WAVELET AUGMENTED REGRESSION PROFILING (WARP): IMPROVED LONG-TERM ESTIMATION OF TRAVEL TIME SERIES WITH RECURRENT CONGESTION *

Alvaro Cabrejas Egea [†]
MathSys Centre for Doctoral Training,
University of Warwick & Alan Turing Institute
CV4 7AL Coventry, NW1 2DB London, UK
a.cabrejas-egea@Warwick.ac.uk

Colm Connaughton [‡]
Warwick Mathematics Institute
University of Warwick
CV4 7AL Coventry, UK
c.p.connaughton@warwick.ac.uk

ABSTRACT

Reliable estimates of typical travel times allow road users to forward plan journeys to minimise travel time, potentially increasing overall system efficiency. On busy highways, however, congestion events can cause large, short-term spikes in travel time. These spikes make direct forecasting of travel time using standard time series models difficult on the timescales of hours to days that are relevant to forward planning. The problem is that some such spikes are caused by unpredictable incidents and should be filtered out, whereas others are caused by recurrent peaks in demand and should be factored into estimates. Here we present the Wavelet Augmented Regression Profiling (WARP) method for long-term estimation of typical travel times. WARP linearly decomposes historical time series of travel times into two components: background and spikes. It then further separates the spikes into contributions from recurrent and residual congestion. This is achieved using a combination of wavelet transforms, spectral filtering and locally weighted regression. The background and recurrent congestion contributions are then used to estimate typical travel times with horizon of one week in an accurate and computationally inexpensive manner. We train and test WARP on the M6 and M11 motorways in the United Kingdom using 12 weeks of link level travel time data obtained from the UK's National Traffic Information Service (NTIS). In out-of-sample validation tests, WARP compares favourably to estimates produced by a simple segmentation method and to the estimates published by NTIS.

1 Introduction

1.1 Background

In the United Kingdom, Highways England (HE) is responsible for most motorways and major roads in England, comprising the Strategic Road Network (SRN). The SRN is monitored by the National Traffic Information System [1], which gathers and processes travel time, speed, flow and headway data in real time using road sensors and special vehicles. The smallest components of the SRN are the "links", segments of motorway ranging from 500 to 20000 metres long. NTIS uses the historic data to assign each link in the network with a traffic profile. Traffic profiles contain, for each minute of a day in a given date, the expected travel time for the link to which it is associated. These profile values are published together with the actually measured travel times.

Travel time profiles of heavily-used roads vary strongly depending on the time of the day and the day of the week. However their inter-week variation remains low, with major changes in usage patterns taking longer (order of months) than the biggest temporal unit considered (weeks). This slow rate of change with respect to the data frequency illustrates that the calculation of travel time profiles is a very different task than short-term forecasting. Since there is no uniformly

accepted definition of traffic profiles (typical travel times), here, we will define it as the collection of values that jointly minimise the Root Mean Squared Error (RMSE) or Mean Absolute Relative Error (MARE), when compared with subsequently measured real travel times.

In this paper, we follow up on previous work in [2], providing a novel algorithm capable of generating one week of expected travel times, using a single parameter and using only publicly available data from NTIS. This approach stems from spectral and statistical analysis of recent historical data and removes the requirement of any prior form of segmentation of times and days into different classes. Instead, the algorithm relies on pattern discovery for intra- and inter-day variability which it computes directly from input data.

1.2 Previous Work

There is an extensive literature on travel times, although recent research is more focused on shorter term forecasting, with far fewer long term estimation studies [3] [4]. Machine learning and statistical analysis methods receive the most attention [5]. In machine learning, neural networks are recently having a high relevance [6] [7]. Other approaches are closer to the methods here presented: making use of historical data [8] [9], differentiating between rush and non-rush hour [10], using spectral methods [11] or Locally Weighted Regressions [12] [13] [14] [15] [16] [17]. The Wavelet Transform has been previously found to be useful in combination with Kalman filters [18], neural networks [19] [20] [21] [24] and statistical analysis [22] [23]; but these either cover short predictions, use the Wavelet analysis of travel times for other purposes such as incident detection, or do not focus on the spectral properties across timescales or are not real-world data based. Comparisons involving some of these studies are performed in [25], [26], [27] and [28], but they either focus on short-term prediction or do not produce an overall best-performer. The method here presented is distinct from the previous work mentioned, since it does not use a sample of individual trips, but all conducted over 12 weeks in multiple sites. The aggregation period and prediction resolution is minutely, and the prediction horizon is larger. Among these methods, the transferability to other locations is not often evaluated, being tuned for a specific location with its own specific conditions. In order to ensure transferability, our method is examined on 39 individual locations across two motorways and independently scored for each site. The work presented here makes use of a combination of Continuous Wavelet Transform [29], tree decisions, spectral analysis, and Locally Weighted Regression (LWR).

2 Travel times in Motorways

From a user perspective, vehicle travel times are the most relevant measure of the state of the traffic flow on a road link. The average travel time in a given minute of the day on a given link will be the average time to travel from the entry to the exit loop sensor for all vehicles that passed through. From Figure 1, it can be observed that most of the time,
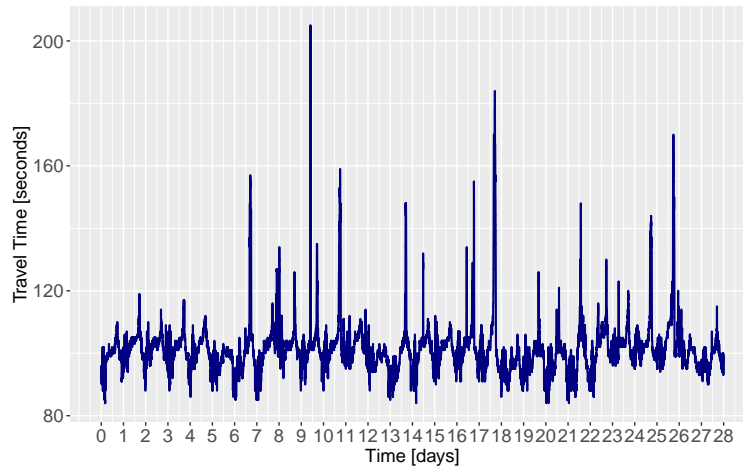

Figure 1: Link 1170079, M6. Travel Time over 28 days of minutely data between 07/Mar/2016 and 04/Apr/2016.

the travel time on a link follows a repeating pattern with minima located at night matching the bounded free flow time (except for speeding drivers). As the morning rush starts, travel time will rise as traffic jams are generated from the collective drivers' behaviour. More effects of these collective dynamics can be observed in the afternoon during the evening rush, normally being possible to find a plateau between these two peaks. Finally, travel times progressively decay towards the night's free-flow regime. In Figure 1, we observe a series of spikes found outside of this normally

bounded yet oscillating behaviour. Travel time in these events can climb up to several times the normal amount. The predictability in terms of duration and amplitude of these spikes is much lower than the periodic component described above. Lastly, in Figure 2, we can see the Autocorrelation Function for the Travel Time, plotted with a maximum lag of
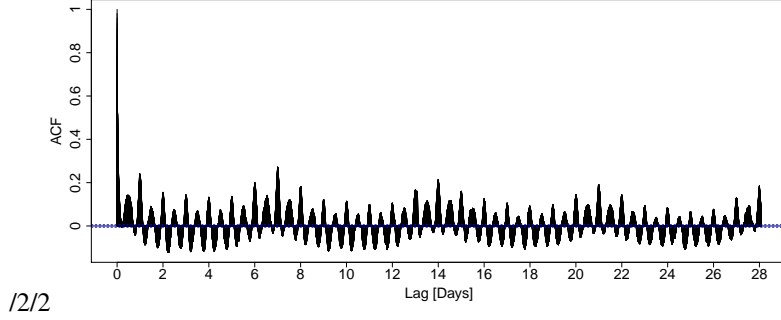

/2/2

Figure 2: Autocorrelation function of Travel Time in figure 1 with a maximum lag of 4 weeks (1440 points/day). Double seasonality can be observed in the weekly and daily periods.

4 weeks. Here we can see a double seasonal pattern with periods of 1 day and 1 week. This regularity seen in time travel time series can and is often used by modellers to approximate and forecast travel times as is explored over the next sections.

3 Basic methods for profile estimation

3.1 Exponentially Weighted Moving Average for Profiles

As explained in [2], a basic approach to estimating profiles is to apply an Exponentially Weighted Moving Average (EWMA) across a given minute of the available days, assuming that similar behaviour is to be expected at similar times on different days. Then, the profile estimation $\hat{x}(i, d + 1)$ for the $i - th$ minute of a date d, controlling our memory parameter $\alpha \in [0, 1]$ to balance the memory of the process and with measured travel time x_i^d, will be:

$$\begin{aligned} \hat{x}_i^{d+1} &= \alpha x_i^d + (1 - \alpha)\hat{x}_i^d \\ &= \alpha x_i^d + \alpha^2 x_i^{d-1} + (1 - \alpha - \alpha^2)\hat{x}_i^{d-2} \\ &= ... \end{aligned} \tag{1}$$

EWMA-based profiles have a main issue: disruptions after large events are generated due to the way in which the memory decays as we can see above in equation 1. Recent measurements receive exponentially greater weights than
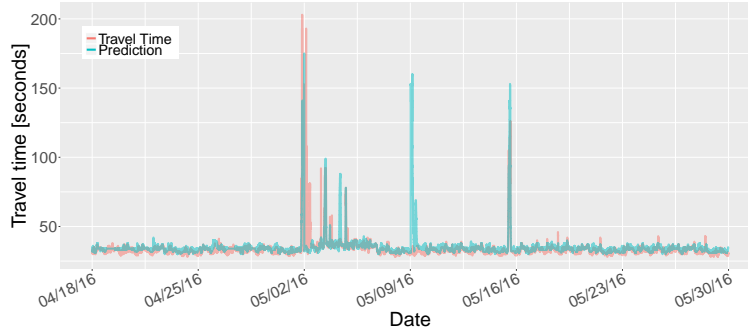

Figure 3: Skewed EWMA predictions after a large deviation.

events farther in the past. If a large deviation from the baseline patter occurs, subsequent estimations will have a bias towards partially replicating the deviation and consistently predicting over-estimates until newer data is included and the effect dissipates.

3

3.2 Time-based Segmentation

Furthermore, to recognize and use the precise characteristics of specially distinct dates, these methods make heavy use of segmentation based on the date. Days are grouped according to pre-defined categories and the EWMA is applied across dates falling in the same category (i.e. Saturdays, weekdays, New Year, ...). This segmentation in classes, joined with the weakness commented in Section 3.1 can generate lasting perturbations, propagating across weeks into future estimations, yet never being reflected on the travel times measurements. A likely instance of such an event is shown in Figure 3. Consequently, significant operational and geographical expertise about specific roads is needed to create a valid segmentation, given that the EWMA approach is effective exclusively for recurrent congestion. These requirements often lead to the generation of legacy systems which become increasingly difficult to maintain as time passes, their usefulness declining over time or requiring additional efforts for continually training new staff to maintain the system. These modelling and operational limitations may make segmentation and EWMA based profiles suboptimal both in performance and operation.

4 Data Selection and Contents

The M6 and M11 motorways in England are chosen due to their high use and their display of both recurrent and outstanding congestion, being key in several heavily used commuting routes.

- The dataset aggregates 90 days (12 complete weeks) minutely entries (07/03/2016-05/06/2016). This length is chosen to minimise the effect of including other wider partially captured periodicities (e.g. natural seasons) during data gathering.
- Links with over 10% of data missing or containing access ramps were discarded.
- The previous condition left 14 different links in the M6 and 25 links in the case of the M11.
- Entries missing for 10 minutes or less were linearly interpolated.
- Entries missing for over 10 minutes were left as missing values.

The algorithm uses 8 weeks of data to predict one entire week ahead. After a week, the oldest week is deleted, and the most recent one is incorporated, producing a new estimate for the subsequent week. This procedure is simulated 4 times. For each link-date pair, the data comprises minutely data, containing: average vehicle travel time in seconds, profile (expected) travel time in seconds, traffic flow in cars/hour and vehicle headway in meters.

5 Background and Spikes

5.1 Characteristics

As described in [2], if we look at the travel times, we find they operate in two different regimes that we call background and spikes. The background is stable with small high-frequency fluctuations around a time-varying mean value. This makes it suitable for seasonal analysis and spectral filtering (smoothing). In contrast, the spikes are zero most of the time but can quickly climb to extreme values. They have much greater amplitude and much lower frequency, creating long reaching effects. Although they are non periodic in the time domain, a non-harmonic seasonality contribution associated with recurrent congestion can be extracted via non-parametric regression.

In this context, if we assume Gaussian noise ξ, and given the additive properties of wavelet decomposition, the decomposition will be of the form:

$$\text{Travel Time}_t = \text{Background}_t + \text{Spikes}_t + \xi \tag{2}$$

The objective is to separate the signals such that the times of smooth non-congested operation, together with the recurring congestion, are captured in the background and passed through spectral smoothing, mitigating the estimation errors created by the high frequency oscillations and achieving a view of what can be daily observed, so seasonal patterns can be extracted in the shorter and longer periods shown in Fig. 2. Meanwhile, the spikes, containing the non-recurring congestion, can be searched for any seasonality left on time scales larger than the period in which the travel times oscillate, as also suggested by Fig. 2. If performed correctly, the remainder after this seasonal extraction step of the decomposition should contain only isolated events with large deviations from the profile and white noise. Wavelets are especially well suited for this since they adapt to the relevant scales and location in time, not being a property that is found in a priori explicit forms (splines, polynomials, etc.).
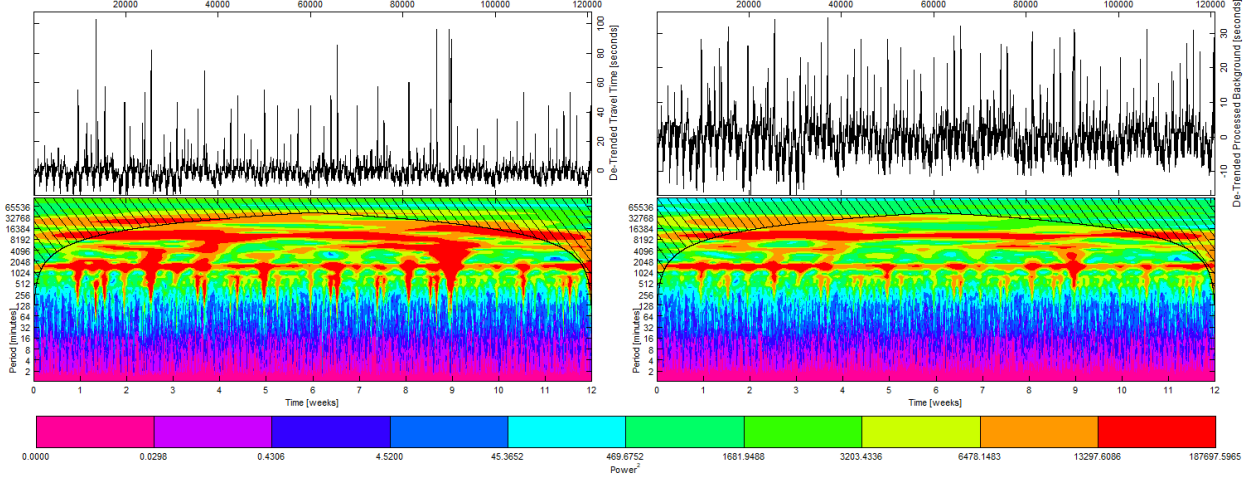
Figure 4: Left: Power of CWT of the original Travel Time series for Link 1170079 of the M6. Right: Power of CWT of the extracted Background $\vec{B}_t$ series, as detailed in Section 5.2, for the same link, displaying more stable and periodic dynamics.

5.2 Wavelet Time Series Decomposition

To perform the decomposition, we will take advantage of the additive properties of the wavelet transform. First the time series $\vec{x_t}$ of length t, and elements x_i with $i \in [0, t]$ is turned into a zero mean series $\vec{x_t} - mean(\vec{x_t})$ and used as input for a Continuous Wavelet Transform (CWT) [30] [31] using a Morse wavelet [32] and 140 timescales levels. The output of this first transform is a complex matrix $\vec{W}_{\text{levels} \times t}$, for the elements of which we calculate their modulus ρ, phase ϕ and power P. A heatmap of P^2 for the original time series can be seen in the first subplot in Figure 4. In the figure we can observe that the most influential dynamics occurring during the series length (x axis) happen at the timescales (y axis) where it was expected based on Figure 2, namely 1 day (1440 minutes) and 1 week (10080 minutes), based on the P^2 values on these bands. In the figure, we also observe that surges in power across different wavelet timescales occur at the same time as the non-recurring congestion, since in order to approximate this signal, the CWT algorithm needs to combine several wavelets with smaller periods than those dominating the recurring part of the series.

In order to isolate this non-recurring component, the series is sequentially assessed over all the time domain by taking a horizontal slice for a single timescale level l, and generating a series $\vec{x_t}^l$. After fixing l, we calculate the Median and Inter Quantile Range of the distribution of values in $\vec{x_t}^l$ to search for outliers. A maximum threshold value is set equal to $T = median(\vec{x_t}) + \alpha * IQR(\vec{x_t}^l)$, with $\alpha \in [0, \inf]$ being a parameter that defines how aggressively we target spikes. Then, the individual elements of $\vec{x_t}^l$, x_i^l are individually evaluated: if found below the limit, they are stored in a background container $B_i^l = x_i^l$; otherwise, the fraction below the threshold will be nonetheless passed on to the background $\vec{B}_i^l = median(\vec{x_t}^l) + \alpha * IQR(\vec{x_t}^l)$, with the remaining part going into the spikes storage $\vec{S}_i^l = x_i^l - B_i^l$.

In the extremes, for $\alpha = 0$ we would find that any deviation from the median is taken into the spikes signal, and with $\alpha = \infty$ only points infinitely distant from the mean would be taken into the spikes. The results produced here were obtained with $\alpha = 1$ to showcase the most basic setup and the effectiveness of even a non-tuned version.

Once all levels have been processed in this manner, we use the previous information about ϕ to reconvert the two series from being characterised in terms of (ρ, ϕ) to the complex components of the DWT. After this step, we apply the Inverse DWT to $\vec{B}_t$ and $\vec{S}_t$, obtaining the two series that can be observed in Figure 5.

5.3 Series recombination and analysis of background

The results of the separation of the background can be seen in Fig. 4. Here it can be observed that the separated background is still characterised by seasonality dominating the weekly (10080 minutes) and daily (1440 minutes) timescales, and keeping a very similar structure in the regions occupied by the faster dynamics (< 120 minutes) to that of the original series. Simultaneously, the surges in power across timescales of the original DWT-transformed series

(left), which are associated with congestion events, have been greatly reduced or eliminated altogether in the second subplot.

This demonstrates the ability to process a time series showing multiple seasonalities and punctuated by large deviations from the normal dynamics of the process into two separate series, one of which will exclusively contain the baseline dynamics of the process, respecting its structure and natural variability, and the other, which will only contain those large events that occur outside of the smooth operational region.
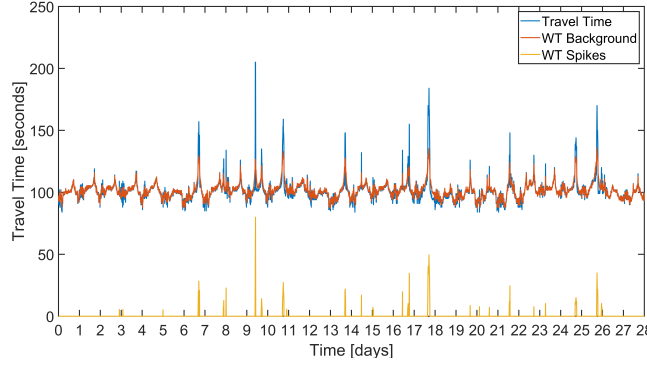

Figure 5: Recombination post Inverse Wavelet Transform

To reduce noise introduction during the inverse transform, a threshold is applied to $\vec{S}_t$, where elements representing spikes less than 3 seconds in amplitude are set to zero when looking at future estimation.

For future prediction steps, an indicator function is defined for every entry in the series, taking value:

$$\delta_i^{spike} = \begin{cases} 1, & \text{if } x_i > \text{Threshold} \\ 0, & \text{otherwise.} \end{cases} \tag{3}$$

6   WARP Travel Time Prediction Algorithm

The objective now is, given the seasonality and separation of the original signal provided above, to generate a time travel prediction model that accounts for the cyclic variations and recurrent congestion but remains resilient to unexpected deviations and rare events.

We aim to provide a robust algorithm that mitigates the propagation of extreme events into the future (unlike EWMA). It must work for all locations and not require the use of time segmentation. The trend term must be nearly flat, based on the difference in timescales for the growth of demand on a motorway level and the seasonalities concerning this paper. Finally, it must have uncorrelated residuals, Gaussian distributed with mean 0. Following these requirements, we introduce the WARP algorithm (Wavelet Augmented Regression Profiling) from section 6.1.

6.1   WARP: Spectral Component

The background signal shows oscillations of high frequency and low amplitude almost ubiquitously. It can be smoothed by discarding, in the frequency domain, the frequencies that in which the oscillations occur and those outside the scope of this study (over 4 weeks and under 4 hours), while keeping those in the information bearing bands by using the Fast Fourier Transform (FFT) [35]. Once this step is performed, and since large events have been removed previously, an EWMA can be applied to the modified weekly power spectra in the FFT, to then compute the Inverse FFT Transform to obtain our background prediction.

6.2   WARP: Seasonal Component

The seasonal component is computed via Seasonal-Trend Decomposition based on LOESS (STL) [33]. STL is resilient to outliers and can manage any combination of seasonalities, allowing to control their change over time as well as the smoothness of the trend [34]. We begin by isolating the daily seasonality S_d from the entire background training series using STL. Trend and remainder are summed and re-analysed for weekly seasonality S_w. This step also produces a trend series T_r and a remainder which should be Gaussian distributed, zero mean. Global seasonality is calculated as

6

$S_g = S_d + S_w$. We then average the seasonality over the training weeks to obtain a value for each minute of the week we are to estimate. Then, T_r, nearly flat as per 6, is linearised to obtain a baseline series B_t, and the Baseline prediction is $B_p = B_t + G_s$. Finally, the spikes are searched for any seasonality left on the weekly level Sp_w, discarding the trend and remainder terms, and obtaining the final seasonal component as $SEASONAL = B_t + Sp_w$.
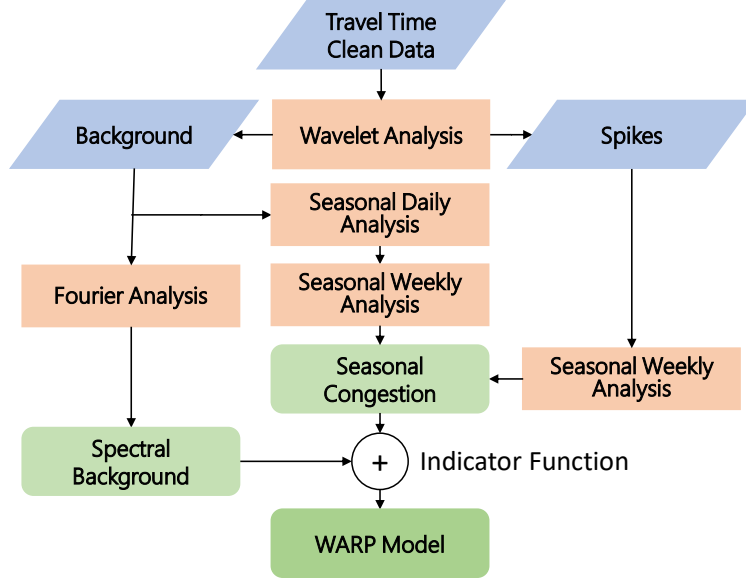
6.3 WARP Hybrid Profile

Figure 6: Flowchart of the data streams in the algorithm

The final WARP profile, containing the spectral and seasonal components, depends on the regime as per Eq. 3:

$$\text{WARP} = \text{Seasonal} * \delta_{\text{spike}} + \text{Spectral} * (1 - \delta_{\text{spike}}) \tag{4}$$

7 WARP model validation

In this section we assess the performance of the WARP model in out-of-sample validation and compare it against a null model based on simple segmentation and against the published NTIS profiles (the details of which are not in the public domain.)

7.1 Simple Segmentation - null model

Our null model for assessment of the performance of the WARP model is a basic segmentation model that applies uniform weights to the training data in a given time interval from previous weeks. On the $i - th$ minute of a week and using the previous n weeks (8 in this paper) as training, the simple segmentation (SS) profile is:

$$\hat{x}(i, n) = \sum_{\text{week}=1}^{n} \frac{x_n^i}{n} \tag{5}$$

Table 1: MARE Distribution in M6 prediction

Profile / MARE	$(> -25\%]$	$(-25\%, -15\%]$	$(-15\%, -5\%]$	$(-5\%, 5\%)$	$[5\%, 15\%)$	$[15\%, 25\%)$	$[> 25\%)$
Published M6	1.58	0.54	5.69	56.21	30.79	3.17	2.02
Wavelet M6	1.57	0.51	3.22	81.73	12.14	0.70	0.13

7.2 Out-of-sample validation

Our out-of-sample validation tests compare the WARP profile values against the subsequently measured travel time. We choose to use the Mean Absolute Relative Error (MARE) to quantify performance since it allows for a fair comparison of links of different lengths. The Root Mean Square Error (RMSE) has also been calculated on a motorway level as the average of the RMSE across its links.
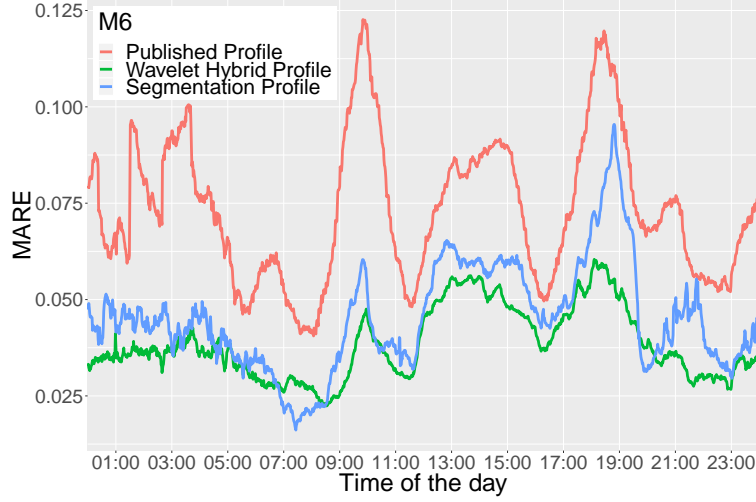

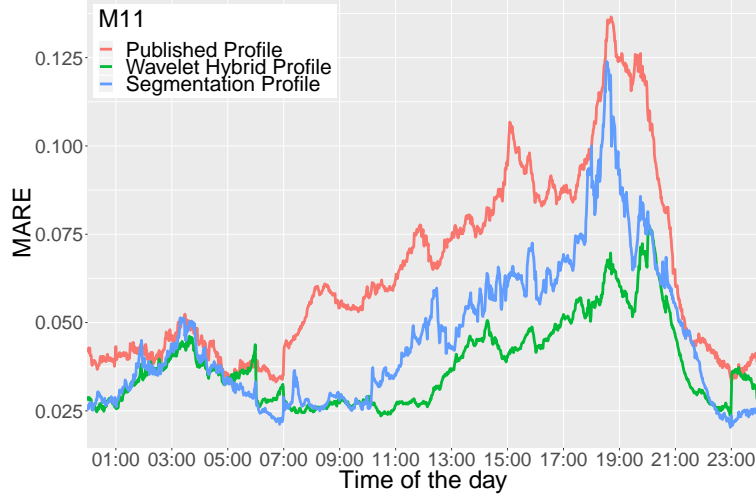Figure 7: MARE over the times of the day for the M6

Figure 8: MARE over the times of the day for the M11

Table 2: MARE Distribution in M11 prediction

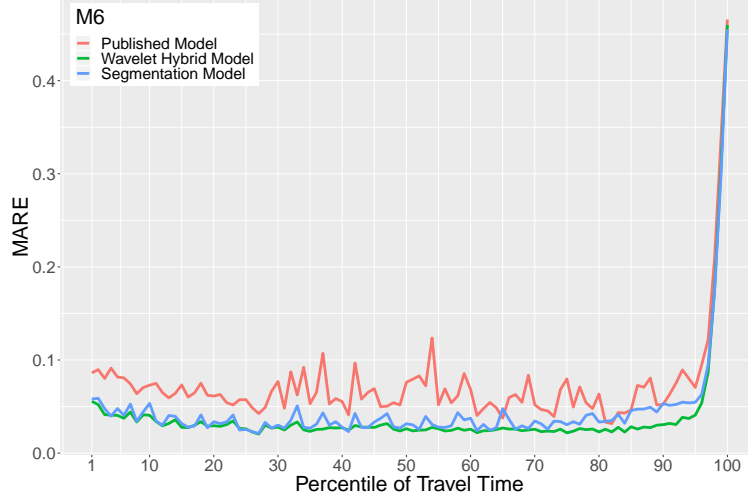| Profile / MARE	$(> -25\%]$	$(-25\%, -15\%]$	$(-15\%, -5\%]$	$(-5\%, 5\%)$	$[5\%, 15\%)$	$[15\%, 25\%)$	$[> 25\%)$
Published M11	0.85	1.15	19.21	62.83	13.21	1.65	1.10
Wavelet M11	0.78	0.32	3.29	81.33	12.47	1.09	0.73

8

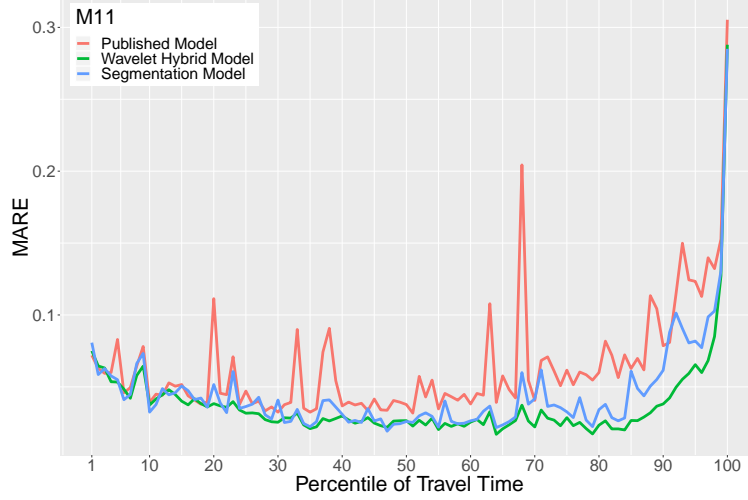Figure 9: MARE across percentiles of travel time for the M6.

Figure 10: MARE across percentiles of travel time for the M11.

In Figures 7 and 8, WARP shows lower predictive error than the published profiles and the SS Model for all times and motorways. This is most relevant for morning and evening rush hours, where the others' predictive error soars, WARP suffers no meaningful performance worsening relative to the plateau in the middle of the day. The error at rush hours is reduced by a minimum of 50% across all cases, reaching as high as 63.4% in the case of the M6 morning rush. In Figs. 9 and 10 it can be observed that the accuracy of the WARP profile is significantly higher than that of the published profiles or the SS Model across nearly all percentiles of travel time. WARP is most competitive in the upper percentiles of the travel time distribution since it explicitly accounts for the predictable contribution of recurrent congestion to travel time spikes. All three models eventually suffer similar errors under the most extreme deviations since these are true outlier and are not amenable to data-driven forecasting.

8   Conclusion and Future Work

This paper has presented an algorithm for separating recurrent and non-recurrent congestion in time travel time series, and generally capable of identifying events following different timescales from those of the baseline dynamic process, and being able to separate them accurately. More sophisticated analysis of the Wavelet Transform series could be achieved, focusing on the information-bearing bands and using adaptive thresholds for the separation. Given the flexibility offered by this method, it could be extended for incident detection or adapted to other problems besides road

Table 3: MARE per Link on M6 and M11

Links M11	MARE
199048301	0.0510
199048701	0.0279
199048801	0.1053
199048901	0.0239
199049002	0.0306
199049101	0.0213
199049402	0.0223
199049501	0.0181
199049702	0.0297
199049801	0.0408
199050002	0.0445
199050101	0.0272
199050202	0.0288
199050901	0.0433
199063301	0.1203
199063701	0.0500
199063801	0.0265
199064203	0.0230
199065202	0.0259
200021668	0.0280
200024801	0.0233
200028639	0.0241
200028641	0.0188
200028645	0.0435
200028648	0.0499

Links M6	MARE
117007401	0.0290
117007501	0.0680
117007601	0.0293
117007801	0.0826
117007901	0.0435
117008401	0.0484
117009102	0.0338
117011901	0.0295
117012001	0.0584
117012101	0.0370
117012201	0.0605
117012301	0.0379
117016001	0.0496
123025901	0.0427

Table 4: Global MARE & RMSE per Motorway

Motorway	MARE	RMSE [s]
M6	0.0385	0.0464
M11	0.0379	0.0484

traffic. The estimation algorithm presented in the last sections meets the requirements described in Section 6 and earlier defined in [2], albeit still using one parameter. In the future, sensitivity analysis could 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] Cabrejas-Egea, A. , de Ford, P., & Connaughton, C. (2018, November). Estimating Baseline Travel Times for the UK Strategic Road Network. In 2018 21st International Conference on Intelligent Transportation Systems (ITSC) (pp. 531-536). IEEE.

[3] 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.

[4] 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

[5] 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.

[6] 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.

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

[8] 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.

[9] 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.

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

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

[12] 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.

[13] 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.

[14] 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.

[15] 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.

[16] 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.

[17] 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.

[18] Li, Sheng. "Nonlinear combination of travel-time prediction model based on wavelet network." Proceedings. The IEEE 5th International Conference on Intelligent Transportation Systems. IEEE, 2002.

[19] Samant, A., and H. Adeli. "Feature extraction for traffic incident detection using wavelet transform and linear discriminant analysis." Computer-Aided Civil and Infrastructure Engineering 15.4 (2000): 241-250.

[20] Ghosh-Dastidar, Samanwoy, and Hojjat Adeli. "Wavelet-clustering-neural network model for freeway incident detection." Computer-Aided Civil and Infrastructure Engineering 18.5 (2003): 325-338.

[21] Jiang, Xiaomo, and Hojjat Adeli. "Dynamic wavelet neural network model for traffic flow forecasting." Journal of transportation engineering 131.10 (2005): 771-779.

[22] Tchrakian, Tigran T., Biswajit Basu, and Margaret O'Mahony. "Real-time traffic flow forecasting using spectral analysis." IEEE Transactions on Intelligent Transportation Systems 13.2 (2012): 519-526.

[23] Yang, Hang, et al. "A hybrid method for short-term freeway travel time prediction based on wavelet neural network and Markov chain." Canadian Journal of Civil Engineering 45.2 (2017): 77-86.

[24] Dharia, Abhijit, and Hojjat Adeli. "Neural network model for rapid forecasting of freeway link travel time." Engineering Applications of Artificial Intelligence 16.7-8 (2003): 607-613.

[25] 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.

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

[27] 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.

[28] 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.

[29] Grossmann, Alexander, and Jean Morlet. "Decomposition of Hardy functions into square integrable wavelets of constant shape." SIAM journal on mathematical analysis 15.4 (1984): 723-736.

[30] Daubechies, Ingrid. Ten lectures on wavelets. Vol. 61. Siam, 1992.

[31] Mallat, Stéphane. A wavelet tour of signal processing. Elsevier, 1999.

[32] Olhede, Sofia C., and Andrew T. Walden. "Generalized morse wavelets." IEEE Transactions on Signal Processing 50.11 (2002): 2661-2670.

[33] 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.

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

[35] 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.