Considering the Impact of Precipitation on the Accuracy of Delay-Function Parameters
Mehdi Bagherian a,b,*, Mark Hickman a and Ahmad Tavassoli a
a: School of Civil Engineering, The University of Queensland, Australia
b: Spatial Data Analyst, MRCagney Pty Ltd, Australia
*Corresponding author: mehdi@mrcagney.com

Abstract
Travel time estimation along a link is an essential component of transport modelling. One common approach in this regard is to use volume-delay function (VDF) to estimate link travel time as a function of traffic condition and link characteristics. Therefore, calibration of such functions is crucial. Available speed flow records for a set of link classes is usually employed to obtain parameters of VDF. One implicit assumption in the calibration process is to ignore the effect of weather conditions on the speed-flow relationship or to treat them as outliers if they exist in the dataset. Consequently, calibrated parameters can only reflect the link behaviour in a dry condition. In spatial-temporal circumstances that precipitation is highly expected, such parameters may cause significant inaccuracy in the model outputs and unable to represent the actual condition anymore.

In this paper, the effect of precipitation on the travel time estimation using delay functions is studied for motorways in South East Queensland, Australia. Using available data extracted from loop-detectors along with the precipitation reports, speed-flow records were firstly clustered based on the level of precipitation. For each link in the database, free flow speed and link capacities were updated accordingly. These clustered were then independently used to calibrate a set of common VDFs. The results confirm the importance of considering the adverse weather conditions on the calibration process as these conditions result in deterioration of roadway network performance.

1. Introduction
Adverse weather refers to environmental conditions that can affect the capacity of a road, as well as driver behaviour and hence traffic flow. The impact of adverse weather conditions, such as rain, snow, ice, fog and storms on freeway capacity was highlighted in the literature (Tsapakis et al., 2013, K. and Bertini, 2010, Watkins and Hallenbeck, 2010). According to the study by Tu (2008), the adverse weather conditions, on average, resulted in twice the Travel Time Variability (TTV) compared with that of normal conditions based on the large empirical data analysis of travel time for various freeways and one year of weather data. In addition, rain and snow had greater effect on TTV than other adverse weather conditions. In another study, Koetse and Rietveld (2009) claimed that temperature and wind have almost no effect on traffic speed. In addition, rain was found to have no substantial effect on free-flow speed; while the intensity of rain and snow had considerable impacts on the traffic speed in already congested roads and during peak periods. Thus, the impacts of adverse weather depend on the local environmental conditions and may vary from place to place. Therefore, weather condition can have significant impact on the defined characteristics of the links such as free-flow speed and capacity (Hou et al., 2013).

Manzo et al. (2014) highlighted the importance of considering the uncertainty of VDF parameters in transport modelling, suggesting to use various parameters to address different
conditions. This study investigated uncertainty in the BPR parameters by generating BPR parameter distributions using the resampling bootstrap technique and performing sensitivity analysis tests on the four-stage Danish national transport model.

To the author's best of knowledge, despite numerous research have been performed addressing the effects of an adverse weather on the traffic flow elements such as free flow speed, link capacity, and fundamental diagrams, there is no research that shows the effect of weather conditions on the delay functions parameters and their assignment outputs. In addition, no attempt has been made to see how exogenous parameters such as adverse weather can affect the traffic flow and potentially urge a more comprehensive approach in BPR model calibration. This is probably mainly due to difficulties in working with such factors and integrating external databases such as meteorology records with traffic data, especially when a huge source of data for a relatively long period is required.

One of the invaluable potentials of having a rich database is the ability of addressing the effect of conditions such as weather and seasonal trends in transport models. On this basis, the aims of this study are to 1) performing a data fusion and forming an integrated relational datasets of network links, loop detectors and weather data, and 2) calibration of volume delay functions for dry and adverse weather conditions. A threshold is defined to classify observed records into either dry or wet condition.

The remaining sections are organised as follows. The next section explains the research methodology, comprising the data description, the preparation and cleaning procedure, and the modelling approach. The results of the numerical example are provided in the third section. Finally, conclusions and suggestions for future work are presented.

2. Methodology

This section elaborates the followed methodology to reflect the effect of precipitation on volume delay function parameters. Firstly, datasets from different sources used in this study are reviewed. Then, a review on the challenges of harnessing such gigantic datasets are outlined and the implemented systematic method to prepare them for analysis is discussed. Then, in the modelling section, the method to classify data based on their characteristics, required adjustments to reflect the rain effect, and delay function estimation method are reviewed. Finally, our approach to analysis the estimated parameters and validate them are described. The proposed methodology to reflect the effect of precipitation in delay estimation model is illustrated in Figure 1.
2.1. **Input Data**

To be able to reflect the effect of precipitation on travel time estimation models, a set of required data, namely speed-flow recorders, network layout, and meteorological data were utilized. In this section a brief review on these datasets and the required attributes of each dataset is presented.
2.1.1. Loop Detector Data

To extract traffic data, Public Transport Data Source (PTDS) data was used. PTDS provides the feeds of traffic condition in three minutes intervals for Brisbane Network. Amongst the data PTDS provides, loop detector data (Speed and Flow rate pairs) were used for this study. Since 2011, these data are continuously collecting at the University of Queensland. To form the speed-flow database, a set of filters were applied to the existing datasets. Firstly, for each link, speed-flow pairs should cover whole range of congestion level to make parameters such as free flow speed, speed at maximum flow, and link capacity estimable. As the aim of this study was to see the effect of precipitation on delay functions, data may not be necessarily limited to any specific day of the week. However, the data were limited to the daylight period to address potential discrepancies of driving behaviour. Furthermore, the main focus of this paper is to address the potential impacts of adverse weather on freeways thus only records related to this road hierarchy were collected. Thirdly, to make the data smooth and more legit, speed-flow pairs were aggregated in 15 minutes intervals.

2.1.2. Meteorological Data

The second dataset received from Bureau of methodology (BOM) is the Meteorological data of thirteen stations across South East Queensland (encompassing Sunshine Coast, Brisbane, and Gold coast), available from 2009 until September 2013. This database is also providing an large amount of weather data in 30 minutes intervals. Regarding the precipitations values, for each weather station the amount of rain since 9am of each day is cumulatively recorded thus the rate of precipitation can be estimated by calculating the precipitation difference in 30 minutes intervals.

2.1.3. Network Data

A set of attributes for each network link is also required to be considered in the estimation process of the VDF parameters. These attributes are the number of lanes, maximum speed, hierarchy and capacity. Such attributes were provided using Brisbane Strategic Transport Model (BSTM) network database. Since BSTM network is not exactly same as PTDS network, a procedure to spatially join these two networks is established and presented in section 2.3.1.

2.2. Data Cleaning

A set of data cleaning procedures were conducted before the modelling process. In this regard, the records with missing attribute and outlier values were excluded from weather data and speed-flow records. Then, the Quantum-Frequency (QF) method (Venkatanarayana et al., 2008) were implemented to exclude outlier data of the speed-flow records, as discussed in the next chapter. In addition, Link types, Maximum speed and the number of lanes that were obtained from BSTM network were also manually validated using Google street view.

2.2.1. Quantum Frequency Analysis

In this study, the Quantum-Frequency (QF) method was implemented to exclude the speed-flow outliers that are not following the normal observed trend. This approach was initially suggested for automated identification of traffic flow patterns in large datasets by assuming that normal data are following similar pattern when compared to abnormal conditions. A 'frequency value of significance' value helps identifying data with insufficient frequency as outliers thus remove them from the dataset(Venkatanarayana et al., 2008). Applying QF to identify daily flow patterns, the similar flows with highest frequency (i.e. flow with the highest frequency along with the flows around it) were selected at each time interval.

Figure 2 shows an example to show how QF algorithm is applied to the speed-flow of a segment in rainy conditions. To fulfil this task, firstly the volume over capacity (v/c) values were obtained using the speed observations and link capacity, obtained from Brisbane Strategic Transport Model(BSTM). These values were then classified into equal intervals (the defined range was optimized for each sample set to assure enough samples are available in
each range). Afterwards, for each v/c interval (0.4-0.5 in this example), the set of speed values were extracted and frequency of each of which were calculated. ‘Normal Data’ were then defined as the observations included in the cluster where the maximum frequency are available. Speed-flow records outside this range were treated as outlier. For each selected range, the mode of speed observations (i.e. highest number of observation in the selected speed range) was calculated as the ‘sampled’ record.

**Figure 2 Using Quantum Frequency Filter to remove records’ outliers**

This process was performed for all the available segments in each weather category. The mode of normal data (i.e. the speed with highest frequency) is introduced as the representation of the defined interval when sampling over the data is required. Consequently, when a sampling method is implemented, in each defined bin only one sample record would be introduced. In this study, a minimum v/c ratio interval of 0.01 is used for all the applied QF filters. Figure 3 shows the selected speed-flow records (normal data as well as sampled ones) after QF was applied to a selected link’s data. Note that the data at v/c < 0.15 were also removed from the dataset since this condition is mostly related to the early morning observations in which heavy trucks are more dominant unlike the AM peak or PM peak time periods.

**Figure 3 the Effect of Quantum Frequency Filter in removing speed-ratio records data**
2.3. Relating Databases

A set of spatial database joining methods were developed to relate STREAMS speed-flow records, BSTM link characteristics, and the weather data. In this regard, to extract the characteristics of each STREAMS link, its corresponding link on BSTM network was identified first. This task was performed using a proposed algorithm that relies on the link vertexes and the directions in both databases and joins them if possible. The weather data were then joined, as explained in this section. The output of this module is a relational database that can be used for further clustering and calibration process of this study.

2.3.1. Spatial Join of BSTM and STREAMS Networks

Joining Transport network from two different databases and identify corresponding link from one network in the other one was required to link speed-flow records to link characteristics. This problem, known as conflation problem is a common task in many projects (e.g. Wallgr et al., 2010, Mustière and Devogele, 2008). Due to the large size of the network (e.g. BSTM and PTDS datasets had 20507 and 33992 links, respectively), manual operation to match the network is quite costly and cumbersome. In addition, in the time of doing this research, no software or module was available to perform the matching process automatically. Consequently, a tool was developed do address the conflation problem. In this algorithm, two links were selected to be joined if:

1. The Distance between start and end points of two links are both within a predefined buffer distance.
2. The Endpoints are within the acceptable buffer distance, and both links are in the same directions.

The inputs of the module are two Shapefiles (target and join) or MapInfo Interchangeable Files (MIF) and the outputs are the matched attributes of the databases. Figure 4 shows the matching steps and the scheme of the applied matching rules.

Figure 4 Developed Method To Match Two Networks
The output of this module is a set of speed-flow records where their corresponding link’s characteristics (free flow speed, capacity, number of lane, and type) is available and thus can be utilized to calibrate the links’ delay functions parameters.

2.3.2. Weather Estimation of Speed-Flow Records

To reflect the effect of inclement weather on the driving behaviour, the precipitation rate of each speed-flow record (that has the link attributes available) was extracted. As discussed in Section 2.1.2, the meteorological data are basically available for a set of stations in regular intervals (30 minutes here). In this study, as will be elaborated in numerical example (section 4), our study area can be covered with three weather stations where the third one has at least 15km distance from the closest link in the database. Consequently, for each link, the closest weather station was reported and the weather data of all the speed-flow records were extracted from the database if available. Considering the scale of available weather data, to estimate the precipitation intensity of a link, the coordination (i.e. latitude and longitude) of its centroids was extracted as the representative of each link.

3. Modelling Approach

This section is dedicated to present three key steps in volume delay function estimation using the prepared database. Firstly, the data was classified based on the level of precipitation and hierarchies. In addition, the effect of inclement weather on link characteristics parameters such as free flow speed and capacity was discussed. Finally, a non-linear regression model was used to fit the selected VDF to the available-cleaned data in different scales. These steps are explained in this section.

3.1. Data Classification

In this study, the records depending on the recorded severity of precipitation, data were classified into “dry” and “rainy” categories. In this regard, a speed-flow record was labelled ‘dry’ if no precipitation was observed in the day the record was observed. Similarly, a precipitation rate of 2mm/h was considered as the minimum threshold to define a record as a ‘wet’ record.

3.2. Adjusting Link Characteristics in Rainy Days

Numerous researches have shown that inclement weather is basically reduce the capacity of the links (e.g. El Faouzi et al., 2010, Maze et al., 2006, Mahmassani et al., 2009, Hooper, 2014). Maze et al. (2006) reported up to 14 percent reduction in freeway capacity as a result of rain precipitation. Similarly, results from a study by Smith et al. (2004) showed a more than 25% reduction in capacity in heavy rain condition. Since link capacities are of the key parameters in serving the assigned flow, it seems necessary to address their reduction and effects on assignment outputs.

Similar to the link capacities, free flow speed values may also be changed in rainy weathers. On the contrary to the link capacity, studies confirmed less sensitivity of free flow speed to the precipitation. In this regard, a handful of studies (e.g. Hooper et al., 2014, Agarwal et al., 2006) reported no change in traffic speed as the result of raining. Having said that, however, Maze et al. (2006) and Hranac et al. (2006) reported a maximum of 9% reduction in free flow speed in inclement weather. Reduction in maximum speed is mainly because of the changes in driver’s behaviour due to the lack of sufficient feasibility and reduction in acceleration and deceleration performance.

In this study, an analysis was performed to see the changes in capacity and free flow speed in rainy weathers for all the studied links. In this regard, 99th percentile value of the flow were calculated as the link capacity value. To obtain free flow speed, two values of speed were considered as feasible values. Firstly, average speed in uncongested segment of the traffic pattern ($v/c= 0$ to $0.4$ in this study) were calculated as a measure of free flow speed. Secondly,
to reflect any noticeable change in speed in the uncongested traffic, a linear regression was also performed over the observed records. The higher value amongst these two measured were then introduced as free flow speed of a link.

### 3.2.1. VDF Parameters Estimation

Over the past few decades, a number of functional forms of VDFs have been proposed. Amongst them, the function developed by the Bureau of Public Roads- BPR(1964), the function suggested by Akcelik (1991), and Conical Functions(Spiess, 1990) have attracted much attention. Regardless the VDF type, having a realistic reflection of the observed behaviours in a demand forecasting study urges a calibration processes. This process can be performed by methods such as minimization of the root-mean-square error (RMSE).

Amongst the available delay functions, BPR function is gained superior attention. This is not only because of its simplicity but also the performance it has shown to predict the observations in numerous studies (e.g. Müller and Schiller, 2014, Neuhold and Fellendorf, 2014, Stevens et al., 2017). It is noteworthy that a set of deficiencies is raised against BPR functions. Spiess (1990) for example has refuted BPR application, especially during the assignment process. Overestimating delay in saturated conditions, negligible difference in lower v/c ratios that may cause equilibrium assignment to be locally trapped into the initial assignment results, and higher computational cost rather than the other functions are the critics they raised against BPR. In addition, they stated that BPR function tends to underestimate the delay at the intersections and shows its appropriateness in freeways only. However, BPR is still the most popular function in traffic assignment modules. This function was formulated by Bureau of Public Roads (1964) as below:

\[ t(v) = t_0 \times (1 + \frac{v}{c})^\beta \]

Where \( t(v) \) is the travel time of the link, \( v \) is the link volume, \( t_0 \) is the free flow time, \( c \) is the link capacity, and \( \alpha, \beta \) are the calibration parameters. With rudimentary operations, BPR function can be used to calculate the ratio of free flow speed to the vehicle speeds as below, where \( f(x) \) is the ratio of free flow speed to the vehicle speed.

\[ f(x) = 1 + \alpha \left( \frac{v}{c} \right) \beta \]

BPR parameters \( (\alpha, \beta) \) should be calibrated to reflect the traffic regime in different levels of congestion. Numerous works exist in the literature, each proposing the extracted parameters that has led to best fit to the data. Depending on the dataset and the study area, these parameters can be in a wide range. In this study, BPR function was selected and calibration of its parameters for both dry and rainy condition was performed.

### 3.2.2. Calibration Method

Calibration of two BPR function parameters \( (\alpha, \beta) \) were performed using Generalized Reduced Gradient (GRG2)(Lasdon et al., 1978) via Microsoft Excel Solver. The problem was aiming to do the regression through least square fitting method thus defined as minimization of sum of the squared residuals(difference between predicted and observed values) where no constraint were defined for Alfa and Beta parameters.

### 3.3. Modelling Analysis

To be able to compare the results in different scenarios, two statistical measures, namely: the coefficient of determination(R-Squared) and the RMSE were utilised. The first measure is the R-squared (R²), as one of the most commonly and widely used measures (Washington et al., 2011), to show the degree of accuracy of the model predictions in comparison to actual data. The higher this measure is, the closer the modelled results are to the actual passengers' behaviours. R² values range from 0 to 1, with higher values indicating less error variance. The
second measure is the root mean square error (RMSE), as shown in the following equation. The lower the value of RMSE, the lower is the variation of the error. Cambridge Systematics (2010).

\[
RMSE = \sqrt{\frac{\sum_{i=1}^{N} (Model_i - Actual_i)^2}{N}}
\]

where \( Actual \) is the observed data, \( Model \) is the predicted value as estimated by the model, and \( N \) is the number of predictions.

4. Numerical Example

This section shows the procedure and results of applying the presented method to reflect the effect of precipitation on South East Queensland’s (SEQ) VDF parameters. Figure 5 shows the map of the study area and the segments analysed in this research. Following the Data cleaning procedure explained in section 2.2, 36 links on Pacific Motorway was selected for this study. It can also be seen that out of 13 available weather stations in available dataset, only two of those are in the proximity of the selected links thus precipitations records were obtained from them.

Figure 5 Study Area, Link Centroids and Weather Stations

4.1. Data managements

To form a database that can reflect the weather impacts, data were collected for both rainy and dry conditions. In this regard, top 100 wet days as well as 100 dry days within a three years period were observed and speed-flow records were extracted for each 15 minutes
interval of the selected corridor. Corresponding weather condition was then extracted from the meteorological data using the link centroid (as location) and timestamp (as time) of each collected record. Weather data were then joined to Speed-flow pairs using their spatial-temporal data. To extract the precipitation records of each link, their centroids were located and considered as representative of the segment geometry. For each timestamp, precipitation intensity was then derived from the database. Table 1 summarizes the available sample size for each level of rainfall intensity in the database.

<table>
<thead>
<tr>
<th>Rainfall Intensity (mm per hour)</th>
<th>number of records</th>
</tr>
</thead>
<tbody>
<tr>
<td>&lt;1</td>
<td>57888</td>
</tr>
<tr>
<td>1-2.6</td>
<td>35709</td>
</tr>
<tr>
<td>2.6 - 7.6</td>
<td>25148</td>
</tr>
<tr>
<td>7.6-40</td>
<td>9853</td>
</tr>
<tr>
<td>Total</td>
<td>128598</td>
</tr>
</tbody>
</table>

4.2. Adjusting Free Flow Speed and Capacity values

The effect of precipitation on maximum speed and capacity of the selected links were examined to improve the modelling results. In this regard, for all 36 selected links in the study area, the method to derive link capacity and maximum speed (see section 3.2) was performed for both dry and rainy datasets. Figure 6 shows the level of reduction on capacity and maximum speed of the links due to the precipitations. It can be seen that in general, the reduction in capacity is around 10% where the maximum speed is slightly decreased (1.5%).

![Figure 6 the Effect of Precipitation on Speed and Capacity Reduction](image)

The link capacity and maximum speed records were updated for rainy conditions using the derived values as the requirements of the VDF calibration process.

4.3. Modelling Results

Individual link analysis were performed on four selected segments along Pacific Motorway with different posted speeds and for each link, calibration parameters were obtained for both dry and rainy conditions. Table 2 shows the results of calibrating BPR parameters for dry and rainy conditions for each separate link. It can be confirmed that in dry conditions, alfa parameter that is a multiplier of congestion level is lower than their corresponding parameters in rainy condition. Besides, Beta parameter where can potentially signify the effect of
congestion is less in rainy condition. These differences is eventually leading to estimation of a higher level of delay for passenger cars in rainy conditions.

Regarding the statistical analyses results, in all selected segments, despite $R^2$ measures are almost close with no pattern in different speeds, RMSE measures in dry conditions show less error than rainy records in all links. Such difference confirms the higher level of uncertainty involved with rainy conditions rather than records in dry condition. In addition, it could be concluded from Beta parameters trend that the higher free flow speed is resulted a higher vulnerability to precipitation where drops in speed is happening in lower levels of road congestion.

Table 2 Individual Link Calibration of BPR parameters

<table>
<thead>
<tr>
<th>Link ID</th>
<th>Posted Speed (kph)</th>
<th>number of included records</th>
<th>Alfa</th>
<th>Beta</th>
<th>$R^2$</th>
<th>RMSE</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td></td>
<td>Dry</td>
<td>Wet</td>
<td>Dry</td>
<td>Wet</td>
<td>Dry</td>
</tr>
<tr>
<td>492950</td>
<td>80</td>
<td>1397</td>
<td>153</td>
<td>0.33</td>
<td>0.57</td>
<td>2.75</td>
</tr>
<tr>
<td>535545</td>
<td>90</td>
<td>2578</td>
<td>538</td>
<td>0.31</td>
<td>0.51</td>
<td>3.28</td>
</tr>
<tr>
<td>203666</td>
<td>95</td>
<td>4838</td>
<td>873</td>
<td>0.30</td>
<td>0.55</td>
<td>4.70</td>
</tr>
<tr>
<td>534340</td>
<td>100</td>
<td>3025</td>
<td>392</td>
<td>0.37</td>
<td>0.59</td>
<td>4.22</td>
</tr>
</tbody>
</table>

Figure 7 depicts the records in both dry and rainy conditions and the fitted delay function curve to each condition for the selected links.

Figure 7 Calibrated BPR functions for dry and rainy conditions for selected links
4.4. Aggregated Model

The current Brisbane Strategic Transport Model (BSTM) is suggesting unique parameters for all the freeways for the planning studies. To extract the parameters as an input of a macroscopic modelling, data extracted from all the links were aggregated and BPR parameters were calibrated for all the results. For each link data set, a QF sampling method (section 2.2.1) was applied along the range of v/c ratio. The data were then aggregated and outliers were removed for final calibration. Figure 8 shows the extracted data as well as aggregated developed BPR functions for both dry and rainy conditions. Following a similar pattern to the individual link analyses, it can be seen that suggested BPR functions for rainy conditions are overall suggesting lower speeds than the estimated speed in dry conditions. While for undersaturated conditions the difference is negligible, a shift in f(x) can be seen for higher v/c ratios. These parameters can be implemented in a network assignment procedure to see the network-wide impact of considering weather effects in developing delay functions.

Figure 8 Calibrated BPR functions for dry and rainy conditions using aggregated data

5. Conclusions and Future Works

In this paper, the effect of precipitation on delay functions parameters were examined using speed-flow records of loop detectors and available meteorological data using a three years data. After a set of data preparation procedures, reductions in free flow speed and capacity of the links were calculated. Then, calibration process were performed and BPR parameters were extracted for freeways in dry and adverse weather conditions. Noticeable difference between the obtained parameters of these two conditions signifies the necessity of introducing weather impacts to volume delay functions parameters if rainy conditions are frequent during a year in the target horizon and area of study.

Several recommendations can be made for continuing the present study. Firstly, the available data for the present work was reliable to see the effects of precipitation on motorways only. The presented method can thus be applied to other road hierarchies such as arterials. Having the same parameters for whole the road hierarchies, travel times estimation methods such as traffic assignment modules can have a more realistic insight of the network status in rainy conditions. Finally, this study classified the data into dry and rainy while the range of light
precipitation were not studied. The future work can fill this gap by considering the severity of rainfall in the model attributes.

6. Acknowledgement

The authors are grateful to the Queensland Department of Transport and Main Roads (DTMR) and Australian Bureau of Meteorology for providing the data for this research. This work is supported by the Transport Academic Partnership, with the Queensland Department of Transport and Main Roads.

7. References


HOOPER, E. 2014. The impact of precipitation on speed–flow relationships along a UK motorway corridor. Theoretical and applied climatology, 117, 303-316.


