ABSTRACT

The object of the present paper is to describe methods for making forecasts of the national traffic between areas in a country or of international traffic between countries.

The performance of two different methods: Kruithof's method and a new weighted least squares method for splitting the aggregated traffic is analyzed. Simulations show that if the structure of the traffic matrix to some extent changes, then the weighted least squares method should be preferred.

A number of forecasting models is examined on aggregated level. The "Airline model" turned out to be the best common model for all time series. For short term forecasting introduction of explanatory variables did not improve the performance.

1. INTRODUCTION

New telecommunication exchanges are able to register not only the traffic through the exchange, but also the volume of traffic to the different destinations. These data form the basis for both short term and longer term traffic forecasts for dimensioning and network planning.

The forecasts can be regarded as a matrix, where each element is the forecast of the traffic from one exchange to another. Usually it is possible to forecast the total incoming and outgoing traffic by some direct method as well, and with a better precision than for the individual point-to-point forecasts. In practical use of the forecasts it is necessary that the forecasts of the matrix elements are consistent with the forecasts of the total incoming and outgoing traffic. It is therefore necessary to adjust the forecasts of the elements and/or the forecasts of the totals so that the row and column sums of the forecast matrix equal the forecasts for the totals.

One way to construct and adjust these forecasts is the well known Kruithof's method [1]. In this paper we propose a weighted least squares (WLSQ) method, as an alternative. The advantage of the WLSQ is its ability to use more information than the Kruithof's method. It will be possible to take special characteristics of each element into consideration. To be able to perform the WLSQ, separate forecasts for each element in the matrix must be available. We discuss some methods for constructing these forecasts.

We are comparing the two methods, to see when one method is superior to the other. The answer depends on the amount of information on each element.

2. THE KRUITHOF'S METHOD

The Kruithof's method is based on forecasts for the total volume of incoming and outgoing traffic in each exchange. Some methods for constructing such forecasts are presented in a separate paragraph. These forecasts are then used to adjust the last known traffic matrix so that the row/column sums are correct. Kruithof's method is defined in the following way:

Let \( \{A_{i,j}^{t}\} \) be the last known traffic matrix and \( B_{i,,t}^{t}, B_{,,j},t \) be forecasts for the row/column sums of the matrix one timestep ahead. The Kruithof's estimates are then defined as the matrix \( \{B_{i,j}^{t}\} \) satisfying

\[
B_{i,j}^{t} = E_{i,t} F_{j,t} A_{i,j}^{t},
\]

\[
B_{i,,t}^{t} = \sum_{j} B_{i,j}^{t} \quad \text{and} \quad B_{,,j}^{t} = \sum_{i} B_{i,j}^{t}
\]

for some arrays \( E \) and \( F \). The Kruithof's method is uniquely defined even though the arrays \( E \) and \( F \) are not. The method of calculating the Kruithof's matrix has been improved [2].

The Kruithof's method is not able to treat elements in the matrix individually. The method is therefore not able to use additional information on separate elements.

Occasionally, some elements are not known in the previous traffic matrix. It is then necessary to insert estimates for the missing elements. But these estimates should not be given the same weight as the known values for the other elements.

Another and more serious disadvantage is that prior knowledge may lead one to expect a larger increase in some elements than in others. This is not possible to handle in Kruithof's method.

Kruithof's method is only using one (usually the last) known traffic matrix. When many previous traffic matrices are known not only one should be used. This is very important if seasonal variation is present.

Kruithof's method uses the previous traffic matrix as a forecast for the next traffic matrix. It will often be possible to
construct better forecasts for each element. At least one will have more confidence in some forecasts than the others. If the uncertainty varies, this information should be taken into consideration.

3. WEIGHTED LEAST SQUARES METHOD

In this paper we propose to use weighted least squares to construct consistent traffic matrix forecasts. Let \( C_{i,j,t} \) be a forecast for the traffic from exchange \( i \) to exchange \( j \), and let \( C_{i,..,t} \) and \( C_{..,j,t} \) be forecasts for the total outgoing traffic from exchange \( i \) and incoming traffic to exchange \( j \), respectively. \( C_{i,..,t} \) and \( C_{..,j,t} \) are equivalent to \( B_{i,..,t} \) and \( B_{..,j,t} \) in the last paragraph. We describe some methods to construct these forecasts in paragraph 5 and 6.

We define the weighted sum of squares \( Q(D,C) \) as

\[
Q(D,C) = \sum_{i,j} a_{i,j} (C_{i,j,t} - D_{i,j,t})^2 + \sum_{i} b_{i} (C_{i,..,t} - D_{i,..,t})^2 + \sum_{j} c_{j} (C_{..,j,t} - D_{..,j,t})^2
\]  
(3.1)

We then define the weighted least squares forecast as \( D = \{D_{i,j,t}\} \) which minimize \( Q(D,C) \) with respect to \( D \) under the constraints

\[
D_{i,..,t} = \sum_{j} D_{i,j,t} \quad \text{and} \quad D_{..,j,t} = \sum_{i} D_{i,j,t}
\]  
(3.2)

A natural choice of the weights \( a_{i,j}, b_{i}, \) and \( c_{j} \) is the inverse of the variances of the estimators \( D_{i,j,t}, D_{i,..,t} \) and \( D_{..,j,t} \). The solution of equation (3.1) under the constraints (3.2) is found using Lagrange's multiplier method.

The advantage of WLSQ is that all available information is used. Some of the information is contained in the forecasts and the knowledge about the forecast precision is accounted for in the weights.

One disadvantage of WLSQ is that if the forecasts for the row/column sum are very far from consistent with the forecasts for each element, and some elements are very small compared to others, the method may give negative forecasts for some elements. We do not believe this to be any practical problem.

4. COMPARISON OF THE TWO METHODS

4.1 Model assumptions

It is difficult to compare the two methods for forecasting traffic matrices. The result will depend heavily on the progress of the "true" matrix from one timepoint to the next, and on how precisely it is possible to forecast each element in the matrix and the row/column sums. The objective of this paper has been to give a rule of thumb as to when one method is superior to the other. The evaluation is done by using numerical simulation.

We have to notice an important difference between the two methods. In the Kruthoff's method the row/column sums remain unaltered during the adjustment procedure. In the weighted least squares method these sums are changed in the same way as the other forecasts according to the uncertainty in each forecast.

In Table 4.1 a typical traffic matrix (5x5) for parts of the Norwegian telephone network is given as a start matrix at time \( t=0 \). Other simulation runs have shown that the conclusions do not depend significantly on the start traffic matrix nor on the size of the matrix.

<p>| | | | | |</p>
<table>
<thead>
<tr>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td>1.2</td>
<td>1.5</td>
<td>3.2</td>
<td>4.2</td>
<td>1.3</td>
</tr>
<tr>
<td>5.8</td>
<td>2.1</td>
<td>10.3</td>
<td>8.5</td>
<td>2.0</td>
</tr>
<tr>
<td>.7</td>
<td>4.8</td>
<td>2.3</td>
<td>5.2</td>
<td>1.0</td>
</tr>
<tr>
<td>3.1</td>
<td>2.0</td>
<td>5.1</td>
<td>8.7</td>
<td>2.0</td>
</tr>
<tr>
<td>1.9</td>
<td>5.6</td>
<td>6.1</td>
<td>7.2</td>
<td>3.1</td>
</tr>
</tbody>
</table>

In the simulations, we generate the "true" traffic matrix one step ahead from a known traffic matrix. Each element is generated by the same stochastic formula. Let \( A_{0} = \{A_{i,j,0}\} \). The traffic matrix one step ahead, at time \( t \), is denoted \( A_{t} = \{A_{i,j,t}\} \). In the simulations \( A_{t} \) is generated by the formula

\[
A_{i,j,t} = \omega_{i,j} + \gamma_{i,j} A_{i,j,0}, \quad \text{where} \quad \gamma_{i,j} = e_{i,j} + \omega_{i,j} \quad (4.1)
\]

Here \( \omega_{i,j} \sim N(0,\sigma_{i,j}) \) and \( \omega_{i,j} \sim N(0,\sigma_{i,j}) \), and \( e \) and \( f \) are fixed arrays. The notation \( N(.,.) \) means that the variable is normally distributed with the given mean and standard deviation.

We then need some forecasts for \( A_{t} \). Because we want to separate the influence of various forecasting methods from the test of the two adjustment methods, we assume that we have available unbiased forecasts \( C_{i,j,t}, C_{i,..,t}, C_{..,j,t} \) for \( A_{i,j,t}, A_{i,..,t}, A_{..,j,t} \) respectively.

3.4B-3-2
More precisely, we assume

\[ C_{i,j} \sim \mathcal{N}(0, A_{i,j} Q \sigma_{i,j}^2) \]

\[ C_{i,j} \sim \mathcal{N}(0, A_{i,j} Q \sigma_{i,j}^2) \]

and

\[ C_{i,j} \sim \mathcal{N}(0, A_{i,j} Q \sigma_{i,j}^2) \]

where \( \sigma \) is equal to 1 or 0 depending on \( \alpha \) and \( \beta \).

In each simulation run no. \( n \), \( n=1,2,\ldots,N \), the "true" traffic matrices

\[ A^n_t = \{A^n_{i,j}\} \]

are generated, while the Kruithof's estimates \( B^n_t = \{B^n_{i,j}\} \) and the WLSQ estimates \( D^n_t = \{D^n_{i,j}\} \) are found.

The norms used in the tests are

\[ \text{KR} = \frac{1}{N} \sum_{n=1}^{N} \sqrt{\frac{1}{M} \sum_{i,j} (A^n_{i,j} - B^n_{i,j})^2} \]

and

\[ \text{LS} = \frac{1}{N} \sum_{n=1}^{N} \sqrt{\frac{1}{M} \sum_{i,j} (A^n_{i,j} - D^n_{i,j})^2} \]

where \( M \) is the number of exchanges in the network. Other norms for the forecasting precision have also been examined, without significantly different results.

In all these tests, we have done at least 100 simulations.

4.2 Simulation results based on exact forecasts for the row and column sums

In the first simulation, test one, we assume that \( \alpha = 1, \beta = 0 \) for all \( i \) and \( j \), so that the marginals are forecasted exactly. The reason for this is the different handling of row/column sums in the two methods. The Kruithof's method does not change the forecast for the row/column sums, while the WLSQ method does. In this way we are able to separate the consequences of the different handling of these forecasts from the rest of the method. When the marginals are assumed to be forecasted exactly, it is natural to change the \( Q(D,C) \) in the least squares method to

\[ Q(D,C) = \frac{1}{2} \sum_{i,j} (C_{i,j} - D_{i,j})^2 \]

Then the parameters are chosen in the following way: \( \sigma = 0 \), \( \alpha = 1 \), \( \beta = 1 \) for all \( i \) and \( j \), and \( \delta = 0 \). The rest of the parameters are given in Table 4.2.

![Table 4.2 Test 1. Forecasting precision of Kruithof's method and WLSQ as a function of \( \sigma_{i,j} \) and \( \sigma_{i,j}^2 \)]

<table>
<thead>
<tr>
<th>( \sigma_{i,j} )</th>
<th>( \sigma_{i,j}^2 )</th>
</tr>
</thead>
<tbody>
<tr>
<td>KR</td>
<td>.04 .08 .12 .16 .20</td>
</tr>
<tr>
<td>( 1.0 )</td>
<td>.02 .04 .06 .08 .10</td>
</tr>
<tr>
<td>( 1.0 )</td>
<td>.08 .12 .14 .16 .20</td>
</tr>
<tr>
<td>LS</td>
<td>.083 .14 .20 .25 .32</td>
</tr>
<tr>
<td>( 1.0 )</td>
<td>.37 .44 .50 .52 .60</td>
</tr>
</tbody>
</table>

The forecasts for each element in the traffic matrix have the same standard deviation. \( \sigma_{i,j}^2 \) means that we are able to construct a forecast for element \( i,j \) which is superior to the value from the previous matrix. The smaller \( \sigma_{i,j}^2 \) is compared to \( \sigma_{i,j}^2 \), the better is WLSQ compared to Kruithof's method. When they are approximately of the same size, Kruithof's method is a little better than WLSQ.

The only change in the parameters from test one to test two is that \( \alpha \) and \( \beta \) are no longer all 1, but is given as in Table 4.3.

![Table 4.3 The e and f arrays in test 2]

<table>
<thead>
<tr>
<th>( i/j )</th>
<th>1</th>
<th>2</th>
<th>3</th>
<th>4</th>
<th>5</th>
</tr>
</thead>
<tbody>
<tr>
<td>( e_i )</td>
<td>1.02</td>
<td>1.04</td>
<td>1.03</td>
<td>1.05</td>
<td>1.03</td>
</tr>
<tr>
<td>( f_j )</td>
<td>1.04</td>
<td>1.07</td>
<td>.99</td>
<td>1.01</td>
<td>1.06</td>
</tr>
</tbody>
</table>

When \( e_i \) and \( f_j \) are not all 1, one should expect Kruithof's method to perform almost exactly as in the previous test. This is because the Kruithof's method is exact for all \( e_i \) and \( f_j \) arrays, when \( \sigma = 0 \). The results in this test confirm this assumption. In this test we got the same results for WLSQ as in test one. The reason for this is that the estimates are the same in both tests.

![Table 4.4 Results from test 2]

<table>
<thead>
<tr>
<th>( \sigma_{i,j} )</th>
<th>( \sigma_{i,j}^2 )</th>
</tr>
</thead>
<tbody>
<tr>
<td>KR</td>
<td>.04 .08 .12 .16 .20</td>
</tr>
<tr>
<td>( 1.0 )</td>
<td>.04 .08 .12 .16 .20</td>
</tr>
<tr>
<td>( 1.0 )</td>
<td>.08 .12 .14 .16 .20</td>
</tr>
<tr>
<td>LS</td>
<td>.083 .14 .20 .25 .32</td>
</tr>
<tr>
<td>( 1.0 )</td>
<td>.37 .44 .50 .52 .60</td>
</tr>
</tbody>
</table>

In the third simulation (see Table 4.5) we have chosen varying precision in the forecasts for each element. Thus \( \sigma_{i,j} \) is chosen at random uniformly in the interval \( (\sigma_{i,j}^2, \sigma_{i,j}^2 + 0.02) \).

![Table 4.5 Results from test 3]

<table>
<thead>
<tr>
<th>( \sigma_{i,j} )</th>
<th>( \sigma_{i,j}^2 )</th>
</tr>
</thead>
<tbody>
<tr>
<td>KR</td>
<td>.04 .08 .12 .16 .20</td>
</tr>
<tr>
<td>( 1.0 )</td>
<td>.02 .04 .06 .08 .10</td>
</tr>
<tr>
<td>( 1.0 )</td>
<td>.12 .14 .16 .20 .22</td>
</tr>
<tr>
<td>LS</td>
<td>.081 .14 .19 .25 .31</td>
</tr>
<tr>
<td>( 1.0 )</td>
<td>.37 .44 .50 .52 .55</td>
</tr>
</tbody>
</table>
The results from the WLSQ method are almost the same as in test 1. This means that when \( u_{i,j} \) is varying uniformly in an interval, the results are almost the same as when \( o_{i,j} \) is constant and equal to the midpoint in the interval.

In simulation four (see Table 4.6) we let \( u_{i,j} \) vary. Because \( u_{i,j} \) is not multiplied with \( A_{i,j,0} \) we use \( o = 1 \).

Then it is possible to compare \( o_{u_{i,j}} \) and \( o_{i,j} \) In this test we let \( o_{i,j} = 0 \). The rest of the parameters are defined in Table 4.6.

Table 4.6 Results from test 4

<table>
<thead>
<tr>
<th>Parameter</th>
<th>0.04</th>
<th>0.08</th>
<th>0.12</th>
<th>0.16</th>
<th>0.20</th>
</tr>
</thead>
<tbody>
<tr>
<td>KR</td>
<td>0.054</td>
<td>0.078</td>
<td>0.10</td>
<td>0.13</td>
<td>0.16</td>
</tr>
</tbody>
</table>

In this test WLSQ is better when \( o_{i,j} < o_{u_{i,j}} \) or the two variables are approximately of the same size. This is a very interesting result. In test 1-3 the stocastic parameter \( o \) was multiplied with \( A_{i,j,0} \).

This is near Kruihof's formula (2.1). In test 1-3 Kruihof's method was a little better, when we had the same uncertainty in the forecasts compared to the precision of \( A_0 \) when used as an estimate for \( A_0 \). In this test the stocastic variable \( o \) is added to \( A_{i,j,0} \).

This is why WLSQ is better in this test when the uncertainty is the same in the forecasts used by WLSQ as when \( A_0 \) is used as a forecast for \( A_0 \).

4.3 Simulation results when the row and column sums are stocastic

We finally perform exactly the same tests as before, with the exception that the row and column sums now are stocastic with \( o_{i,j} = o_{u_{i,j}} = 0.04 \). This is more realistic than the previous test, as there are no reason to believe that the total future traffic will be known exactly.

Table 4.7 Test 5. Forecasting precision of Kruihof's method and WLSQ as a function of \( o_u \) and \( o_{i,j} \).

<table>
<thead>
<tr>
<th>Parameter</th>
<th>0.04</th>
<th>0.08</th>
<th>0.12</th>
<th>0.16</th>
<th>0.20</th>
</tr>
</thead>
<tbody>
<tr>
<td>KR</td>
<td>0.25</td>
<td>0.30</td>
<td>0.39</td>
<td>0.47</td>
<td>0.56</td>
</tr>
<tr>
<td>( o_{i,j} )</td>
<td>0.02</td>
<td>0.04</td>
<td>0.06</td>
<td>0.08</td>
<td>0.10</td>
</tr>
<tr>
<td>LS</td>
<td>0.14</td>
<td>0.22</td>
<td>0.28</td>
<td>0.35</td>
<td>0.39</td>
</tr>
</tbody>
</table>

The error is naturally a little larger in this test than in test 1, where the exact row and column sums were known. But the difference is decreasing with increasing \( o_u \) and \( o_{i,j} \).

In this test, as in test 1, WLSQ is the better when \( o_{i,j} < o_u \).

Table 4.8 Results from test 6

<table>
<thead>
<tr>
<th>Parameter</th>
<th>0.04</th>
<th>0.08</th>
<th>0.12</th>
<th>0.16</th>
<th>0.20</th>
</tr>
</thead>
<tbody>
<tr>
<td>KR</td>
<td>0.24</td>
<td>0.30</td>
<td>0.37</td>
<td>0.45</td>
<td>0.54</td>
</tr>
</tbody>
</table>

In test 6, the e and f arrays are not equal to 1. Just as in test 1 and 2, we see that the Kruihof's method is exactly as good as when the e and f arrays are equal to 1.

Table 4.9 Results from test 7

<table>
<thead>
<tr>
<th>Parameter</th>
<th>0.02</th>
<th>0.04</th>
<th>0.06</th>
<th>0.08</th>
<th>0.10</th>
<th>0.12</th>
<th>0.14</th>
<th>0.16</th>
</tr>
</thead>
<tbody>
<tr>
<td>LS</td>
<td>0.17</td>
<td>0.23</td>
<td>0.28</td>
<td>0.35</td>
<td>0.40</td>
<td>0.45</td>
<td>0.50</td>
<td>0.58</td>
</tr>
</tbody>
</table>

In test 7, \( o_{i,j} \) is chosen at random from a uniform distribution on the interval \((-0.02,0.02)\). Again, the results are very close to the results from test 5. This means that when the standard deviation is varying uniformly in an interval, then WLSQ is approximately as good as when the standard deviation is fixed in the middle of the interval.

Table 4.10 Result from test 8

<table>
<thead>
<tr>
<th>Parameter</th>
<th>0.04</th>
<th>0.08</th>
<th>0.12</th>
<th>0.16</th>
<th>0.20</th>
</tr>
</thead>
<tbody>
<tr>
<td>KR</td>
<td>0.22</td>
<td>0.22</td>
<td>0.23</td>
<td>0.25</td>
<td>0.26</td>
</tr>
<tr>
<td>( o_{i,j} )</td>
<td>0.02</td>
<td>0.04</td>
<td>0.06</td>
<td>0.08</td>
<td>0.10</td>
</tr>
<tr>
<td>LS</td>
<td>0.12</td>
<td>0.18</td>
<td>0.20</td>
<td>0.23</td>
<td>0.25</td>
</tr>
</tbody>
</table>

In test 8 the stocastic element is added to each element in \( A_0 \). When \( o_u \) and \( o_{i,j} \) are small then the dominating error is in the row and column sums. The Kruihof's method, which does not take the uncertainty in each forecast into consideration, is not able to use this information. The error is almost as large for \( o_u = 0.04 \) as for 0.12.

In addition, we notice that WLSQ is better when it is possible to construct a better forecast for each element than the corresponding element in the previous matrix.

4.4 Conclusion

In this paper we have presented a weighted least squares method (WLSQ) for constructing consistent forecasts for traffic matrices. We have compared it with Kruihof's method, using simulation techniques. The simulations show that if we are able to give forecasts for each element which are better forecasts than the
value in the last known matrix, then we should use weighted least squares. If we are unable to construct forecasts which are better than the values in the previous traffic matrix, then we should use the Kruithof's method when we expect the change to be close to the Kruithof's formula (4.5), else WLSQ should be used.

5. TRAFFIC FORECASTING MODEL BASED ON TIME SERIES ANALYSIS

5.1 Different methods

A number of different methods can be used to forecast the traffic on aggregated and local levels. The most relevant methods are:

- Time series models
- Kalman filter models
- Regression models
- Econometric models

The Kalman filter models are used by administrations in USA and are described in a number of papers, among others in [3], [4] and [5]. In this paper only the time series models will be examined.

Identification and estimation of the univariate models resulted in different model structures in some cases. This is a drawback in practical situations. Therefore analysis were carried out to find the best common model for all time series.

The paper presents a comparison between the best univariate model, the best common univariate model and models based on number of main stations (subscribers) as explanatory variables.

5.2 The best univariate model

In the Norwegian telephone network traffic measurements are carried out four times a year. Quarterly observations of incoming and outgoing traffic from six groups exchanges for the last eight years have been studied. By using time series analysis the best univariate forecasting models were developed.

Let \( Y_t \) be the traffic at time \( t \), \( B \) the backward shift operator, \( \omega \) and \( \alpha \) parameters and \( \epsilon_t \) white noise at time \( t \). Then the best univariate models are expressed by:

**Outgoing traffic**

<table>
<thead>
<tr>
<th>Exchange</th>
<th>( \omega_1 )</th>
<th>( \alpha_4 )</th>
<th>( Q_{22} )</th>
</tr>
</thead>
<tbody>
<tr>
<td>Sandnes</td>
<td>0.72</td>
<td>1.03x</td>
<td>37.6</td>
</tr>
<tr>
<td>Narvik</td>
<td>0.76</td>
<td>0.61</td>
<td>12.3</td>
</tr>
<tr>
<td>Mandal</td>
<td>0.55</td>
<td>0.76</td>
<td>20.3</td>
</tr>
<tr>
<td>Sandefjord</td>
<td>0.43</td>
<td>0.73</td>
<td>16.9</td>
</tr>
<tr>
<td>Tvedestrand</td>
<td>0.69</td>
<td>0.91</td>
<td>16.2</td>
</tr>
<tr>
<td>Kragers</td>
<td>0.32</td>
<td>0.46</td>
<td>19.3</td>
</tr>
</tbody>
</table>

\( x \) Insignificant value.

Only in a few of the models the log transform of the data turned out to give a better fitting.

5.3 The best common univariate model

The data base for network planning consists of a large number of time series. It would be extremely time-consuming to analyse every time series in order to find the best model. For such analysis it is also necessary to use special experts in this field. The most relevant solution in a practical situation will be to identify the best common model for the time series and then estimate the parameters in the models separately.

Analysis have shown that the so-called "Airline model" [6]:

\[
Y_t = \omega + (1-B)(1-B^4)\alpha \epsilon_t
\]  

(5.3)

is the best common model for outgoing traffic. The results are similar to forecasting models developed in [7]. Table 5.1 contains the estimated parameter values and the Portmanteau test for the different time series. The Portmanteau test \( Q \) is based on estimation of 24 autocorrelation coefficients and has 22 degrees of freedom.

**Table 5.1 Outgoing traffic. Estimated parameter values and the Portmanteau test**

\[
Y_t = (1-B)(1-B^4)\omega + (1-B)(1-B^4)\alpha \epsilon_t
\]  

(5.2)

As for outgoing traffic the best common model for incoming traffic is the Airline model given by equation (5.3). The estimated parameters and the Portmanteau test are given in Table 5.2.
Table 5.2 Incoming traffic. Estimated parameter values and the Postmanteau test

<table>
<thead>
<tr>
<th>Exchange</th>
<th>$\theta_1$</th>
<th>$\theta_4$</th>
<th>$Q_{22}$</th>
</tr>
</thead>
<tbody>
<tr>
<td>Sandnes</td>
<td>-0.45</td>
<td>0.67</td>
<td>12.7</td>
</tr>
<tr>
<td>Narvik</td>
<td>0.41</td>
<td>0.55</td>
<td>20.8</td>
</tr>
<tr>
<td>Mandal</td>
<td>0.43</td>
<td>0.81</td>
<td>35.6</td>
</tr>
<tr>
<td>Sandefjord</td>
<td>0.55</td>
<td>0.86</td>
<td>19.8</td>
</tr>
<tr>
<td>Tvedestrand</td>
<td>0.61</td>
<td>0.63</td>
<td>15.7</td>
</tr>
<tr>
<td>Kragerø</td>
<td>0.79</td>
<td>0.17</td>
<td>11.3</td>
</tr>
</tbody>
</table>

5.4 Traffic models based on number of main stations

Explanatory variables like telephone tariffs, consumers price index, number of main stations may improve the forecasting models. In Norway the number of main stations within an exchange area has been the most significant variable during the last years. The telephone tariffs and the consumers price index have not affected the traffic in the same way because their relative increase is about the same.

On aggregated level the number of main stations is included as an explanatory variable. The first step has then been to make one model for the traffic per main station and then another for the number of main stations within the exchange area. The last model is used as an input to the first one.

Analysis have shown that the best univariate models are expressed by:

\[
\begin{align*}
\text{Outgoing traffic per main station} \\
\text{Sandnes: } & Y_t = (1-\theta_1 B)(1-\theta_4 B^4) a_t \\
\text{Narvik: } & Y_t = (1-\theta_1 B)(1-\theta_4 B^4) a_t \\
\text{Mandal: } & Y_t = (1-\theta_1 B)(1-\theta_4 B^4) a_t \\
\text{Sandefjord: } & Y_t = (1-\theta_1 B)(1-\theta_4 B^4) a_t \\
\text{Tvedestrand: } & Y_t = (1-\theta_1 B)(1-\theta_4 B^4) a_t \\
\text{Kragerø: } & Y_t = (1-\theta_1 B)(1-\theta_4 B^4) a_t
\end{align*}
\]

\[
\begin{align*}
\text{Incoming traffic per main station} \\
\text{Sandnes: } & (1-B)(1-B^3) Y_t = (1-\theta_1 B)(1-\theta_4 B^4) a_t \\
\text{Narvik: } & (1-B)(1-B^3) Y_t = (1-\theta_1 B)(1-\theta_4 B^4) a_t \\
\text{Mandal: } & (1-B) Y_t = (1-\theta_1 B)(1-\theta_4 B^4) a_t \\
\text{Sandefjord: } & (1-B) Y_t = (1-\theta_1 B)(1-\theta_4 B^4) a_t \\
\text{Tvedestrand: } & (1-B) Y_t = (1-\theta_1 B)(1-\theta_4 B^4) a_t \\
\text{Kragerø: } & (1-B)(1-B^4) Y_t = (1-\theta_1 B)(1-\theta_4 B^4) a_t
\end{align*}
\]

Also for outgoing and incoming traffic per main station analysis have shown that the best common model is the Airline model given in (5.3).

The development of the number of main stations in an exchange area is well fitted by the Airline model. However, in some cases there are only annual observations available. The missing intermediate observations are then estimated by an interpolation procedure which obviously may affect the modelling. The best common model in such situations has turned out to be:

\[
(1-B)^2 x_t = (1-\theta_1 B^4) a_t
\]

where $x_t$ is the number of main stations at time $t$.

6. COMPARISON OF THE VARIOUS FORECASTING METHODS

6.1 Methods for comparison

For evaluation of the different forecasting methods it is necessary to introduce a comparison measure. The root mean square error (RMSE) compares the forecasts with the true observations.

In all time series the last $m=4$ observations are removed before the modelling procedure. Then the forecasting model is used to forecast the traffic $m$ steps ahead in order to compare the forecasts with the true observations. The root mean square error is given by:

\[
\text{RMSE} = \sqrt{\frac{1}{m} \sum_{k=1}^{m} (Y_{n+k} - Y_{n+k}^*)^2}
\]

where $n$ is the actual time before forecasting, $Y_{n+k}$ the true value at time $n+k$, and $Y_{n+k}^*$ the forecast $k$ steps ahead.

6.2 Forecasting models for outgoing traffic

The root mean square error is calculated for:

- the best univariate models (BU), given by equation (5.1)
- the best common univariate model (BC), given by equation (5.3) and Table 5.1
- the best univariate model per main station given by equation (5.4), combined with the best common model for number of main stations (BUM)
- the best common univariate model per main station given by equation (5.3) and Table 5.1, combined with the best common model for number of main stations (BCM).

The results are presented in Table 6.1.
Table 6.1 Root mean square error for different forecasting models for outgoing traffic in Erlang

<table>
<thead>
<tr>
<th>Exchange Model</th>
<th>Sandnes</th>
<th>Narvik</th>
<th>Mandal</th>
<th>Sandefjord</th>
<th>Tvedestrand</th>
<th>Kragerø</th>
<th>Mean</th>
</tr>
</thead>
<tbody>
<tr>
<td>BU</td>
<td>3.9</td>
<td>3.0</td>
<td>1.1</td>
<td>3.5</td>
<td>2.8</td>
<td>9.8</td>
<td>4.0</td>
</tr>
<tr>
<td>BC</td>
<td>13.4</td>
<td>3.1</td>
<td>1.1</td>
<td>3.5</td>
<td>2.6</td>
<td>9.8</td>
<td>5.6</td>
</tr>
<tr>
<td>BCM</td>
<td>7.8</td>
<td>3.7</td>
<td>1.8</td>
<td>8.0</td>
<td>5.3</td>
<td>7.1</td>
<td>5.6</td>
</tr>
<tr>
<td>BCM</td>
<td>6.4</td>
<td>2.3</td>
<td>1.4</td>
<td>8.0</td>
<td>3.5</td>
<td>0.9</td>
<td>5.0</td>
</tr>
</tbody>
</table>

6.3 Forecasting models for incoming traffic

The root mean square error is calculated for:
- the best univariate models (BU), given by equation (5.2)
- the best common univariate model (BC), given by equation (5.3)
- the best univariate model per main station given by equation (5.5), combined with the best common model for number of main stations (BUM)
- the best common univariate model per main station given by equation (5.3), combined with the best common model for number of main stations (BCM).

The results are presented in Table 6.2.

Table 6.2 Root mean square error for different forecasting models for incoming traffic in Erlang

<table>
<thead>
<tr>
<th>Exchange Model</th>
<th>Sandnes</th>
<th>Narvik</th>
<th>Mandal</th>
<th>Sandefjord</th>
<th>Tvedestrand</th>
<th>Kragerø</th>
<th>Mean</th>
</tr>
</thead>
<tbody>
<tr>
<td>BU</td>
<td>37.3</td>
<td>0.3</td>
<td>0.9</td>
<td>6.6</td>
<td>1.4</td>
<td>4.5</td>
<td>8.5</td>
</tr>
<tr>
<td>BC</td>
<td>33.0</td>
<td>1.3</td>
<td>0.8</td>
<td>2.4</td>
<td>1.4</td>
<td>4.5</td>
<td>7.3</td>
</tr>
<tr>
<td>BCM</td>
<td>35.2</td>
<td>1.6</td>
<td>1.3</td>
<td>4.7</td>
<td>1.0</td>
<td>4.6</td>
<td>8.2</td>
</tr>
<tr>
<td>BCM</td>
<td>35.6</td>
<td>1.6</td>
<td>1.1</td>
<td>4.1</td>
<td>1.0</td>
<td>4.6</td>
<td>8.1</td>
</tr>
</tbody>
</table>

6.4 Evaluation of the forecasting models

Comparison between the best univariate models and the best common univariate model for both outgoing and incoming traffic shows that there is a small difference in the behaviour of the models. Both types of models perform very well except for Sandnes group exchange, where the traffic increased unexpectedly for the last four observations. Since there is only a small difference in the performance between the models, the conclusion is that the best common model is used for forecasting.

The root mean square error is based on four observations representing a time period of one year. Hence, the conclusions are valid for short term forecasting.

The results in Tables 6.1 and 6.2 show that introduction of the number of main stations as an explanatory variable does not improve the performance of the forecasting models to any great extent. This means that the best common univariate model, the Airline model, given by equation (5.3), is preferred for short term forecasting of outgoing and incoming traffic.

The outgoing and incoming traffic represents the column sum and the row sum respectively, in the traffic matrix. Concerning the traffic elements within the matrix, it is not recommended to have more complicated models for short term forecasting. The traffic elements will among other things depend on the geographical distance between the exchanges and the number of main stations in the exchange areas. But the fact that models without explanatory variable are preferred on higher levels indicates that models of no more than the same complexity should be used on the lower levels for short term forecasting.

For long term forecasting additional studies have to be carried out in order to examine the effect of using explanatory variables.

7. ACKNOWLEDGEMENT

The authors thank Eivind Damsleth for helpful discussions and exchange of view in preparing this paper.

REFERENCES