ELSEVIER

International Journal of Forecasting 13 (1997) 63–72

international journal of forecasting

Short-term forecasting based on a transformation and classification of traffic volume time series

Dieter Wild[1]

Technische Expertensysteme und Robotik, Forschungszentrum Informatik (FZI), Haid-und-Neu-Str. 10–14, D-76131 Karlsruhe, Germany

Abstract

This article describes a short-term forecasting method for traffic volumes at cross-sections. The prediction is based on classified historical patterns. Continuously collected time series are transformed into a representation with objects. These objects are interpreted as polylines and support the qualitative and quantitative interpretation of the data. The historical patterns are classified in great detail, including environmental conditions and the shape of the patterns. An evaluation based on collected data from a field trial in Köln is presented. The results of the new approach are compared with those from two simple time series predictors. The sensitivity analysis shows the differences of using different levels of classification.

Keywords: Short-term forecasting; Classification; Pattern matching; Adaptive systems

1. Introduction

The short-term traffic forecasting method described in this article was developed in the EC DRIVE project KITS (V2039) (Knowledge-Based Intelligent Traffic Control Systems) (Boero et al., 1993, 1994). Using knowledge-based techniques, KITS is designed to complement real-time traffic control systems in assisting the traffic manager at a strategic level. KITS enables the user to define situations and problems which are matched on-line to current as well as to predicted traffic states. These traffic states are specified by a set of traffic parameters, which may be estimated using information from the traffic detectors. Measured traffic volumes from the detectors are used locally to forecast the future volumes. The requirements for such knowledge-based procedures have been as follows: the forecasting horizon should not be fixed but flexible and the prediction method should be open to the user which means one should avoid the use of black-box approaches.

The forecasting method developed in this article is entirely empirical and makes no use of traffic theoretic relationships; it may be seen as a successor to predictors that use the averages of smoothed historical time series directly. Those predictors have shown better results than pure time series models (e.g. Nicholson and Swann,

[1] Present address: PTV Consult GmbH, Stumpfstr. 1, D-76131 Karlsruhe, Germany. Tel.: 49721 9651 177; fax: +49 721 9651 199.

PII S0169-2070(96)00701-7

1974; Hogberg, 1976; Ahmed and Cook, 1979; Ahmed, 1983; Okutani and Stephanedes, 1984; Stephanedes et al., 1990) and those combined with historical average data (e.g. Kreer, 1976; Stephanedes et al., 1981) when the traffic flow is recurrent and where the forecasting horizon is not too short (more than 5 min).

Improvements can be expected if the historical patterns are classified not only on a day-of-the-week basis, but also use both the whole available environmental information and a pattern matching criterion to reach a site-specific adaptive set of typical cases. By transforming the traffic volume time series into a doubly connected list of objects and using a classification based on day groups, environmental conditions and pattern matching, it is now possible to use the old natural idea of forecasting on historical time-series in a flexible and self-adapting way.

The presented technique can also be used to forecast other features with day-specific time series, such as in the traffic domain, the speed and journey, and in the electrical power or information networks, the load.

2. Forecasting method

The prototype method has three main parts (see Fig. 1): pattern transformation, pattern classification and the choice of a suitable comparison pattern including forecasting. Additionally, there are two knowledge bases. The first contains the basic knowledge for pattern classification using weekday groups and additional site-specific information. The second contains acquired and transformed typical patterns. The forecasting method is based on traffic volumes measured at a specific location (the data source) and on classification knowledge.

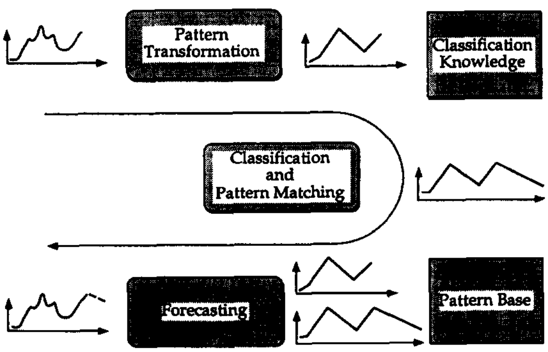


Fig. 1. System modules and knowledge bases.

2.1. Pattern transformation

Data on traffic volumes are received from sensors, usually loop detectors, as a time series, cycle by cycle. In general predictions are not computed every cycle. Therefore, one operation involves the storage of all the information generated by the detectors and another checks for possible missing volumes due to detector malfunction. The original time series are then compressed into series with a longer cycle length to remove all the high frequency variations and to smooth the pattern. However, the resulting cycles should be short enough to keep the possibility of recognizing trend changes. Cycle lengths around 10 min seem to be a good compromise between these opposing goals. For the field trials in KITS, the measured series have been transformed into 9-min time series by adding up 6 values of the 90-s cycles of Köln.

Missing values are interpolated using an average of the adjacent values. However, when there are too many missing values, the compression phase could produce insignificant results, and the forecasting method may not be able to predict, yielding a failure warning.

The resulting 9-min time series is transformed (see Figs. 2 and 3) using a smoothing algorithm to keep only the essential and typical aspects of the data. The critical patterns from the point of view of prediction are trend changes and peak values. Therefore, the algorithm first determines all the extrema of the time series and interprets them as a piece-wise linear trend line (polyline). In subsequent passes unimportant short and steep peaks are cut, but some intermediate points are included to avoid large discrepancies between the time series and the new polyline.

The polyline or pattern is represented as a doubly connected list of objects, as shown in Fig. 3. Each object corresponds to one change point

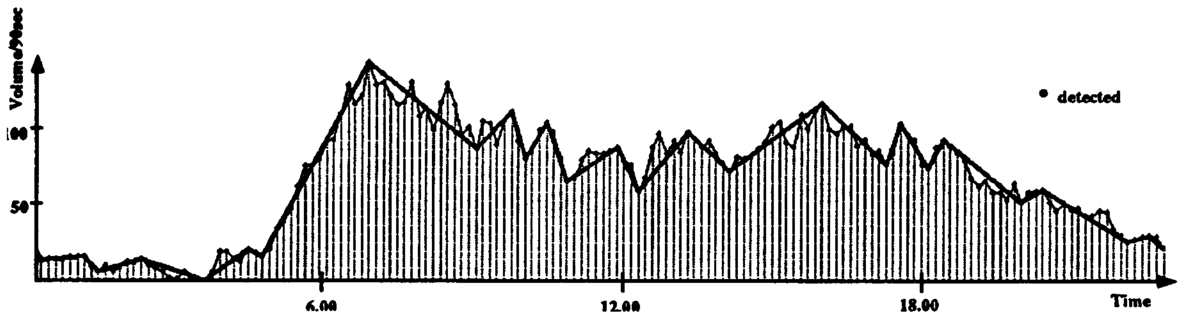


Fig. 2. Transformation of a 9-min time series into a polyline.

of the time series. It is possible to include useful information about the pattern itself into the objects. In contrast to conventional smoothing or approximation methods, like splines or kernel regression, which search an approximation for all the points of a pattern, the algorithm favours the main trend changes and the extrema. The characteristic changes of the traffic flow are then represented as pattern-point objects and, therefore, they are directly accessible. Experience derived from the comparisons (matching) and the forecasts can be integrated directly into these objects.

The algorithm adapts to the general traffic volume of the time series by changing internal thresholds. Otherwise patterns from busy locations would have too many pattern points. With this adaptation the graphical representations of two patterns differing only in the general volume size may be approximated in the same way by visual inspection.

2.2. Classification and pattern matching

Adaptive classification and similarity-based pattern matching (Whiteley and Davis, 1993) are used to obtain a wide set of representative patterns and to restrict the total number of historical patterns. Transformed patterns representing typical situations are held in the pattern knowledge base. In contrast with the ATHENA model (Danech-Pajouh and Aron, 1991) which divides volume profile vectors into four classes independently from the day, there is no initial

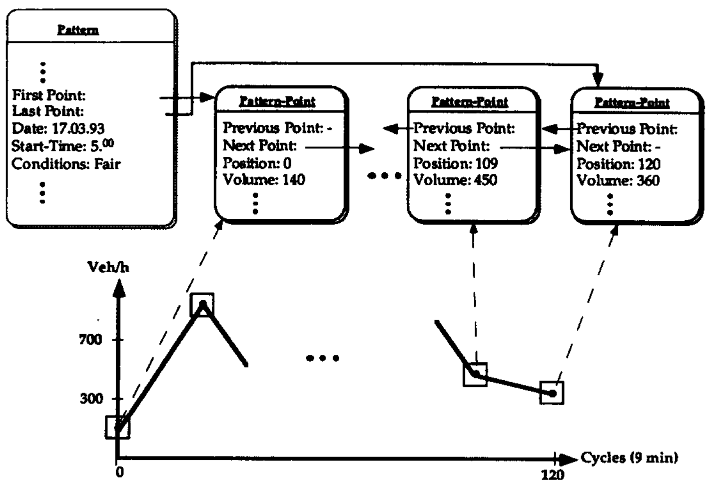


Fig. 3. Pattern representation with objects.

restriction on the number of classes. The starting point for the process is the classification knowledge explicitly stored in the model. This site-specific information indicates which days of the week have similar traffic and which possible conditions affect the traffic in a serious way. These conditions subsume every factor which recurrently influences traffic and which is available as information, such as weather condition, events at local places like fair-grounds, gymnasiums, sports fields, theatre and so on. Fig. 4 shows the day groups and the conditions from the Köln field trial.

If a new pattern is acquired, it will be tested to see whether a group (see Fig. 4) showing the day and the conditions already exists. If such a group does not exist, it will be created and a new cluster will be filled with this archetypical pattern. If a group exists, the new pattern will be compared with all present patterns from the group (each stands for a cluster). This pattern matching is based on the maximal vertical difference between patterns and can be reached in calculating the differences only at the pattern-point objects. If the maximal difference to the best matching pattern is small enough to satisfy the clustering criterion, the patterns belong to the same cluster, some characteristics of the new pattern are included in the objects of the representative pattern for the cluster and the new pattern is discarded. If there are significant differences, a new cluster is created with the new pattern as representative.

The first pattern of a cluster is always the archetypical structure to be used for pattern matching and is never changed.

The pattern knowledge base should be maintained either automatically or by hand. Similar clusters can be merged and old single-element clusters can be discarded in order to prevent fragmentation. All the tasks are usually done at night, when traffic is low and no other computations are performed.

Patterns are bound to their sources. While two patterns from one detector belong to the same cluster, it is possible that the corresponding patterns from another detector have large differences (perhaps being close to an event). Therefore, the groups built are not the same for all detectors.

In the current prototype all detectors use the same day groups and conditions. This could easily be changed to gain flexibility, giving each detector its own knowledge base.

The length of the patterns is not fixed. The general size is a daily pattern, even though it is possible to manage data referring only to parts of a day.

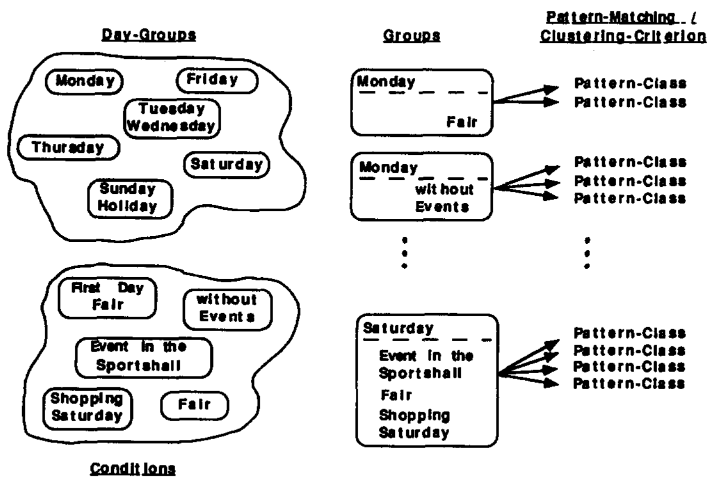


Fig. 4. The pattern classification with day groups and conditions from the Köln field trial.

2.3. Comparison pattern selection and forecasting

In order to predict using a historical pattern, the detected and transformed pattern must be matched with the historical patterns in the data base. Current conditions are used, coming either from the operator or from an event calendar, to restrict the initial search to the corresponding group. Then the same similarity-based pattern matching as used for classification, based on the maximum difference in volume, performs the final choice. If no pattern is found, a warning is given to the operator indicating that the traffic flow is different than usual.

During the acquisition of new detected values, the chosen comparison pattern is continuously checked. If large deviations occur, a new search in the pattern base is started to retrieve a more suitable pattern for the comparison and if no similar one is found, an incident warning (Teale, 1980; Ahmed, 1983) is given to the operator.

The forecasting method provides a quantitative estimate of the expected traffic volume. The forecast horizon is flexible and given as a parameter. The internal results are based on 9-min intervals and must be reduced to the original cycle-length (90 s for Köln and therefore division by 6) before the output. The calculation is shown in Table 1.

Table 1
Equations for the calculation of forecasts with the new approach

New forecasting

$$\text{Pred: } \hat{x}_{t+\text{horizon}} = \frac{X^{\text{historical}}_{t+\text{horizon}} + k(X^{\text{historical}}_{t} - X_t)}{6}$$

with

$k = -0.52\text{H} + 0.842$ if $0 < H \leq 1.6$

$k = 0$ if $H > 1.6$

and

H = forecasting horizon in hours

$X_t = \sum_{k=0}^{5} x_{t-k(90\,\text{s})}$ = measured volume in [vehicles per 9 min]

x_t = measured volume in [vehicles per 90 s]

$X^{\text{historical}}_t$ = volume in [vehicles per 9 min] out of the historical pattern

The forecasting procedure determines the actual (t) and the forecasted (t + horizon) location in the comparison pattern and multiplies the actual shift in traffic volume with the factor k to get an adaptation to recent changes. The value of k depends on the forecasting horizon, see Table 1.

The predicted values maintain the same trend as the ones used for the comparison (see Fig. 5). Having several possible comparison patterns, it is possible to predict using each of them or to follow a best-fit strategy. For the evaluation in Köln only the best-fit strategy was used.

If the actual pattern is different from all the patterns in the pattern base, the algorithm currently uses a moving average method to forecast.

The quality of the prediction depends on the completeness of the pattern base, on the clustering criterion and on the recurrence of the traffic flow. A steady decrease in the average prediction error is, therefore, expected after the initialization of the method, as more patterns are included in the data base.

3. Evaluation

The pattern-based forecasting method was tested on the field trial of KITS in Köln and the main results are presented below. The evaluation was done off-line in the laboratory with recorded data, using for most tests an isolated version of the forecasting software in order to reach the necessary performance for the many predictions.

3.1. Description of the data

The test site has nine locations, some of them detecting on several lanes. A detection cycle is 90 s and recorded time-series have two different lengths: 18 h from 5 a.m. to 11 p.m. and 24 h starting at midnight. Patterns from 42 days have been used to fill the pattern base: 28 days in 1993 and 14 in the year 1994. The day groups and conditions from the classification knowledge-

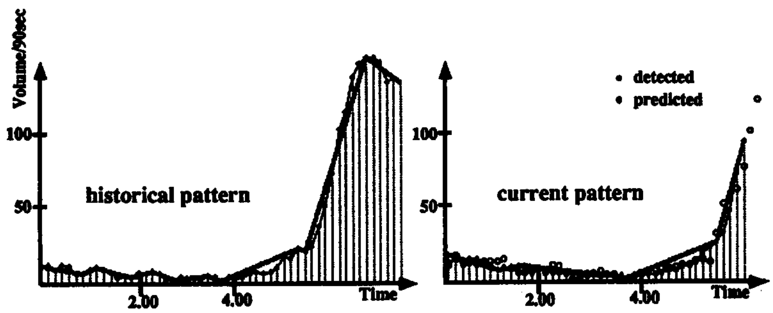


Fig. 5. Forecasting results compared to the chosen historical pattern.

base are represented in Fig. 4. Events are known for all the days considered.

The following selections have been used in the evaluation: 10 days in March 1994 (inclusive of Saturdays and Sundays), six detectors (M8, M9, M10, M16, M27 and M28) and three forecasting horizons (270 s, 900 s or 15 min and 3600 s or 1 h).

3.2. Quality mesurements

Quality measurements have been carried out for the new forecasting approach (Pred) described above and for two simple prediction methods (see Table 2), namely trivial forecasting (Triv) and moving average (MVA). The applied error measures are shown in Table 3.

The results of the quality measurements are summarized in Table 4. Each value represents the average of the measurements with the six detectors during the 10 days, which are about 35.000 single forecasts.

Table 2
Equations for two simple prediction methods

Trivial forecasting	Moving average forecasting
Triv: $\hat{x}_{t+\text{horizon}} = x_t$	MVA: $\hat{x}_{t+\text{horizon}} = \dfrac{\sum_{k=0}^{2} x_{t-k(90\,\text{s})}}{3}$
with	
x_t = measured volume in [vehicles per 90 s]	

Table 3
Error measures

Mean absolute error	Root mean squared error	Mean relative error	Root mean squared error proportional
$\text{MAE} = \dfrac{\sum_{i=1}^{n} \lvert x_i - \bar{x}_i \rvert}{n}$	$\text{RMSE} = \sqrt{\dfrac{\sum_{i=1}^{n} (x_i - \bar{x}_i)^2}{n}}$	$\text{Rel} = \dfrac{\sum_{i=1}^{n} \dfrac{\lvert x_i - \bar{x}_i \rvert}{x_i}}{n} \times 100$	$\text{RMSEP} = \dfrac{\sqrt{n \sum_{i=1}^{n} (\bar{x}_i - x_i)^2}}{\sum_{i=1}^{n} x_i}$

Table 4
Mean errors for three forecasting methods and three horizons

	Pred	MVA	Triv
MAE			
270-s horizon	1.56	1.47	1.38
900-s horizon	2.05	2.07	2.04
3600-s horizon	2.66	3.73	3.69
RMSE			
270-s horizon	2.07	1.92	1.84
900-s horizon	2.67	2.74	2.72
3600-s horizon	3.49	5.11	5.07
Rel			
270-s horizon	14.60	13.44	12.76
900-s horizon	18.70	17.98	17.73
3600-s horizon	23.63	27.23	26.94
RMSEP			
270-s horizon	0.15	0.14	0.13
900-s horizon	0.19	0.18	0.18
3600-s horizon	0.24	0.33	0.33

The new forecasting procedure produces better results than the trivial and MVA methods as the forecasting horizon increases. They are about the same for a horizon of 900 s. The results are as expected, since previous predictors using historical averages have yielded similar conclusions. It is difficult to beat the simple time-series predictors for horizons shorter than about 10–15 min (predicting short cycles of about 60 or 90 s).

In the Appendix there are three graphical representations of the forecasting results, one for each forecasting horizon. In particular, the last one (Fig. A3), with a horizon of 1 h, shows the huge influence of the chosen historical comparison patterns. For example, between 12:00 and 14:00 there is a significant difference between the actual and the historical values. But also in Fig. A2 with a horizon of 15 min there is a difference at 12:30. Each differing peak leads to large errors and without the comparison pattern(s), it is possible to guess the course of the graph. Remember that the comparison pattern changes if another one fits better during the day; only if the actual and historical patterns are similar, e.g. in Fig. A3 between 15:00 and 17:00, are the results very good compared with the long horizon.

The quality of the forecasting results depends on the pattern base. Better quality is achieved when the base includes a historical pattern very similar to the new one. Looking at the classification and the pattern base (see Table 5, Case A) it can be seen that there are some groups with only one pattern. This is an indication that the number of acquired patterns should be increased in order to improve the results.

Table 5
Sensitivity of the classification

	A	B	C	D
Detectors	9	9	9	9
Groups	197	197	9	9
Patterns	324	272	188	59
Max. matched patterns	4	4	15	24
Pattern points	7647	6376	4534	1442
MAE				
270-s horizon	1.55	1.58	1.62	1.63
900-s horizon	2.05	2.13	2.13	2.27
3600-s horizon	2.66	2.78	3.10	3.73
RMSE				
270-s horizon	2.07	2.10	2.17	2.20
900-s horizon	2.67	2.78	2.87	3.05
3600-s horizon	3.49	3.76	4.55	5.42

3.3. Sensitivity measures

Tests have been performed to analyze the sensitivity of the classification and the clustering criterion. The same tests as above have been performed using a different classification each time and the results are presented in Table 5. For a complete view, the classification of the basic Case A is included.

Case A: This is a case as described above in Section 2 (Fig. 4). Classification is based on day groups, on conditions and on a clustering criterion which has the same size as the tolerance for the pattern transformation. Having 366 patterns for nine detectors, 197 groups are built with 324 patterns (clusters) in total. The remaining 42 patterns were then incorporated into the existing ones. There is at least one pattern including three others, yielding a maximum of four matched patterns (max. groupsize = 4). This case has the largest number of pattern points and shows the best forecasting results.

Case B: This case differs from Case A in terms of the clustering criterion, which is set to allow more clustering. The total number of patterns is reduced to 272, with the consequence that the forecasting errors for middle and long horizons have slightly increased.

Case C: What happens if the clustering is used without day groups and conditions? Answers are given in C and D. In C a narrow clustering criterion as in Case A has been used. Only one group is created for each detector and in total 188 different clusters have been constructed. There is at least one cluster with 15 similar patterns. The pattern base shrinks to less than two thirds of that of base Case A.

Case D: Widening the clustering criterion (as in Case B), only about one sixth of the acquired patterns are left. The results for all horizons are the worst of the four cases.

The evaluation plan for Köln also involved the comparison of the predicted traffic states with those derived from measured values, and the comparison of situations and problems based on predicted and measured volumes. The results have been mostly satisfactory, but this evaluation included several other modules from KITS. Direct assessment of the forecasting procedure is, therefore, difficult. In order to get feedback from the traffic experts, they completed a questionnaire and some of the replies are mentioned in the conclusion.

4. Conclusion

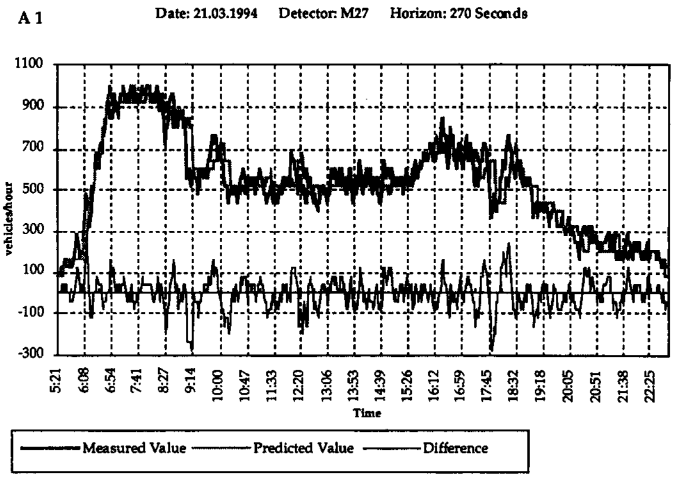
The developed pattern-based forecasting procedure has several advantages compared with conventional approaches. Perhaps the most important, not only for knowledge-based environments, are the openness and the explanatory capability. Using on-screen diagrams the traffic operators always have the option to compare the current situation with the best matching historical one, without having to search the data base. They have also the option to inspect the pattern base in order to analyse the traffic. The realized representation indicates directly the actual trend and the forthcoming trend changes in the volumes. Furthermore, pattern acquisition is automized and adaptive, including patterns with new forms and discarding those whose forms already exist. As a consequence, the pattern base adapts in a short time to new situations like medium-term roadworks or changes in the infrastructure. The patterns recorded generate new clusters, which can be used for prediction for subsequent days. There is not much advance effort necessary for model building. Only conditions have to be acquired and their accessibility must be established.

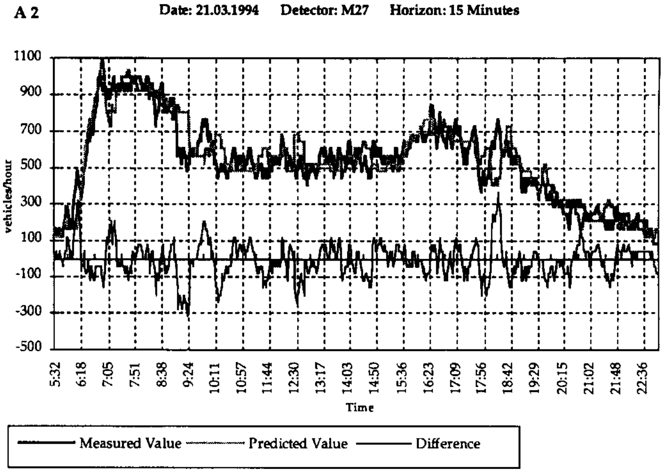
Besides the flexible forecasting horizon there is also the possibility of using the pattern base for simulation. Instead of reading values from the detectors, values can be derived from patterns and the user can preview outcomes; for example the expected traffic situations and the problems arising from a Sunday fair, simply by selecting the corresponding group. This feature is available in KITS.

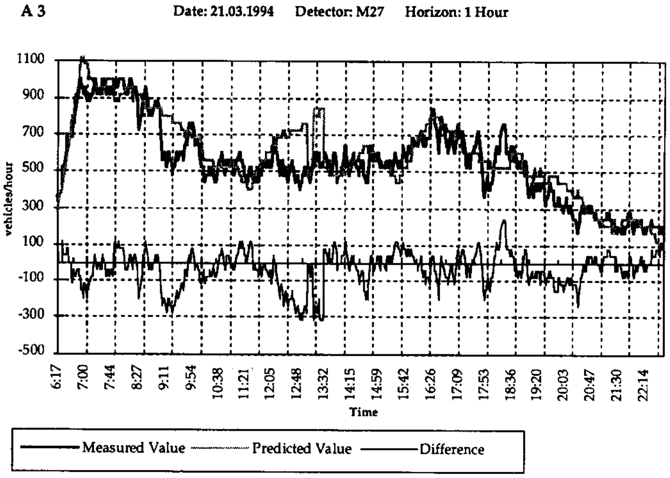
In the future more emphasis could be given to the pattern-point objects. Pattern selection and forecasting could be improved by including matching as well as forecasting statistics into the local objects.

Acknowledgements

This work was carried out at the Forschungszentrum Informatik (FZI), Karlsruhe, in the research group Technische Expertensysteme und Robotik led by Prof. Dr.-Ing. U. Rembold. It was part of the KITS (Knowledge-Based Intelligent Traffic Control Systems) project, partly funded by the EEC under DRIVE contract V2039.







Appendix A

Figs. A1–A3 in this appendix show the measured and predicted volumes together with their difference for three forecasting horizons: 270 s, 900 s and 3600 s.

References

Ahmed, S.A., 1983, Stochastic processes in freeway traffic, *Transportation Engineering and Control*, 24, 306–310.

Ahmed, M.S. and A.R. Cook, 1979, Analysis of freeway traffic time-series data by using Box–Jenkins Techniques, *Transportation Research Record*, 722, 1–9.

Boero, M., J. Cuena, H. Kirschfink and C. Krogh, 1993, KITS: A general approach for knowledge-based traffic control models, in advanced transport telematics, in: *Proc. of the Technical Days*, Bruxelles, 8–10 March, Vol. II.

Boero, M., J. Cuena, H. Kirschfink, H. Traetteberg and D. Wild, 1994, The role of knowledge-based models in traffic management and their design, in: *Proc. of the First World Congress on Applications of Transport Telematics and Intelligent Highway Systems*, Paris (Artech House, London) 572–579.

Danech-Pajouh, M. and M. Aron, 1991, ATHENA: A method for short-term inter-urban motorway traffic forecasting, *Recherche Transports Sécurité*, English issue, (6).

Hogberg, P., 1976, Estimation of parameters in models for traffic prediction: A non-linear regression approach, *Transportation Research*, 10, 263–265.

Kreer, J.B., 1976, Factors affecting the relative performance of traffic responsive and time-of-day traffic signal control, *Transportation Research*, 10, 75–81.

Nicholson, H. and C.D. Swann, 1974, The prediction of traffic flow volumes based on spectral analysis, *Transportation Research*, 8, 533–538.

Okutani, I. and Y.J. Stephanedes, 1984, Dynamic prediction of traffic volume through Kalman filtering theory, *Transportation Research*, 18b, 1–11.

Stephanedes, Y.J., P.G. Michalopoulos and R.A. Plum, 1981, Improved estimation of traffic flow for real-time control, *Transportation Research Record*, 795, 28–39.

Stephanedes, Y.J., E. Kwon and P. Michalopoulos, 1990, On-line diversion prediction for dynamic control and vehicle guidance in freeway corridors, *Transportation Research Record*, 1287.

Teale, G., 1980, Procedures for analysis of traffic flow patterns using detector loops, in: *Proc. 10th ARRB Conf.*, Part 4.

Whiteley, J.R. and J.F. Davis, 1993, Qualitative interpretation of sensor patterns, *IEEE Expert*, 8(2).

Biography: Dieter WILD is a Ph.D. candidate at the Computer Science faculty of the Technical University of Karlsruhe. He received his Diploma in Computer Science at the mentioned faculty in 1988. He worked as a scientist in the research group Technische Expertensysteme und Robotik at the Forschungszentrum Informatik (FZI) in Karlsruhe and is presently employed as a systems engineer in the research and consultancy department of PTV Planungsbüro Transport und Verkehr GmbH, a leading German group for consultancy and software in traffic and transport. His research interests include traffic management, traffic forecasting, knowledge-based systems and distributed problem-solving.