Traffic Estimation and Prediction for Urban Road Networks, Application to GrandLyon

Laura Wynter
IBM Research Singapore Collaboratory, Singapore
lwynter@sg.ibm.com

Barry M Trager, Yichong Yu
IBM Research, Yorktown Heights, NY USA

Yiannis Kararianakis
Arizona State University, Tempe, Arizona, USA

Saif Jabari
New York University-Abu Dhabi, Abu Dhabi, UAE

Jean Coldefy, Dimitri Marquois
ITS and Mobility services unit, GrandLyon, Lyon, France

Thomas Baudel
IBM France Lab & Efficacity Institutem France

ABSTRACT

This paper describes joint work done by IBM Research (development of the solution) and GrandLyon (assessment of the solution) for real-time traffic estimation and prediction on the urban road network of Lyon Metropolis, France. The methods are described and a new feature called congestion thresholds is presented which enhances prediction accuracy on the highly-variable conditions on an urban road network such as the one in Lyon, France. The methods were deployed in a pilot study on live real-time traffic data from Lyon. Numerical results and analysis are provided. The novelty of the work lies in three main features: the application of real-time traffic prediction to an urban road network; the definition and use of congestion thresholds as part of the prediction scheme; and the live piloting of the prediction technology with an urban traffic management center.

1. INTRODUCTION

Accurate short-term forecasting of traffic is essential for intelligent transportation systems applications, such as real-time route guidance, adaptive traffic control and advanced traveler information systems. Traffic forecasting models are usually designed for and evaluated on data from expressways, which are admittedly less variable than data from urban networks and not subject to the effects of traffic lights. In urban networks, neighborhood relationships and the definitions of spatial weight matrices for space-time parametric frameworks are not straightforward; some locations may not be clearly upstream or downstream a given location. Furthermore, detectors can be dense in some parts of an urban network but not others, so that locations with useful predictive information may be hard to identify; this again affects the construction of spatial weight matrices used in space-time modeling schemes such as those proposed in Min and Wynter (2011) and Kamarianakis, Shen and

Wynter (2012). Erroneous and missing data are expected to be more frequent in urban networks, which makes essential the implementation of robust estimation procedures.

The methods presented were developed by IBM Research and calibrated and tested during the course of the Optimod'Lyon project (http://www.optimodlyon.com/en/). The Optimod'Lyon project consisted of a consortium that tested, among other tasks, multiple solutions from several companies for various forms of traffic prediction. This paper presents the assessment by GrandLyon of the methodology developed by IBM Research on the Lyon Metropolis road network. The novelty of the work lies in three main features: the application of real-time traffic prediction to an urban road network; the definition and use of congestion thresholds as part of the prediction scheme; and the live piloting of the prediction technology with an urban traffic management center.

In the next section we present key features of the prediction method. In section 3 we describe the estimation procedure for locations with no detectors. In section 4 we present results from the estimation and prediction on the Lyon road network. Section 5 concludes with future work.

2. URBAN TRAFFIC PREDICTION AND CONGESTION THRESHOLDS

For short-term vehicular traffic forecasting, the proposed method is based on piecewise-linear, parametric spatial time-series models. The adopted modeling framework avoids issues which can be challenging in real-world traffic forecasting applications that involve numerous measurement locations. First, it does not require a model building procedure that uses computationally expensive location-specific sequential statistical tests to identify the number of traffic regimes or the statistically significant predictors within each regime. Second, the modeling framework does not demand the a-priori specification of spatial weight matrices for the representation of neighboring associations among measurement locations; construction of such matrices and statistical evaluation of their alternative forms can be a tedious step in the model-building procedure. Details of the prediction method can be found in Kamarianakis, Shen and Wynter (2012).

The principal traffic data on the Lyon, France road network comes in the form of occupancies ("TO"), which have been found in practice to be of greatest use for intelligent transport applications such as real-time route guidance, adaptive traffic control and advanced traveler information systems. However, on urban road networks, under very congested situations, the occupancy rate values are too volatile to be directly exploitable. One of the key contributions of the proposed method is to automatically compute a congestion threshold for each detector, based on the reconstruction of the fundamental diagram (plot of occupancy vs. flow). Using a polynomial regression on a sufficient history of the detector, a maximum flow is determined for each link. The occupancy rate for this maximum flow is used to determine a congestion threshold specific to that link. Occupancy reports beyond that value are considered as indicating that the link is in a "congested state", and that the actual value is of little significance. For instance, if a link has a congestion threshold of 35, occupancies reported of 80% or 60% are to be considered as similar: the link is highly congested, and a large gap between the forecast value and the actual observed value is actually insignificant.

Determination of the per-detector congestion threshold is done by a non-linear regression on the fundamental diagram flow vs. occupancy rate:

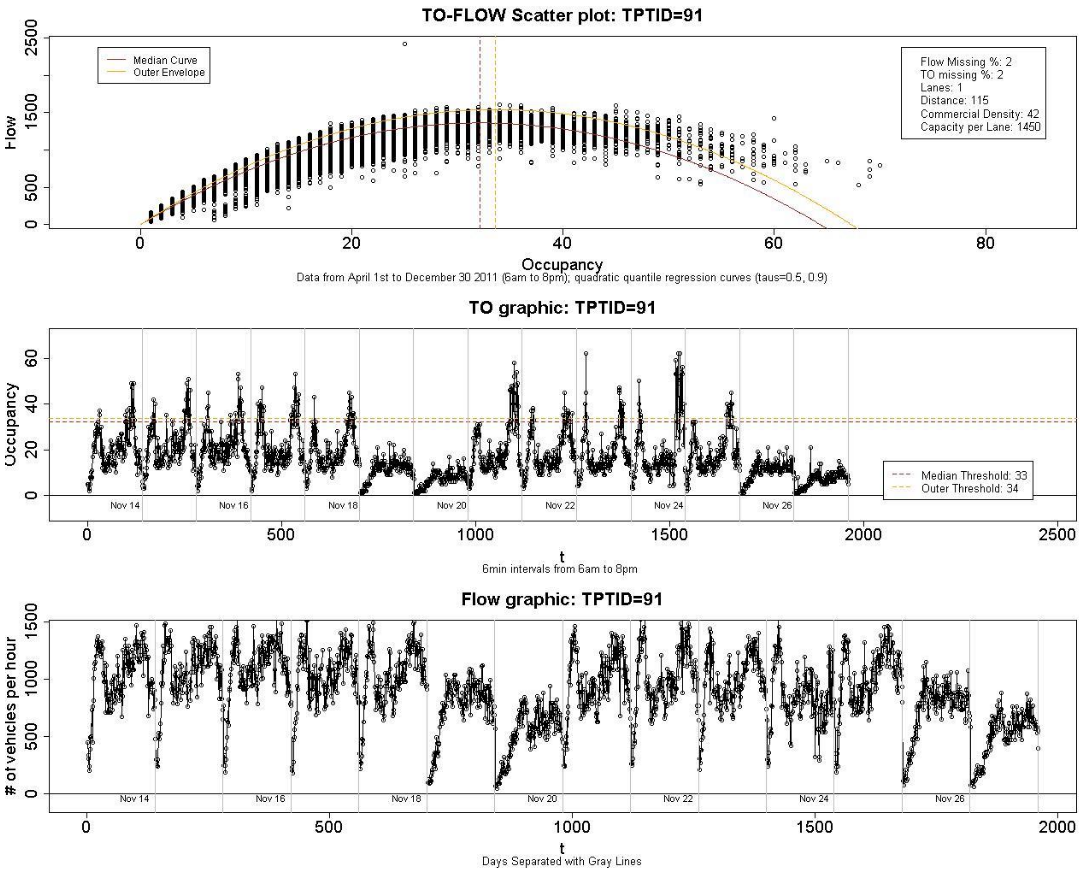

Figure 1. Methodology for determination of detector-specific congestion thresholds.

The congestion threshold determination and occupancy rate corrections are performed using the following method:

- Determine detector-specific congestion thresholds, w_i, for detector, *i*.
- Determine optimal quadratic function fit *F* to the TO vs Volume curve for detector *i*
- Compute the maximum of F. Assign this value as the "congestion" or "red" threshold
- Project all congested TO as follows: for each detector, *i*, for $TO_i > w_i$, where w_i is the congestion threshold for detector *i*, $TO_i^{'} = \text{Min}\{TO_i , w_i\}$.

Hence, congested traffic predictions are up to the level of the congestion threshold; beyond that the detector is "congested" rather than at a specific occupancy level of TO.

3. DATA EXPANSION ALGORITHM FOR TRAFFIC ESATIMATION

This section describes a nearest neighbor approach called Data Expansion Algorithm (DEA) to expand real-time data which is available at detectors to nearby un-instrumented road segments that possess similar characteristics (e.g., number of lanes, speed category, and function). The underlying hypothesis is that under general traffic conditions, traffic levels are similar for urban road segments with similar characteristics in appropriate geographic areas. Such a conjecture is reasonable for regular (incident-free) traffic conditions, but cannot be applied to estimate traffic conditions along roads impacted by incidents. For this reason, a different, companion, methodology was developed to handle incident conditions through a model of the dynamics of congestion build-up and dissipation.

DEA computes an estimate of the occupancy along road segment j for time interval t as the average occupancy over instrumented road segments in the neighborhood of j over the time intervals t and t-1. The key task is the definition of appropriate neighborhoods for instrumented and un-instrumented links. In order to determine neighborhoods of instrumented links for all sensors in the network, a correlelogram is generated and the 10 most highly correlated sensors for each sensor are selected based on the correlation coefficients of their occupancy measurements. In order to determine the neighborhoods of the un-instrumented links, an m-step breadth first search in both the upstream and downstream directions for each aggregated link in the network was first performed (m was chosen to be 5). This results in a connected sub-graph for each link in the network. That is, the neighborhood of each un-instrumented link j is determined by finding similar links which are instrumented in the sub-graph of j and then including all instrumented links which are in the k-nearest neighborhood of j.

4. RESULTS OF ESTIMATIONAND PREDICTION ON THE LYON ROAD NETWORK

The prediction methodology including the use of congestion thresholds was tested on the Lyon road network and results are presented below. The figure below provides average absolute error, measured in the absolute value of the occupancy error, ABS(true occupancy value – predicted occupancy value) over all detectors and all time periods on all days during one week and broken down by forecast horizon from 6 minutes to 1 hour into the future. Note that average errors are from 1.5 to just over 2.5 in occupancy, which is on a scale of 1-100.

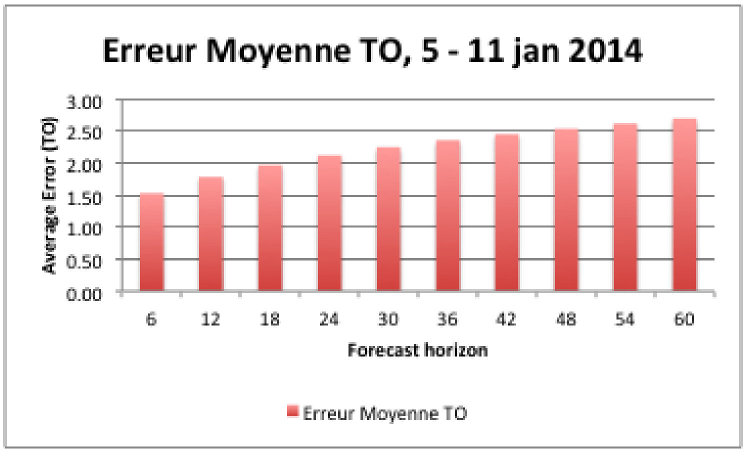

Figure 2. Average error of the traffic prediction method per forecast horizon over one week.

The next figure examines 30-minute-ahead predictions by day over a 10-day period. This period was examined to assess the prediction accuracy on the holiday, 1 January. While higher than on other non-holiday days, the average error remains very low at just over 3.

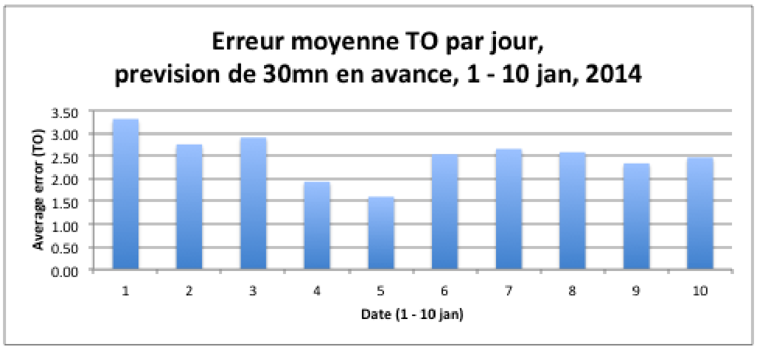

Figure 3. Average error of 30-minute-ahead traffic prediction over a 10-day period including a holiday.

The following figure examines 6-minutes-ahead forecasts again over the period 1 to 10 January, by time-of-day: morning peak, afternoon peak, mid-day periods, and night and early morning. As expected, the afternoon peak is the most volatile and hence presents the largest forecasting error, but the error overall remains very low.

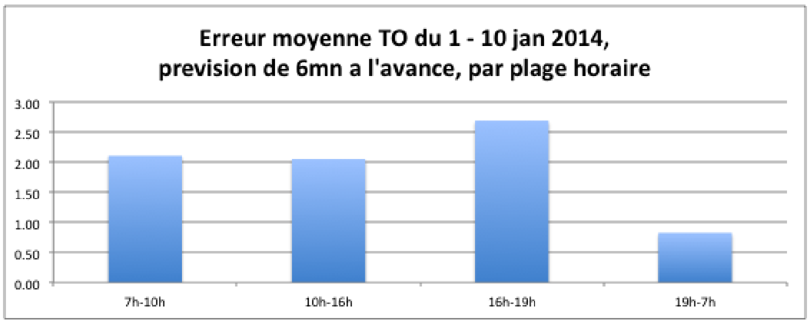

Figure 4. Average error of 6-minute-ahead traffic prediction over a 10-day period including a holiday, broken down by time-of-day.

The next figure illustrates the accuracy of the online forecasts in the live prediction system on the Lyon road network. Three curves are presented once the user selects either occupancy or flow. The curves correspond to the predicted quantity over the pre-defined time and date range for the pre-selected detector, as well as the actual quantity, and the gap, which is the (unsigned) difference between the two. Since the gap is unsigned, it is always zero or positive.

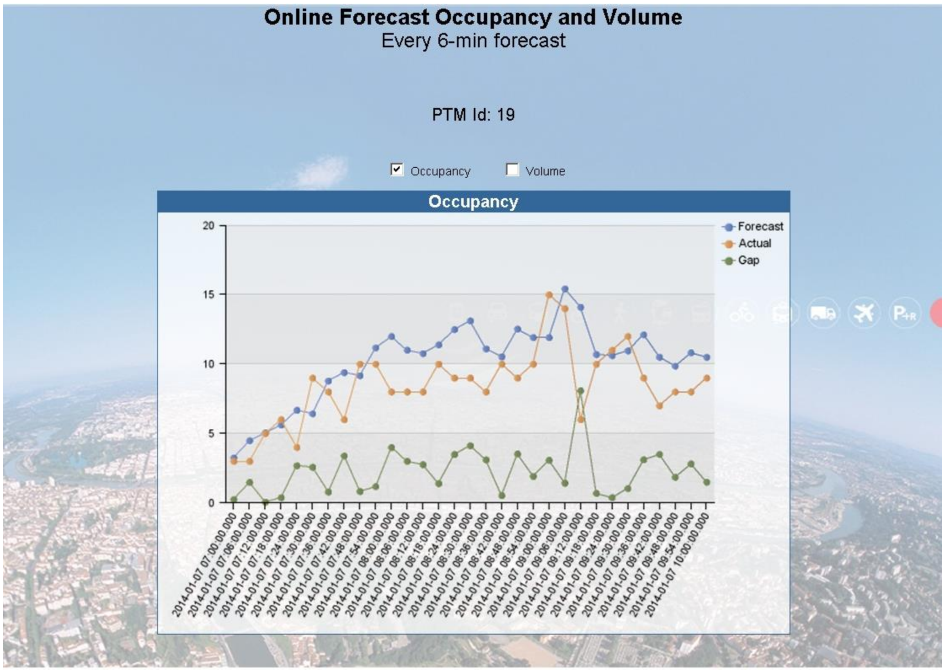

Figure 5. Accuracy analysis in the live real-time system for a particular link occupancy.

The following figure shows the online accuracy of volume predictions. In this case, the gap is not illustrated, because for volume a different error metric is typically used, which is the percent accuracy (i.e., the gap divided by the actual value). This quantity cannot be illustrated on the same graph as the actual and predicted values as the scale is different.

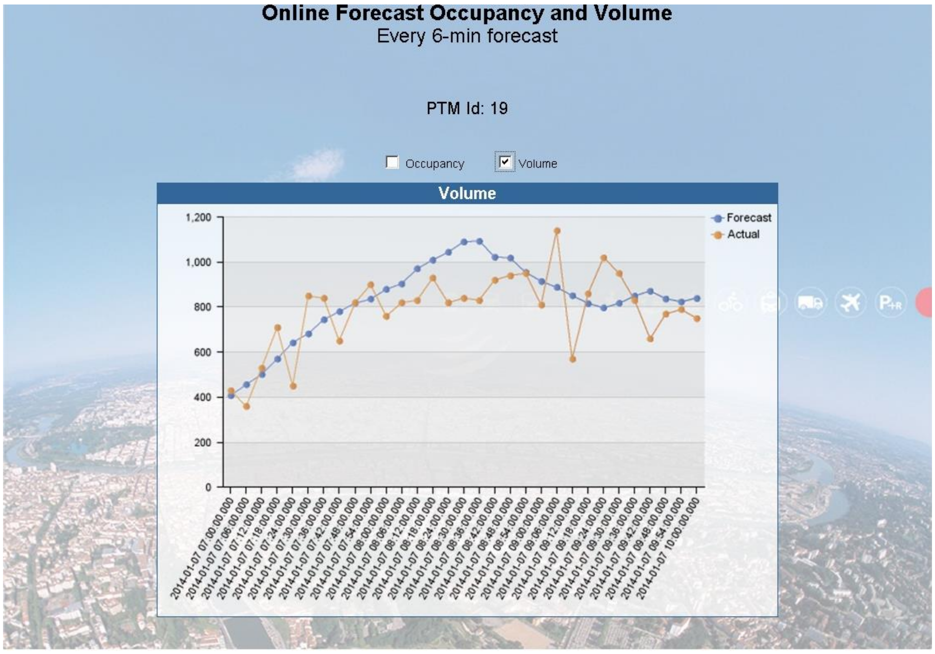

Figure 6. Accuracy analysis in the live real-time system for a particular link volume.

The following figure shows the online forecast reliability in pie chart format. For a given time and date range and a given forecast horizon, here set to 6-minutes-ahead, the reliability is defined as the percentage of detectors whose predictions in this range and for the given forecast horizon are within the target level of an error of no more than 5. Hence the majority (86%) of the detectors fall into this category. The remaining 14% of the detectors, i.e. 50 detectors, fall into the range of predictions with a gap of greater than 5. The cutoff level of 5 can be adjusted if desired.

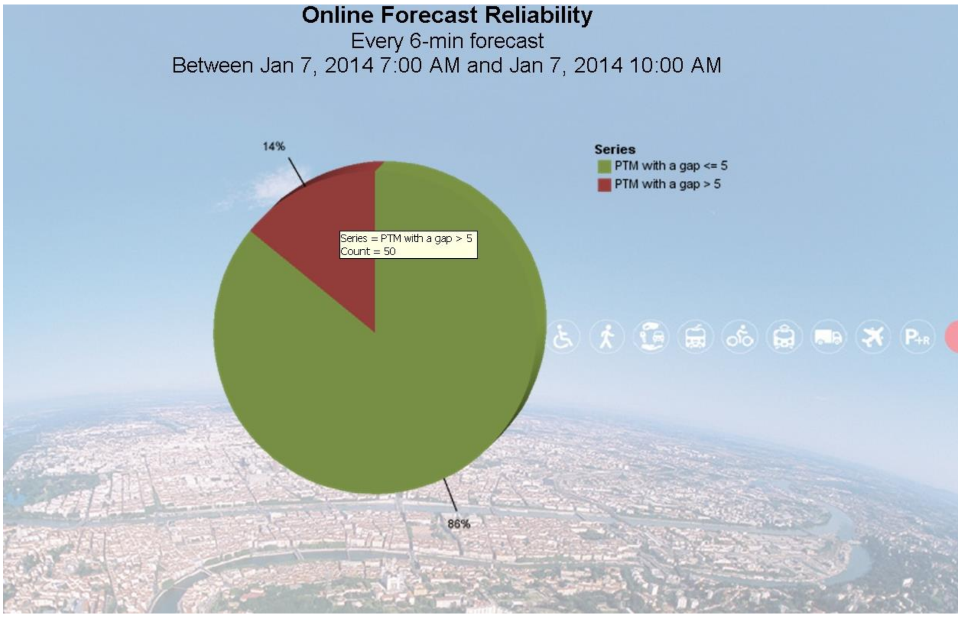

Figure 7. Accuracy analysis in the live real-time system of network-wide occupancy predictions.

The next figure provides a drill-down into the results presented above in pie chart form. It can be seen that the input quality is across the board green, therefore good data availability, that is at least 70% available. However, there are 5 detectors with orange-level forecast reliability (i.e. between 5 and 10) and one with a gap of greater than 10. The detector IDs are provided to facilitate further investigation.

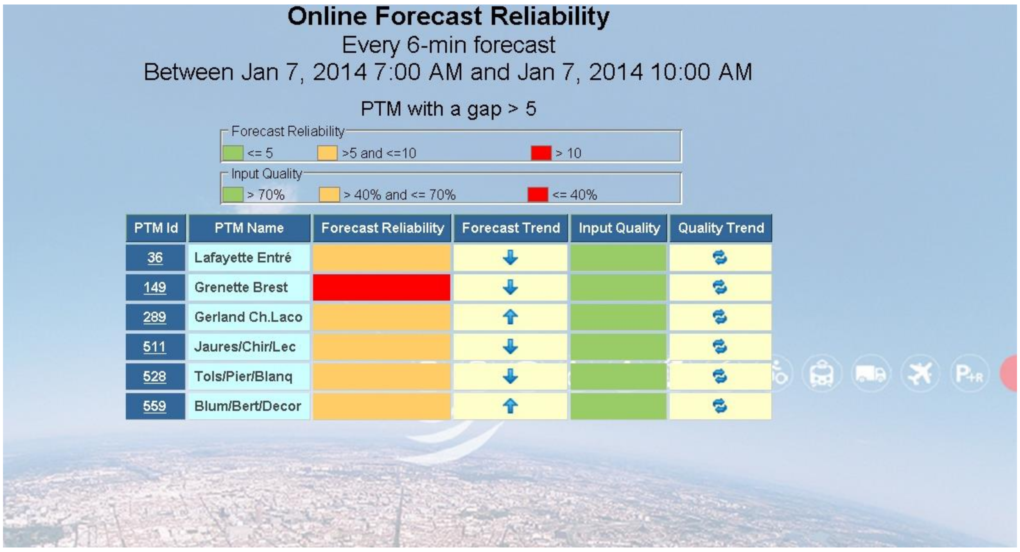

Figure 8. Drill-down accuracy analysis in the live real-time system of occupancy predictions.

Similarly, the following figure shows the same wherein the detectors in the table all have a forecast reliability of a gap of no more than 5.

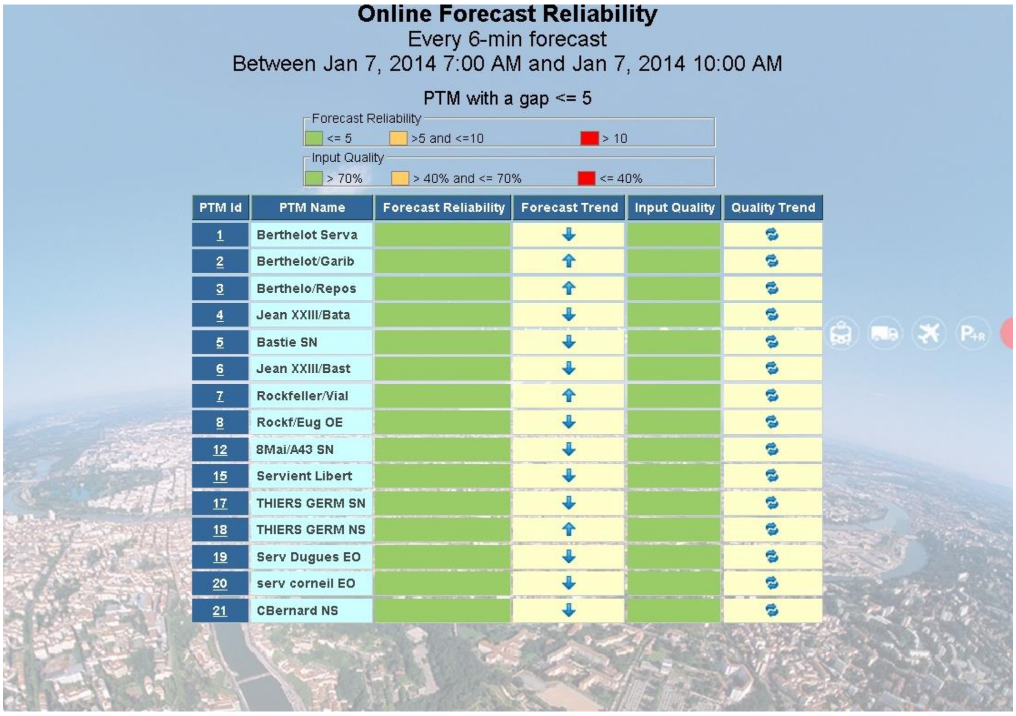

Figure 9. Drill-down accuracy analysis in the live real-time system of occupancy predictions.

The testing network for DEA on the Lyon road network consists of major road segments within the ring road as shown in Figure 4. Minor streets/side streets in the figure are depicted green and all other road segments are depicted in blue. The blue segments constitute the testing network. To reduce the computational burden, small links were merged to form aggregated road segments and the bi-directional aggregated segments were divided into two one-directional segments.

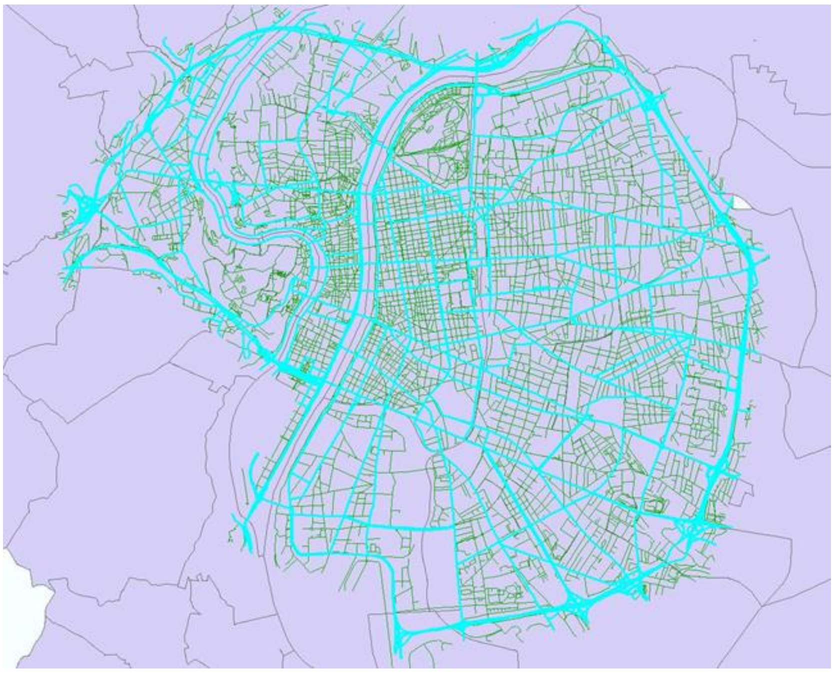

Figure 10. Lyon metropolis road network

The number of road segments (both the blue and the green segments in Fig. 1) is 14,346 bi-directional segments (roughly 4,864 blue and 9,482 green). After aggregating and dividing the bi-directional links into one-directional links, the number of road segments in the aggregated network is 2,278. The road segments were aggregated in such a way that each link in the road network has uniform characteristics (numbers of lanes, speed limit, and function class). It was ensured that each aggregated link maps to one sensor at most. Aggregated links were then mapped to CRITER segments. Approximately 45% of the mapped CRITER segments are instrumented, while 55% are not. The green segments are those which were used for purposes of DEA testing.

Figure 11. Lyon CRITER segments; green segments are those for which sensor data is available (instrumented), red segments are un-instrumented, grey segments to minor un-mapped streets.

Accuracy analysis was performed by running DEA on each CRITER segment with a detector but suppressing the occupancy for that detector only, and then comparing the DEA accuracy with the suppressed occupancy value. The procedure is repeated for all time steps and for all detectors. (Only one detector is suppressed each iteration). The error metric used to assess the quality of the algorithm is absolute error in occupancy for segment *j* over 6-minute time period *t*.

Absolute errors less than 10% are considered to be good. Table 1 summarizes the frequencies of errors within certain thresholds over the testing period.

Percentage of points (time x detector) with error < 10% (good)	89%
Percentage of points (time x detector) with error < 7.5% (very good)	83%
Percentage of points (time x detector) with error < 5% (excellent)	73%

Table 1. Summary of DEA accuracy

The following figures show percentages of CRITER segments with absolute errors less than the three thresholds shown in Table 1, over a month for weekday morning peak period (8:00 – 10:00am), weekday afternoon peak (5:00-7:00pm), for Saturdays, and Sundays, respectively.

Figure 12. DEA error analysis for weekday morning peak, by category of error

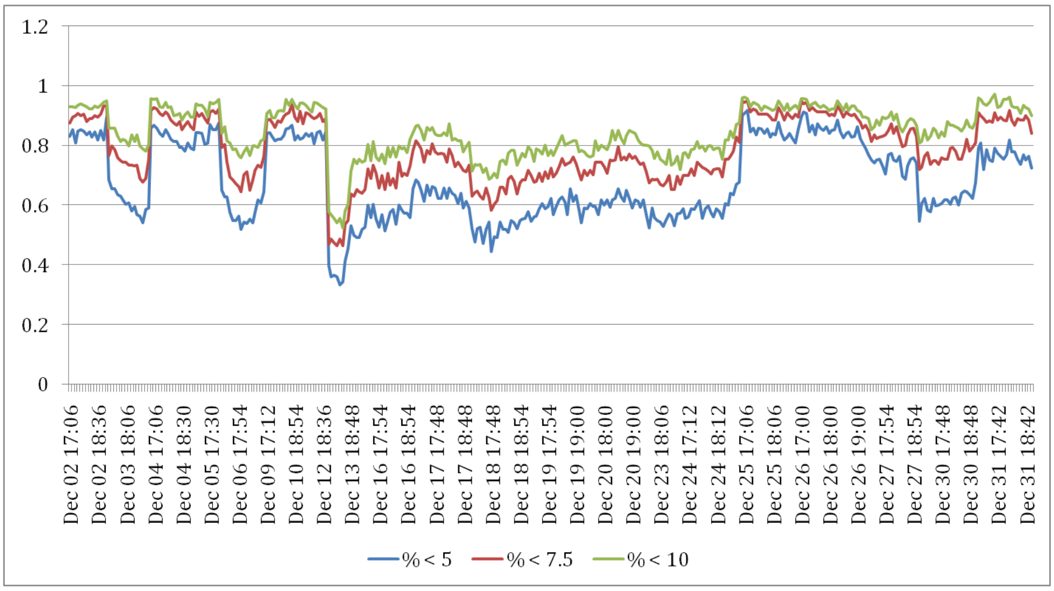

Figure 13. DEA error analysis for weekday evening peak, by category of error

Figure 14. DEA error analysis for Saturdays, by category of error

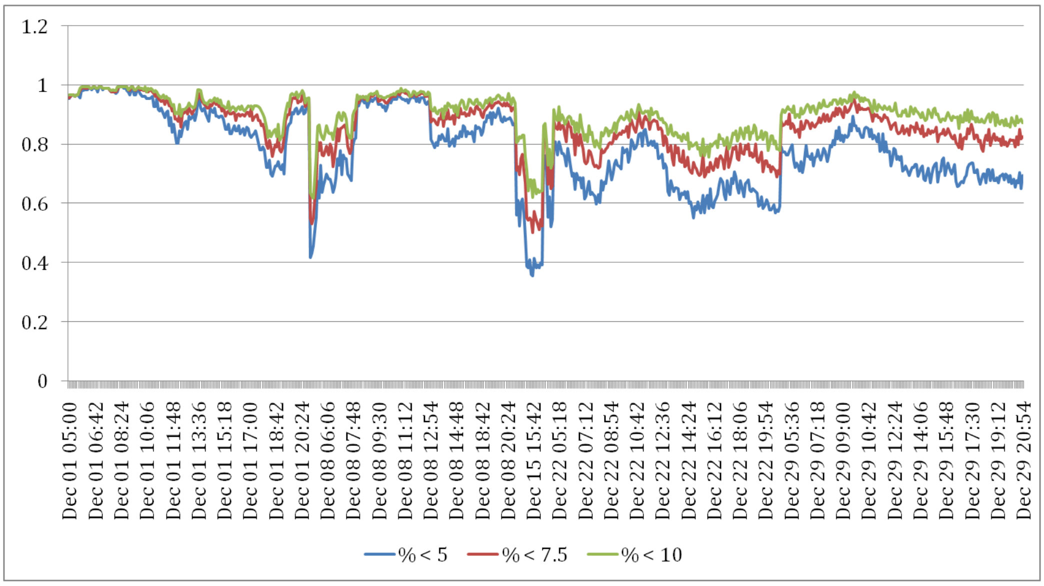

Figure 15. DEA error analysis for Sundays, by category of error

An interesting analysis is provided in figure 15 below which illustrates a comparison between traffic estimation/prediction on un-instrumented links versus prediction on instrumented links using the prediction method presented here. We observe that the accuracy loss obtained by using DEA in place of TPT for predictions is negligible. It must be noted that DEA makes use of the TPT as input, and performs an intelligent aggregation of the most relevant TPT predictions to fill-in the gaps on the non-equipped links.

The fact that the accuracy loss is on the order of less than 0.5 (in units of occupancy) implies that to some extent, capital expenditures on physical sensors can be offset or delayed through the use of intelligent analytics. This is the case once a critical mass of detectors is already present, as is the case in numerous areas of the Lyon network.

It should however be noted that this analysis was limited to the existing CRITER detectors which are in a relatively dense configuration. Other parts of the greater Lyon network with little to no sensors would likely not fare as well. Nonetheless the analysis clearly points to the benefit of augmenting and in some cases deferring capital investment through analytics.

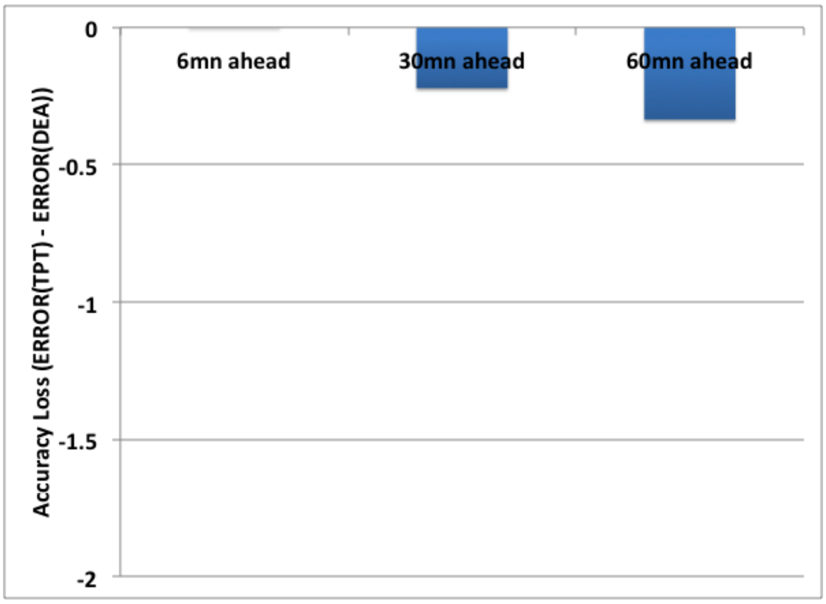

Figure 16. Comparison between traffic estimation/prediction on un-instrumented links versus prediction on instrumented links.

5. CONCLUSIONS

In this paper we present an overview of the IBM Research methods for traffic estimation and prediction calibrated and tested for use as part of the Optimod'Lyon project (http://www.optimodlyon.com/en/). Future work involves the additional calibration to take into account holidays and planned events as well as integration with prediction components developed for handling incident conditions.

REFERENCES

W. Min and L. Wynter, 2011. Real-time road traffic prediction with spatio-temporal correlations, Transportation Research Part C: Emerging Technologies, 19(4), 606–616.

Y. Kamarianakis, W. Shen, L. Wynter, 2012. Real-time road traffic forecasting using regime-switching space-time models and adaptive LASSO, Applied Stochastic Models in Business and Industry 28 (4), 297-315.