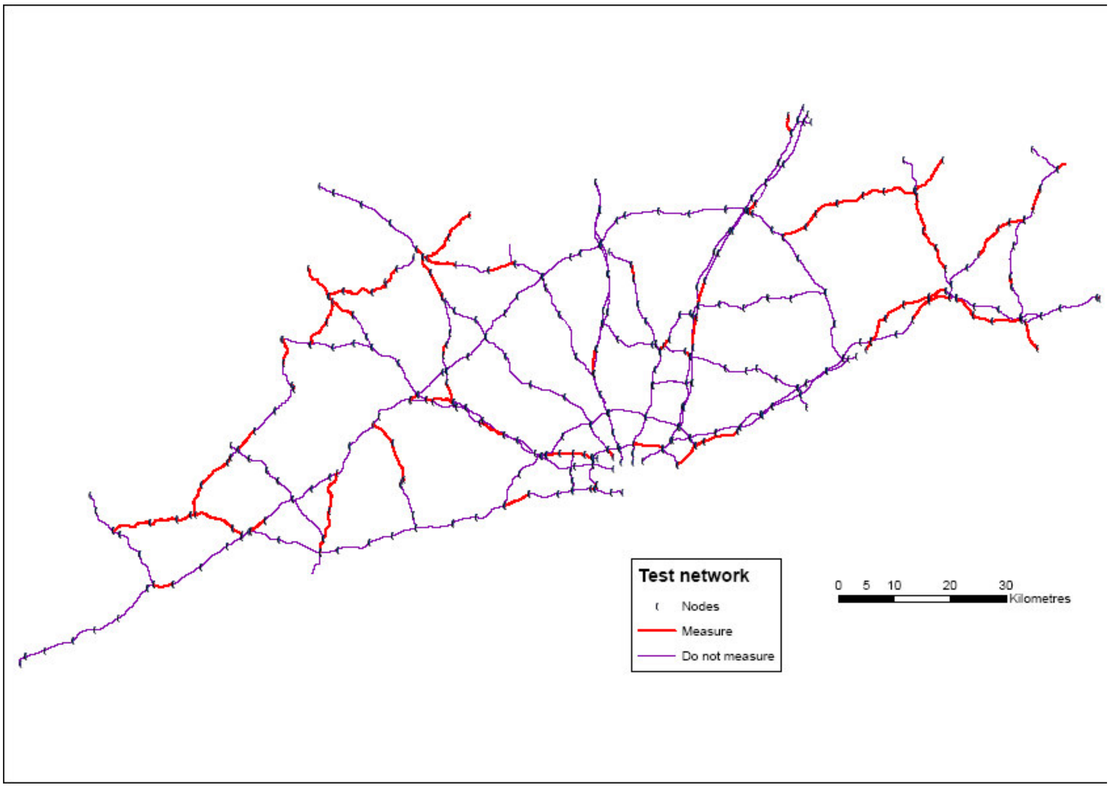

Figure 7. Optimised solution for roughness measurements on main and regional roads of Uusimaa Road District in Finland.

[7] www.ampl.com

From Figure 7 it can also be observed that the optimised measurement route is not connected. If the visual assessment of connectivity had not been possible, it could have been checked by applying Equation (9). Indeed, we calculated the sum of access matrices after first removing empty rows from the access matrix to reduce the required number of matrix multiplications to 178, which is the number of routes measured or driven in the optimised solution. Naturally, the sum of access matrices reveals the same fact as the map - that the route is not connected. This is due to one shortcoming of our example, namely that one-carriageway roads are measured in one direction only. We used the same expected gain for both directions. This implies that in the optimised solution, for most measured arcs, both directions are measured, and the balancing constraints for all nodes are satisfied even when the route is not connected.

The next problem is how to find a solution where the measurement route is connected. One solution might be to incorporate the access matrix as a constraint to the optimisation model presented in Equation (10). This, however, is proposed for further inquiry. At this moment, we can conclude that the current measurement policy could be altered so that both directions are measured on all roads. This can be done by reallocating the prevailing measurement budget because, in the light of our results, some of the roads now measured every year, could be measured every other year. Shifting the current policy has the added benefit of further improving the efficiency of data collection by reducing driving on without measuring.

5 Conclusions

Decision-makers concerned with road maintenance activities face the question about which road sections to measure as input to the maintenance management process. In this context it is important to predict the likelihood of each section to be reclassified as in need or not of maintenance. We apply a Bayesian analysis for developing the idea of gain from measurement as the expected length of reclassified road after measurement activities. This is then used as the objective function in an integer linear programming problem to maximise the expected gain.

We develop the linear optimisation model into a route optimisation model by introducing balancing constraints for each node in the network. The connectivity of an optimised solution is evaluated by using an access matrix, whose elements indicate the existence of connection between the nodes in the network. These two advances enable the use of an integer programming model for a route optimisation problem, where the actual routes to be measured (arcs) are selected based on the expected benefits of new measurements not yet taken. In this way the measurement budget is allocated as effectively as possible so as to enhance our decision-making process.

A simple stylised example, which shows the relevant aspects of our optimal selection and routing method was presented and applied to a more complicated real-life network. A cost-benefit analysis of measuring and modelling accuracies reveals the interesting result that if a road is measured in the previous year, it is more beneficial to model the current condition of the road from previous measurements than to measure it again this year. If no data is available from the preceding year, it is more beneficial to measure than to use a model. According to this calculation, the roads should be measured every other year.

The major limitations of our study relate to the available data, which has been collected mostly in one direction only. We therefore used the same expected gain from new measurements to be taken for both directions. This results in a situation where in the optimised solution most arcs that are measured are measured in both directions. The condition of lanes in different directions is correlated, but certainly not equal. Measuring in both directions would allow us to evaluate the expected gains of an optimised measurement program more realistically. This can be done by reallocating the prevailing measurement budget because, in the light of our results, some of the roads now measured every year, could be measured every other year.

Unfortunately, the resulting optimal solution could possess the property of non-connectivity, i.e. there is no connection from some nodes to other nodes on the selected route. This can happen despite the balancing condition of the number of incoming arcs equalling the number of leaving arcs. The access matrix concept provides a tool for checking the consistency of the route, and it can also be possibly incorporated as a constraint in our optimisation model. This is something that could be done at a later stage of the development work.

6 Acknowledgments

The work of Antti Ruotoistenmäki was partly funded by the Finnish section of the Nordic Road Association and partly by his previous employer, Pöyry Infra Oy. Mr. Pasi Porkka of the Helsinki School of Economics, kindly provided the platform for optimisation using the AMPL software. Ms Tiina Kiuru of Destia drew the maps of the test network in the Uusimaa Road District.

7 References

Berger, J O (1985). Statistical Decision Theory and Bayesian Analysis, 2nd ed. Springer-Verlag, New York. Third corrected printing 1993.

Ruotoistenmäki, A., Seppälä, T. & Kanto, A. (2006). Comparison of Modeling and Measurement Accuracy of Road Condition Data. Journal of Transportation Engineering, Vol. 132, No 9, pp. 715-721.

Sayers, M. W., Gillespie, T. D. and Queiroz, C. A. V. (1986a). *The International Road Roughness Experiment: Establishing Correlation and Calibration Standard for Measurement.* World Bank Technical Paper Number 45. 453 p.

Sayers, M. W., Gillespie, T. D. and Paterson, W. D. O. (1986b). *Guidelines for Conducting and Calibrating Road Roughness Measurements.* World Bank Technical Paper Number 46. 87 p.

UMTRI (2007). "Road roughness home page." *The University of Michigan Transportation Research Institute (UMTRI) and The Great Lakes Center for Truck and Transit Research.* <http://www.umtri.umich.edu/erd/roughness/index.html> (October 11, 2007).