An Early Look into Spectral Techniques for Travel Demand Modeling

Joseph J. Flood, Indianapolis Metropolitan Planning Organization
Contact Information
Telephone: (317)327-5646
Fax: (317)327-5950
Email: Joseph.Flood@indympo.org

Catherine Kostyn, Indianapolis Metropolitan Planning Organization
Contact Information
Telephone: (317)327-5142
Fax: (317)327-5950
Email: catherine.kostyn@indympo.org

Suzanne Childress, Puget Sound Regional Council
Contact Information
Telephone: (206)971-3282
Email: (206)587-4825
Email: schildress@psrc.org

Andrew Swenson, Indianapolis Metropolitan Planning Organization
Contact Information
Telephone: (317)327-5132
Fax: (317)327-5950
Email: andrew.swenson@indympo.org

Word Count: 4718
Number of Tables/Figures: 8
Abstract

Current practice in static traffic assignment typically involves dividing the day into multi-hour time periods. When this approach is used, one cannot determine anything about changes in conditions for sub-intervals within such time periods. Given that such models try to represent a typical weekday, the daily travel patterns they try to emulate can be assumed to be periodic. As a result, daily travel patterns can be represented by Fourier series, which decompose complicated periodic functions into the summation of simple waves. A procedure for obtaining such waves using static assignment is derived, and applied using the Indianapolis Regional Travel Demand Model. The results suggest potential in further developing such techniques for travel demand modeling, including the possibility of using spectral techniques to enhance Dynamic Traffic Assignment.
INTRODUCTION

Standard practice in static traffic assignment (STA) accounts for daily temporal variation in travel demand by dividing the day into discrete time periods, and assigning all trips within that time period to be made at the same time (1). A major limitation of this approach is that network attributes are assumed to be constant over the time period. Further, if a planner would like modeled data over time periods other than those that were specified in the model, such information is unavailable. Shorter time intervals can be modelled, but that would mean that more time and computational resources would be required to run the model.

Dynamic traffic assignment (DTA) is another assignment technique that is becoming used more and more in practice. DTA uses microsimulation to assign vehicles to various routes, and accounts for phenomena such as differing behavior between lanes and effects of queues on route choice. This allows for much higher temporal resolution in models, but transitioning from STA to DTA is a major investment (2), and finer resolutions necessitate more information to be stored in memory.

Because travel demand models explain daily behavior, such behavior can be assumed to be periodic with a period of one day in the eyes of the model. Because of this, network attributes, such as flow, can be represented as Fourier series (3):

\[ q(t) = \sum_{k=-\infty}^{\infty} Q[k] e^{\frac{2\pi kt}{D}} = q_{avg} + \sum_{k=1}^{\infty} 2\text{Re}(Q[k]) \cos\frac{2\pi kt}{D} + 2\text{Im}(Q[k]) \sin\frac{2\pi kt}{D} \]  

(1)

Where \(D\) represents the length of the day using the same units as \(t\). The values of the sequence \(Q\) can be estimated from observed data using the discrete Fourier transform (DFT) (4, 5). There are well-known algorithms that can compute the DFT very efficiently (5, 6). Whereas the variable \(q\) lives in the time domain, showing how flows change over time, the sequence \(Q\) resides within the
frequency domain, showing the contribution of waves of frequencies. The $k$-th element of $Q$
describes the contribution of the wave with a frequency of $k$ cycles per day (6).

A major benefit to performing static assignment in the frequency domain is that the results
are continuous in time, as a Fourier series is a discrete representation of continuous functions.
When performing assignment over discrete time periods, the results are constrained to their time
periods. For example, if a model assigns 2000 trips to a link during a three-hour period, one cannot
know the distribution of trips over 15-minute periods within that three-hour period. When
temporally continuous functions are outputs of the model, those functions can be integrated over
any time interval to get the number of trips during that time period.

Frequency-domain analysis of transportation data is not new. In 1969, a group of
researchers identified a strong contribution of waves with frequencies of one cycle every four
minutes ($\approx$4 mHz) in oscillatory traffic behavior (7). An analysis of stop-and-go traffic on freeways
in the Puget Sound region found contributions of waves at 2-3 mHz to be important, even when
analyzing data over different time windows (8). While both papers found significance in
oscillations on the order of one millihertz, lower frequencies would probably be more important.
The present study is primarily concerned with frequencies ranging from 10-130 $\mu$Hz. Due to the
fact that travel models only cover a period of one day, the lowest frequency that can be measured
is one cycle per day (11.57 $\mu$Hz). The other frequencies that can be measured are integer multiples
of 11.57 $\mu$Hz, which will be referred to further in this paper as circadian harmonics.

THEORETICAL BACKGROUND

Derivation of Continuous-Time User Equilibrium

In a famous transportation paper, Wardrop defined user equilibrium (UE) as satisfying the
following two conditions (9):
i. Travel times for all possible route choices will be equal.

ii. The average travel time is minimized.

By letting $d_j(q_j)$ representing the volume-delay function for link $j$, this would mean that when a traveler has a choice between $n$ routes, UE would be satisfied with the minimum travel time $T$ that satisfies the following equation:

$$T = d_1(q_1) = d_2(q_2) = \cdots = d_n(q_n)$$  \hspace{1cm} (2)

By the definition of a volume-delay function, the travel time $T$ on a link when the volume is equal to $q$ is given by the following equation:

$$T = d(q)$$ \hspace{1cm} (3)

When the volume changes from $q_1$ to $q_2$ halfway through the temporal duration of the trip, the equation for travel time then becomes:

$$T = \frac{1}{2} d(q_1) + \frac{1}{2} d(q_2)$$ \hspace{1cm} (4)

If the number of time segments in Equation 4 is increased to $n$, the equation then becomes:

$$T = \sum_{i=1}^{n} \frac{1}{n} d(q_i)$$ \hspace{1cm} (5)

It can be noted that Equation 5 is a Reimann sum, meaning that as $n$ approaches infinity, the equation becomes an integral (and the volume becomes a continuous function of time):

$$T = \int_{t_d}^{t_a} d(q(t)) dt$$ \hspace{1cm} (6)

Where $t_d$ and $t_a$ are the departure and arrival times of the trip, respectively. Because $t_a = t_d + T$, the travel time can be shown to solve the following equation:
\[ T = \int_{t_d}^{t_d+T} d(q(t))dt \] (7)

The travel times on \( n \) route choices would then be equal if the following equation is satisfied:

\[ \int_a^{a+T} d_1(q_1(t))dt = \int_a^{a+T} d_2(q_2(t))dt = \cdots = \int_a^{a+T} d_n(q_n(t))dt \] (8)

It can be shown that Equation 8 is satisfied for all possible trip start times and durations if and only if:

\[ d_1(q_1(t)) = d_2(q_2(t)) = \cdots = d_n(q_n(t)) \forall t, t_d, T \in \mathbb{R} \] (9)

Further, if the quantities in Equation 9 are minimized, then so are the quantities in Equation 8 (the travel times). Thus, if a continuous-time network were at UE, then Wardrop’s conditions need to be satisfied at all times.

A convenient property of the Fourier transform is that it is linear, which means that if one function is the summation of two other functions, then its Fourier transform with be the sum of the two other functions’ Fourier transforms. This is useful when analyzing travel behavior on a network. For example, say that the flow on Link 1 is represented by the equation \( q_1(t) \). Link 1 ends at an intersection, where drivers have the option of traveling on Link 2, Link 3, or Link 4, whose flows are represented by the equations \( q_2(t), q_3(t), \) and \( q_4(t), \) respectively. Then the following is true:

\[ q_1(t) = q_2(t) + q_3(t) + q_4(t) \] (10)

Let \( \hat{q}_i(\omega) \) represent the Fourier transform of \( q_i(t) \) for Link \( i \). The linearity of the Fourier transform then makes the following statement true:

\[ \hat{q}_1(\omega) = \hat{q}_2(\omega) + \hat{q}_3(\omega) + \hat{q}_4(\omega) \] (11)
Note that Equations 10 and 11 do not account for the time lag due to travel on Link 1. For every minute added of such time lag, it can be shown that harmonic \( k \) is shifted by \( \frac{k}{4} \) degrees. One potential topic of future research could be into whether or not such a time lag is significant and how to deal with it if it is.

**Spectral Travel Modeling with Static Assignment**

Due to the fact that Fourier series representing periodic and differentiable functions converge in the \( L^2 \)-norm, continuously changing network attributes can be represented with any desired accuracy by a finite number of Fourier coefficients (3). It is in fact possible to estimate such coefficients using outputs of STA. The procedure for doing so is fairly simple and straightforward, and does not require a large addition to common practice in modeling with STA. When approximating the network attributes using \( K \) circadian harmonics, the procedure is as follows:

1. Estimate the first \( K \) Fourier coefficients of the demand profile. This can be done in multiple ways.
2. Apply the inverse DFT to the shortened array to get the inputs to static assignment (quasi-trip tables). Some of these points may have negative values, so adjustment would be required to obtain reasonable results.
3. Perform static assignment at each of the \( 2K + 1 \) time points.
4. Apply the DFT to all of the network attributes to get the Fourier coefficients, and save them to memory.
5. The Fourier coefficients can then be used to reconstruct the network attributes as a function of time at any resolution.

**Error Issues**
When using a spectral approach, there are new sources of error that need to be considered. One new source of error is created by the fact that the result is a low-frequency approximation to the actual travel behavior. Fortunately, an upper bound on this error can be computed from observed data. From Plancherel’s theorem, it can be shown that (3):

\[ \|q(t)\|^2 = \int_D |q(t)|^2 \, dt = \sum_{k=-\infty}^{\infty} |Q[k]|^2 = \|Q\|^2 \quad (12) \]

Where \(D\) represents one day. Next, we can define two different functions, \(q_{mod}\) for modeled data and \(q_{obs}\) for observed data. A function describing error over time, \(q_e\) can be defined as \(q_{obs} - q_{mod}\). Plugging this into Equation 12 and taking advantage of the linearity of the Fourier transform results in the following equation for the square of the 2-norm of the error:

\[ \|q_e(t)\|^2 = \sum_{k=-\infty}^{\infty} |Q_e[k]|^2 = \sum_{k=-\infty}^{\infty} |Q_{obs}[k] - Q_{mod}[k]|^2 \quad (13) \]

When \(|k| > K\), then the corresponding value of \(Q_{mod}[k]\) is equal to zero, meaning that Equation 13 can be restated in the following way:

\[ \|q_e(t)\|^2 = \sum_{k=-\infty}^{-K-1} |Q_{obs}[k]|^2 + \sum_{k=-K}^{K} |Q_{obs}[k] - Q_{mod}[k]|^2 + \sum_{k=K+1}^{\infty} |Q_{obs}[k]|^2 \quad (14) \]

The fact that \(q_e(t)\) is real-valued for all possible times means that \(Q[k] = Q[\overline{-k}]\), allowing the equation to be further simplified:

\[ \|q_e(t)\|^2 = |Q_{obs}[0] - Q_{mod}[0]|^2 + 2 \sum_{k=1}^{K} |Q_{obs}[k] - Q_{mod}[k]|^2 + 2 \sum_{k=K+1}^{\infty} |Q_{obs}[k]|^2 \quad (15) \]

Such an expression can be used to compute the root-mean square error (RMSE) for the continuous function. Because Fourier series converge, the total error can be estimated by taking the limit as \(k\) approaches infinity. However, issues such as noise and travel survey time-reporting response bias...
can make convergence difficult. Erroneous contributions from high-frequency components may need to be filtered out before comparing observed data to modeled data. Constraints like excessive runtime may make increasing the number of modeled circadian harmonics unreasonable.

### Data Storage

Data representing an entire transportation network requires a lot of space, even when saving only Fourier coefficients. The amount of data can be greatly reduced by using a smaller data format, such as converting 64-bit floating point numbers to 16-bit floating point numbers or integers. While this reduces the storage space, there is loss of precision, resulting in more error. However, it is possible to put an upper bound on such an error.

The DFT of size $N$, as well as its inverse, can be expressed as a matrix-vector multiplication $(4, 5)$:

\[
\begin{pmatrix}
q(t_0) \\
q(t_1) \\
\vdots \\
q(t_N)
\end{pmatrix}
= F^{-1} Q 
\begin{pmatrix}
1 & 1 & 1 & \cdots & 1 \\
1 & \bar{w} & \bar{w}^2 & \cdots & \bar{w}^{N-1} \\
1 & \bar{w}^2 & \bar{w}^4 & \cdots & \bar{w}^{2(N-1)} \\
\vdots & \vdots & \vdots & \ddots & \vdots \\
1 & \bar{w}^{N-1} & \bar{w}^{2(N-1)} & \cdots & \bar{w}^{(N-1)^2}
\end{pmatrix}
\begin{pmatrix}
Q[0] \\
Q[1] \\
Q[2] \\
\vdots \\
Q[K] \\
0 \\
0 \\
\vdots \\
Q[K] \\
0 \\
\bar{w} = e^{-\frac{2\pi i}{N}}
\end{pmatrix}, \quad w = e^{-\frac{2\pi i}{N}} \quad (16)
\]

Due to the fact that only $K$ circadian harmonics are assumed to contribute to the travel behavior, $Q[k] = 0$ for $|k| > K$. This means that equation 16 is equivalent to:
\[
\begin{pmatrix}
q(t_0) \\
q(t_1) \\
\vdots \\
q(t_N)
\end{pmatrix}
= F_{K-1} Q_K
\]

\[
= \begin{pmatrix}
1 & 1 & \cdots & 1 & 1 & \cdots & 1 \\
1 & \tilde{w} & \cdots & \tilde{w}^K & \tilde{w}^{N-K} & \cdots & \tilde{w}^{N-1} \\
\vdots & \vdots & \ddots & \vdots & \vdots & \ddots & \vdots \\
1 & \tilde{w}^{N-1} & \cdots & \tilde{w}^{K(N-1)} & \tilde{w}^{(K-1)(N-1)} & \cdots & \tilde{w}^{(N-1)^2}
\end{pmatrix}
\begin{pmatrix}
Q[0] \\
Q[1] \\
\vdots \\
Q[K] \\
\vdots \\
Q[1]
\end{pmatrix}
\tag{17}
\]

The absolute condition number $\hat{\kappa}$ of a matrix-vector multiplication is the norm of the matrix \((10)\). When looking for the maximum possible change, the infinity norm, which is the maximum absolute value row sum of the matrix, is used. If normalization is performed on the DFT instead of the inverse, then the absolute value of every term in $W_c$ is equal to one, as each term is a root of unity, meaning that all row sums (and thus the maximum) are equal to $2K + 1$.

If one were to round all of the Fourier coefficients to the nearest integer, then the maximum change in absolute value of one of them would be equal to $\frac{1}{\sqrt{2}}$. After the inverse (IDFT) is applied (depending on the normalization scheme), the maximum possible change in the network attributes would be $\frac{2K+1}{\sqrt{2}}$. If replacing 64-bit floating point numbers with 16-bit floating point numbers, this number becomes $(2K + 1)2^{\text{ceil}(\log_2(Q_{\text{max}})) - 10}$.

**DEMONSTRATION**

**Model**

The Indianapolis Travel Demand Model (ITDM) was used to test out a spectral approach and compare it with a non-spectral approach. The current version of the ITDM is a trip-based 4-step model using TransCAD 7.0. Trip generation is performed with cross-classification, and trip distribution was recently updated as a destination choice model. Time-of-day (TOD) modeling
involves placing a percentage of trips within each time period. This is done separately for auto and transit trips, with 5 auto time periods and 4 transit time periods. The auto time periods are AM (6-9AM), midday (9AM-3PM), PM (3-6PM), evening (6-9PM), and night (9PM-6AM). TOD factors are calculated by trip purpose and direction (production to attraction vs attraction to production).

**Conversion to Spectral Model**

In this study, conversion to a spectral model involved only changing the TOD factors. More steps may be needed to increase the validity of such a model, which will be discussed further later in the paper. Whereas the non-spectral TOD factors all added up to one over the five time periods, the spectral TOD factors represent five points on a continuous function that integrates to one over the course of a day, so they should not be expected to add up to one.

Calculation of the spectral TOD factors first involved approximating the temporal distributions of trips by purpose and direction. This was accomplished by creating continuous temporal distributions using circular moments. The equation for the \( n \)-th circular moment \( m_n \) of a probability distribution on a circular domain \( p(\theta) \) is (11):

\[
m_n = \int_{\Gamma} p(\theta)e^{in\theta} d\theta
\]  

(18)

Where \( \Gamma \) represents any interval of length \( 2\pi \). Because \( p(\theta) \) is periodic, it can be represented by a Fourier series. This leads to the following derivation:

\[
m_n = \int_{\Gamma} \left( \sum_{k=-\infty}^{\infty} P[k]e^{ik\theta} \right) e^{in\theta} d\theta = \sum_{k=-\infty}^{\infty} P[k] \int_{\Gamma} e^{ik\theta} e^{in\theta} d\theta
\]  

(19)

Because complex exponentials are orthogonal functions on periodic domains, the following is true:

\[
\int_{\Gamma} e^{ik\theta} e^{in\theta} d\theta = \begin{cases} 
2\pi & \text{if } k = -n \\
0 & \text{if } k \neq -n
\end{cases}
\]  

(20)
Plugging equation 20 into equation 19 results in the following expression:

\[ m_n = 2\pi P[-n] = 2\pi \overline{P[n]} \iff P[n] = \frac{m_n}{2\pi} \]  \hspace{1cm} (21)

In order to estimate the circular moments, the trip start times were converted from number of hours since the day started to the angle that the earth rotated since the day started \((\theta_n)\) by multiplying the time by \(\frac{\pi}{12}\). The following formula was then used to approximate the circular moments (the summations were over all auto trips of matching purposes and directions):

\[ m_n \approx \frac{\sum w_n e^{i\theta_n}}{\sum w_n} \]  \hspace{1cm} (22)

Where \(w_n\) represents trip \(n\)’s expansion weight. Only two circular moments were calculated, as the ITDM uses five time periods, and thus only two harmonics can be modeled. The moments were then converted to Fourier coefficients using Equation 21. They were then arranged into the following vector: \([P[0], P[1], P[2], \overline{P[2]}, \overline{P[1]}]\), of which the IDFT was applied to in order to get the spectral TOD factors, which represented values of the temporal trip distribution at five equidistant points throughout the day (12:00 AM, 4:48 AM, 9:36 AM, 2:24 PM, 7:12 PM).

An issue was discovered at this point that was caused by the use of low-frequency approximations to the temporal trip distributions: some of the resultant TOD factors were negative. For this study, the negative TOD factors were set to zero, and the others were adjusted so they still represented a function that integrates to one over the course of a day. It remains to be seen if this is the best way to deal with this issue.

The model run produced flows at five time points throughout the day. These flows were converted to Fourier coefficients (and then circular moments) and stored in memory using a Python package for spectral and circular statistical analysis of transportation data that is currently in development.
Results

Spectral model results, as well as the results of a base-year model run, were compared with 2-day traffic counts conducted by the Indiana Department of Transportation (INDOT). The counts were conducted as early as 2013 and as late as 2016. Results were compared in both directions in 13 locations, as shown in Figure 1. Table 1 contains detailed information about the locations and traffic counts. Four locations were in the interior of Indianapolis, and the other nine were outside of the city. This was done to reduce the effect of results being correlated. The most recent 48-hour counts between 12:00 AM Monday and 5:00 PM Friday were used.

FIGURE 1 Analysis Locations (Imagery from Open Street Map)
TABLE 1 Analysis Locations

<table>
<thead>
<tr>
<th>Location</th>
<th>Highway</th>
<th>Bounds</th>
<th>City/Town</th>
<th>Inner/Outer</th>
<th>Start Time</th>
<th>End Time</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>I-70</td>
<td>Harding St to West St</td>
<td>Indianapolis</td>
<td>Inner</td>
<td>6/17/2014 2:00 PM</td>
<td>6/19/2014 2:00 PM</td>
</tr>
<tr>
<td>2</td>
<td>I-65</td>
<td>Raymond St to I-70</td>
<td>Indianapolis</td>
<td>Inner</td>
<td>6/17/2014 2:00 PM</td>
<td>6/19/2014 2:00 PM</td>
</tr>
<tr>
<td>3</td>
<td>I-70</td>
<td>Keystone Ave to I-65</td>
<td>Indianapolis</td>
<td>Inner</td>
<td>6/13/2016 11:00 AM</td>
<td>6/15/2016 11:00 AM</td>
</tr>
<tr>
<td>4</td>
<td>I-65</td>
<td>38th St to MLK Drive</td>
<td>Indianapolis</td>
<td>Inner</td>
<td>6/6/2016 2:00 PM</td>
<td>6/8/2016 2:00 PM</td>
</tr>
<tr>
<td>5</td>
<td>I-70</td>
<td>SR 39 to SR 267</td>
<td>Plainfield</td>
<td>Outer</td>
<td>6/24/2013 12:00 PM</td>
<td>6/26/2013 12:00 PM</td>
</tr>
<tr>
<td>6</td>
<td>SR 37</td>
<td>Banta Rd to SR 144</td>
<td>Waverly</td>
<td>Outer</td>
<td>8/26/2014 8:00 AM</td>
<td>8/28/2014 8:00 AM</td>
</tr>
<tr>
<td>7</td>
<td>I-65</td>
<td>Main St to Co Line Rd</td>
<td>Greenwood</td>
<td>Outer</td>
<td>6/17/2014 12:00 PM</td>
<td>6/19/2014 12:00 PM</td>
</tr>
<tr>
<td>8</td>
<td>I-74</td>
<td>Pleasant View Rd to Acton Rd</td>
<td>Indianapolis</td>
<td>Outer</td>
<td>4/12/2016 12:00 PM</td>
<td>4/14/2016 12:00 PM</td>
</tr>
<tr>
<td>9</td>
<td>I-70</td>
<td>SR 9 to Mt Comfort Rd</td>
<td>Mount Comfort</td>
<td>Outer</td>
<td>7/8/2013 1:00 PM</td>
<td>7/10/2013 1:00 PM</td>
</tr>
<tr>
<td>10</td>
<td>I-69</td>
<td>SR 13 to Campus Pkwy</td>
<td>Noblesville</td>
<td>Outer</td>
<td>3/16/2015 12:00 PM</td>
<td>3/18/2015 12:00 PM</td>
</tr>
<tr>
<td>11</td>
<td>US 31</td>
<td>191st St to SR 38</td>
<td>Westfield</td>
<td>Outer</td>
<td>9/17/2013 6:00 PM</td>
<td>9/19/2013 6:00 PM</td>
</tr>
<tr>
<td>12</td>
<td>I-65</td>
<td>SR 267 to Whitestown Pkwy</td>
<td>Whitestown</td>
<td>Outer</td>
<td>7/6/2016 12:00 AM</td>
<td>7/8/2016 12:00 AM</td>
</tr>
<tr>
<td>13</td>
<td>I-74</td>
<td>Jeff Gordon Blvd to SR 267</td>
<td>Brownsburg</td>
<td>Outer</td>
<td>4/19/2016 12:00 PM</td>
<td>4/21/2016 12:00 PM</td>
</tr>
</tbody>
</table>

Analysis on the observed traffic counts suggest that the first two circadian harmonics contribute to 50-70% of the daily variation in flow at the analysis locations, suggesting that more harmonics may need to be modelled to accurately capture the daily variation in flow. All of the variation was captured with 11 harmonics (expected due to there being analysis at each hour in the day). However, a significant amount of the contributions of higher frequencies is likely noise. Traffic counts over more than 2 days may be needed to reduce such noise and better gauge how many circadian harmonics to model.
The results were compared in multiple ways. First, the percent root mean square error (%RMSE) was calculated for the results spectral and nonspectral runs for volumes in each of the ITDM’s five time periods. The volumes in the spectral run were computed by integrating the resultant continuous functions over these time periods using the following equation:

\[ q_{t_e,t_s} = \int_{t_s}^{t_e} q(t) dt \]

\[ = \int_{t_s}^{t_e} \left( Q[0] + Q[1]e^{\frac{\pi}{12}it} + Q[2]e^{\frac{\pi}{6}it} + Q[2]e^{-\frac{\pi}{6}it} + Q[1]e^{-\frac{\pi}{12}it} \right) dt \]

\[ = (t_e - t_s)Q[0] + \frac{12}{\pi}i \left( e^{\frac{\pi}{12}it_e} - e^{\frac{\pi}{12}it_s} \right) Q[1] + \frac{6}{\pi}i \left( e^{\frac{\pi}{6}it_e} - e^{\frac{\pi}{6}it_s} \right) Q[2] \]

\[ - \frac{6}{\pi}i \left( e^{-\frac{\pi}{6}it_e} - e^{-\frac{\pi}{6}it_s} \right) Q[2] - \frac{12}{\pi}i \left( e^{-\frac{\pi}{12}it_e} - e^{-\frac{\pi}{12}it_s} \right) Q[1] \]

Where \( t_s \) and \( t_e \) represent the start and end times of the time period, respectively. These results are shown in Table 2. The %RMSE was then calculated for hourly counts as well, assuming each hour within a time period has the same volume. These are shown in Table 3. Additionally, continuous %RMSE values were computed for the spectral data comparing it to 2-frequency and 11-frequency approximations of the observed traffic counts using Equation 15. These results are shown in Table 4. Plots comparing the results of different modeling techniques on westbound Interstate 70 at the White River with observed counts for the time periods and each hour of the day are shown in Figures 2 and 3.
FIGURE 2 Number of vehicles crossing the White River on westbound Interstate 70 as computed by both modeling techniques as well as observed counts.

FIGURE 3 Hourly results of both modeling techniques compared with observed counts along with a 2-harmonic approximation of the observed counts.
### TABLE 2 %RMSE for Spectral and Nonspectral Runs Based on Five Time Periods

<table>
<thead>
<tr>
<th>Location</th>
<th>Inbound Spectral</th>
<th>Nonspectral</th>
<th>Improvement</th>
<th>Outbound Spectral</th>
<th>Nonspectral</th>
<th>Improvement</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>48.7%</td>
<td>40.6%</td>
<td>-8.0</td>
<td>33.4%</td>
<td>19.2%</td>
<td>-14.2</td>
</tr>
<tr>
<td>2</td>
<td>45.9%</td>
<td>33.8%</td>
<td>-12.1</td>
<td>59.5%</td>
<td>41.5%</td>
<td>-18.0</td>
</tr>
<tr>
<td>3</td>
<td>32.8%</td>
<td>15.4%</td>
<td>-17.4</td>
<td>28.1%</td>
<td>20.6%</td>
<td>-7.5</td>
</tr>
<tr>
<td>4</td>
<td>37.1%</td>
<td>33.2%</td>
<td>-3.9</td>
<td>28.7%</td>
<td>22.3%</td>
<td>-6.4</td>
</tr>
<tr>
<td>5</td>
<td>29.8%</td>
<td>23.1%</td>
<td>-6.7</td>
<td>35.2%</td>
<td>31.1%</td>
<td>-4.0</td>
</tr>
<tr>
<td>6</td>
<td>59.7%</td>
<td>40.2%</td>
<td>-19.5</td>
<td>67.5%</td>
<td>64.6%</td>
<td>-2.9</td>
</tr>
<tr>
<td>7</td>
<td>31.4%</td>
<td>22.5%</td>
<td>-9.0</td>
<td>46.5%</td>
<td>36.0%</td>
<td>-10.5</td>
</tr>
<tr>
<td>8</td>
<td>30.9%</td>
<td>30.0%</td>
<td>-0.9</td>
<td>28.7%</td>
<td>35.3%</td>
<td>6.7</td>
</tr>
<tr>
<td>9</td>
<td>31.4%</td>
<td>32.9%</td>
<td>1.5</td>
<td>25.5%</td>
<td>30.2%</td>
<td>4.7</td>
</tr>
<tr>
<td>10</td>
<td>30.3%</td>
<td>24.7%</td>
<td>-5.6</td>
<td>35.7%</td>
<td>32.8%</td>
<td>-2.9</td>
</tr>
<tr>
<td>11</td>
<td>40.6%</td>
<td>47.1%</td>
<td>6.5</td>
<td>36.5%</td>
<td>48.7%</td>
<td>12.3</td>
</tr>
<tr>
<td>12</td>
<td>39.9%</td>
<td>47.4%</td>
<td>7.4</td>
<td>35.1%</td>
<td>45.2%</td>
<td>10.1</td>
</tr>
<tr>
<td>13</td>
<td>32.4%</td>
<td>22.1%</td>
<td>-10.3</td>
<td>34.1%</td>
<td>28.8%</td>
<td>-5.2</td>
</tr>
</tbody>
</table>

The volumes from the spectral run were more accurate than the nonspectral run in 3 of the 13 inbound locations and 4 of the outbound locations. None of the locations in which the spectral run was more accurate were inner locations, and in general, the differences were much greater for the inner locations. This could be due to differences in characteristics between urban and rural travel behavior patterns.

One possibility is that the larger errors are caused by limiting the results to be a two-frequency approximation to the data. To test this, the improvement in the %RMSE from the nonspectral run to the spectral run was compared with the percentage that the third circadian harmonic of the observed counts contributes to the behavior (when compared to the first eleven). To compute the latter statistic, the DFT was applied to the observed hourly counts at each location to approximate the Fourier coefficients associated with the observed counts. The absolute value of the coefficient associated with the third harmonic was divided by the sum of the first eleven. If higher third-harmonic contributions are associated with lower improvement when adopting a spectral approach, there should be a negative correlation with the two variables.
The analysis provided marginally significant support for this hypothesis (Spearman’s $\rho = -0.378, p = 0.057$). A scatterplot of the two variables (shown in Figure 4) suggests a possible linear relationship between the two variables, which was significant ($r = -0.471, p < 0.05$). If this relationship is indeed true, then it would be of great use to practitioners considering conversion to a spectral model.

![Image of scatterplot](image)

**FIGURE 4** Relationship between contribution of third harmonic and improvement in %RMSE. The negative relationship suggests that the lack of improvement observed at most of the locations could be due to modelling an insufficient number of harmonics.

When looking at hourly counts, improvements were observed in 3 of the 13 inbound and outbound locations. Nonspectral hourly counts were estimated by dividing the number of vehicles assigned to the location in each time period by the number of hours in the time period.

**TABLE 3 %RMSE for Spectral and Nonspectral Runs Based on Hourly Traffic Counts**

<table>
<thead>
<tr>
<th>Location</th>
<th>Inbound Spectral</th>
<th>Nonspectral</th>
<th>Difference</th>
<th>Outbound Spectral</th>
<th>Nonspectral</th>
<th>Difference</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>51.9%</td>
<td>46.9%</td>
<td>-5.0</td>
<td>36.9%</td>
<td>23.8%</td>
<td>-13.1</td>
</tr>
<tr>
<td>2</td>
<td>46.8%</td>
<td>38.8%</td>
<td>-8.0</td>
<td>60.6%</td>
<td>43.5%</td>
<td>-17.1</td>
</tr>
<tr>
<td>3</td>
<td>36.1%</td>
<td>23.0%</td>
<td>-13.1</td>
<td>32.3%</td>
<td>26.2%</td>
<td>-6.1</td>
</tr>
</tbody>
</table>
Equation 15 discussed a way to measure the RMSE using Fourier coefficients that reflects what the RMSE would be for continuous data. Table 4 shows values of this metric comparing the spectral results with 2-harmonic and 11-harmonic approximations to the observed traffic counts. It can be observed that the two values are close together, reflecting most of the variation being captured within the first two circadian harmonics.

<table>
<thead>
<tr>
<th>Location</th>
<th>Inbound 2-Harmonic</th>
<th>11-Harmonic</th>
<th>Outbound 2-Harmonic</th>
<th>11-Harmonic</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>43.0%</td>
<td>44.1%</td>
<td>43.7%</td>
<td>44.5%</td>
</tr>
<tr>
<td>2</td>
<td>44.6%</td>
<td>45.6%</td>
<td>46.9%</td>
<td>47.9%</td>
</tr>
<tr>
<td>3</td>
<td>45.3%</td>
<td>46.6%</td>
<td>45.0%</td>
<td>45.7%</td>
</tr>
<tr>
<td>4</td>
<td>45.7%</td>
<td>47.0%</td>
<td>46.1%</td>
<td>47.3%</td>
</tr>
<tr>
<td>5</td>
<td>43.1%</td>
<td>43.8%</td>
<td>44.8%</td>
<td>45.2%</td>
</tr>
<tr>
<td>6</td>
<td>46.0%</td>
<td>47.2%</td>
<td>46.7%</td>
<td>47.5%</td>
</tr>
<tr>
<td>7</td>
<td>44.3%</td>
<td>45.2%</td>
<td>43.6%</td>
<td>44.1%</td>
</tr>
<tr>
<td>8</td>
<td>43.9%</td>
<td>44.9%</td>
<td>44.8%</td>
<td>45.3%</td>
</tr>
<tr>
<td>9</td>
<td>43.3%</td>
<td>43.7%</td>
<td>43.6%</td>
<td>43.8%</td>
</tr>
<tr>
<td>10</td>
<td>44.7%</td>
<td>46.0%</td>
<td>45.0%</td>
<td>45.6%</td>
</tr>
<tr>
<td>11</td>
<td>46.8%</td>
<td>47.6%</td>
<td>46.1%</td>
<td>47.7%</td>
</tr>
<tr>
<td>12</td>
<td>43.5%</td>
<td>43.9%</td>
<td>44.5%</td>
<td>44.9%</td>
</tr>
<tr>
<td>13</td>
<td>44.7%</td>
<td>45.9%</td>
<td>46.4%</td>
<td>47.1%</td>
</tr>
</tbody>
</table>

In this study, the only thing that was changed to make the model spectral was adjusting the TOD factors. It is entirely possible that after additional calibration, the errors associated with the spectral run could decrease significantly.
Effects on Other Parts of the Model

Due to the fact that the ITDM repeats all of its steps a few times, the adjusted TOD factors influenced other parts of the model. For example, the number of transit trips increased from 20,758 to 21,340, a 2.7% increase, even though the transit TOD factors were not adjusted. This could be due to increased congestion that was created during the peak periods.

FUTURE RESEARCH

If spectral travel modeling techniques were to further develop, there are many areas of potential research on the topic. One potential topic of investigation would be to investigate if spectral techniques could be used to model travel behavior patterns that exhibit discrete changes, such as transit or reversible lanes. A lot of harmonics are required to accurately represent sudden changes in functions, so that could lead to accuracy issues when modeling such elements. Additionally, it might be possible to extract Fourier coefficients from modeled flows using Equation 23 without changing anything about how existing models work, meaning that the only part of the model that would change would be additional post-processing.

Thus far this paper has only looked at flows. Other network attributes exist. It is well established that speed and flow have a nonlinear relationship (12). This would mean that speed might require more Fourier coefficients to model than flow, which can be shown by the convolution theorem. It is extremely important to understand the relationship between speed and flow in the frequency domain. It is possible that the Fourier coefficients of flow are such that higher-frequency coefficients of speed would converge to zero when repeated convolution occurs, so this would need to be looked into.

Recent advancements in computer technology are making DTA practical. In fact, DTA has been run at resolutions as low as five minutes (2). However, the computational effort involved in
simulating traffic behavior at such a high temporal resolution may not be necessary. If changes in the transportation network are not very rapid, it might only be necessary to run DTA at a much lower resolution over the course of a day, estimate the Fourier coefficients, and interpolate to any desired resolution. If that were to be done, then one could be able to model phenomena such as queue buildup and discharge and lane-by-lane flows using a spectral approach, which is beyond the capability of the techniques outlined in this paper (2). However, a lot of research needs to be accomplished to identify practicality and potential caveats resulting in the use of such an approach.

ACKNOWLEDGEMENTS

The authors would like to thank Jim Lam from Caliper Corporation for providing technical support on running the model for this study.

REFERENCES


