2016 IEEE 19th International Conference on Intelligent Transportation Systems (ITSC)
Windsor Oceanico Hotel, Rio de Janeiro, Brazil, November 1-4, 2016

Synthesizing Route Travel Time Distributions Considering Spatial Dependencies

Isaac K Isukapati
The Robotics Institute
Carnegie Mellon University
Pittsburgh, PA 15213
Email: isaack@cs.cmu.edu

George F List
Department of Civil Engineering
North Carolina State University
Raleigh, NC 27695
Email: gflist@ncsu.edu

***Abstract*—Estimation of route-level travel time distributions (or travel rates) from segment-level data is of great interest today. This paper shows how the random variable properties of comonotonicity and independence can be used in combination to develop such distributions for a wide range of operating conditions. Efficacy of the technique is illustrated using Bluetooth-based travel times collected from individual vehicles on I-5 in Sacramento. The technique offers a way that agencies and other service providers can provide credible travel time (travel rate) distributions to route guidance devices, freight dispatchers and other users who want to know about the reliability of travel time estimates for the routes they utilize.**

I. Introduction

The estimation of route-level travel time distributions (or travel rates) based on segment-level data is a topic of great interest today. Traditionally, travel times have been estimated and reported as historic averages. While this information is useful to understand general trends in travel times, cumulative density functions (CDFs) provide richer information on travel time reliability. Even though developing travel time CDFs or PDFs at a segment level is a fairly straightforward task, constructing defensible route travel-time distributions using segment level observations can be complicated by strong correlations between travel times on adjacent segments.

This paper shows how proportional sampling, comonotonicity and independence can be used to synthesize route-level distributions for a wide range of operating conditions. Comonotonicity is a condition where random variables are quantile-additive while independence is where they are uncorrelated. The efficacy of the technique is illustrated using individual vehicle travel times collected on I-5 in Sacramento by Bluetooth readers. The technique offers a way for service providers to generate credible travel time (travel rate) distributions for route guidance devices, freight dispatchers and other users who are concerned about understanding the reliability of travel times (travel rate) estimates for the routes they utilize.

To illustrate the basic idea, assume travel times t_{ab} and t_{bc} are for adjacent segments ab and bc. Next, let CDF_1 be the distribution of travel times for route ac formed by quantile addition, that is, the n^{th} percentile value of CDF_1 is identified by adding the nth percentile value for tab to the n^{th} percentile value for t_{bc}. Then, let CDF_2 be the random variable formed by adding t_{ab} to t_{bc} assuming t_{ab} and t_{bc} are independent. Finally, let

$$z = (1-\alpha) * CDF_1 + \alpha * CDF_2. \quad (1)$$

This paper demonstrates that carefully chosen values of α make it possible to synthesize a very close approximation to the true distribution of z from the distributions of t_{ab} and t_{bc}.

II. Prior Research

A limited amount of research has focused on the estimation of route-level travel time distributions based on various types of data. Jintanakul et al, [1] experimented with using Bayesian mixture models to estimate freeway travel time distributions based on small samples of probe vehicles on multiple days. Guo et al, [2] experimented with using a multi-state (multi-mode) model to portray the distributions of travel times being observed. Susilawati et al, [3] have experimented with the use of multi-modal Burr density functions. This work shows promise and is intriguing because the analytical functions are simple and the density function allows for skew.

Recent efforts continue to advance these lines of investigation. Mahmassani et al, [4] explored ideas about robust relationships for reliability analysis. Barkley et al, [5], as a result of involvement in SHRP-2 L02 explored the use of multi-state models to develop relationships between travel time reliability and non-recurrent events. Ernst et al, [6] developed sampling guidelines for using probe vehicles to characterize arterial travel times. Ramezani et al, [7] explored the use of Markov chains to estimate arterial route travel time distributions; Guo et al, [8] extended their investigations of using multi-state models to represent travel time distributions by exploring the use of mixed skewed models, and Feng et al, [9] explored the use of Bayesian models to construct arterial travel time distributions based on data from GPS-equipped probe vehicles.

III. Field Data

It is customary in many papers to present the proposed methodology at this point, but in this case, examining the field data that provided the inspiration seems like a more effective way to motivate the idea. Therefore, field data material will be presented next; and the methodology will be presented in Section 4. For approximately three months (January to April) in 2011, CalTrans collected Bluetooth data at four monitoring stations on I-5 in Sacramento, CA. Figure 1 shows where the sensors were located. Data were collected both in the northbound and southbound directions. The stations (going southbound) are 39, 9, 10, and 11. These numbers can be treated as labels. The important information is the locations of the stations and the three freeway segments they create: 39-9 (segment-1), 9-10 (segment-2), and 10-11 (segment-3). The lengths of the segments are 3.6 miles, 1.1 miles, and 0.9 miles, respectively.

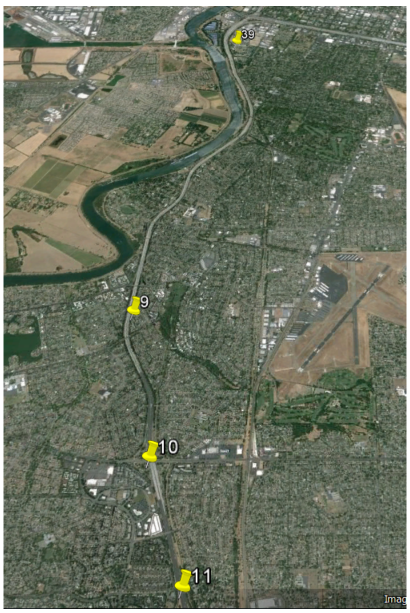

Fig. 1. I-5 in Sacramento

A. Obtaining Travel Times from Bluetooth Data

Before the issue of synthesizing route-level travel time distributions is considered, distributions of the segment-level travel times have to be developed. Several intermediate steps are involved in doing this. Because other researchers may want to test the synthesis idea being presented, describing these intermediate steps is useful.

As the reader probably recognizes, the Bluetooth data are not direct observations of travel times; rather they are records of vehicles passing reader locations at specific times. In this case, each record includes a MAC ID, a time stamp, and a location, plus several other data items such as signal strength. The hope is that if vehicle X was observed passing reader A at time t_a and then reader B at time t_b, it was in the process of making a trip. If this is true, and it traveled from A to B directly, then time t_{ab} is a travel time. But this might not be the case if vehicle X stopped somewhere in-between. Then it could represent a trip time (from A to B) but not a travel time. On the other hand, if t_{ab} is very large, it might simply reflect the fact that the next sensor location that vehicle X passed after reader A was reader B, as in the case of being parked somewhere overnight in-between. Unlike observations from an AVL system, it is not possible to tell what vehicle X did in-between the reads.

In order to develop travel time observations from the raw data, preliminary data processing is required. To address that need, the authors developed a three-step process.

The first step sorts observations into sequence by MAC ID and timestamp. The second creates matched pairs from sequential observations. The third filters paired observations to identify the travel times. Matching is done sequentially with or without an upper bound on the interval between time stamps. The filtering step extracts paired observations that are likely to represent travel times (as opposed to trip times that involve intermediate stops). Truncation is not employed because very short travel times (e.g., from low-flying aircraft) would not be filtered out; and legitimate long travel times, (e.g., due to incidents), might be missed.

Instead, for the filtering step (step 3), the authors developed a special algorithm. The authors have demonstrated [10] that this technique is quite effective in identifying the evolving pattern of travel times under congested and incident conditions while not missing significant transients.

B. Correlation Between Adjacent Segments

A basic premise of the synthesis methodology is that specific values of α can be employed to combine CDF_1 and CDF_2 depending on the operating conditions present on the two segments. (See the introductory paragraphs.) Binning the data based on operating conditions is the next step.

In the text that follows, we refer to operating condition bins as regimes. Each regime is a combination of two segment-level attributes: 1) the congestion level (demand to capacity ratio) that was extant at the time of the observation (absent the influence of any non-recurring event) and 2) extant non-recurring event (including none).

When travel times are binned in this manner, two patterns become apparent. They help inform the determination of α.

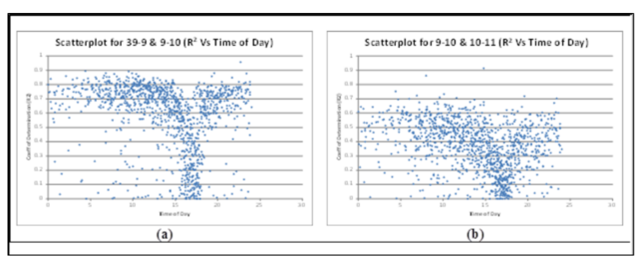

Fig. 2. Scatterplot of R^2 Vs. Time of day

Patterns in travel time variability for short and long periods of time have an implication for how travel time distributions are correlated on adjacent segments (of course, by direction). Since in general similar traffic loadings produce similar travel times for the same capacity and operating conditions, and many of the same drivers are involved, the travel rate distributions are also similar (and also the travel times). As the remainder of the paper demonstrates, this similarity in behavior is most apparent when the facilities are uncongested. It is less observable when the facilities are congested or a non-recurring event is affecting capacity.

To gain more insight into the correlations between travel time distributions on upstream and downstream segments, the individual vehicle travel time data from I-5 can be examined using the following technique. Let:

$$\tau_{j+1}^{i} = \theta_0 * \tau_j^i + \theta_1; \quad (2)$$

where τ_{j+1}^{i}, τ_j^i are downstream and upstream travel times of vehicle 'i', and θ_0, and θ_1 are regression parameters.

Two analyses can be conducted based on the equation above and the time stamps for stations 39, 9, 10, and 11. The first is to predict segment 9-10 travel times using segment 39-9 data and the second is to predict segment 10-11 travel times based on segment 9-10 data. The main point is to see what coefficients of determination (R^2) exist for each pair of adjacent segments under different operating conditions and what values of α are most appropriate for these conditions.

To conduct this analysis, Bluetooth travel time observations have to be screened to identify vehicles that passed all four detectors. Then to compute R^2 values, non-overlapping chronological sets of travel times for the same 50 vehicles on the upstream and downstream segments can be subjected to a regression analysis. At the same time, the time span for the 50 observations can be determined by noting the difference in time from when the first vehicle passed the first station until the 50th vehicle passed that same station.

Figure 2 shows scatterplots of the R^2 values for both 39-9 versus 9-10 and 9-10 versus 10-11 versus time of day. The x axis value is the point in time when the first vehicle (in each set of 50) passed the upstream reader: station 39 in the case of the 39-9 versus 9-10 plots and station 9 in the case of the 9-10 versus 10-11 plots. One can see that R^2 values decrease as the afternoon peak approaches and increase after the peak has passed. This pattern is more apparent for the first pair of segments (39-9 versus 9-10) than the second. In both cases, however, a pattern is observable, and it strongly suggests that the degree of correlation diminishes as congestion increases.

IV. Methodology For Generating Route Travel Time Distributions

Building on the ideas presented so far, this section describes a procedure based on the statistical properties of comonotonicity and independence that has been developed to synthesize route-level travel time density functions from segment-level distributions.

It is important to note that the authors have already demonstrated that the property of comonotonicity can be used to synthesize route-level travel time distributions under uncongested conditions. However, in that research, the authors also found that the comonotonicity-based procedure has difficulty matching travel time distributions when congestion is high.

This paper extends that previous research by combining comonotonicity with independence. Algorithm 1 presents the logic behind the methodology. Three main steps are involved: 1) develop a comonotonicity-based route-level distribution, 2) develop an independence-based route-level distribution, and 3) sample appropriately from these two distributions to synthesize a composite route-level distribution using α as the percentage of independence-based route-level observations to employ.

Algorithm 1: Computing Composite Distribution

```
acquire raw data;
use filter algorithm to extract legitimate travel times;
sort travel times in chronological order;
for temporally sequential 50 pairwise observations do
    Compute R^2 value for adjacent segments
end
Use K-Means clustering for regime classification;
for Each cluster do
    generate comonotoncity-based route density
      (CDF_1);
    generate independence-based route density
      (CDF_2);
    for α ∈ [0, 1] do
        compute composite CDF of (CDF_1, CDF_2);
        α+ = 0.01
    end
    score each composite CDF using Max-deviation
      test;
    pick CDF with maximum score and lowest α
      value
end
```

A. Comonotonicity-based distribution

The first sub-step involves constructing a route-level travel time density function based on comonotonicity. It implicitly assumes the drivers experience very little or no impedance from other users of the network. Under these conditions, drivers have the ability to achieve their desired speeds; therefore strong correlations exist among travel time distributions for adjacent segments. Theoretical constructs exist to compute joint density function for two or more random variables that are correlated. However, as demonstrated in our previous paper Isukapati et al,[10], a simple concept like comonotonicity [11] can be used to compute the joint CDF.

Comonotonicity is widely used in calculating portfolio risk. Random variables X and Y are comonotonic if there is a third random variable Z as well as increasing functions f and g such that X = f(Z) and Y = g(Z). It implies that individual percentile values from each of the distributions can be added to obtain the percentile value for the distribution of the sum.

B. Independence-based distribution

The second sub-step involves constructing a route-level travel time density function while assuming the travel times on adjacent segments are independent. Operationally, this implies that drivers lose their ability to achieve their desired speeds, and the travel times they experience are derived from their interactions with other vehicles. Monte Carlo simulation is used to convolve segment level travel time distributions.

C. Create the composite distribution

The third sub-step involves synthesizing a route-level travel time distribution that proportionally combines the distributions described in Sections 4.1 (CDF_1) and 4.2 (CDF_2). This is done using stratified sampling in conjunction with Monte Carlo simulation. That is, for a sample size (N) used for creating such a composite distribution, the sampling fraction is n_1 for CDF_1 and n_2 for CDF_2 where:

1) $n_1 = (1 - \alpha) * N$
2) $n_2 = \alpha * N$
3) α = reflection of the degree of independence between the segment-level distributions

Of course, the correspondence between values of α and the operating conditions have to be identified from similar facilities and/or prior experience, but once that correspondence is understood, synthesized versions of the route-level travel time distributions can be obtained. As Section 5 demonstrates, this technique is effective across a wide range of operating conditions for both two- and three-segment routes.

V. FIELD DATA-BASED TESTS

This section demonstrates the efficacy of the proposed synthesis technique for both two- and three-segment routes. It also illustrates the notions of regimes and combinations of regimes for successive segments.

Inasmuch as this section is aimed at demonstrating the efficacy of the procedure, the material presented is of three types. The first type shows synthesized candidate distributions based on various values of α,using the method described in the previous section. The second compares these synthesized distributions with the actual distributions observed (i.e., the ground truth) to find the synthesized distribution that best matches the actual distribution. The third assesses the quality of the best matches found.

A. Find the best value of α

Finding the best value of α involves a search. Many search techniques are possible, but the authors have elected to use exhaustive enumeration looking at two decimal values of α ranging from 0 to 1 (101 values). The test for the best α value focuses on comparing percentile-specific travel times for the synthesized distribution with those from the actual distribution. Acceptable values of α are those that produce maximum absolute differences that are less than a threshold value. The best value of α minimizes the sum of these absolute deviations for all 101 percentiles (from the $0 - 100^{th}$ percentile). This sum is called the test-score (s_{max}). If more than one α produces the same test-score (s_{max}), the lowest α value is selected.

B. Why Kolmogorov Smirnov (K-S) test is not used

Initially the procedure used a Kolmogorov-Smirnov test to select the best value of α.The K-S test values were computed for each candidate α and then the α value was selected that yielded the best results from the K-S test. However, the K-S test struggled to elucidate the differences between the distributions produced by various α values. The K-S test focused only on the maximum deviation, not the sum of the deviations. Figure 3 illustrates this result. The figure shows four travel time CDFs: the actual distribution is shown in red while three candidate distributions are shown in green, violet, and blue. The KS test indicated that the statistical differences between the composite and actual CDF were all insignificant. However, as one can quickly see, CDF-1 matches the observed field data the best.

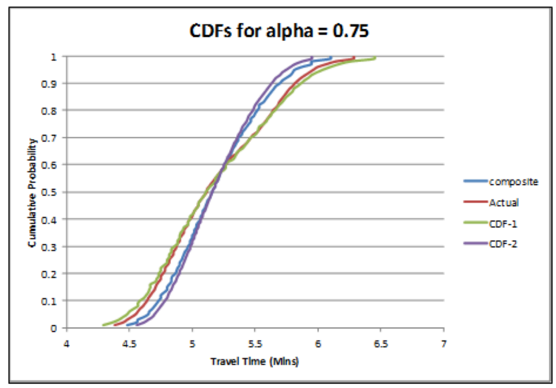

Fig. 3. K-S test Limitations

C. Two segment routes

Routes with two segments are a good place to start illustrating the methodology. The steps involved are to select chronological blocks of 50 vehicles; and then for each block: develop upstream and downstream segment-level CDFs as well as the overall route CDF (because in this instance it has been observed), develop the comonotonicity- and independence-based distributions (as described earlier); sweep through the values of α from 0 to 1 (in 0.01 increments) and for each create the candidate synthesized route-level distribution, find the value of α that produces a synthesized route-level distribution that best matches the actual distribution, and record the value of α, the value of the test score, and any other information that helps create an understanding of how the value of α depends on the operating conditions, etc. Repeat the analysis for every block of 50 vehicles and summarize the findings.

It is important to realize that in two-segment analyses, there are regimes involved (not just one) because the upstream segment is in one operating condition (regime) and the downstream segment is in another. These regimes may be very similar (or effectively identical), but there is no reason why they should be exactly the same.

To characterize the regimes, the authors have demonstrated that a statistical measure called the semi-standard deviation (SSD) is very effective. The semi-standard deviation is the square root of the semi-variance. The semi-variance is a one-sided metric that uses a reference value r (instead of the mean) as the basis for calculating the sum of the squared deviations. Moreover, only observations t_i that are greater than (or less than) that reference value r are included in the calculation.

For one of the blocks of 50 vehicle observations, Figure 4 presents a comparison between the synthesized route-level travel time density function and the actual one. This happens to be for a situation where the SSD on segment A was 1.5 and the SSD on segment B was 1.7, meaning both segments were moderately congested. The α value that provides the best match is 0.07. This means there is a high correlation between travel times on the two segments. The α = 0.07 result implies that the best synthesized route-level travel time distribution is 93% when drawn from the comonotonicity-based travel times and only 7% when drawn from the independence-based travel times.

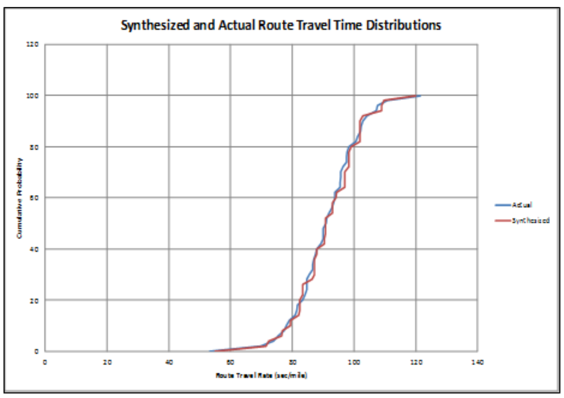

Fig. 4. Synthesized Vs Actual travel time CDF

The remaining issue relates to how the α value varies as the operating regimes vary. Figure 5 shows that the α values tend to be low when the R^2 value is high and vice versa. This is consistent with the ideas already presented: that when the correlation is high, the best value of α tends to be at or near zero (with the comonotonicity-based density function being dominant) and when the correlation is low, the best value of α tends to be high (with the independence-based density function being dominant).

D. Three segment routes

Three segment routes introduce the added complexity that three operating regimes are extant, one for each segment. At least in theory, the selection of an appropriate value of α is more difficult. However, the analysis about to be presented shows that the selection process is relatively straightforward.

As with the previous two-segment cases, the dataset underlying this analysis is a record of vehicles observed at all four Bluetooth stations (39, 9, 10, and 11). This means all of the segment travel times were observed as well as the overall route travel time. Hence, any route-level distribution that is synthesized can be compared with the

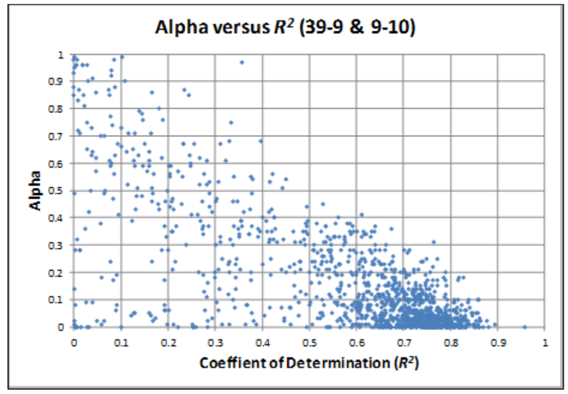

Fig. 5. Trends in α values versus R^2

actual, observed distribution to evaluate the quality of the synthesized distribution.

Moreover, since the previous section showed that the values tend to vary inversely with the R^2 values, the analysis presented here uses the R^2 values for the segment-to-segment correlation analyses as the basis for analyzing the α value variations.

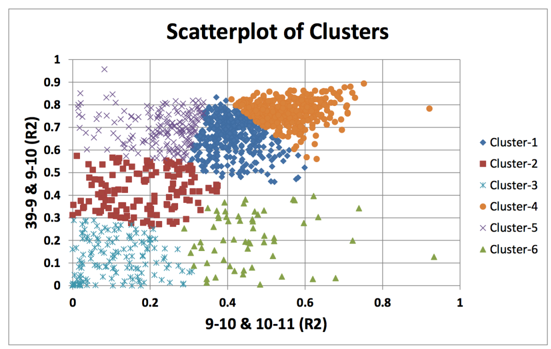

Fig. 6. Regime Classification

1) Computing R^2 values for the analysis: Since this analysis involves three segments, there are R^2 values that relate the regime of segment 39-9 to 9-10 as well as R^2 values that related the regime of 9-10 to that of 10-11. Moreover, since these R^2 values seem to provide a good indication of the pairwise combinations of regimes for the upstream and downstream segments, analysis of the three-segment case can be predicated by an analysis of the combinations of R^2 values that pertain to the upstream and downstream pairs of segments (39-9 and 9-10 being the first pair and 9-10 and 10-11 being the second).

Therefore, a preprocessing step involved grouping the 1250 chronologically ordered vehicle observations into groups of 50 and then computing the R^2 values for both pairs of segments for each group. These combinations of R^2 values were recorded and stored as ordered pairs $[R^2_{(1,i)}, R^2_{(2,i)}]$. $R^2_{(1,i)}$ is the coefficient of determination for the first pair of adjacent segments and the i^{th} set of 50 pairwise observations; and $R^2_{(2,i)}$ is the coefficient of

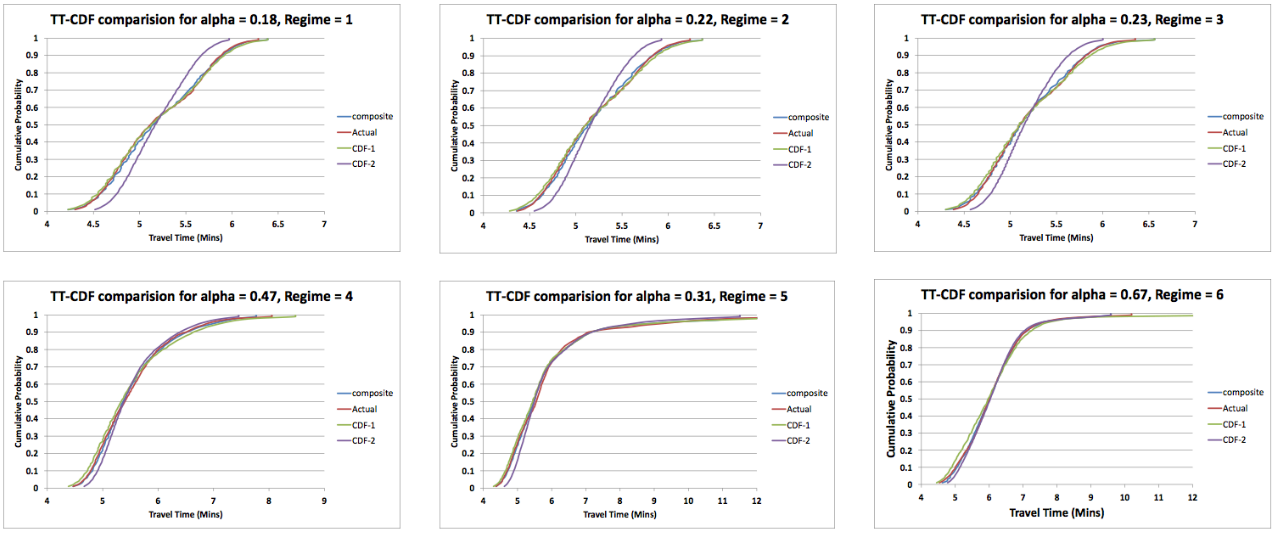

Fig. 7. Composite CDF's that passed max-deviation test

determination for the second pair of adjacent segments and the i^{th} set of 50 pairwise observations.

2) Clustering the R^2 pairs: Each pair of R^2 values $[R^2_{(1,i)}, R^2_{(2,i)}]$ provides information about the level of congestion on the overall network for i^{th} set of 50 pairwise observations. For example, if $R^2_{(1,i)}$ is high, and $R^2_{(2,i)}$ is low, then the first two segments must be uncongested while the third is congested.

K-Means clustering was used to classify the pairwise R^2 values into six groups. Code for K-Means clustering was developed in Python.

3) Results from K-Means clustering: Figure 6 presents a scatterplot of the R^2 pairs as clustered by the clustering algorithm. The $R^2_{(2,i)}$ value for the second pair of adjacent segments is plotted on X-axis and the $R^2_{(1,i)}$ value for the first pair of segments is plotted on Y-axis. Each cluster represents a different combination of regimes among the three segments. Below are some observations about them.

1) **Cluster 1**: Both $R^2_{(1,i)}$, and $R^2_{(2,i)}$ values tend to be quite high, which suggests that all three segments in the network are operating under uncongested conditions
2) **Cluster 2**: values for $R^2_{(1,i)}$ tend to be high, whereas $R^2_{(2,i)}$values tend to be moderate, which suggests that segment-1 and segment-2 might be operating under uncongested conditions, while segment-3 might be operating under "low-congestion" conditions.
3) **Cluster - 3**: values for $R^2_{(1,i)}$ tend to be very high, while $R^2_{(2,i)}$ values are moderate, which suggested that segment-1, and segment-2 are uncongested, and segment-3 is moderately congested
4) **Cluster 4**: values for both $R^2_{(1,i)}$, and $R^2_{(2,i)}$ are not too high nor not too low, which suggests that all three segments are moderately-congested
5) **Cluster 5**: values for $R^2_{(1,i)}$ are moderate, while $R^2_{(2,i)}$ values tend to be high, suggesting that the first two segments are moderately congested, while segment-3 is uncongested
6) **Cluster 6**: Very low values for both $R^2_{(1,i)}$, and $R^2_{(2,i)}$ suggest that all three segments are operating under high congestion conditions.

4) Finds for α: Using the maximum deviation test described earlier, test-scores were computed (for $\delta_{tot} \in [-1, 1]$) for each of the 101 composite CDFs in a given clusters (regimes). Table 1 summarizes these results. Column one represents cluster number, columns 2-4 represent test scores for CDF_1, CDF_2, and composite. Column 6 is the corresponding α for the composite that passed the test. Please notice that test-scores for the composite are consistently higher than those for either CDF_1 or CDF_2. Furthermore, test-scores for all of the clusters are in the range of 90, except for cluster-5. However, when δ_{tot} changed from [-1, 1] to [-1.5, 1.5] the test-score increased to 85.

TABLE I. MAX-DEVIATION TEST SCORES

Cluster #	Test-Score			Alpha
	CDF-1	CDF-2	Composite	
1	91	10	95	0.18
2	87	15	98	0.22
3	80	17	94	0.23
4	39	33	95	0.47
5	28	28	64	0.31
6	33	74	90	0.67

Figure 7 presents travel time CDF comparisons for each of the regimes. As one might notice, the composite CDF matches the actual CDF better than either CDF_1 or CDF_2

In summary, the results presented here demonstrate that the authors' suggested methodology for synthesizing route-level travel time distribution works very well.

VI. Conclusions

This paper shows how proportional sampling, comonotonicity and independence can be used to synthesize route-level distributions for a wide range of operating conditions. The methodology involves three main steps: 1) develop a comonotonicity-based route-level distribution, 2) develop an independence-based route-level distribution, and 3) sample appropriately from these two distributions to synthesize a composite route-level distribution using α as the percentage of independence-based route-level observations to employ. Furthermore, a new statistical test called max-deviation test is proposed to select a composite travel time distribution that has closest statistical properties as those in actual travel time distribution. Efficacy of the technique is illustrated using individual vehicle travel times collected on I-5 in Sacramento by Bluetooth readers.

References

[1] K. Jintanakul, L. Chu, and R. Jayakrishnan, "Bayesian mixture model for estimating freeway travel time distributions from small probe samples from multiple days," *Transportation Research Record: Journal of the Transportation Research Board*, no. 2136, pp. 37–44, 2009.

[2] S. Park, H. Rakha, and F. Guo, "Multi-state travel time reliability model: impact of incidents on travel time reliability," in *Intelligent Transportation Systems (ITSC), 2011 14th International IEEE Conference on*. IEEE, 2011, pp. 2106–2111.

[3] S. Susilawati, M. A. Taylor, and S. V. Somenahalli, "Distributions of travel time variability on urban roads," *Journal of Advanced Transportation*, vol. 47, no. 8, pp. 720–736, 2013.

[4] H. Mahmassani, T. Hou, and J. Dong, "Characterizing travel time variability in vehicular traffic networks: deriving a robust relation for reliability analysis," *Transportation Research Record: Journal of the Transportation Research Board*, no. 2315, pp. 141–152, 2012.

[5] T. Barkley, R. Hranac, and K. Petty, "Relating travel time reliability and nonrecurrent congestion with multistate models," *Transportation Research Record: Journal of the Transportation Research Board*, no. 2278, pp. 13–20, 2012.

[6] J. Ernst, C. Day, J. Krogmeier, and D. Bullock, "Probe data sampling guidelines for characterizing arterial travel time," *Transportation Research Record: Journal of the Transportation Research Board*, no. 2315, pp. 173–181, 2012.

[7] M. Ramezani and N. Geroliminis, "On the estimation of arterial route travel time distribution with markov chains," *Transportation Research Part B: Methodological*, vol. 46, no. 10, pp. 1576–1590, 2012.

[8] F. Guo, Q. Li, and H. Rakha, "Multistate travel time reliability models with skewed component distributions," *Transportation Research Record: Journal of the Transportation Research Board*, no. 2315, pp. 47–53, 2012.

[9] Y. Feng, J. Hourdos, and G. A. Davis, "Bayesian model for constructing arterial travel time distributions using gps probe vehicles," in *Transportation Research Board 91st Annual Meeting*, no. 12-3490, 2012.

[10] I. Isukapati, G. List, B. Williams, and A. Karr, "Synthesizing route travel time distributions from segment travel time distributions," *Transportation Research Record: Journal of the Transportation Research Board*, no. 2396, pp. 71–81, 2013.

[11] J. Dhaene, M. Denuit, M. J. Goovaerts, R. Kaas, and D. Vyncke, "The concept of comonotonicity in actuarial science and finance: theory," *Insurance: Mathematics and Economics*, vol. 31, no. 1, pp. 3–33, 2002.