Estimating Baseline Travel Times for the UK Strategic Road Network*

Alvaro Cabrejas Egea[1], Peter de Ford[2] and Colm Connaughton[3]

Abstract— **We present a new method for long-term estimation of the expected travel time for links on highways and their variation with time. The approach is based on a time series analysis of travel time data from the UK's National Traffic Information Service (NTIS). Time series of travel times are characterised by a noisy background variation exhibiting the expected daily and weekly patterns punctuated by large spikes associated with congestion events. Some spikes are caused by peak hour congestion and some are caused by unforeseen events like accidents. Our algorithm uses thresholding to split the data into background and spike signals, each of which is analysed separately. The the background signal is extracted using spectral filtering. The periodic part of the spike signal is extracted using locally weighted regression (LWR). The final estimated travel time is obtained by recombining these two. We assess our method by cross-validating in several UK motorways. We use 8 weeks of training data and calculate the error of the resulting travel time estimates for a week of test data, repeating this process 4 times. We find that the error is significantly reduced compared to estimates obtained by simple segmentation of the data and compared to the estimates published by the NTIS system.**

I. Introduction

The UK collects and processes data from its Strategic Road Network (SRN) in real time and makes it available through the National Traffic Information Service (NTIS) [1]. NTIS collects speed, flow and travel time data using sensors on the road and in vehicles. The basic building blocks of the SRN are called links - segments of motorway between 500 and 20000 metres in length. NTIS data is used to assign a travel time profile to each link. Profiles should represent the typical time to traverse a link at a given time of day on a given day of the week and are the topic of study of this paper.

Profiles clearly vary with day of the week and time of the day but should be relatively stable from one week to the next. This stability over time means calculating profiles is different from the problem of short-term forecasting. While there is no a-priori mathematical definition of a typical travel time, we take it to mean the value that minimises the mean absolute relative error (MAPE) with respect to subsequently measured travel times. NTIS publishes profile values alongside measured travel times although the methodology used to calculate them is not in the public domain. In this paper, we present a novel method for generating profiles for a complete week. Our approach is based on statistical analysis of previous data and does not require any a-priori segmentation of days into classes. Rather patterns of intra- and inter-day variability are learned directly from the data.

The available literature on travel times is extensive but most recent research focuses on short-term forecasting and with fewer studies looking into the long term [2] [3]. From a methodological point of view statistical methods and machine learning take most of the attention [4]. Within the last group, neural networks are getting a level or relevance [5] [6] . Others take closer approaches to this paper: using historical data [7] [8], differentiating between peak and non-peak [9], performing spectral analysis [10] or Locally Weighted Regressions [11] [12] [13] [14] [15] [16]. Comparisons between some of these studies and others can be found in [17], [18], [19] and [20]. Most of these methods have been specifically tuned for the conditions on which they have been developed and transferability to other sites is often not evaluated. To assess the transferability of our method, it is tested on three different motorways and scored independently in each of them. The method here presented, uses a combination of spectral analysis, tree decisions and Locally Weighted Regression (LWR).

II. Examples of travel times

The most direct variable for measuring the state of traffic over a length of road is the vehicles' travel times. The instantaneous travel time for a given segment of road is the average time that the vehicles currently in it are taking since they enter the segment until they exit it.

[1]Alvaro Cabrejas Egea is with MathSys Centre for Doctoral Training, University of Warwick, CV4 7AL, Coventry, United Kingdom `a.cabrejas-egea@warwick.ac.uk`

[2]Peter de Ford with the Geospatial Department, McKinsey & Company, Costa Rica `peter_de_ford@mckinsey.com`

[3] Colm Connaughton with the Warwick Mathematics Institute & Centre for Complexity Science, University of Warwick, CV4 7AL, Coventry, United Kingdom `b.d.researcher@ieee.org`
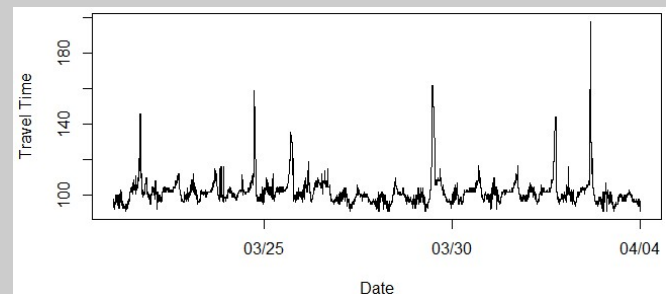
Fig. 1.   M6 travel times (seconds), link 117007401, over three weeks.

As it can be seen in Figure 1 the travel time remains within a vaguely predictable pattern most of the days, with bounded minima during nights corresponding with the free flow time and with some outliers below this value corresponding to speeding drivers. Travel time will meaningfully rise as people leave to work and add load to the motorways. This collective behaviour will create the morning traffic jams, which in our data, are partially replicated during the evening rush hour, normally finding a plateau in between. After this, the travel times slowly decay towards the night period of free-flow regime. In Figure 1, it can also be seen that there are a series of excursions out of this oscillating yet bounded typical behaviour. In these, the travel time can increase several times fold the usual values. These extreme oscillations are much less predictable both in intensity and inter-oscillation period than the recurrent congestion described previously.

III. Basic methods for profile estimation

A. Use of Exponentially Weighted Moving Average (EWMA)

The most basic approach to profile estimation is to apply an EWMA on the same minute of every day, with the implicit assumption that similar behaviour is expected at the same time of the day. In this approach, the estimated profile $\hat{x}(i, d + 1)$ on the $i - th$ minute of a given day d, for a memory parameter $\alpha \in [0, 1]$ and with measured travel time x_i^d, will be:

$$\hat{x}_i^{d+1} = \alpha * x_i^d + (1 - \alpha) * \hat{x}_i^d \tag{1}$$

The main problem with EWMA is the manner in which the memory decays. Recent measurements are weighted more heavily than events in the past. Thus if an extreme fluctuation occurs, the following profile estimates will be biased, partially replicating this event and over-estimating the travel times until enough new measurements arrive to dissipate this effect.

B. Segmentation

In addition, to acknowledge the specific differences between some special dates in the year, and the difference between days of the week, this family of methods requires the use of heavy date based segmentation. The EWMA will be applied across dates which fall in the same category (i.e. Mondays, weekends, Christmas Day, ...). If this is combined with the shortcomings presented in Section III-A, some long reaching effects are generated, which can propagate for weeks into the future predictions, but do not have any reflection on the observed travel times. A possible instance of this effect can be seen in Figure 2. Furthermore, in order to generate a valid segmentation, an experienced team is necessary, since the needs of the process can geographically vary, given that the EWMA approach tends to better approximate endemic congestion. This dependence can lead to the creation of legacy systems which may not be well understood after a few years, decreasing their usability over time unless extra effort is put into transmitting this knowledge and continually train new staff.
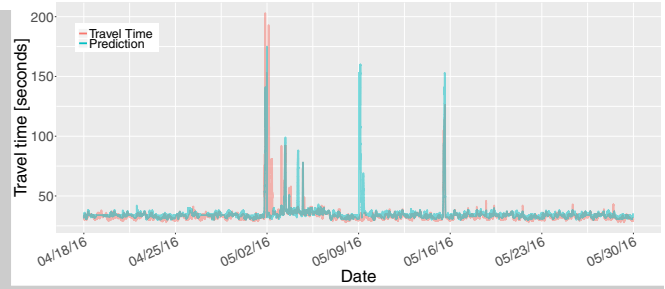
Fig. 2. A single large random spike in travel time can lead to over-estimation of subsequent profile estimates.

IV. Data

A. Data Gathering and Selection

Data was gathered from 3 different Motorways in England. The M6 and M11 were selected due to their high usage and combination of recurrent and outstanding congestion. The M25 was selected on the base that it is the most used Motorway in England on a daily basis, suffering from chronic congestion.

- The dataset for M6 comprises 90 days (12 complete weeks) of data (07/03/2016-05/06/2016) across 14 links.
- The dataset for d M11 comprises 90 days (12 complete weeks) of data (07/03/2016-05/06/2016) across 25 links.
- The dataset for M25 comprises 75 days (10 complete weeks) of data (07/04/2017-20/06/2017) across 61 links.
- Links were discarded if they had more than 10% of missing data or if they contained any entry or exit ramps.
- Whenever missing data was detected for 10 or less minutes, it was linearly interpolated.
- Whenever missing data was detected for over 10 minutes, it was left as missing values.

In the case of the M6 and the M11 the first 8 weeks are used to predict one complete week ahead. One week later, the process is repeated, deleting the oldest week of training data and incorporating measurements of the week predicted in the previous step. This is performed 4 times. In the case of the M25, 6 weeks of data were used for training and the process as described above is performed 3 times.

B. Data Contents

For each link on a specific date, the required input data consists of one entry per minute, containing:

- Measured travel time
- Profile (expected) travel time

V. Travel Time Prediction Algorithm

Given the cyclic nature of traffic, the aim was a prediction algorithm that could account for the periodic variations and endemic congestion while being resilient to fluctuations and rare events. This algorithm also should:

- Be robust, mitigating the propagation of isolated events into future forecasts, unlike methods using EWMA.
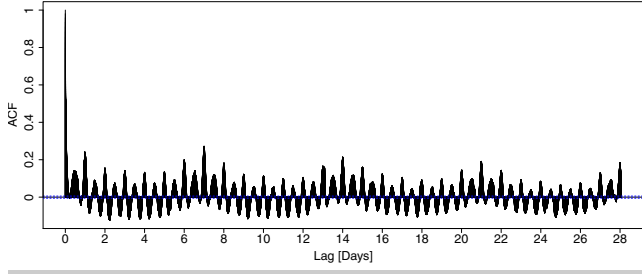
Fig. 3. Autocorrelation function of a link in the M6 over a period of 4 weeks, showing seasonal patterns on the daily and weekly periods.

- Not require the use of time segmentation and be valid for regular and "special" dates.
- Be location agnostic, the internal parameters should be set algorithmically based on the data.
- Have Gaussian, mean 0, uncorrelated residuals.
- Near flat Trend term, given the different time scales between seasonal cycles and changes in the general motorway flow.

A. Naive Segmentation

To obtain an accurate comparison of the performance of the algorithm developed in the following subsections, an example of basic segmentation was coded. This involved a weighted combination of the training data points using uniform weights. In this way, for the $i-th$ minute of a week and using a training set composed of the previous of n weeks, the Naive Segmentation (NS) profile is:

$$\hat{x}(i,n) = \sum_{\text{week}=1}^{n} \frac{x_n^i}{n} \qquad (2)$$

B. Decomposition in Background and Spikes

During exploratory data analysis we found that, from the point of view of travel times, traffic operates in two differentiated regimes that we denominated Background and Spikes.

- Background:
  - Stable around a mean value.
  - Oscillates with small amplitude and high frequency.
  - Suitable for spectral filtering.
- Spikes:
  - Zero most of the time. Quickly go to extreme values.
  - Oscillates with great amplitude and low inter-oscillation frequency, creating far reaching effects.
  - Suitable for seasonal decomposition.

In this context, assuming Gaussian noise ξ, and that the components operate in an additive way, generated a signal of the form:

$$\text{Travel Time}_t = \text{Background}_t + \text{Spikes}_t + \xi \qquad (3)$$

The objective was to separate them in such a way that the moments of smooth and normal operation are captured
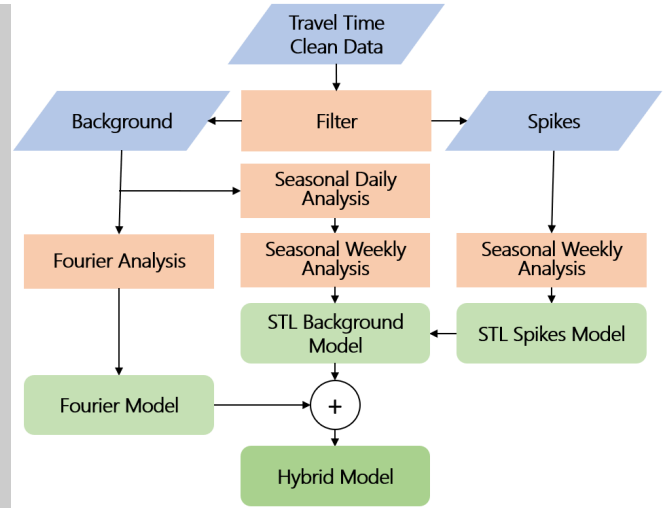
Fig. 4. Schematic of data streams in the algorithm.

as part of the Background and used for spectral analysis, attempting to mitigate the prediction error induced by the high frequency oscillations and obtaining a basic view of what can be daily observed. Meanwhile, the spikes, including the recurring and non-recurring congestion, were treated separately searching for seasonality in larger time scales than those in which the travel time signal oscillates, as suggested by Fig. 3. Ideally, after this final step of the decomposition, the remainder should only contain isolated large rare events deviating from the profile and white noise.

To prevent the differing lengths of the links from affecting this decomposition, for this step, all travel times were normalized according to their corresponding link's free flow time, defined as the time to transverse the length of the link at the maximum legal speed allowed by the motorway. However, this step only mitigates the non-regularity of the time series, since there are drivers who do not follow these limits. A threshold was heuristically set, as seen in Table I to separate the two components in the different regimes from the normalised travel times on a per-motorway basis, although better results can be obtained by setting each link individually. Intuitively, this threshold scales with the amount of recurring congestion in a link. Whenever a data point was above the threshold, it was flagged as belonging to a spike.

TABLE I
TRAVEL TIME NORMALISATION THRESHOLDS

Motorway	Threshold
M6	1.1
M11	1.2
M25	1.4

For this purpose, an indicator function was defined, taking for every minute the value:

$$\delta_t^{spike} = \begin{cases} 1, & \text{if } x_t > \text{Threshold} \\ 0, & \text{otherwise} \end{cases} \qquad (4)$$
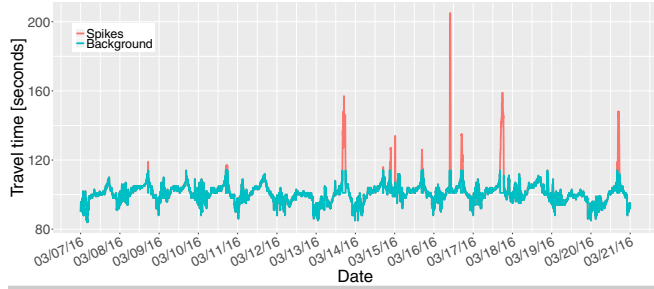
54 pt
0.75 in
19.1 mm

54 pt
0.75 in
19.1 mm

Fig. 5.   Decomposition time series in background and spikes

C. Spectral Component

The main difficulty when dealing with the Background signal is the low amplitude, high frequency fluctuations that can be found almost ubiquitously. Signal smoothing can be performed by completely removing a range of frequencies while the information bearing bands are retained. For this task, the Fast Fourier Transform (FFT) [21] was used as follows:

1) Calculate Background Power Spectrum using FFT
2) Remove frequencies corresponding to periods under 4 hours and over 1 week
3) Repeat for all n weeks in training set
4) Apply EWMA to the modified weekly Power Spectra
5) Compute the Inverse FFT Transform

D. Seasonality Component

Seasonal-Trend Decomposition based on LOESS (STL) [22] was the chosen algorithm for the seasonality analysis since it can handle any type of seasonality, allowing the user to control how it changes over time as well as the smoothness of the trend-cycle while being highly robust to outliers [23].

Below, the sequence of steps taken to extract the Seasonality Components that can be observed in Fig. 3 is described in line with the schematic data streams represented in Fig. 4. Note that this should be applied to the n training weeks as a single time series.

1) Decomposition of Background for daily seasonality
2) Extract and sum the series corresponding Trend and Remainder from Step 1
3) Decomposition of output from 2 for weekly seasonality
4) Add daily and weekly Seasonal components from Step 1 and Step 3 to obtain global seasonality
5) Average seasonality across training weeks
6) Linearise Trend term from output of Step 3
7) Add linearised trend to seasonality obtained in Step 6
8) STL Decomposition of Spikes for weekly seasonality
9) Extract Spike's Seasonality corresponding to the number of weeks for forecasting
10) Add Spike's Seasonal component to the output of 8

To ensure a successful decomposition, after performing Step 3, the Background's Remainder should be zero mean, Gaussian distributed. Background's Trend should have a near zero slope. These outcomes should also hold if performed again after Step 8, with the addition that the trend should also be close to zero in absolute value, since the time scales in which such a global trend can meaningfully vary should be much greater than the prediction scope of this algorithm.

E. Seasonal-Spectral Hybrid Profile

In order to create the final Seasonal-Spectral Hybrid profile (Hybrid), one of the two forecasts generated in the previous points is taken, depending on what is the identified regime:

$$\text{Hybrid} = \text{Seasonal} * \delta_{\text{spike}} + \text{Spectral} * (1 - \delta_{\text{spike}}) \quad (5)$$

VI. ACCURACY RESULTS

In this section the accuracy of the algorithm described above is compared against the Published Profiles and the NS Model. For each temporal point i, the Mean Average Relative Error (MARE) is defined below:

$$\text{MARE} = \frac{\left(\sum_{i=1}^{n} \frac{\|x_i - \hat{x_i}\|}{x_i}\right)}{n} \quad (6)$$

For complete Motorways, the Root Mean Squared Error (RMSE) has been calculated for each temporal point i as:

$$\text{RMSE} = \sqrt{\frac{\sum_{i=1}^{n}(x_i - \hat{x_i})^2}{n}} \quad (7)$$

In both cases the error for a Link is defined as the average MARE or RMSE across all prediction points. The error for a Motorway is defined as the average of the error across all its links.
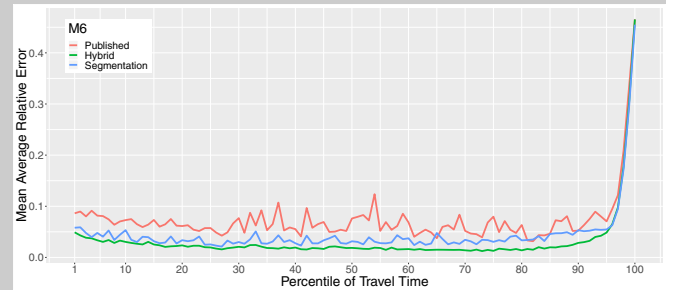
A. Accuracy by Quantile

Fig. 6.   Average accuracy results in M6 across percentiles of travel time.

Here, the accuracy of the algorithm is compared against the Published Profiles and the NS Profile across all percentiles of travel time. As it can be seen in Figs. 6, 7 and 8, the Hybrid Profile has a higher accuracy than the Published Profiles and the NS Model for all percentiles of travel time except for the most extreme values where they all perform poorly.

The most meaningful difference occurs between percentiles [50 − 95], where the Published Profile starts to suffer from higher inaccuracy.

54 pt
0.75 in
19.1 mm

54 pt
0.75 in
19.1 mm

54 pt
0.75 in
19.1 mm

Margin requirements for the other pages
Paper size this page US Letter
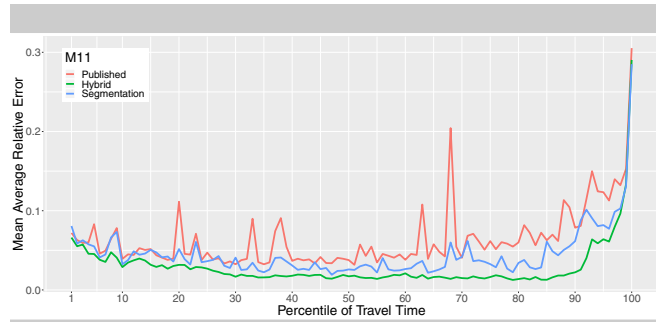54 pt
0.75 in
19.1 mm

Fig. 7. Link average accuracy results in M11 across percentiles of travel time.
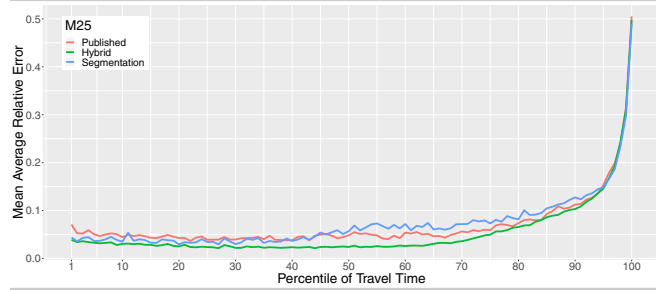

Fig. 8. Link average accuracy results in M25 across percentiles of travel time.
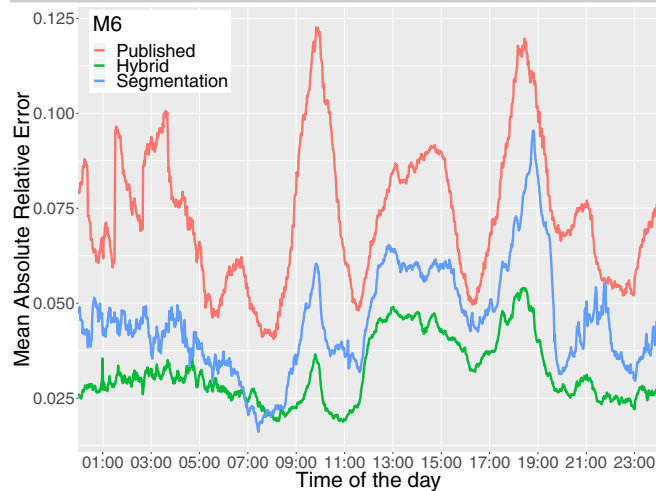
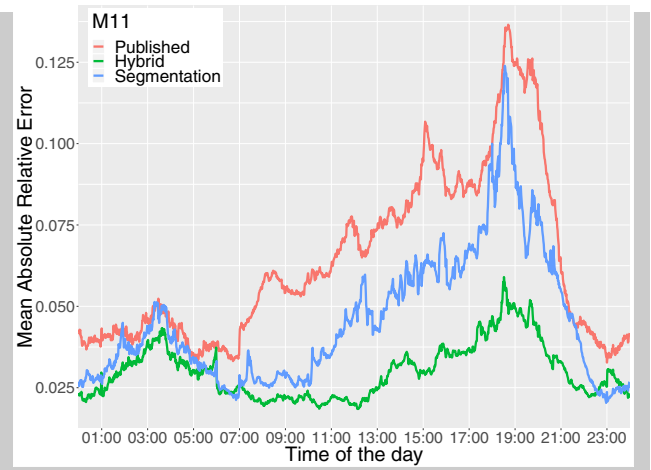Fig. 9. Link average accuracy results in M6 across times of the day.

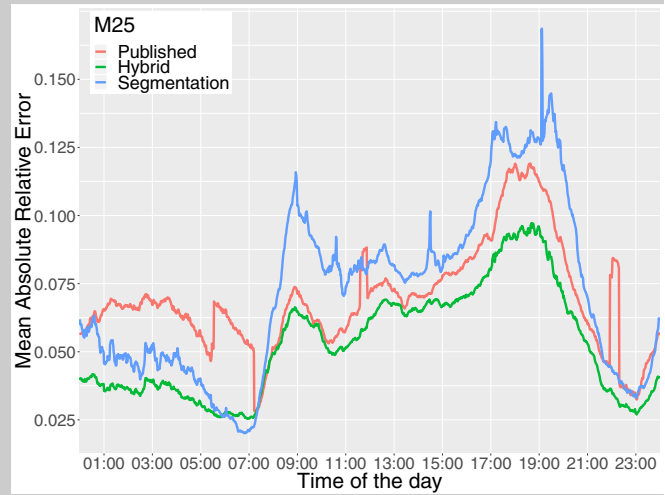Fig. 10. Link average accuracy results in M6 across times of the day.

Fig. 11. Link average accuracy results in M25 across times of the day.

B. Daytime Error

Here, the accuracy of the algorithm is compared against the Published Profiles and the NS Profile across the times of the day. As it can be seen in Figures 9, 10 and 11, the Hybrid Profile has a higher accuracy than the Published Profiles and the NS Model for all times of the day, all locations and training lengths. The most relevant improvement occurs during the morning and evening peak hours, where the algorithm presented in this paper does not accuse meaningful performance worsening relative to the morning plateau when compared with the other two profiles.

In the case of the M6 and M11, where the training set is

TABLE II

MARE DISTRIBUTION PER MOTORWAY

% rel. error	<-25%	-25% to -15%	-15% to -5%	-5% to 5%	5% to 15%	15 % to 25%	>25%
M6	1.58	0.60	3.77	88.01	5.97	0.06	0.01
M11	0.80	0.35	4.07	86.15	7.97	0.49	0.15
M25	3.85	2.73	10.42	75.12	7.29	0.33	0.28

richer, the error at peak times is reduced by at least 50% in all cases, reaching as much as 68.7% in the case of the M6 morning rush. In the case of the M25, which is congested on a regular basis, the errors in the Published Profile during peak times are slightly lower than on the other cases, indicating that, given the use of EWMA for its calculation, recurrent congestion is better captured by it. Even in this case, the proposed algorithm performs significantly better than any of the other two, except for a brief window between 6-7AM when it is outperformed by the NS Model.

TABLE III

MARE PER LINK ON M6

Link	MARE
117007401	0.0211
117007501	0.0421
117007601	0.0246
117007801	0.0480
117007901	0.0308
117008401	0.0331
117009102	0.0269
117011901	0.0235
117012001	0.0446
117012101	0.0262
117012201	0.0431
117012301	0.0275
117016001	0.0360
123025901	0.0189

TABLE IV

GLOBAL MARE & RMSE PER MOTORWAY

Motorway	MARE	RMSE [s]
M6	0.0318	0.0368
M11	0.0300	0.0355
M25	0.0552	0.0667

VII. FUTURE WORK

The algorithm presented above meets the requirements described in Section V except for the need of an heuristically set threshold to separate the original Travel Time signal into Background and Spikes. One potential way of reaching compliance with these requirements is to perform the decomposition by applying a Wavelet Transform and separating background from spikes based on statistical analysis of the transformed time series in terms of how their Wavelet coefficients fluctuate over time within a scale level. In the future a sensitivity analysis should be conducted to explore the limits of the algorithm in terms of minimum training data set, as well as maximum performance with increased training.

REFERENCES

[1] The Highways Agency, "National Transport Information System Publish Services", Technical Report, 2011.

[2] J. Moreira, A. Jorge, J. Sousa, C. Soares 2012. Comparing state-of-the-art regression methods for long term travel time prediction. Intelligent Data Analysis. 16. 10.3233/IDA-2012-0532.

[3] G. Klunder, P. Baas, F. op de Beek. 2018. A long-term travel time prediction algorithm using historical data. 14th World Congress on Intelligent Transport Systems, ITS 2007; Beijing; China; October 2007 through 13 October 2007, 1191-1198

[4] Kirby, H. R., Watson, S. M., Dougherty, M. S. 1997. Should we use neural networks or statistical models for short-term motorway traffic forecasting?. International Journal of Forecasting, 13(1), 43-50.

[5] Lee, Y. 2009. Freeway travel time forecast using artifical neural networks with cluster method. In Information Fusion, 2009. FUSION'09. 12th International Conference on (pp. 1331-1338). IEEE.

[6] Park, D. et al. 1999. Spectral basis neural networks for real-time travel time forecasting. Journal of Transportation Engineering, 125(6).

[7] Rice, J., Van Zwet, E. 2004. A simple and effective method for predicting travel times on freeways. IEEE Transactions on Intelligent Transportation Systems, 5(3), 200-207.

[8] Chien, S. I. J. et al. 2003. Dynamic travel time prediction with real-time and historic data. Journal of transportation engineering, 129(6), 608-616.

[9] Zhang, Y. et al. 2011. Analysis of peak and non?peak traffic forecasts using combined models. Journal of Advanced Transportation, 45(1).

[10] Nicholson, H., Swann, C. D. 1974. The prediction of traffic flow volumes based on spectral analysis. Transportation Research, 8(6), 533-538.

[11] Williams, B., Durvasula, P., Brown, D. 1998. Urban freeway traffic flow prediction: application of seasonal autoregressive integrated moving average and exponential smoothing models. Transportation Research Record: Journal of the Transportation Research Board, (1644), 132-141.

[12] Sun, H., Liu, H. X., Xiao, H., He, R. R., Ran, B. 2003. Short term traffic forecasting using the local linear regression model. In 82nd Annual Meeting of the Transportation Research Board, Washington, DC.

[13] Zhong, M., Sharma, S., Lingras, P. 2005. Refining genetically designed models for improved traffic prediction on rural roads. Transportation planning and technology, 28(3), 213-236.

[14] Chowdhury, N. K., Nath, R. P. D., Lee, H., Chang, J. 2009. Development of an effective travel time prediction method using modified moving average approach. In International Conference on Knowledge-Based and Intelligent Information and Engineering Systems (pp. 130-138). Springer, Berlin, Heidelberg.

[15] Dell'Acqua, P., Bellotti, F., Berta, R., De Gloria, A. 2015. Time-aware multivariate nearest neighbor regression methods for traffic flow prediction. IEEE Transactions on Intelligent Transportation Systems, 16(6), 3393-3402.

[16] Kumar, S. V., Vanajakshi, L. 2015. Short-term traffic flow prediction using seasonal ARIMA model with limited input data. European Transport Research Review, 7(3), 21.

[17] Nikovski D, Nishiuma N, Goto Y, Kumazawa H. Univariate short-term prediction of road travel times. Intelligent Transportation Systems, 2005. Proceedings. 2005 Sep 13 (pp. 1074-1079). IEEE.

[18] van Hinsbergen, C.P.IJ., van Lint, J.W.C., Sanders, F. M. Short Term Traffic Prediction Models. ITS World Congress, Beijing, China. 2007.

[19] Mori U, Mendiburu A, Alvarez M, Lozano JA. A review of travel time estimation and forecasting for Advanced Traveller Information Systems. Transportmetrica A: Transport Science. 2015 Feb 7;11(2):119-57.

[20] Lana I, Del Ser J, Velez M, Vlahogianni EI. Road Traffic Forecasting: Recent Advances and New Challenges. IEEE Intelligent Transportation Systems Magazine. 2018;10(2):93-109.

[21] J. W. Cooley and J. W. Tukey, "An algorithm for the machine calculation of complex Fourier series", Mathematics of computation, 1965, vol. 19, no 90, p. 297-301.

[22] R. B. Cleveland, W. S. Cleveland, J. E. McRae and I. Terpenning, "STL: A Seasonal-Trend Decomposition Procedure Based on Loess", Journal of Official Statistics, vol. 6, 1990, pp. 3-73.

[23] R. J. Hyndman and G. Athanasopoulos, "Forecasting: Principles and Practice", Otexts, 2013.