Software and Services for the Quality Management of Traffic Data (QUATRA)

**Duration:** 01/10/2011 – 30/10/2013 (25 months)

**Budget:** EUR 290,000

**Coordinator:** nast consulting ZT GmbH, Austria
Friedrich Nadler
e-mail: office@nast.at

**Partner:** TRANSVER GmbH, Germany
Thomas Heinrich
e-mail: heinrich@transver.de
Content of the presentation

1. QUATRA project objectives
2. online evaluation system for freeway data
3. offline evaluation system for urban road data
4. results of the field tests
5. business concept
6. benefits
QUATRA objective

• system development for quality management of traffic data

1. **Freeway tool**: focus on evaluation of the quality of incoming traffic data **online** for quick response in case of abnormal data/traffic conditions (e.g. for purpose of traffic control)

2. **Urban roads tool**: focus on a similar process for cities **offline** e.g. for efficient scheduling of repairs of faulty detectors

• QUATRA combines a diversity of quality indicators and statistical approaches to check the quality of the traffic data
Freeway tool (online)
Local indicators
• for each detector: counting of intervals with e. g. missing data, error messages generated by detector itself, alerts from statistical model

Global indicators
• for each measurement cross section: comparison of the counting with previous cross section

Plausibility indicators
• for each detector or measurement cross section: counting of intervals with alerts from 22 plausibility tests, e. g. threshold tests, if a speed value has been measured → at least one vehicle has to be counted, more hgv on the main lane than on the overtaking lanes
Statistical model – approach

- A general regression model has been implemented to check whether detector measurements are likely to be faulty
- Statistical information of historical data is generated and used

- Prediction intervals are generated
- Value prediction based on combination of available variables of the data set
Statistical model - identification of detector failure

- collected data (number of vehicles for a given time interval at a specific detection site) is analysed and compared with the predicted number of vehicles from historical information

Identification of abnormal data is based on two assumptions:
1) the measured number of vehicles must lie within its corresponding prediction interval (regression model)
2) the measured number of vehicles must lie within the range of the 2.5% and 97.5% percentile (normal range) of the historical information
Software

- extension of existing offline software tool LOTRAN-DQ 2 for quality management of traffic data
- online calculation of all indicators (for floating window)
- implementation of statistical model
- calculation and visualisation of new indicators
System architecture LOTRAN-DQ 2

**Offline**

- **Server**
  - LOTRAN-DQ2-Comp
  - calculate quality indicators

- **Central Data Base**
  - store quality indicators

- **LOTTRAN-DQ2-Plugin**
  - display quality indicators

- **Client** (with GUI)
  - read indicators

**Online**

- **Server**
  - LOTRAN-DQ2-Plugin
  - display quality indicators

- **Client** (with GUI)
  - read indicators
System architecture LOTRAN-DQ 2.5 QUATRA

Online

LOTRAN-DQ2-Comp
calculate quality indicators for moving time frames

Central Data Base
store quality indicators

LOTTRAN-DQ2-Plugin
display current quality indicators

Control unit
Client (with GUI)

Server

① write indicators cyclic

② notify display

③ read indicators

Data Information

LOTTRAN-DQ2-Plugin
display current quality indicators

Client (with GUI)

System architecture LOTRAN-DQ 2.5 QUATRA

Online

LOTRAN-DQ2-Comp
calculate quality indicators for moving time frames

Central Data Base
store quality indicators

LOTTRAN-DQ2-Plugin
display current quality indicators

Control unit
Client (with GUI)
Implementation of the Statistical Model

- implementation of pattern matching (as basis for online calculation of statistical indicators)
- implementation of online calculation of statistical indicators

Requirements for the pattern matching

- problem: sequential matching of data from all detectors for a long period (e.g. one year) → high computing resources required
- solution: calculation for corridors
GUI

road section

indicator visualisation

freeway (online)
Urban tool (offline)
Local indicators

- for each detector: counting of intervals with e.g. missing data, error messages generated by detector itself, alerts from statistical model, stagnant detector

Global indicators

- comparison of the countings of neighboring cross section only for bounded stretches/networks

Plausibility indicators

- for each detector or measurement cross section: counting of intervals with alerts from plausibility tests depending on the available traffic variables (e.g. counting, speed, classification)
Statistical model

- contrary to freeways urban traffic networks are open systems which allow vehicles to “disappear” or “appear”
- they are equipped less densely with traffic detection sites
- within the statistical approach for urban roads the information of spatial connected detection sites and historical data is used to calculate pair wise correlations
- correlations are calculated for each pair of available detection sites
- step 1: if the correlation is verified as being outside of a defined verification interval both traffic sites are treated providing inadequate data
- step 2: once all pairs are validated the detection sites are identified that have the highest number of mis-matching correlations. The data of all remaining traffic sites is treated as normal again
Statistical model

Example for spatial connected traffic detection sites

![Diagram of spatial connected detection sites](image)

Data example for spatial connected detection sites

<table>
<thead>
<tr>
<th>Hour</th>
<th>$q_{all}$/h</th>
</tr>
</thead>
<tbody>
<tr>
<td>0</td>
<td>100</td>
</tr>
<tr>
<td>5</td>
<td>200</td>
</tr>
<tr>
<td>10</td>
<td>300</td>
</tr>
<tr>
<td>15</td>
<td>400</td>
</tr>
<tr>
<td>20</td>
<td>500</td>
</tr>
<tr>
<td>25</td>
<td>600</td>
</tr>
</tbody>
</table>

Pairwise correlation of the spatial connected detection sites:

- M1 / M2: 0.9968
- M1 / M3: 0.9136
- M1 / M4: 0.9932
- M2 / M3: 0.9065
- M2 / M4: 0.9941
- M3 / M4: 0.8876

Number of pairwise correlations outside of the verification interval:

- M1: 1
- M2: 1
- M3: 3
- M4: 1

High likelihood providing inaccurate data.
# GUI

![Screenshot of the roadERA.net application](image)

**Indicator Visualisation**

## Absolute Values

<table>
<thead>
<tr>
<th>MQ_NAME</th>
<th>KnotenNr</th>
<th>Spur1</th>
<th>m1</th>
<th>m2</th>
<th>m3</th>
<th>m4</th>
<th>m5</th>
<th>m6</th>
<th>m7</th>
<th>m8</th>
<th>m9</th>
<th>m10</th>
<th>m11</th>
<th>m12</th>
<th>m13</th>
<th>m14</th>
<th>m15</th>
<th>m16</th>
<th>m17</th>
<th>m18</th>
</tr>
</thead>
<tbody>
<tr>
<td>MA46_Schleuse_1101_R1</td>
<td>1211</td>
<td>1</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td></td>
</tr>
<tr>
<td>MA46_Schleuse_1101_R2</td>
<td>1211</td>
<td>2</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td></td>
</tr>
<tr>
<td>MA46_Schleuse_1101_R3</td>
<td>1211</td>
<td>3</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td></td>
</tr>
<tr>
<td>MA46_Schleuse_1101_R4</td>
<td>1211</td>
<td>4</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td></td>
</tr>
<tr>
<td>MA46_Schleuse_1102_R1</td>
<td>21</td>
<td>21</td>
<td>50</td>
<td>00</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td></td>
</tr>
<tr>
<td>MA46_Schleuse_1102_R2</td>
<td>21</td>
<td>50</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td></td>
</tr>
<tr>
<td>MA46_Schleuse_1102_R3</td>
<td>21</td>
<td>50</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td></td>
</tr>
<tr>
<td>MA46_Schleuse_1102_R4</td>
<td>21</td>
<td>50</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td></td>
</tr>
<tr>
<td>MA46_Schleuse_1103_R1</td>
<td>22</td>
<td>22</td>
<td>10</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td></td>
</tr>
<tr>
<td>MA46_Schleuse_1103_R2</td>
<td>22</td>
<td>10</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td></td>
</tr>
<tr>
<td>MA46_Schleuse_1103_R3</td>
<td>22</td>
<td>10</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td></td>
</tr>
<tr>
<td>MA46_Schleuse_1103_R4</td>
<td>22</td>
<td>10</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td></td>
</tr>
<tr>
<td>NCE_2201_R1</td>
<td>2201</td>
<td>2201</td>
<td>1</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td></td>
<td></td>
</tr>
<tr>
<td>NCE_2201_R2</td>
<td>2201</td>
<td>2201</td>
<td>1</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td></td>
<td></td>
</tr>
<tr>
<td>NCE_2201_R3</td>
<td>2201</td>
<td>2201</td>
<td>1</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td></td>
<td></td>
</tr>
<tr>
<td>NCE_2201_R4</td>
<td>2201</td>
<td>2201</td>
<td>1</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

---

**Note:** The screenshot shows a part of the roadERA.net application interface, specifically the visualisation of indicators for traffic flow and other road-related data. The table and diagram provide detailed data for specific points and times, which can be useful for traffic management and analysis.
Software

- adaption of existing offline software tool LOTTRAN-DQ 2 for quality management of traffic data
- implementation of statistical model
- calculation and visualisation of new indicators
Field tests – method

- online implementations and tests of freeway tool/statistical model in Austria and Bavaria
- Relevance: manual check of a sample of alerts from the statistical model if there is an anomaly at the traffic data
- Significance: manual check of a sample of known anomalies of traffic data if there is an alert from the statistical model
Relevance – low number of hgv (undercounting?)

no speed drop
Relevance – high number of vehicles (overcounting)

Quality Indicators

Speed

no speed drop
Significance – high number of vehicles (overcounting?)

m12: $Q_{\text{Kfz}} > \text{threshold}$
m13: $Q_{\text{Pkw}} > \text{threshold}$
Significance – high number of vehicles (overcounting?)

Quality Indicators

Speed

no speed drop
Significance – speed drop (congestion)
Significance – speed drop (congestion)

out1 detects high flows but not speed drop
Field tests – results

- anomalous countings have been reliably detected mainly by statistical indicator out1 (prediction intervals)
- anomalous speeds have not yet been detected by statistical indicators

→ use of statistical indicator out1 for quality assessment of traffic counting (together with or instead of difficult to interpret global indicators)
→ indicators show (indicate) that there is an anomaly
→ further analyses required to find the reason of the anomaly
→ further improvement of the detection of anomalies by optimisation of the parameters of the statistical model expected
Business concept

• QUATRA software system core will be provided at no charge
• Installation/setup and maintenance fees
• Individual negotiations with each client regarding priorities, data format and configuration preferences
• Most of the work that needs to be carried is setting up the data interfaces and calibrating the statistical model
• Error messaging standards will need to be identified for each client
Benefits

QUATRA is able to be used by various telematic operators to offer user friendly data quality assessment services with the required standards.

- accurate quality management tool for the assessment of traffic data on freeways and in urban traffic environments
- robust software which is able to import and use different kinds of data streams and formats
- the system core can be modified and equipped with additional tools and statistical modules e.g. to impute faulty data and check other variables such as speed and density according to the clients’ needs
Thank you for your attention
Appendix
Workshops with authorities

- Toll system data could be used for evaluation purposes
- Check and substitution of speed data vital for traffic guidance
- User friendliness of the product very important
- Automated expert system could provide major benefit
- Labelling of erroneous data for statistics very important (also in case of road works, accidents...)
- Online focus very important
- Interface for road works database would be beneficial
<table>
<thead>
<tr>
<th>WP 1</th>
<th>WP 2</th>
<th>WP 3</th>
<th>WP 4</th>
<th>WP 5</th>
<th>WP 6</th>
</tr>
</thead>
<tbody>
<tr>
<td>Project Management</td>
<td>State-of-the-art analysis</td>
<td>Model Development</td>
<td>Software Development</td>
<td>Field trial</td>
<td>Evaluation</td>
</tr>
<tr>
<td>1.1 Kick Off Meeting</td>
<td>2.1 literature review of data quality</td>
<td>3.1 define general indicators and criteria</td>
<td>4.1 define a software system</td>
<td>5.1 first test</td>
<td>6.1 evaluation and business concept</td>
</tr>
<tr>
<td>Steering Committee</td>
<td>for detectors on freeways</td>
<td>and timetables</td>
<td>and service</td>
<td>5.2 adaption and analysis</td>
<td></td>
</tr>
<tr>
<td>1.2 project structure plan</td>
<td>2.2 literature review of data quality</td>
<td>3.2 define indicators specific for freeways</td>
<td>4.2 development online-freeway tool</td>
<td>5.3 second test</td>
<td></td>
</tr>
<tr>
<td>with project budget and timetables</td>
<td>for detectors in urban environment</td>
<td></td>
<td>4.2 development urban - offline tool</td>
<td></td>
<td></td>
</tr>
<tr>
<td>1.3 Meetings with Road Authorities</td>
<td>3.3 define indicators specific for urban networks</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>1.4. coordination meetings</td>
<td>3.4 development freeway model</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>1.5. finalisation</td>
<td>3.5. development urban service model</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

- **finished**
- **ongoing**
Implementation of the Statistical Model
Pattern matching

- Measurements
- Aggregation
- Statistic Model
- Central Data Base
- Corridor Patterns

Corridor definitions, Device filter etc.
Implementation of the Statistical Model
Online calculation of statistical indicators

Measurements → Aggregation

Corridor patterns, Device filter etc. → Statistic Model

Central Data Base → Statistical Indicators
Statistical model – examples of abnormal data conditions
Statistical model – example of linear regression