Method of Parameter Calibration for Error Term in Stochastic User Equilibrium Traffic Assignment Model

Xiang Zhang, David Rey, S. Travis Waller

Abstract—Stochastic User Equilibrium (SUE) model is a widely used traffic assignment model in transportation planning, which is regarded more advanced than Deterministic User Equilibrium (DUE) model. However, a problem exists that the performance of the SUE model depends on its error term parameter. The objective of this paper is to propose a systematic method of determining the appropriate error term parameter value for the SUE model. First, the significance of the parameter is explored through a numerical example. Second, the parameter calibration method is developed based on the Logit-based route choice model. The calibration process is realized through multiple nonlinear regression, using sequential quadratic programming combined with least square method. Finally, case analysis is conducted to demonstrate the application of the calibration process and validate the better performance of the SUE model calibrated by the proposed method compared to the SUE models under other parameter values and the DUE model.

Keywords—Parameter calibration, sequential quadratic programming, Stochastic User Equilibrium, traffic assignment, transportation planning.

I. INTRODUCTION

In transportation planning, in order to support current network evaluation and potential network modification, analysis and estimation of the traffic system performance should be conducted through determining travel decisions [1], [2]. For travel decisions, the Four-step model is adopted in most cases. The four steps refer to trip generation, trip distribution, mode choice and traffic assignment, which correspond to the decisions of whether to travel, where to go, by which mode and which path to select. Among all the four steps, traffic assignment is acknowledged to be the most mathematically complex step as it aims to determine how to assign origin-destination (OD) travel demands onto a network [3]. In the traffic assignment process, it is assumed that every traveler seeks the path with the minimum travel time. As a result, the travel demand can be loaded onto the network until it reaches an equilibrium state, known as User-Equilibrium (UE) [4]. In other words, the objective of traffic assignment is to find the UE flow pattern, given a transportation network and travel demand.

UE conditions for traffic assignment can be categorized into Deterministic User Equilibrium (DUE) and Stochastic User Equilibrium (SUE) [5], [6]. Under DUE conditions, no traveler can improve his or her travel time by unilaterally changing routes since it is assumed that every traveler has perfect knowledge of network performance and selects the minimum travel time route. However, in real world, travelers may not consistently make the correct decisions concerning route choices and can be assumed to make errors. In particular, the perceived travel time by a traveler is not necessarily the same as the real travel time. Thus, every traveler will only select the route with minimum perceived travel time. If so, under the equilibrium condition, no one believes that he can shorten this travel time by unilaterally changing routes, which is stated by SUE principle [7]. By comparison, SUE has been found to better represent users travel behavior [8].

DUE and SUE conditions can be formulated into two mathematical models. The SUE model requires that a random variable representing the error term be determined, i.e. the difference between perceived travel time and actual one [9], [10]. For the distribution of the random variable, the Gumbel Distribution and Multivariate Normal Distribution have been tested, corresponding to Logit and Probit route choice model respectively [11], [12]. Logit-based SUE has been more widely used by researchers due to its relatively low computational cost [13], [14]. The variance of the random variable is indicated by the parameter \( \sigma \) of the error term in the SUE model. It is quite an essential parameter for SUE model since it scales the error term and describes the accuracy of the perceived travel time. Therefore, this parameter of the error term can have a significant impact on the SUE-derived flow results. A proper parameter value can lead to high-quality flow results which approach closely realistic flow patterns. On the other hand, an inappropriate parameter value may result in poor performance of the SUE model [15]. The value of this parameter is typically determined empirically, according to the extent to which travelers’ perceptions disperse [16]. To the best of our knowledge, no systematic method has been focused to quantify the value of this parameter.

This study is intended to develop a dedicated method to calibrate the parameter of the error term in the SUE traffic assignment model. The rest of this paper is organized as follows. Section II gives a brief review of the SUE model as well as a motivating example to demonstrate the significance of parameter calibration for the SUE model. In Section III, we use a Logit-based route choice model and develop a calibration method for the parameters of the error term in the SUE model. Section IV demonstrates the validity of the proposed calibration method based on a network and compares the performance of the calibrated SUE model with other SUE models and the DUE model. Finally, outcomes of
this research are discussed and some directions for future work are presented in Section V.

II. MOTIVATION

In this section, general background of the SUE traffic assignment model and its mathematical representation are analyzed. Then a numerical example is presented to demonstrate the significance of parameter calibration for SUE model.

A. Notations

Notations used throughout the paper are listed in Table I unless otherwise specified.

| \( (i,j) \) | Link with upstream node \( i \) and downstream node \( j \) |
| \( c_{ij} \) | Capacity of link \( (i,j) \) |
| \( x_{ij} \) | Flow on link \( (i,j) \) |
| \( t_{ij} \) | Travel time on link \( (i,j) \) |
| \( t_{ij}^f \) | Free flow travel time on link \( (i,j) \) |
| \( r \) | Origin (O) node index |
| \( s \) | Destination (D) node index |
| \( d^{rs} \) | Travel Demand from origin \( r \) to destination \( s \) |
| \( \pi \) | Path index |
| \( h^n \) | Flow on path \( \pi \) |
| \( t^n \) | Travel time on path \( \pi \) |
| \( \delta^n_{ij} \) | Link-path incidence coefficient, whose value is equal to one if link \( (i,j) \) belongs to path \( \pi \), zero otherwise |
| \( A \) | Link set |
| \( Z^2 \) | OD pair set |
| \( \Pi \) | Set of all the paths among a network |
| \( \Pi^n \) | Set of paths connecting origin \( r \) and destination \( s \) |

Table I: Notations

B. Stochastic User Equilibrium Model

For the traffic assignment under SUE condition, Fisk proposed the first mathematical programming model, which is formulated as follows [17]:

Model (I) – SUE model:

\[
\min_{x_{ij}} \sum_{(i,j) \in A} \int_{0}^{x_{ij}} t_{ij}(x) \, dx + \frac{1}{2} \sum_{(r,s) \in Z^2} \sum_{\pi \in \Pi} h^n \cdot (t^n - t^n)^2
\]

subject to

\[
\begin{align*}
\sum_{(i,j) \in A} \delta^n_{ij} \cdot h^n & = d^{rs} & \forall (r,s) \in Z^2 \\
\sum_{(i,j) \in A} \delta^n_{ij} \cdot h^n & = \sum_{\pi \in \Pi} h^n & \forall \pi \in \Pi \\
x_{ij} & = \sum_{\pi \in \Pi} \delta^n_{ij} \cdot h^n & \forall (i,j) \in A \\
h^n & \geq 0 & \forall \pi \in \Pi
\end{align*}
\]

The SUE model relaxes the assumption of the DUE model that all travelers have full knowledge of travel conditions and choose the path with minimum travel time. Mathematically, it can be described as follows:

\[
T^n = t^n
\]

where \( T^n \) is the perceived travel time.

In contrast, the SUE model accounts for the difference between perceived and actual travel time by adding an error term:

\[
\begin{align*}
T^n &= t^n + \xi^n \\
\xi^n &= \frac{1}{\theta} \xi^n
\end{align*}
\]

In (3), \( \xi^n \) is the error term in the SUE model. \( \xi^n \) is an independent and identical random variable for a specified network. When \( \xi^n \) adheres to Gumbel distribution with zero mean and a unit standard deviation, the Logit-based SUE condition can be reached. \( \theta \) is the parameter of the error term in the SUE model. In this paper, we focus on parameter \( \theta \), which can be seen as a scaling parameter ranging from zero to infinity. In the next section we present the motivation for the proposed calibration method.

C. Motivating Analysis

In terms of (3), the greater the value of \( \theta \) is, the smaller perception error in terms of travel time is expected to made by travelers. Thus, as \( \theta \) tends to infinity, travelers are assumed to choose the shortest path, as in the DUE traffic assignment. On the other hand, small \( \theta \) indicates large perception variance among all network users. Therefore, error term parameter \( \theta \) in the SUE model represents the accuracy level of users’ perception with regards to the actual travel time. For the two extreme circumstances:

1) When \( \theta \) approaches infinite, it means that every traveler in the network is capable of choosing the shortest path on a trip. In this case, the error term \( \xi^n \) can be discarded and the SUE model is equivalent to the DUE model.

2) When \( \theta \) is close to zero, it signifies that information of network performance is hardly accessible to the public and traveler’s route choice decision process is literally no better than guessing. In this case, any path linking the same OD pair has the same probability to be selected regardless of travel costs.

This shows that the flow pattern resulting from the SUE traffic assignment model are highly dependent to the value of \( \theta \). Consider the following numerical example depicted in Fig. 1. The network performance under conditions of SUE with different \( \theta \) values is summarized in Table II.

Fig. 1 Numerical Network Example
Table II shows that travel demand is widely spread among the existing paths for smaller values of θ; hence link flows change significantly with θ. Thus, SUE-originated traffic assignment results are highly correlated to the parameter θ. On the other hand, in the real world, the traffic performance will be approaching the traffic assignment result derived from SUE with certain θ, 2.00 for example. Now assume that we conduct a SUE traffic assignment working with a parameter θ valued at 0.25. If so, in terms of link flows shown in Table II, the difference between flow patterns derived from our adopted SUE model (with parameter θ 0.25) and the SUE model representing the reality (with parameter θ 2.00) is bigger than that between flow patterns derived from the DUE model and the SUE model representing the reality. Consequently, the estimation accuracy of the adopted SUE is worse than that of DUE. By contrast, if we choose an appropriate θ around 2.00, such as 3.0, the SUE model might perform better than the DUE model with regards to the reality.

In theory, the SUE model is a more developed traffic assignment model than DUE since it accounts for the difference between the actual travel time and perceived travel time by travelers. However, SUE requires that its error term parameter be accurately calibrated, as illustrated in the above motivating example; where if the value of θ is misestimated, the SUE model may provide less precise and less reliable results compared to the DUE model.

### III. METHODOLOGY

In this section, a Logit-based route choice model is explored and a method for parameter calibration for the error term in the SUE model is proposed.

#### A. Logit-Based Route Choice Model

The objective function of the SUE model is a strictly convex function, which implies uniqueness of its optimal solution. Although the formulation itself does not have explicit physical meaning, it is determined by working backward from the optimality conditions of the Logit-based SUE condition. The Logit-based route choice model is formulated as follows:

Model (II) – Logit-based Route Choice Model

\[
\begin{align*}
P^n &= \frac{\exp(-\theta \cdot t^n)}{\sum_{n'} \exp(-\theta \cdot t^{n'})} \\
R^n &= d^{rs} \cdot P^n
\end{align*}
\]

where \(P^n\) is the probability that the perceived travel time of path \(\pi\) for OD \(r-s\) is minimal, indicating the probability that the path \(\pi\) is chosen by a traveler.

By combining the two equations of (4), a formula with respect to path flow can be achieved:

\[
R^n = d^{rs} \cdot \frac{\exp(-\theta \cdot t^n)}{\sum_{n'} \exp(-\theta \cdot t^{n'})}
\]

It seems plausible that paths flow can be gained directly from (5). However, it is not the fact since in most cases the variables of path travel costs on the right-hand side of (5) depend on the flows:

1) First, path travel costs can be calculated as link travel costs based on path and link travel cost relationships:

\[
t^n = \sum_{(i,j) \in \pi} d_{ij} \cdot t_{ij}
\]

2) Second, link travel costs are closely related to link flows, which can be calculated trough link performance functions. Among all these functions, the BPR function is widely used and can be stated in (7):

\[
t_{ij} = t_{ij}^0 \cdot \left(1 + a \left(\frac{q_{ij}}{c_{ij}}\right)^\beta\right)
\]

where \(a\) and \(\beta\) are parameters of the BPR function.

Therefore, (5) is actually an implicit function of flow solutions under SUE condition, which is derived from Model (II). In other words, SUE-derived flow results cannot be obtained straightforwardly only through Logit-based route choice model described as (4) or (5). Despite this, Logit-based route choice model can help with the calibration of parameter \(\theta\) in the SUE model, as illustrated in the next section.

#### B. Parameter Calibration Method for SUE Model

Given a specified transportation network, the parameter \(\theta\) can be determined using the method summarized as the following five steps:

1. **Data Collection**
   - Collect the flow pattern \([q_{ij}]\), i.e. for traffic flow on every link \((i,j)\) among the network;
   - Estimate the path flow proportion, \([P^n]\), of trips from the origin \(r\) to the destination \(s\) assigned to each acyclic path \(\pi\), for the \(k\) th selected OD pair \((r,s)\) in \(Z^2\), where \(k=1,2,\ldots,K\) and \(J=1, J^*=1\).

   Both link flows and path flow proportions can be obtained by traffic field investigations [18], [19]. Regarding path flow proportions, \((k + f)\) OD pairs are selected in total, among which the first \(K\) OD pairs are used for parameter calibration while the second \(J\) OD pairs for parameter evaluation. Every path flow proportion \([P^n]\) represents the probability in the Logit-based route choice model that path \(\pi\) is selected among all the alternative paths connecting the \(k\) th OD pair.

2. **Path Travel Time Calculation**
   - Calculate a travel cost on each link by a link performance function using (7);
- Determine each path travel cost $t^{\pi,k}$ using (6).

3. Determination of the Dependent Variable and Explanatory Variable Set

- [$P^x_k$] serves as a dependent variable.
- Every dependent variable [$P^x_k$] corresponds to a set of explanatory variables, denoted as [$X^x_k$].

The size $s$ of the set [$X^x_k$] is determined as follows:

- If there is only one OD pair selected, the size $s$ is equal to $|P^x_k|$, the number of acyclic paths connecting the OD pair.
- If there exist more than one OD pair, the size $s$ is equal to $\max\{|P^x_k|\}$, the maximum number of acyclic paths linking an OD pair among all the $K$ selected pairs.

The component values of the set [$X^x_k$] are determined as follows:

- If there is only one OD pair selected, the component values are derived from (8).

$$X^{\pi,k} = \exp(-\pi^{\pi,k}) \quad k = 1; \quad \pi = 1, 2, ..., |P^x_k|$$  

- If there are more than one OD pairs selected, the component values are derived from (9).

$$X^{\pi,k} = \exp(-\pi^{\pi,k}) \quad \text{for the first} \quad |P^x_k| \quad \text{components}$$  

$$X^{\pi,k} = 0 \quad \text{otherwise}$$

4. Parameter Calibration

Use sequential quadratic programming combined with least square method to conduct parameter calibration [20], [21], based on the variables for the first $K$ OD pairs, $(r,s)_k \in Z^2$, where $k = 1, 2, ..., K$. The model to be calibrated is shown as follows:

Model (III):

$$[P^{x,k}] = \frac{(x^{x,k})^\theta}{\sum_{\alpha \in \Lambda} (x^{x,k})^\theta}$$

Indeed, Model (III) is a multiple nonlinear regression (MNR) model equivalent to (1) of Logit-based route model (II).

Through sequential quadratic programming, a quadratic program is established at every iteration in order to determine the direction of optimization. Then for each direction, the estimated parameter is inserted into the loss function to calculate the loss. The procedure will terminate when the loss function reaches its minimum.

The loss function can be formulated as the sum of squared residuals. In this case, the objective of the sequential quadratic program is to minimize the sum of squared residuals, which is similar to the least square method. A residual is the difference between an actual value and the corresponding estimated one. Each proportion $[P^x_k]$ collected from traffic field investigations serves as an actual value while its corresponding estimated value $[P^{x,k}]$ can be calculated with the explanatory variable set [$X^x_k$] and the estimated parameter $\theta$ using the following formula derived from (1):

$$p^{\pi,k} = \frac{(x^{x,k})^\theta}{\sum_{\alpha \in \Lambda} (x^{x,k})^\theta}$$

Accordingly, the sum of squared residuals (SSR) can be calculated as follows:

$$SSR = \sum_{k=1}^{K} \sum_{\pi \in \Pi^x} (p^{\pi,k} - [P^{x,k}])^2$$

5. Parameter Evaluation

The calibrated parameter should be verified through estimation error based on the data not used in the calibration process, i.e. the variables for the OD pair $(r,s)_k \in Z^2$, where $k = K + 1, K + 2, ..., K + J$. The estimation error is used to compare the difference between the estimated dependent variables and those in the real world. It can be indicated by the average gap described in (13):

$$\varepsilon = \frac{1}{\sum_{k=1}^{K} |[r,s]_k|} \sum_{k=1}^{K} \sum_{\pi \in \Pi^x} |p^{\pi,k} - [P^{x,k}]|$$

If the relative gap does not exceed the permissible maximum one, the parameter $\theta$ derived from the previous step is acceptable. Otherwise, adjust the arguments in the sequential quadratic programming algorithm (i.e. maximum iterations, step limit, optimality tolerance, function precision and infinite step size), and return to Step 4 to conduct parameter calibration once again until the termination criteria are satisfied.

IV. CASE ANALYSIS

This section demonstrates how to implement the parameter calibration for the SUE model using the proposed method. The proposed method is evaluated throughout the following case study; the network to be considered is shown in Fig. 2.

<table>
<thead>
<tr>
<th>TABLE III</th>
<th>LINK INFORMATION</th>
</tr>
</thead>
<tbody>
<tr>
<td>ID</td>
<td>Link $(l_j)^\alpha$</td>
</tr>
<tr>
<td>1</td>
<td>(1,2)</td>
</tr>
<tr>
<td>2</td>
<td>(2,3)</td>
</tr>
<tr>
<td>3</td>
<td>(1,4)</td>
</tr>
<tr>
<td>4</td>
<td>(4,5)</td>
</tr>
<tr>
<td>5</td>
<td>(5,3)</td>
</tr>
<tr>
<td>6</td>
<td>(2,5)</td>
</tr>
<tr>
<td>7</td>
<td>(4,6)</td>
</tr>
<tr>
<td>8</td>
<td>(5,7)</td>
</tr>
<tr>
<td>9</td>
<td>(3,8)</td>
</tr>
<tr>
<td>10</td>
<td>(6,7)</td>
</tr>
<tr>
<td>11</td>
<td>(7,8)</td>
</tr>
<tr>
<td>12</td>
<td>(7,9)</td>
</tr>
<tr>
<td>13</td>
<td>(8,10)</td>
</tr>
<tr>
<td>14</td>
<td>(9,10)</td>
</tr>
<tr>
<td>15</td>
<td>(6,9)</td>
</tr>
</tbody>
</table>

a. Origin node No.11 and destination nodes No.12, 13 and 14 are connected to nodes No.1, 8, 10 and 3 respectively, by links of free travel time.
The calibration process is conducted step by step according to the proposed method in Section III. Additionally, the permissible average gap in terms of path choice probability is set to be 10%.

For the first two steps, the collected data and calculated path travel time are summarized in Tables III and IV. Data for OD pairs (11,14) and (11,13) are used for parameter calibration, and (11,12) for evaluation. In order to represent the real world, the data used in this example are generated by the macroscopic traffic simulation software TransCAD (4.5 Version). Alternatively, these sorts of data can be collected through traffic field investigation in a project.

For step 3, the dependent variables has been listed in the very right column in Table IV and the sets of explanatory variables can be obtained using (9) with the path costs calculated in Table IV.

For step 4, the nonlinear regression is conducted using sequential quadratic programming with least square method incorporated, based on the data for OD pairs (11,14) and (11,13). This process is performed using SPSS 19.0 (Statistical Product and Service Solutions, 19.0Version) [22]. The analysis results are summarized in Table V:

Table V shows that after 7 major iterations, the sum of squared residuals for the MNR Model (III) reach the minimum. As a result, the estimated parameter value along with the standard error and 95% confidence interval is given in Table V. It shows that the standard error is only approximate 7% of the estimated parameter and the 95% confidence interval is relatively small. Meanwhile, the adjusted R squared reaches 0.973, which indicates a significant goodness of fit. Thus, the estimated parameter value 3.82 is statistically significant.

For step 5, the estimated probability \( p^{\theta,k} \) for OD pair (11,12) is calculated using (11). Then via (13), the average gap between estimated probability \( p^{\theta,k} \) and investigated \( [p^{\theta,k}] \) in the real world for the OD pair (11, 12) is achieved. The error turns out to be 4.48%, which is lower than the permissible one, 10%. Thus, the calibration process can be terminated and the result is acceptable.

Furthermore, the estimated path choice probabilities derived from the DUE model and the SUE models with other values of parameter \( \theta \) are summarized in Table VI.

It shows in Table VI that SUE Model with value of \( \theta \) 3.82
TABLE VI
RESULTS FROM DUE AND SUE WITH DIFFERENT PARAMETERS

<table>
<thead>
<tr>
<th>Path ID</th>
<th>Estimated Probability SUE Model</th>
<th>Actual Probability DUE Model</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Estimated Probability $P_{\theta}$</td>
<td>Actual Probability $P_{\theta}$</td>
</tr>
<tr>
<td>0.50</td>
<td>1.00</td>
<td>1.00</td>
</tr>
<tr>
<td>2.00</td>
<td>3.82</td>
<td>0.00</td>
</tr>
<tr>
<td>5.00</td>
<td>10.00</td>
<td>1.00</td>
</tr>
<tr>
<td>1</td>
<td>0.343</td>
<td>0.340</td>
</tr>
<tr>
<td>2</td>
<td>0.342</td>
<td>0.500</td>
</tr>
<tr>
<td>3</td>
<td>0.315</td>
<td>0.160</td>
</tr>
<tr>
<td>4</td>
<td>0.104</td>
<td>0.051</td>
</tr>
<tr>
<td>5</td>
<td>0.104</td>
<td>0.039</td>
</tr>
<tr>
<td>6</td>
<td>0.096</td>
<td>0.030</td>
</tr>
<tr>
<td>7</td>
<td>0.106</td>
<td>0.040</td>
</tr>
<tr>
<td>8</td>
<td>0.098</td>
<td>0.010</td>
</tr>
<tr>
<td>9</td>
<td>0.081</td>
<td>0.006</td>
</tr>
<tr>
<td>10</td>
<td>0.095</td>
<td>0.012</td>
</tr>
<tr>
<td>11</td>
<td>0.088</td>
<td>0.007</td>
</tr>
<tr>
<td>12</td>
<td>0.072</td>
<td>0.005</td>
</tr>
<tr>
<td>13</td>
<td>0.156</td>
<td>0.800</td>
</tr>
<tr>
<td>14</td>
<td>0.177</td>
<td>0.250</td>
</tr>
<tr>
<td>15</td>
<td>0.176</td>
<td>0.160</td>
</tr>
<tr>
<td>16</td>
<td>0.163</td>
<td>0.190</td>
</tr>
<tr>
<td>17</td>
<td>0.180</td>
<td>0.130</td>
</tr>
<tr>
<td>18</td>
<td>0.166</td>
<td>0.060</td>
</tr>
</tbody>
</table>

V. CONCLUSION AND FUTURE STUDIES

This study proposes a method to calibrate the parameter in the error term of the SUE model. The main contributions of this study include:

1) It demonstrates the significance of the parameter in the SUE model through a numerical example.

2) It develops the parameter calibration method for SUE model, based on a Logit-based route choice model, using sequential quadratic programming with least square method.

3) It validates the efficacy of the proposed calibration method as well as the advantages of calibrated SUE models.

In perspective of application, the appropriately calibrated SUE model can be used for transportation systems analysis and planning.

Additionally, the methodology and outcome presented in this paper stress the need to carry out traffic field investigations. For future studies, the SUE condition accounting for demand volatility and incident interference will be investigated. Accordingly, parameter calibration methods for these corresponding traffic assignment models deserve further exploring.

REFERENCES


