frontiers | Frontiers in Future Transportation

ORIGINAL RESEARCH
published: 15 July 2022
doi: 10.3389/ffutr.2022.885967

Path Choice Models in Stochastic Assignment: Implementation and Comparative Analysis

Massimo Di Gangi and Antonio Polimeni*

Dipartimento di Ingegneria, Università degli Studi di Messina, Messina, Italy

In assignment models, a key role is played by the path choice simulation that evaluates the path chosen by users in relation to the perceived paths and relative costs. This study deals with the effects of the implementation of some most adopted path choice models (Logit, Weibit, Probit, and Gammit) within a Stochastic User Equilibrium assignment procedure. Some considerations on parameters needed to make results comparable and the method used to estimate them are also suggested some extensions based on Weibit model are proposed. Results obtained both on a test network and on a real one are reported.

Keywords: path choice behavior, discrete choice models, RUM, SUE, Logit, Weibit, Probit, Gammit

OPEN ACCESS

Edited by:
Ghim Ping Ong, National University of Singapore, Singapore

Reviewed by:
Hooi Ling Khoo, Universiti Tunku Abdul Rahman, Malaysia
De Zhao, Southeast University, China

****Correspondence:***
Massimo Di Gangi
mdigangi@unime.it

Specialty section:
This article was submitted to Transportation Systems Modeling, a section of the journal Frontiers in Future Transportation

Received: *28 February 2022*
Accepted: *01 June 2022*
Published: *15 July 2022*

Citation:
Di Gangi M and Polimeni A (2022) Path Choice Models in Stochastic Assignment: Implementation and Comparative Analysis. Front. Future Transp. 3:885967. doi: 10.3389/ffutr.2022.885967

1 INTRODUCTION

In this study, the simulation of the users' path choice behavior is evaluated by considering different random utility models (RUM).

Path choice is simulated once the origin and destination of the travel, the departure time, and the transport mode have been defined. It is a component of assignment models that can be formulated as a combination of a demand model and a supply model that, in the search for equilibrium between demand and supply, leads to a fixed-point problem (Cantarella, 1997). It is generally assumed that users perceive trip time or cost in a random form. Different authors have considered the statistical distributions of the perceived trip times as belonging to different families.

Starting from the generalized extreme value (GEV) class of models, proposed by McFadden (1978), many models for path choice have been formulated, such as multinomial Logit, C-Logit, path-size Logit, nested Logit, cross-nested Logit, and link-nested models (Manski and McFadden, 1981; Ben-Akiva and Lerman, 1985; Cascetta et al., 1996; Ben-Akiva and Bierlaire, 1999), where Gumbel-distributed random costs are considered. The main advantage of assuming the GEV distributions is that these models are closed under maximization, which means a simple and tractable choice function. Castillo et al. (2008) introduced a closed-form expression for the choice probabilities in the case of independent Weibull-distributed random costs.

By introducing a more general structure of the covariance matrix of the joint distribution of the random residuals (and of the utilities), Daganzo and Sheffi (1977), Sheffi (1985), Rosa and Maher (2002), Yai et al. (1997), and Sheffi and Powell (1982) analyzed the probit model that assumes a normal distribution. Cantarella and Binetti (2002) analyzed the gammit model that assumes a gamma distribution in order to avoid positive perceived utility values, as allowed by a normal distribution. Recently, considering the lack of information on one or more alternatives, other classes of path choice models have been proposed, such as quantum utility models (QUM) (Vitetta, 2016; Di Gangi and Vitetta, 2018) and fuzzy utility models (De Maio and Vitetta, 2015).

Aim of this study is to analyze the effects of the practical implementation of different path choice models based on random utility theory (Logit, Weibit, Probit, and Gammit) within a stochastic user equilibrium (SUE) assignment procedure. Particular attention was paid to the parameters needed to make results from different models comparable. The path choice models are crucial for the assignment models, both for private vehicles (Cantarella and Fiori, 2022; Wang et al., 2019; Di Gangi et al., forthcoming) and transit (Nuzzolo and Comi, 2016; Nuzzolo and Comi, 2018). There are two main approaches: explicit path enumeration or implicit path enumeration (Quattrone and Vitetta, 2011). In implicit field can be cited (Russo and Vitetta, 2003; Antonisse et al., 1989) while in explicit field can be cited (Fosgerau et al., 2013; Mai et al., 2015; Comi and Polimeni, 2022). A review on path choice models is reported in (Prashker and Bekhor, 2004) and (Prato, 2009).

The main original contributions of this study are as follows:

- extensions of a stochastic loading procedure based on path costs following a Weibull distribution with the implementation of a weibit loading procedure that does not require explicit path enumeration;
- some considerations on how to define model parameters in order to make results comparable;
- a comparison of the performances of the considered models obtained both on a test and on a real network.

The advantage in using a Weibull distribution is twofold: 1) the choice probability can be calculated in a closed form, and 2) the dependence of the variance on the path cost allows having different variance values for different o/d pairs, overcoming the issue of Logit models where the variance is the same for all o/d pairs.

Considering the structure of this article, in **Section 2**, after the description of some requirements for path choice models (2.1), a short summary of the considered RUM (2.2) and some operational considerations regarding the implementation of the considered path choice models (2.3) are described. In **Section 3**, which concerns stochastic network assignment, an extension based on the weibit model is reported, particularly a loading procedure that does not require explicit path enumeration. In **Section 4**, some results obtained by carrying out tests both on a simple network (4.1) and on a real network (4.2) are presented. In this last case, results are compared by considering the performances obtained on a real system. Finally, **Section 5** contains a summary of the obtained results and some indications for further developments.

2 MODELING PATH CHOICE BEHAVIOR

2.1 Requirements for Path Choice Models

Following (Cantarella et al., 2020), some requirements useful in classifying path choice models are, for the convenience of the reader, summarized in the following, where the classification of the requirements is carried out both from a mathematical and modeling point of view. It is worth noting that these requirements hold whatever is the theory behind the path choice models. In this paper, path choice models derived from the theory of random utilities will be considered.

2.1.1 Mathematical Requirements

Under the assumptions of linear utility functions, mathematical requirements allow effectively modeling any choice behavior. Considering, in particular, a path choice model, the main requirements are the continuity and monotonicity of the utility function. Then, the model can be specified by a function if the values and random residuals of the perceived utility are assumed distributed as continuous random variables with a nonsingular covariance matrix. In order to guarantee that small changes of path costs induce small changes of choice probabilities, the *continuity* of the path choice model must be assured (note that if the model is also differentiable, the Jacobian is continuous). This feature, assured by commonly used joint probability density functions, guarantees the continuity of the resulting arc flow function. Thus, it is useful to state the existence of SUE. The assurance that an increase in cost of a path corresponds to a decrease of its choice is given by the *monotonicity* of the path choice function. More generally, the path choice function should be nonincreasing monotone with respect to path costs. This feature guarantees the monotonicity of the resulting arc flow function. Thus, it is useful to state the uniqueness of the solution of SUE. If any change of the scale of the utility does not affect the model, the choice model has the *independence from linear transformations of utility* property.

2.1.2 Modeling Requirements

Modeling requirements are also useful to effectively simulate path choice behavior. Considering the *similarity of perception of partially overlapping paths* allows us to avoid counter-intuitive results. In the case of two partially overlapping paths, a positive covariance between them can simulate their similarity because they are likely not perceived as two totally separated paths. This covariance can be specific to the distribution (e.g., Probit and Gammit) or can be induced through modifications of utilities (e.g., C-Logit, path-size Logit, and C-Weibit; the latter is introduced later in this article).

Considering the network model, if an arc can be divided into subarcs, redefining arc costs so that path costs are not affected, this does not affect the perceived utility or random residual distribution of paths and, consequently, choice probabilities. Because this requirement makes reference to the features of the distribution of the sum of random variables, two approaches can be identified to specify perceived utility distribution: 1) direct formulations of probabilistic path choice models, where the distribution of path-perceived utility or random residuals is explicitly specified, and 2) indirect formulations, where path-perceived utility or random residual distribution is specified as a linear combination of arc-perceived utilities (specifying their distribution), even though the analysis of path choice behavior is still carried out at the path level. The independence from arc segmentation mainly requires the use of reproductive random variables. For instance, the sum of several independently distributed normal random variables is still a normal random variable, with the mean given by the sum of the means and variance by the sum of the variances. This feature is also shown by independently distributed gamma random variables with the same variance-to-mean ratio. In both cases, if the arcs in a path are further segmented, provided that the mean path cost and the variance are not affected by segmentation, the resulting path-perceived utility distribution (and then the choice probability) is not affected by segmentation. Generally, the arc flow function with any arc-formulated choice model can be easily computed, if arc-perceived utilities are assumed independently distributed, through Monte Carlo techniques (introduced by Burrel, 1968; see also Sheffi, 1985). Traveling along a path, users perceive a cost that is usually considered a negative utility. The *negativity of perceived utility* assures that no user perceives a positive utility to travel along any path. This feature can be assured by assuming lower bounded random distributions (for instance log-normal or gamma). If this feature is not presented, it means that a nonelementary path may be a better choice than the elementary path within it, possibly leading to unrealistic situations (a part from some algorithmic drawbacks).

2.1.3 Notations

For the convenience of the reader, **Table 1** reports the symbols (grouped by type) used throughout the article (in any case, the meaning of each symbol is recalled each time it is used).

TABLE 1 | Symbols.

Symbol	Explanation
Paths	
p_k, j	Probability of choosing path k
K_j	Perceived choice set related to *od* pair *j*
N_k	Choice set cardinality
G_k, j	Disutility of path *k* in the *od* pair *j*
$\mathbf{G}_i$	Path-perceived utility vector for users of class *i*
g_k, j	Expected value of G_k, j
ϵ_k	Random residual
$\mathbf{V}$	Systematic utility vector
$\mathbf{U}$	Perceived utility vector
$\mathbf{h}_i$	Path flow vector for user class *i*
$\mathbf{v}_i$	Path systematic utility for user class *i*
$\mathbf{p}_i$	Path choice probability vector for user class *i*
H_r	Path elongation ratio
Δ	Path cost multiplier
Arcs	
σ_a^2	Variance of the cost on arc *a*
$\bar{c}_a$	Reference cost of arc *a*
c_a	Arc *a* cost
w_a	Arc *a* disutility
M	Expected value of arc cost
Σ	Standard deviation of arc cost
c_{0a}	Free flow arc cost
$\mathbf{F}$	Arc flow vector
$\mathbf{f}^*$	Equilibrium arc flow vector
$\mathbf{c}^*$	Equilibrium arc cost vector
S_f	Feasible arc flow set
Ω_{uv}	Binary variable (equal to 1 if arc *uv* belongs to a reasonable path; 0, otherwise)
$W(a)$	Arc *a* weight
Nodes	
L_r	Ordered list of nodes
$FS(u)$	Forward star of node *u*
$BS(u)$	Backward star of node *u*
Others	
Θ	Gumbel scale parameter
Φ	Euler's constant
α_k^j	Weibull scale parameter
β_j	Weibull shape parameter
ξ_j	Weibull location parameter
$\Gamma(.)$	Gamma function
ξ_j^0	Estimated value of parameter ξ_j
$\hat{\xi}_j$	Estimated value of parameter ξ_j
$\mathbf{\Sigma_g}$	Covariance matrix
Cv	Variation coefficient
$G(N, A)$	Graph
N	Set of nodes
A	Set of arcs
$\mathbf{d}_i$	Demand flow for user class *i*

2.2 Random Utility Path Choice Models

Given an origin/destination (o/d) pair *j*, the analyst evaluates the probability $p_{k,j}$ of choosing path *k* belonging to the perceived choice set of paths K_j. Disutility $G_{k,j}$ can be expressed as follows:

$$G_{k,j} = g_{k,j} + \epsilon_k \tag{1}$$

where $g_{k,j} = E[G_{k,j}]$ is the expected value of the disutility (cost) of path *k* and ϵ_k is the random residual.

Possible models for the evaluation of probabilities (deriving from the hypotheses on random variable $G_{k,j}$ distribution) are described in the following. Even if they can be classified on the basis of the requirements described in **Section 2.1**, for the sake of simplicity, description is here conducted on the basis of the existence (or not) of a closed form to define choice probabilities.

2.2.1 Closed-Form Probability Formulation

Multinomial Logit and Weibit model assumptions entail that the covariance matrix of the joint distribution of the random residuals (and of the utilities) is a diagonal matrix with nonzero entries. From a mathematical point of view, Multinomial Logit and Weibit models satisfy the *continuity* and *monotonicity* requirements, but *independence from linear transformations of utility* requirements are not satisfied. From a modelistic point of view, the requirements regarding the *similarity of perception of partially overlapping paths* are not satisfied because of the general structure of the covariance matrix. In addition, *independence from arc segmentation* is not assured, and perceived utility distribution is specified *in path*. The *negativity of perceived utility* requirement is formally satisfied by multinomial weibit but not by multinomial logit because random residuals, following a Gumbel distribution, can assume negative values (that is a positive disutility), which implies a decrease in path cost.

2.2.1.1 Multinomial Logit

In the Multinomial Logit model, it is assumed that the path costs $G_{k,j}$ are identical and independently distributed as a *Gumbel* random variable.

In particular, utility can be expressed as $G_{k,j} = -(g_{k,j} + \epsilon_k)$, where $-g_{k,j} = E[G_{k,j}]$ is the expected value of the utility (cost) of path *k*, and random residuals -ϵ_k are independently and identically distributed (i.i.d.) as a Gumbel random variable of zero mean and scale parameter θ (Ben-Akiva and Lerman, 1985; Domencich and McFadden, 1975).

The marginal probability distribution function of each random residual can be written as follows Cascetta (2009):

$$F_{\epsilon_k}(x) = exp[-exp(-\theta x - \Phi)] \quad (2)$$

where Φ is the Euler's constant. For multinomial logit, it is assumed that the mean and variance of the Gumbel distribution are, respectively, as follows:

$$E[\epsilon_k] = 0 \quad ; \quad Var[\epsilon_k] = \frac{\pi^2}{6 \cdot \theta^2} \qquad \forall j \quad (3)$$

The specification of probability for Multinomial Logit is as follows:

$$p_{k,j} = \frac{exp\left(-\theta_j \cdot g_{k,j}\right)}{\Sigma_{h \in K_j} exp\left(-\theta_j \cdot g_{h,j}\right)} \quad (4)$$

where θ_j is the distribution parameter related to the choice set K_j.

2.2.1.2 Multinomial Weibit

The Weibit model assumes that the path costs $G_{k,j}$ are independently distributed as a $Weibull(\xi_j, \alpha_k^j, \beta_j)$ random variable, and its probability density function is as follows:

$$f(x) = \frac{\beta_j}{\alpha_k^j}\left(\frac{x - \xi_j}{\alpha_k^j}\right)^{(\beta_j - 1)} \cdot exp\left[-\left(\frac{x - \xi_j}{\alpha_k^j}\right)^{(\beta_j)}\right] \quad (5)$$

where

- α_k^j is a scale parameter;
- β_j is a shape parameter;
- ξ_j is a location parameter.

The mean and variance are, respectively, as follows:

$$E\left[G_{k,j}\right] = \xi_j + \alpha_k^j \cdot \Gamma\left(1 + \frac{1}{\beta_j}\right) = g_{k,j} \quad (6)$$

$$Var\left[G_{k,j}\right] = \left(\alpha_k^j\right)^2\left[\Gamma\left(1 + \frac{2}{\beta_j}\right) - \Gamma^2\left(1 + \frac{1}{\beta_j}\right)\right] \quad (7)$$

where $\Gamma(.)$ is the gamma function.

The specification of probability for Multinomial Weibit is (Castillo et al., 2008) as follows:

$$p_{k,j} = \frac{\left(g_{k,j} - \xi_j^0\right)^{-\beta_j}}{\Sigma_{h \in K_j}\left(g_{h,j} - \xi_j^0\right)^{-\beta_j}} \quad g_{k,j} \geq \xi_j^0 \quad \forall k \in K_j \quad (8)$$

where ξ_j^0 is an estimated value of parameter ξ_j.

2.2.2 Not Closed-Form Probability Formulation

From a mathematical point of view, Probit and Gammit models satisfy the *continuity*, *monotonicity*, and *independence from linear transformations of utility* requirements. From a modelistic point of view, the requirements regarding the *similarity of perception of partially overlapping paths* are satisfied because they allow for the general structure of the covariance matrix. In addition, *independence from arc segmentation* is assured, and perceived utility distribution is specified *in arc*. Because a closed form is not available for the choice behavior model, an unbiased estimate of arc flows can be computed through a Monte Carlo technique by successively averaging several loading to the shortest paths (Sheffi, 1985; Burrel, 1968). The Monte Carlo technique is used as a numerical tool to compute the path choice probabilities or to enhance the corresponding arc flows. To apply this approach, by virtue of *independence from arc segmentation*, the path-perceived utility distribution is derived from independently distributed arc random costs. For practical purposes, to apply the Probit or the Gammit model, path-perceived utilities $\mathbf{G}_i$ can be specified through arc utilities $\mathbf{g}$, whose expected values are the opposite of arc costs, $E[\mathbf{g}] = -\mathbf{c}$, say: $\mathbf{G}_i = \mathbf{B}_i^T\mathbf{g}$, where $\mathbf{B}_i$ is the arc-path incidence matrix for user class i. If the arc-perceived utilities are independently Normal or Gamma distributed with diagonal covariance matrix $\mathbf{\Sigma_g}$, the resulting path-perceived utilities are still Normal or Gamma distributed with covariance matrix $\mathbf{\Sigma} = \mathbf{B}_i^T\mathbf{\Sigma_g}\mathbf{B}_i$ with nonnull covariance for each pair of partially overlapping paths. *Negativity of perceived utility* is assured only for gammit because, for probit, random residuals, following a normal distribution, can assume negative values (that is a positive disutility), which implies a decrease in path cost.

2.2.2.1 Probit

The Probit model or multinomial probit model (MNP) results when the random residuals in **Eq. 1** are assumed to be multivariate normally distributed with zero mean and arbitrary covariance matrix (Daganzo, 1983). The utility vector $\mathbf{U}$ (index of o/d pair is omitted to simplify notation) of dimension N_k is therefore $MVN(\mathbf{V}, \mathbf{\Sigma})$, and its probability density function is therefore the following:

$$f(\mathbf{V}, \mathbf{\Sigma}) = \left[(2\pi)^{N_k} \cdot |\mathbf{\Sigma}|\right]^{-\frac{1}{2}} exp\left[-\frac{1}{2}(\mathbf{U} - \mathbf{V})^T\mathbf{\Sigma}^{-1}(\mathbf{U} - \mathbf{V})\right] \quad (9)$$

where $\mathbf{V}$ is the systematic utility vector.

With the MNP model, the choice probability of path k belonging to the perceived choice set of paths K_j of cardinality N_k is expressed by the following:

$$p_{k,j} = \int_{U_1 < U_k} \ldots \int_{U_k = -\infty}^{\infty} \ldots \ldots \int_{U_{N_k} < U_k} \frac{exp\left[-\frac{1}{2}(\mathbf{U} - \mathbf{V})^T\mathbf{\Sigma}^{-1}(\mathbf{U} - \mathbf{V})\right]}{\left[(2\pi)^{N_k} \cdot |\mathbf{\Sigma}|\right]^{1/2}} dU_1 \ldots dU_{N_k} \quad (10)$$

2.2.2.2 Gammit

The Gammit model is obtained by assuming that perceived disutilities are jointly distributed as a nonnegative "shifted" multivariate gamma, with the mean equal to the path costs $g_{k,j}$ and path covariance matrix $\mathbf{\Sigma}$. Path formulations of the gammit model may be undetermined because the covariance matrix alone does not completely define the joint probability density function of a multivariate gamma random variable. The adopted formulation was proposed by Cantarella and Binetti (2002). It is an arc formulation that overcomes this drawback and, at the same time, assures independence of arc segmentation and rule out positive perceived utilities for any paths. As in Cantarella and Binetti (2002), let $\bar{c}_a$ be the (strictly positive) reference cost on arc

a. It is assumed that $0 < \bar{c}_a \le c_a$, with c_a being the arc cost. The perceived disutility of arc a ($-w_a$) is assumed distributed as a non-negative shifted gamma variable with the mean given by the arc cost, c_a, variance by $\sigma_a^2 = \theta \bar{c}_a$, proportional to reference arc costs $\bar{c}_a > 0$, and shifting factor obtained as the difference between the arc cost and reference cost, $(c_a - \bar{c}_a)$:

$$-w_a \sim (c_a - \bar{c}_a) + Gamma\left(\alpha_a = \frac{\bar{c}_a}{\theta},\ \beta = \theta\right) \tag{11}$$

where α_a and β are, respectively, the shape parameter and the rate parameter of the gamma function.

This means that the arc-perceived disutility is the sum of a nonnegative deterministic term that can depend on arc flows and a nonnegative stochastic term independent from arc flows. The assumption on arc reference costs yields that the corresponding reference cost $\bar{g}_k$ on path k is strictly positive and not greater than the path cost, $0 < \bar{g}_k \le g_k$. Finally, the perceived disutility on path k is marginally distributed as a nonnegative shifted gamma variable:

$$-u_{k,j} \sim \left(g_{k,j} - \bar{g}_{k,j}\right) + Gamma\left(\alpha_{k,j} = \frac{\bar{g}_{k,j}}{\theta},\ \beta = \theta\right) \tag{12}$$

2.3 Implementation of Path Choice Models

This section focuses on some operational considerations regarding the implementation of the considered path choice models. At the end, an application to a simple test network is conducted in order to compare the performances of the considered models. The parameters used to perform the application depend on the considered model. Then, in this section, the parameters needed to implement each model are identified.

2.3.1 Logit

Logit models are based on the Gumbel distribution and, considering **Section 2.2.1.1**, are characterized only by the parameter θ, which is estimated starting from the variance of path costs G_k as in **Eq. 3**. In particular, if $Sdev\,[G_k]$ is the standard deviation of path cost, a variation coefficient Cv can be introduced as follows:

$$Cv = \frac{Sdev[G_k]}{E[G_k]} \tag{13}$$

It is possible to define the parameter of the Gumbel distribution from **Eq. 3** as follows:

$$Var[G_k] = (Cv \cdot E[G_k])^2 \quad \Rightarrow \theta = \frac{\pi}{\sqrt{6} \cdot Cv \cdot E[G_k]} \tag{14}$$

Thus, for the implementation of a logit model, the independent variable to be assumed is Cv.

2.3.2 Weibit

Weibit models are based on a Weibull distribution characterized by three parameters ξ_j, α_k^j, β_j (see **Section 2.2.1.2**).

An estimate of the parameter ξ_j ($\hat{\xi}_j$), suggested in Castillo et al. (2008), is the minimum possible travel time ξ_j^0 associated with the o/d pair j, that is, $\hat{\xi}_j = \xi_j^0$.

Starting from **Eq. 6**, parameter α_k^j can be written as follows:

$$\alpha_k^j = \frac{g_{k,j} - \hat{\xi}_j}{\Gamma\left(1 + \frac{1}{\beta_j}\right)} \tag{15}$$

and replacing **Eq. 15** in **Eq. 7** yields the following:

$$Var[G_k] = \left(\frac{g_{k,j} - \hat{\xi}_j}{\Gamma\left(1 + \frac{1}{\beta_j}\right)}\right)^2 \left[\Gamma\left(1 + \frac{2}{\beta_j}\right) - \Gamma^2\left(1 + \frac{1}{\beta_j}\right)\right] \tag{16}$$

Once the value of variance is assumed, the value of β_j can be calculated from **Eq. 16**. Once β_j is known, replacing its value in **Eq. 15**, it is possible to compute the value of α_k^j. One example of the procedure for the estimation of the value of β_j is shown in Algorithm 1, where the secant method, which does not require the calculation of derivatives, is used.

Referring to Algorithm 1, the initialization involves to set: 1) the variables b_0 and b_1, 2) the threshold value ι to stop the algorithm, and 3) the initial value of objective function F_n. The procedure is iterated while the objective function F_n is less than ι. At each iteration, a value of β_j is obtained, is calculated the variance corresponding to such a value and when the difference with assumed variance is less than ι the procedure is stopped.

Algorithm 1. Procedure for the estimation of the Weibull distribution parameters.

```
procedure WEIBETA((ξ, Cost, Var))
    ι ← 1e − 9
    b0 ← 0.2
    b1 ← 0.1
    Fn ← 1
    while Fn < δ do
        α0, V0 ← VARWEIBULL(ξ, b0, Cost)
        α1, V1 ← VARWEIBULL(ξ, b1, Cost)
        F0 ← V0 − Var
        F1 ← V1 − Var
        βn ← b1 − ((β1 − β0)/(F1 − F0)) ∗ F1
        αn, Vn ← VARWEIBULL(ξ, bn, Cost)
        Fn ← Vn − Var
        β0 ← β1
        β1 ← βn
    end while
    return βn
end procedure
function VARWEIBULL((ξ, β, Cost))
    α ← ((Cost − ξ)/Γ(1 + (1/β)))
    g1 ← Γ(1 + (2/β))
    g2 ← Γ(1 + (1/β))
    Var ← α^2 ∗ (g1 − g2^2)
    return α, Var
end function
```

For a practical purpose, owing to a numerical problem that can occur if the difference $(g_{k,j} - \xi_j^0)$ is equal to zero, that is, k is the minimum path, a fraction of the minimum cost is considered by means of a multiplier $\delta < 1$, so that:

$$\hat{\xi}_j = min\left(g_{i,j}\right) \cdot \delta \quad \forall i \in od_j \tag{17}$$

Because weibit variance depends, by means of parameter α_k^j, on the value of path cost, in order to estimate a value of β_j common for all paths making up the choice set of o/d pair j, the arithmetic average of the costs of the paths belonging to the choice set has been considered an estimate of expected value $E[G_k]$, so that:

$$Var[G_k] = \left(Cv \cdot \frac{1}{|K_j|} \sum_{k \in K_j} g_{k,j} \right)^2 \quad (18)$$

where K_j is the choice set related to the o-d pair j.

Thus, for the implementation of a weibit model, the independent variables to be assumed are Cv and δ.

2.3.3 Probit

As discussed in 2.2.2, the path-perceived utility distribution is derived from independently distributed arc random costs that, for the Probit model, follow a normal distribution characterized by the two parameters μ and σ. For a specific case, the expected value and standard deviation can be expressed as follows:

$$\mu = c_a \quad ; \quad \sigma = Cv \cdot c_{0a} \quad (19)$$

where

c_a is the arc cost;
c_{0a} is the free flow arc cost;
Cv is the variation coefficient.

As discussed above, the Monte Carlo technique is implemented by successive averaging several loading to the shortest paths, so a Nit number of samples have to be carried out. Thus, for the implementation of a probit model, the independent variables to be assumed are Cv and Nit.

2.3.4 Gammit

As discussed in 2.2.2, the path-perceived utility distribution is derived from independently distributed arc random costs that, for the Gammit model, follow a gamma distribution. Following the formulation adopted in 2.2.2.2, a gamma distribution is based on two parameters, which are the following: $\alpha > 0$ (shape parameter) and $\beta > 0$ (rate parameter), and the expected value and variance can be expressed as follows:

$$E[X] = \alpha / \beta \quad ; \quad Var[X] = \alpha / \beta^2 \quad (20)$$

For a specific case, the expected value and variance can be expressed as follows:

$$E[X] = c_a \quad ; \quad Var[X] = (Cv \cdot c_{0a})^2 \quad (21)$$

where

c_a is the arc cost;
c_{0a} is the free flow arc cost;
Cv is the variation coefficient.

Thus, the distribution parameters can be estimated as follows:

$$\alpha = \frac{(c_a)^2}{(Cv \cdot c_{0a})^2} \quad ; \quad \beta = \frac{c_a}{(Cv \cdot c_{0a})^2} \quad (22)$$

In addition, in this case, the Monte Carlo technique is implemented by successive averaging several loading to the shortest paths, so a Nit number of samples have to be carried out. Thus, for the implementation of a gammit model, the independent variables to be assumed are Cv and Nit.

3 STOCHASTIC NETWORK ASSIGNMENT

Transportation supply models express how user behavior affects network performances. They are usually based on congested network models, that is, a graph G (N, A) with a transportation cost c_a and a flow f_a associated to each arc a in set A. Let $\mathbf{B}_i$ be the arc-path incidence matrix for user class i with entries $b_{ak} = 1$ if arc a belongs to path k, $b_{ak} = 0$ otherwise; $\mathbf{d}_i \geq 0$ be the demand flow for user class i; $\mathbf{h}_i \geq 0$ be the path flow vector for user class i, with entries h_k, $k \in K_i$; $\mathbf{f} \geq 0$ be the arc flow vector, with entries f_a, $a \in A$; $\mathbf{c}$ be the arc cost vector, assumed below with nonnegative entries c_a, $a \in A$; $\mathbf{g}_i \geq 0$ be the path cost vector for user class i, with entries w_k, $k \in K_i$; the following three equations completely describe the transportation supply:

$$\mathbf{f} = \Sigma_i \mathbf{B}_i \mathbf{h}_i \quad (23)$$

$$\mathbf{c} = \mathbf{c}(\mathbf{f}) \quad (24)$$

$$\mathbf{g}_i = \mathbf{B}_i^T \mathbf{c} \quad \forall i \quad (25)$$

The function in **Equation 24** is defined as the *arc cost function*. Path choice behavior can be simulated by assuming that users' perception of path costs, for each user class, can be expressed by the perceived utility vector U_i modeled as a random variable given by the sum of the expected value, or systematic utility, v_i and a random residual following the random utility theory. The cost attributes associated to a path allow specifying the path systematic utility:

$$\mathbf{v}_i = -\mathbf{g}_i \quad (26)$$

The probability of choosing path k for user class i is given by the probability of path j being the maximum perceived utility one. Hence, the choice probability vector $\mathbf{p}_i$ depends on the systematic utility (and the parameters of the random residual distribution):

$$\mathbf{p}_i = \mathbf{p}_i(\mathbf{v}_i) \quad (27)$$

Path flows are thus:

$$\mathbf{h}_i = \mathbf{d}_i \mathbf{p}_i \quad (28)$$

A probabilistic choice model, derived from the random utility theory, specifies **Equation 27**. Thus, relation $\mathbf{p}_i$ ($\mathbf{v}_i$) is a function.

Combining the above equations yields the path-flow model:

$$\mathbf{h}_i = \mathbf{d}_i \mathbf{p}_i(\mathbf{g}_i) \quad (29)$$

Under mild assumptions, utility distribution parameters (apart the mean) do not depend on systematic utility values (invariant or additive choice models), and this function can be proved monotone increasing with symmetric positive semidefinite Jacobian with respect to path systematic utilities (Cantarella, 1997).

All usually adopted probabilistic choice functions give strictly positive probabilities and are continuous and continuously differentiable with respect to systematic utility. Moreover, if the parameters of the perceived utility pdf do not depend on systematic utility values, the resulting choice function, called invariant, is monotone increasing with respect to systematic utility with symmetric (semidefinite positive) Jacobian (Cantarella (1997); Cascetta (2009)) and choice probabilities depend on differences between systematic utility values only. The (stochastic) arc flow function with a constant demand is obtained by combining supply model (23) and (25) with the path-flow model (29):

$$\mathbf{f} = \Sigma_i \mathbf{d}_i \mathbf{B}_i \mathbf{p}_i \left(\mathbf{B}_i^T \mathbf{c}\right) \tag{30}$$

that gets values in the feasible arc flow set S_f that is nonempty (if the network is connected), compact (since closed and bounded, if only elementary paths are considered), and convex. The arc flow function (30) is a general model of stochastic assignment to uncongested networks, or SUN for short, and the solution of the SUN depends on the considered choice model. In the case of the logit and weibit family of choice models, either path choice set should be explicitly defined (path enumeration), or for a logit family of path choice model, an implicit procedure such as Dial's algorithm (Dial, 1971) can be adopted. In general, considering the probit or gammit family of choice models, the computation of the arc flow function (30) requires the well-known Monte Carlo algorithm (Cascetta, 2009). In the following of the paper, some original extensions of the SUN procedure based on path costs following a Weibull distribution are described. Explicit path enumeration can be avoided by considering existing algorithms. Dial's algorithm (Dial, 1971) is one of the most effective and popular procedures for a logit-type stochastic traffic assignment because it does not require path enumeration over a network. Leaving out the problem associated to the definition of "efficient paths", which sometimes produces unrealistic flow patterns (Bell, 1995; Leurent, 2005; Si et al., 2010), the attention is here focused on a way to compute the weight associate to each arc considering path costs following a Weibull distribution.

As stated in Castillo et al. (2008), *assuming that the costs are independent Weibull, the variance for different costs is different. The variance is, however, functionally dependent on the location parameter so that higher mean cost results in higher variance of the cost. This property may seem quite natural in many applications. But because of the functional dependence, it is not possible to choose the variance independent of the mean for the different choice alternatives. So by representing the costs in a suitable logarithmic form in a logit model or in linear form in a Weibull model lead to the same choice probabilities.*

Considering Leurent (2005), in a similar way, it is possible to define the impedance $A(a)$ to be associated to arc $a = \{u, v\}$ from the expression of probability (8) as follows:

$$A(a) = \left(g_{k,j} - \xi_j^0\right)^{-\beta_j} \tag{31}$$

To evaluate arc impedance, it is necessary to define a path containing the arc $a = \{u, v\}$ in order to estimate the values of parameters β and α; a path is chosen including arc $a = \{u, v\}$ between the shortest path connecting origin r with the tail node u and shortest path connecting the head node v with destination s. Thus, the value of the path cost to be considered in computing impedance is given by $g_{k,j} = C_r(u) + c_{uv} + C_s(v)$, where $C_r(u)$ is the cost of the shortest path from origin r to node u, c_{uv} is the cost of arc $a = \{u, v\}$, and $C_s(v)$ is the cost of the shortest path from node v to destination s. Considering the procedure introduced in point 2.3.2, introducing a variation coefficient $Cv = \sigma/\mu$, the variance can be estimated as $\sigma^2 = (Cv \cdot g_{k,j})^2$ and ξ_0 is estimated as a fraction of the cost of the shortest path connecting o/d pair rs. Once the variance is known, it is possible to estimate β and α, as seen in 2.3.2, and to calculate arc impedance $A(a)$ following **Equation 31**. Thus, once the shortest paths from origin node r to all nodes n, yielding the reference access costs $C_r(n)$, and the shortest paths from destination node s to all nodes n, yielding the reference access costs $C_s(n)$ (based on the reference arc travel costs c_{uv}), are computed, the adaptation of the steps of the loading procedure (Dial, 1971; Leurent, 2005), referred to o/d pair rs, can be expressed as shown in Algorithm 2, where H_r is the elongation ratio (a cost multiplier) and Ω_{uv} is a binary variable (equal to 1 if arc uv belongs to a reasonable path; 0, otherwise), introduced in order to obtain efficient paths in network loading (Leurent, 2005).

The weibit network loading summarized in Algorithm 2 allows obtaining the arc flow without the explicit path enumeration. In step 0, after the variable initialization and established an origin r and a destination s, the impedance of each arc a is evaluated. Moreover, the belonging of the arc to a reasonable path is evaluated. Step 1 consists of a forward pass where the weight of each arc is calculated. Finally, step 2 consists of a backward pass allowing the calculation of the flow on each arc.

Algorithm 2. Implicit weibit loading procedure.

```
Step 0 - Reasonable paths, initialization of arc flow and impedance.

• Define the ordered list L_r of the nodes n in the order of increasing access cost from r.
• Estimation of parameter of Weibull distribution

Var ← (Cv * C_r(s))^2                              ▷ compute variance of path cost
ξ_0 ← ε * C_r(s)                                   ▷ compute an estimate of parameter ξ_0
β ← WeiBeta(ξ_0, C_r(s), Var)                      ▷ compute an estimate of parameter β
for each link a = {u, v} do
    X_A(a) ← 0
    C_a ← C_r(u) + c_uv + C_s(v)                   ▷ compute the cost of the path crossing link a
    A(a) ← (C_a − ξ_0)^(−β)                        ▷ compute the impedance of the arc
    if (1 + H_r) · (C_r(v) − C_r(u)) ⩾ c_uv > 0 then
        Ω_uv = 1
    else
        Ω_uv = 0
    end if
end for
Step 1. Forward pass.
for each link a = {u, v} do                        ▷ Initialize the arc and node weights
    W_A(a) ← 0
    W_N(u) ← 0
    W_N(v) ← 0
end for
W_N(r) ← 1
for u ∈ L_r taken in increasing order do
    for v ∈ FS(u) do
        if Ω_uv = 1 then
            W_A(a) ← A(a) · W_N(u)                 ▷ compute the weight of the arc
            W_N(v) ← W_N(v) + W_A(a)               ▷ update the weight of the node
        end if
    end for
end for
Step 2. Backward pass.
for each node n do
    X_N(n) ← 0
end for
X_N(s) ← d_rs                                      ▷ Initialize flow entering the destination node s
for v ∈ L_r taken in decreasing order do
    for u ∈ BS(v) do
        if Ω_uv = 1 then
            X_A(a) ← X_N(v)·W_A(a) / W_N(u)
            X_N(u) ← X_N(v) + X_A(a)
        else
            X_A(a) ← 0
        end if
    end for
end for
```

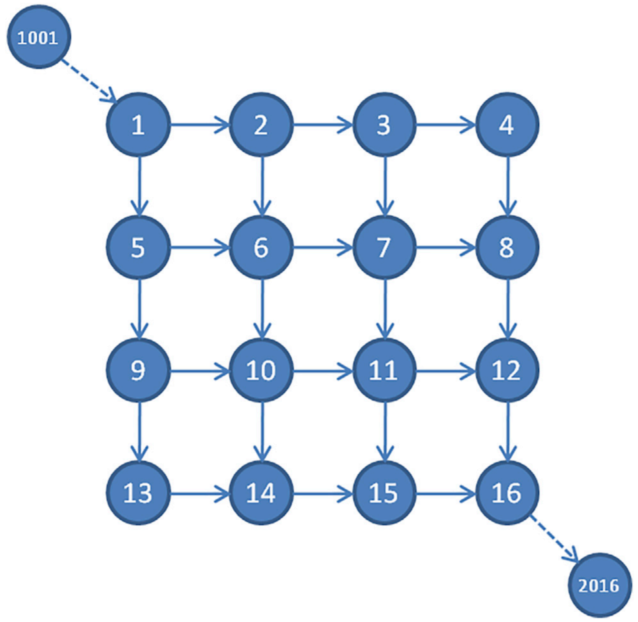

FIGURE 1 | Considered test network.

The path choice models here considered are also used to model congested situations taking into account the variability of costs as described in the following.

SUE assignment can effectively be expressed by fixed-point models given by the arc flow function (30) and the arc cost function **Eq. 24**:

$$\mathbf{f}* = \mathbf{f}(\mathbf{c}*,\ \mathbf{d}) \quad \in S_f \tag{32}$$

$$\mathbf{c}* = \mathbf{c}(\mathbf{f}*) \quad \in \mathbf{c}\left(S_f\right) \tag{33}$$

where S_f is the feasible arc flow set.

Algorithms based on the method of successive averages (MSA) (Cantarella, 1997) and (Cascetta, 2009) are the most used ones to solve fixed-point models for SUE assignment because they can accommodate any choice model based on random utility theory and are suitable for large-scale applications.

Their basic iteration requires the computation of the cost function **Eq. 24** to get arc costs from arc flows and the computation of the arc flow function **Eq. 30** to get arc flows from arc costs.

Applying the MSA leads to the MSA-FA solution algorithm based on the following recursive equation:

$$\mathbf{f}^k = \mathbf{f}^{k-1} + \alpha\left(\mathbf{f}\left(\mathbf{c}\left(\mathbf{f}^{k-1}\right) - \mathbf{f}^{k-1}\right)\right). \tag{34}$$

Convergence may be proved if the Jacobian of the arc cost function is symmetric (Cantarella, 1997).

4 EXPERIMENTS

4.1 Application to a Test Network

The first experiment on a test network was carried out with the aim to compare the performance of the considered models in terms of path choice probabilities in a SUN context. The considered network was a 4 × 4 square network, and in order to simplify the interpretation of the results, only one o/d pair was considered. It is depicted in **Figure 1**, and the characteristics of the arcs are shown in **Table 2**. Paths for the o/d pair 1001–2016 were considered, and the list of paths is shown in **Table 3**, where the first six paths in terms of cost are considered. The choice of the number of considered paths derives from considerations on the number of paths potentially used by the users in order to maximize the coverage between the generated and used paths. From literature (e.g., Cascetta et al., 1996), the number of paths was less than 8. Moreover, the analyses reported in (Cascetta et al., 1996) highlight that the use of the first six paths is a reasonable compromise.

Tests were carried out by considering a set of values for the independent variables above defined for each considered model. Considered values of independent variables are shown in **Table 4**.

It is worth noting that *Cv* is used to compute the variance of: path costs → logit and weibit; arc costs → probit and gammit.

4.1.1 Models With Explicit Path Enumeration

The path set considered in this application is obtained with an explicit path enumeration and a selective approach. Then, among all available paths, the first k ones with respect to the path cost are considered. In this application, we set k = 6. Reported results are at first a comparison among choice probabilities obtained for the shortest path for each of

TABLE 2 | Characteristics of the arcs of the test network.

Arc	Length [M]	Speed [m/s]	Arc	Length [M]	Speed [m/s]	Arc	Length [M]	Speed [m/s]
1–2	100	2	6–7	125	2	11–12	100	2
1–5	105	2	6–10	125	2	11–15	105	2
2–3	100	2	7–8	100	2	12–16	100	2
2–6	120	2	7–11	125	2	13–14	105	2
3–4	100	2	8–12	100	2	14–15	105	2
3–7	120	2	9–10	115	2	15–16	105	2
4–8	100	2	9–13	105	2	1,001–1	100	5
5–6	115	2	10–11	125	2	16–2016	100	5
5–9	105	2	10–14	105	2			

TABLE 3 | Considered paths of the test network.

Path	Cost	Nodes
0	340.0	[1,001, 1, 2, 3, 4, 8, 12, 16, 2016]
1	350.0	[1,001, 1, 2, 3, 7, 8, 12, 16, 2016]
2	355.0	[1,001, 1, 5, 9, 13, 14, 15, 16, 2016]
3	360.0	[1,001, 1, 5, 9, 10, 14, 15, 16, 2016]
4	375.0	[1,001, 1, 2, 6, 7, 11, 12, 16, 2016]
5	380.0	[1,001, 1, 5, 6, 10, 11, 15, 16, 2016]

TABLE 4 | Considered values of the independent variables.

Model	Variable	Set of Considered Values
Common	*Cv*	[0.01, 0.05, 0.20, 0.50]
Weibit	δ	[0.995, 0.975, 0.925, 0.900]
Probit, Gammit	*Nit*	[50, 100, 500]

the considered models. Then, the probabilities obtained for each one of the six paths are compared. The explanation of the notation adopted in the figures for the considered model is the following:

- Logit → Logit model.
- WMin*yyy* → Weibit model where, to estimate β, variance is computed considering the minimum value of path costs and *yyy* indicates the value of δ (0.995, 0.975, 0.925, and 0.900).
- WMed*yyy* → Weibit model where, to estimate β, variance is computed considering the mean value of path costs and *yyy* indicates the value of δ (0.995, 0.975, 0.925, and 0.900).
- Prob*xxx* → Probit model where *xxx* indicates the number of carried out samples *Nit* (010, 050, 100, and 500).
- Gam*xxx* → Gammit model where *xxx* indicates the number of carried out samples *Nit* (010, 050, 100, and 500).

Comparing logit with weibit models, as shown in **Figure 2**, the lower the value of parameter δ is, the closer are weibit models to logit one.

Considering probit and gammit (**Figure 3**), their behavior is quite similar and a large number of iterations *Nit* is useful if the variance of costs, that is, the value of *Cv*, increases.

In **Figure 4** all models are compared considering WeibMed model with $\delta = 0.995$ and gammit and probit with $Nit = 500$. It can be seen how probit and gammit practically coincide and that the weibit model is quite closer to probit and gammit ones.

This last comparison has been conducted among all considered paths in **Figure 5**, and it can be seen that at the increase of variance, differences diminish.

4.1.2 Weibit Model Without Explicit Path Enumeration

In this case, tests have been carried out by considering a value of $\delta = 0.995$ and the set of variation coefficient $Cv = \{0.05, 0.20, 0.50\}$.

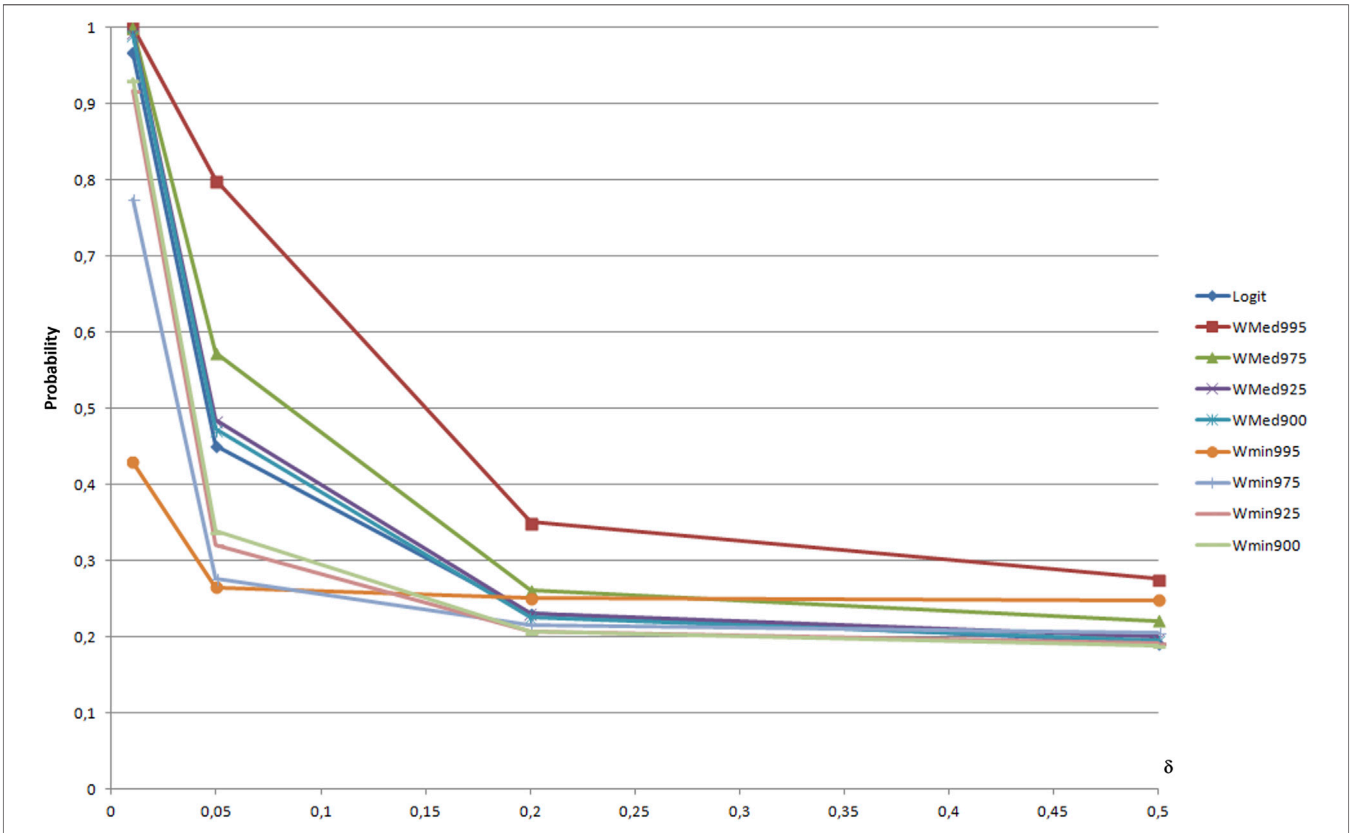

FIGURE 2 | Path choice probabilities depending on *Cv* for path 0: Logit vs. Weibit models.

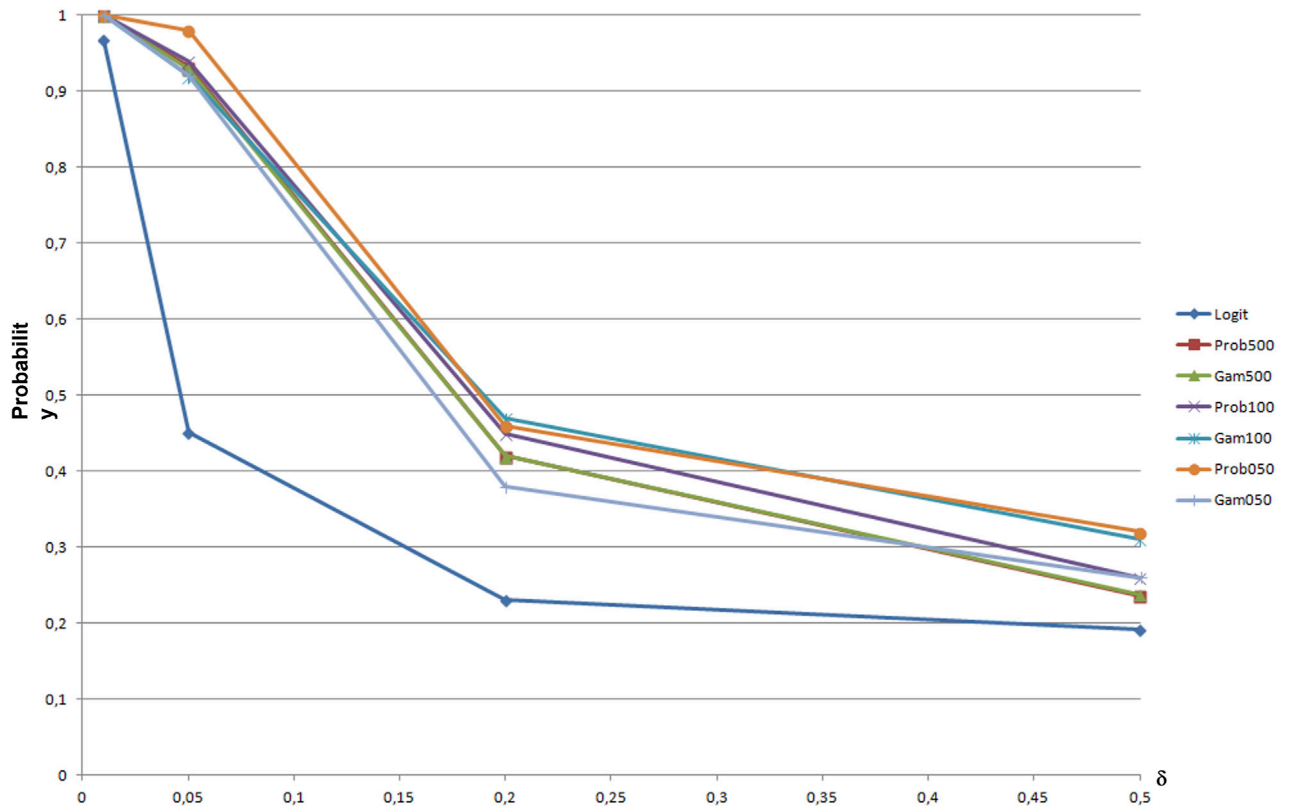

FIGURE 3 | Path choice probabilities depending on *Cv* for path 0: Logit vs. Probit and Gammit models.

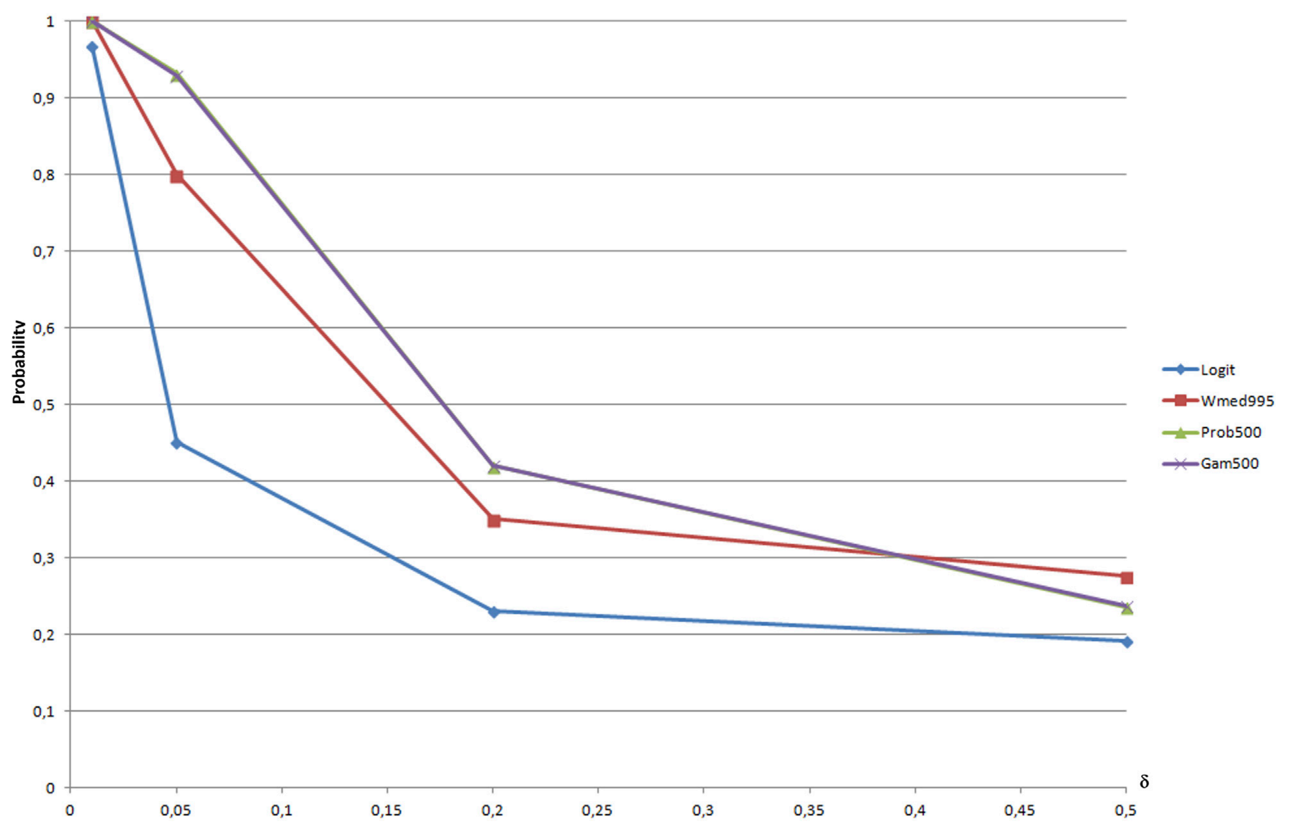

FIGURE 4 | Path choice probabilities depending on *Cv* for path 0: Logit vs. Weibit, Probit, and Gammit best models.

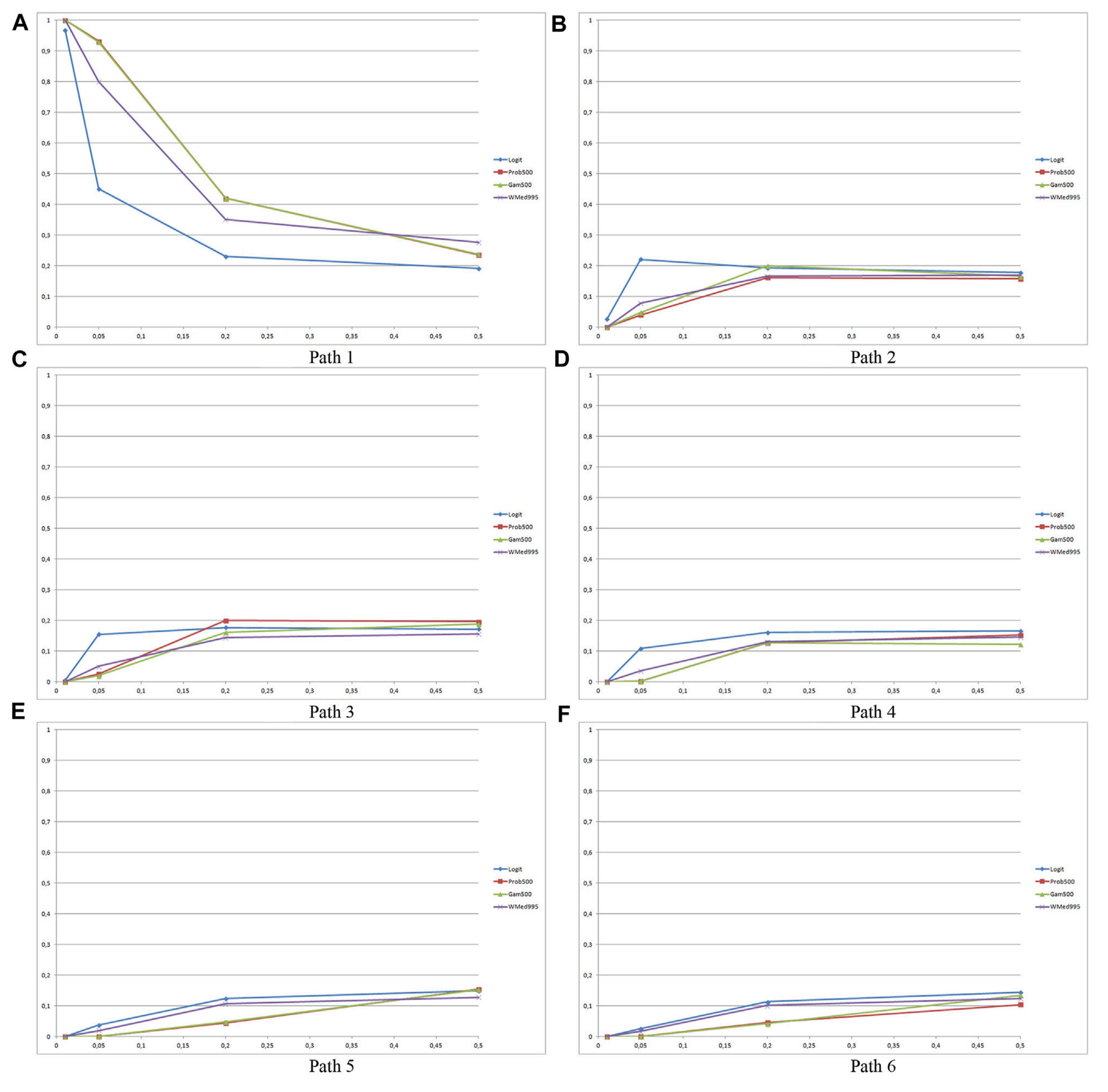

FIGURE 5 | Path choice probabilities depending on *Cv* for each path: Logit vs. Weibit, Probit, and Gammit best models.

A first attempt has been conducted to define the value of the cost multiplier H so that the algorithm considered the largest number of arcs in the network and to make a reasonable comparison between explicit and implicit versions of Weibull path choice. For the considered network, a value of H = 0.4 has been considered because, for lower values, only 50% of links result with flows greater than zero, as shown in **Figure 6**.

Results are shown by means of a scatter diagram (dots are connected by lines to improve readability) reported in **Figure 7**, where for each link indicated in the horizontal axis, the values of flows obtained for the two weibit models, with path enumeration (explicit) and without path enumeration (implicit).

To compare the values of flows obtained by the explicit Weibit model (f_i) with the implicit one ($\hat{f}_i$), the values of some statistical indicators have been considered: mean squared deviation (MSD), root-mean-square deviation (RMSD), normalized mean-square deviation (NMSD), and normalized root-mean-square deviation (NRMSD).

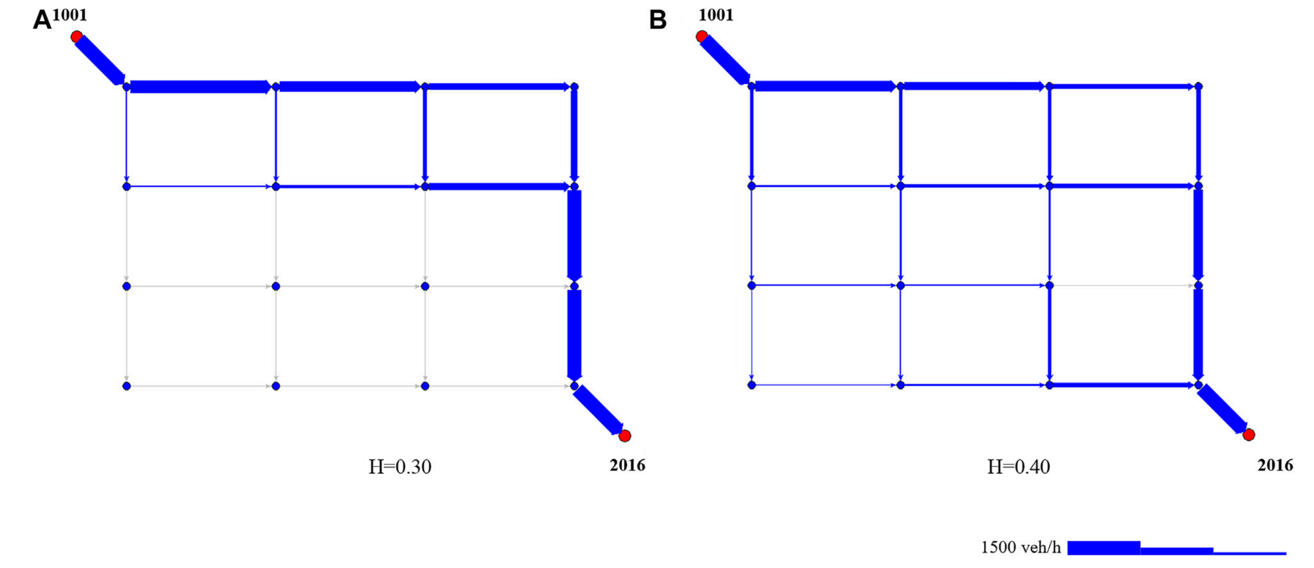

FIGURE 6 | Link considered in the implicit loading procedure for values of H =0.30 **(A)** and H =0.40 **(B)**.

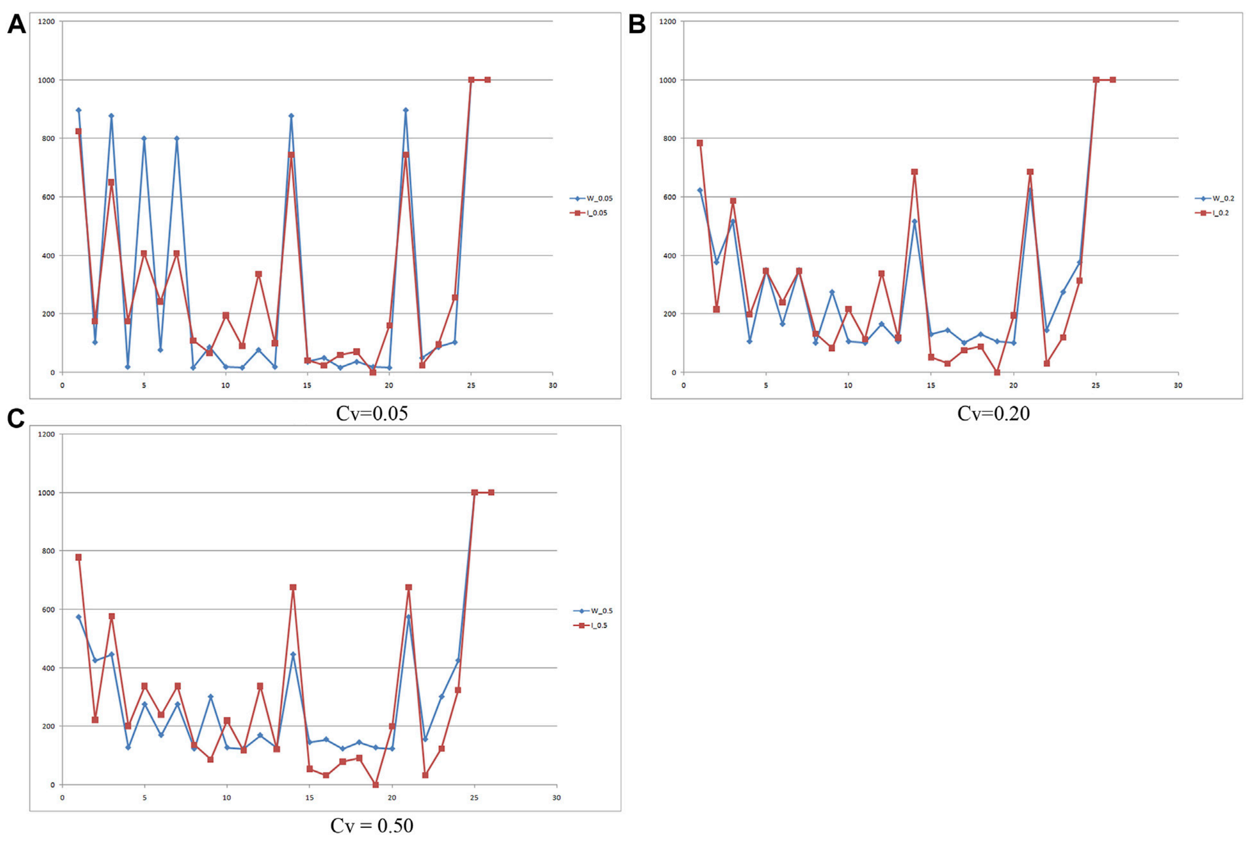

FIGURE 7 | Comparison of link flows for values of **(A-C)** Cv = 0.05, 0.20, 0.50 explicit vs. implicit weibit models.

$$MSD = \frac{1}{n}\Sigma_{i=1,n}\left(f_i - \hat{f}_i\right)^2 \quad (35)$$

$$RMSD = \sqrt{\frac{1}{n}\Sigma_{i=1,n}\left(f_i - \hat{f}_i\right)^2} \quad (36)$$

TABLE 5 | Comparison of flows explicit weibit vs. implicit weibit.

Cv	MSD	RMSD	NMSD	NRMSD
0.05	24,282.37	155.83	12.2641	3.5020
0.20	10,217.57	101.08	0.3079	0.5549
0.50	14,303.80	119.60	0.2711	0.5206

$$NMSD = \frac{1}{n}\Sigma_{i=1,n}\left(\frac{f_i - \hat{f}_i}{f_i}\right)^2 \tag{37}$$

$$NRMSD = \sqrt{\frac{1}{n}\Sigma_{i=1,n}\left(\frac{f_i - \hat{f}_i}{f_i}\right)^2} \tag{38}$$

Numerical values are reported in **Table 5**, where lower values indicate less residual variance.

4.2 Application to a Real Network

This section concerns the comparison of results obtained, in a real context, considering the above-described models applied in a SUE framework in order to explore the capability of reproducing counted flows. Application to a real network allows evaluating model performance on the field by comparing estimated flows with measured ones. Experiments here described have been conducted by considering the road network of Salerno, a town of about 130,000 inhabitants located in southern Italy. The road network has been schematized by means of a graph with 526 nodes and 1,147 arcs connecting 61 internal zones and 13 external ones. Comparisons have been performed by considering observed flows obtained by means of surveys conducted on 69 survey sections. In the adopted stochastic network loading, within the MSA procedure, for those models requiring explicit enumeration of paths, paths have been generated using De la Barra procedure (De La Barra et al., 1993). For each o/d pair, a maximum number of 2,000 iterations have been conducted, considering at most 10 paths not exceeding the 25% of the shortest paths free flow time. Reported results refer to an application considering the values of the independent variables shown in **Table 6**.

In the MSA-FA algorithm, among the indicators eligible for a stop test (Sheffi, 1985), it has been considered the maximum percentage deviation of flows on arc i, between flow assigned at iteration k (f_i^k) and average flow of previous iteration ($\bar{f}_i^{k-1}$)

$$\frac{f_i^k - \bar{f}_i^{k-1}}{f_i^k} \tag{39}$$

To compare the values of SUE flows obtained by the models (f_i) with the measured ones ($\hat{f}_i$) in the n survey sections, the values of the statistical indicators described in 4.1.2 have been considered. The observed vs. simulated flows reported in **Figure 8** summarize the accuracy of the simulation in relation to the type of model used. The goodness of the forecast is represented by any point of the bisector (for which the simulated flow is the same as the observed one).

TABLE 6 | Considered values of the independent variables in the application to the network of Salerno.

Model	Variable	Value
Common	*Cv*	0.10
Weibit	ϵ	0.995
Probit, Gammit	*Nit*	100
Implicit Weibit	*H*	0.40

The points on the graph are slightly scattered, and this scattering varies according to the type of model. The goodness of fit is quite high for each model, with values ranging from 0.65 (Logit) to 0.79 (implicit Weibit). Note that the results obtained with probit and gammit are very similar in value.

Moreover, **Table 7** reports, for each model, the frequencies of the bias grouped into intervals with an amplitude of 250 vehicles/h. It emerges that most of the errors are overestimation errors. In particular, the interval 0–500 contains most of the errors with the implicit Weibit that has about 43% of the errors in the interval [0–250) and the Logit that has about 44% of the errors in the interval [250–500). Note that the Logit and the implicit Weibit are the worst in flows underestimation. In this experiment, the worst model in flows overestimation is the Logit model.

Another graphical analysis (**Figure 9**) of the obtained results is performed using the Tukey mean-difference plot (also known as Bland–Altman plot), which provides a visual assessment of the shift between two different distributions of values (in this case, between the observed and simulated flows) (Cleveland, 1993). Referring to **Figure 9**, the y-axis reports the differences (bias) between the observed and simulated flows whereas the x-axis reports the means between such values. The continuous line represents the average value of the bias (Δ), whereas the two dashed lines are the boundaries of the confidence interval. By considering a significance of 95%, the lower limit (l_l) and the upper limit (l_u) are calculated as follows:

$$\begin{aligned} l_l &= \Delta - 1.96 \cdot \sigma \\ l_u &= \Delta + 1.96 \cdot \sigma \end{aligned} \tag{40}$$

Considering the logit model (**Figure 9A**) emerges that the bias ranges from about −1,250 vehicles/h (underestimation) to about 515 vehicles/h (overestimation), although there are errors in under- and overestimation, as discussed above, most of the values fall within the confidence interval (three values are less than l_l). The bias for Probit and Gammit models (**Figures 9B,D**) presents some points less than l_l and some ones greater than l_u. The weibit model (**Figure 9C**) suffers from some minor underestimation errors, but most points are within the confidence interval. The implicit weibit (**Figure 9E**) has values less than l_l and other ones greater than l_u.

The results obtained in terms of NRMSD for each MSA-FA experiment are reported in **Table 8** for the considered path choice models, where *Iter* indicates the number of iteration of the MSA

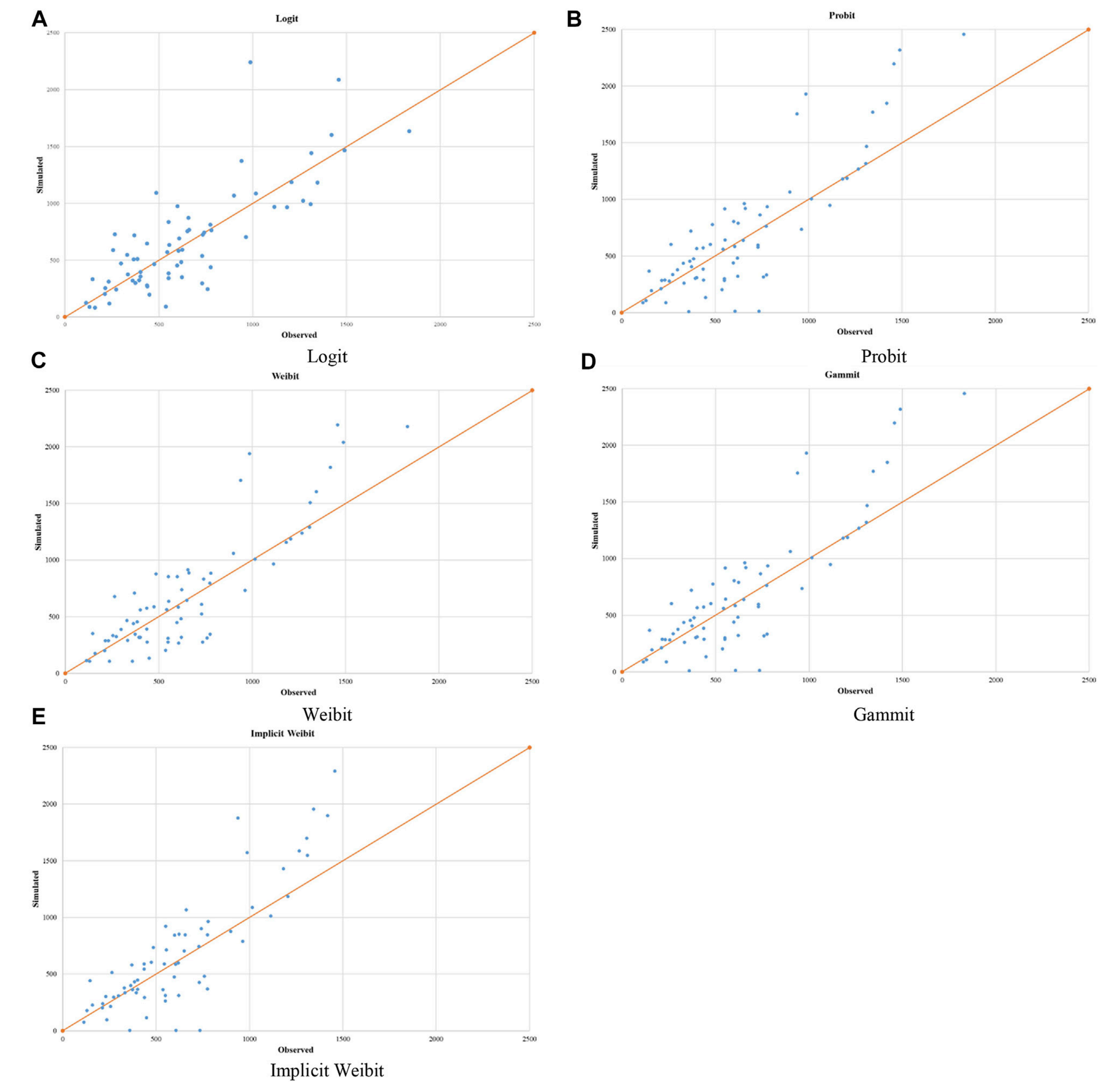

FIGURE 8 | Comparison of link flows: observed vs. simulated. **(A)** Logit **(B)** Probit **(C)** Weibit **(D)** Gammit **(E)** Implicit Weibit.

procedure and *Time* is the time in second needed to carry out the SUE procedure.

5 CONCLUSION AND FUTURE DEVELOPMENTS

Some short comments can be made on the basis of the obtained results. In the Monte Carlo technique, the greater the variance is, the more is the number of samples to be carried out. Practically, a value of Nit = 100 gives acceptable results. Considering the weibit model, a value of parameter δ closer to one (confirming the estimate of the parameter ξ_j, suggested in Castillo et al. (2008)) provides values close to those obtained with models that consider covariance such as probit and gammit, especially in case of a very low variance of path costs. The results obtained with the probit and gammit models practically do not differ from each other considering both the test network and the real one. Considering the minimum path (but also other paths for which it is the same, as shown with regard to the test network in **Figure 5**), choice probabilities obtained with the weibit model are closer to those obtained with the probit/gammit models than to those obtained with the logit models. About the real network, the best

TABLE 7 | Bias frequencies (%).

Bias Range	Logit	Probit	Weibit	Gammit	Weibit (Implicit)
≥-1,000 and < -750	1.47	0.00	0.00	0.00	1.47
≥-750 and < -500	0.00	4.41	2.94	4.41	4.41
≥-500 and < -250	2.94	2.94	2.94	2.94	2.94
≥-250 and < 0	8.82	11.76	13.24	11.76	08.82
≥0 and < 250	30.88	33.82	32.35	33.82	42.65
≥250 and < 500	44.12	32.35	35.29	32.35	26.47
≥500 and < 750	10.29	11.76	13.24	11.76	10.29
≥750 and < 1000	1.47	2.94	0.00	2.94	2.94

results are obtained by considering the C-weibit model by taking a time that differs by two orders of magnitude from that taken with probit. It is clear that the analysis conducted cannot be considered exhaustive or definitive and that the results, although indicative, also depend on the considered application to reality. The main purpose of this work was to illustrate the hypotheses to be considered for a practical application of the considered models and to provide an example of the results that can be obtained by indicating also an order of magnitude of the bias that can be expected and of the necessary computing resources, in terms of time, considering a standard situation. A study involving other

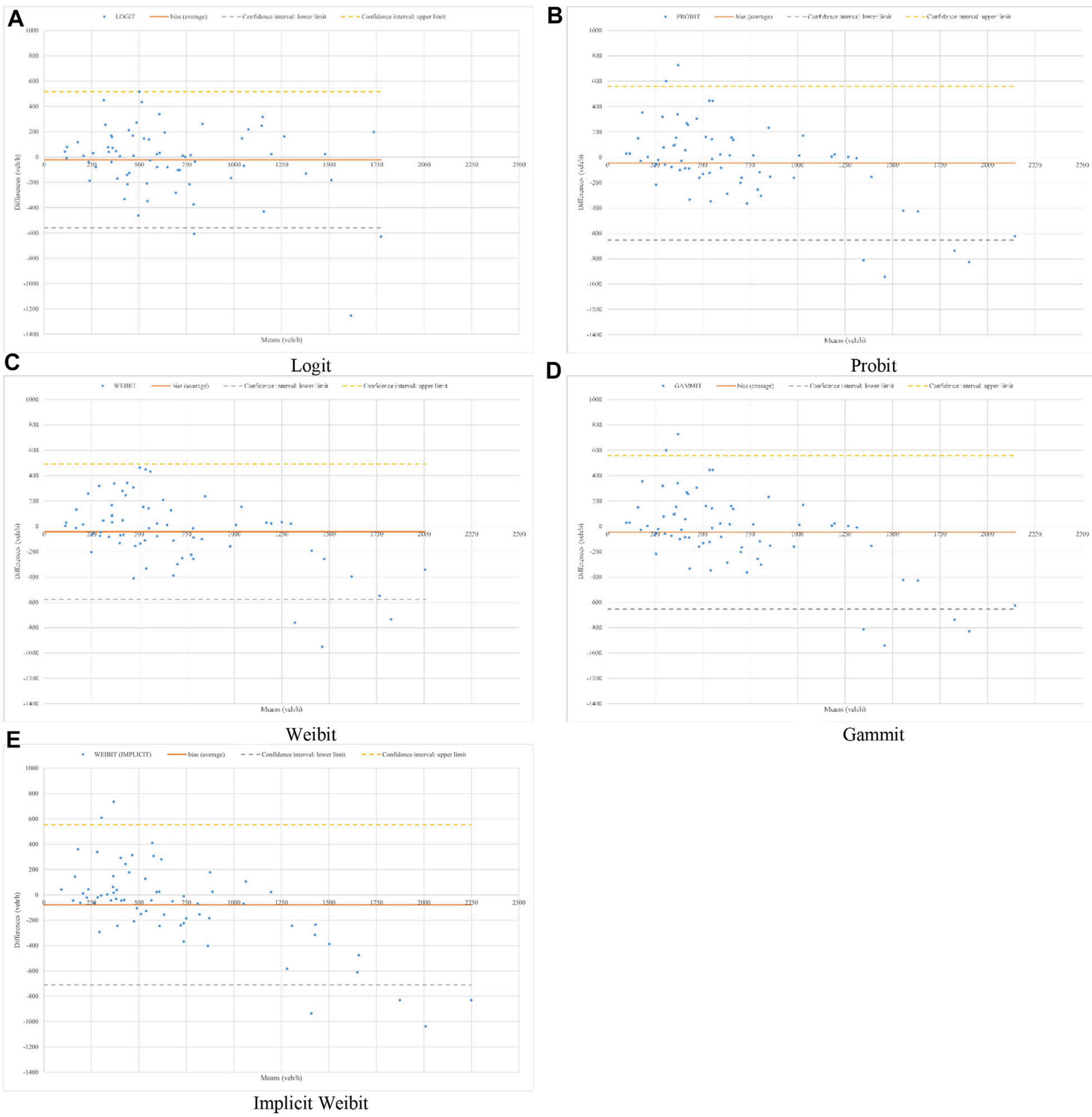

FIGURE 9 | Comparison of link flows: Bland–Altman plot.

TABLE 8 | Results for MSA-FA experiments on a real network of Salerno.

Model	Iter	Time	MSD	RMSD	NMSD	NRMSD
LOGIT	3	24.79	74,508.7373	272.9629	0.2505	0.5005
WEIBIT	13	120.51	75,028.9412	273.9141	0.2079	0.4559
PROBIT	38	19,039.65	96,405.9265	310.4930	0.2328	0.4825
GAMMIT	50	32,782.10	96,418.8084	310.5138	0.2328	0.4825
WEI-IMP	26	2429.10	108,557.9770	329.4814	0.2380	0.4878

types of models, which is under development, consists of extending a similar analysis taking into account other Logit and Weibit derived models (i.e., path-size logit, C-Logit, mixed Logit, Logit–Probit, and Weibit–Gammit) and QUM. An extension to multitype vehicle assignment to consider autonomous vehicles is also under development.

DATA AVAILABILITY STATEMENT

The original contributions presented in the study are included in the article. Further inquiries can be directed to the corresponding author.

AUTHOR CONTRIBUTIONS

MDG and AP contributed to the conception and design of the study. MDG performed the statistical analysis and wrote the first draft of the manuscript. MDG and AP wrote sections of the manuscript. All authors contributed to manuscript revision, read, and approved the submitted version.

ACKNOWLEDGMENTS

The authors wish to thank the reviewers for their suggestions, which were most useful in revising the article.

REFERENCES

Antonisse, R. W., Daly, A. J., and Ben-Akiva, M. (1989). Highway Assignment Method Based on Behavioral Models of Car Drivers' Route Choice. *Transp. Res. Rec.* 1220, 1–11.

Bell, M. G. H. (1995). Alternatives to Dial's Logit Assignment Algorithm. *Transp. Res. Part B Methodol.* 29, 287–295. doi:10.1016/0191-2615(95)00005-X

Ben-Akiva, M., and Bierlaire, M. (1999). "Discrete Choice Methods and Their Applications to Short Term Travel Decisions," in *Handbook of Transportation Science. International Series in Operations Research & Management Science.* Editor R. W. Hall (Boston, MA: Springer US), 5–33. doi:10.1007/978-1-4615-5203-1_2

Ben-Akiva, M., and Lerman, S. R. (1985). "Discrete Choice Analysis: Theory and Application to Travel Demand," in *Transportation Studies* (Cambridge, MA, USA: MIT Press).

Burrel, J. (1968). "Multiple Route Assignment and its Application to Capacity Restraint," in *Proceedings of the 4th International Symposium on the Theory of Road Traffic Flow* (Karlsruhe, Germany: Leutzbach W, Baron P), 229–239.

Cantarella, G. E. (1997). A General Fixed-Point Approach to Multimode Multi-User Equilibrium Assignment with Elastic Demand. *Transp. Sci.* 31, 107–128. doi:10.1287/trsc.31.2.107

Cantarella, G. E., and Binetti, M. G. (2002). Stochastic Assignment with Gammit Path Choice Models. *Transp. Plan.* 64, 53–67. doi:10.1007/0-306-48220-7_4

Cantarella, G. E., and Fiori, C. (2022). Multi-vehicle Assignment with Elastic Vehicle Choice Behaviour: Fixed-point, Deterministic Process and Stochastic Process Models. *Transp. Res. Part C Emerg. Technol.* 134, 103429. doi:10.1016/j.trc.2021.103429

Cantarella, G. E., Watling, D. P., de Luca, S., and Di Pace, R. (2020). *Dynamics and Stochasticity in Transportation Systems.* Amsterdam, Netherlands: Elsevier.

Cascetta, E., Nuzzolo, A., Russo, F., and Vitetta, A. (1996). "A Modified Logit Route Choice Model Overcoming Path Overlapping Problems. Specification and Some Calibration Results for Interurban Networks," in *Proceedings of the Thirteenth International Symposium on Transportation and Traffic Theory* (Lyon, France: Lesort, J.B.), 697–711.

Cascetta, E. (2009). "Transportation Systems Analysis," in *Springer Optimization and its Applications.* 2 edn (Boston, MA: Springer US). doi:10.1007/978-0-387-75857-2

Castillo, E., Menéndez, J. M., Jiménez, P., and Rivas, A. (2008). Closed Form Expressions for Choice Probabilities in the Weibull Case. *Transp. Res. Part B Methodol.* 42, 373–380. doi:10.1016/j.trb.2007.08.002

Cleveland, W. S. (1993). *Visualizing Data*. Troy, Ohio: Hobart Press.

Comi, A., and Polimeni, A. (2022). Estimating Path Choice Models through Floating Car Data. *Forecasting* 4 (2), 525–537. doi:10.3390/forecast4020029

Daganzo, C. F., and Sheffi, Y. (1977). On Stochastic Models of Traffic Assignment. *Transp. Sci.* 11, 253–274. doi:10.1287/trsc.11.3.253

Daganzo, C. F. (1983). Stochastic Network Equilibrium with Multiple Vehicle Types and Asymmetric, Indefinite Link Cost Jacobians. *Transp. Sci.* 17, 282–300. doi:10.1287/trsc.17.3.282

de la Barra, T., Perez, B., and Anez, J. (1993). "Multidimensional Path Search and Assignment," in *PTRC Summer Annual Meeting, 21st, 1993* (Manchester, England: University of Manchester), Vol. P 366, 307–320.

De Maio, M. L., and Vitetta, A. (2015). Route Choice on Road Transport System: A Fuzzy Approach. *Ifs* 28, 2015–2027. doi:10.3233/IFS-141375

Di Gangi, M., and Vitetta, A. (2018). "Specification and Aggregate Calibration of a Quantum Route Choice Model from Traffic Counts," in *New Trends in Emerging Complex Real Life Problems: ODS, Taormina, Italy, September 10–13, 2018. AIRO Springer Series.* Editors P. Daniele and L. Scrimali (Cham: Springer International Publishing), 227–235. doi:10.1007/978-3-030-00473-6_25

Di Gangi, M., Polimeni, A., and Belcore, O. M. (forthcoming). C-Weibit Discrete Choice Model: A Path Based Approach. *International Conference on Optimization and Decision Science (ODS2022), Firenze (Italy), August 30-September 2, 2022. Springer*.

Dial, R. B. (1971). A Probabilistic Multipath Traffic Assignment Model Which Obviates Path Enumeration. *Transp. Res.* 5, 83–111. doi:10.1016/0041-1647(71)90012-8

Domencich, T. A., and McFadden, D. (1975). *Urban Travel Demand: A Behavioral Analysis*. Amsterdam: North-Holland Publishing. Number: Monograph.

Fosgerau, M., Frejinger, E., and Karlstrom, A. (2013). A Link Based Network Route Choice Model with Unrestricted Choice Set. *Transp. Res. Part B Methodol.* 56, 70–80. doi:10.1016/j.trb.2013.07.012

Leurent, F. (2005). Contributions to the Logit Assignment Model. *Transp. Res. Rec.* 1493, 207–212.

Mai, T., Fosgerau, M., and Frejinger, E. (2015). A Nested Recursive Logit Model for Route Choice Analysis. *Transp. Res. Part B Methodol.* 75, 100–112. doi:10.1016/j.trb.2015.03.015

Manski, C. F., and McFadden, D. (1981). *Structural Analysis of Discrete Data with Econometric Applications*. Cambridge: The MIT Press.

McFadden, D. (1978). "Modeling the Choice of Residential Location," in *Spatial Interaction Theory and Residential Location*. Editors A. Karlqvist, L. Lundqvist, and J. W. Weibull (Amsterdam: North-Holland), 72–77.

Nuzzolo, A., and Comi, A. (2018). A Subjective Optimal Strategy for Transit Simulation Models. *J. Adv. Transp.* 2018, 1–10. doi:10.1155/2018/8797328

Nuzzolo, A., and Comi, A. (2016). Individual Utility-based Path Suggestions in Transit Trip Planners. *IET Intell. Transp. Syst.* 10, 219–226. doi:10.1049/iet-its.2015.0138

Prashker, J. N., and Bekhor, S. (2004). Route Choice Models Used in the Stochastic User Equilibrium Problem: A Review. *Transp. Rev.* 24, 437–463. doi:10.1080/0144164042000181707

Prato, C. G. (2009). Route Choice Modeling: Past, Present and Future Research Directions. *J. Choice Model.* 2, 65–100. doi:10.1016/S1755-5345(13)70005-8

Quattrone, A., and Vitetta, A. (2011). Random and Fuzzy Utility Models for Road Route Choice. *Transp. Res. Part E Logist. Transp. Rev.* 47, 1126–1139. doi:10.1016/j.tre.2011.04.007

Rosa, A., and Maher, M. (2002). "Algorithms for Solving the Probit Path-Based Stochastic User Equilibrium Traffic Assignment Problem with One or More User Classes," in *Transportation and Traffic Theory in the 21st Century*. Editor M. A. P. Taylor (Bingley, UK: Emerald Group Publishing Limited), 371–392. doi:10.1108/9780585474601-019

Russo, F., and Vitetta, A. (2003). An Assignment Model with Modified Logit, Which Obviates Enumeration and Overlapping Problems. *Transportation* 30, 177–201. doi:10.1023/A:1022598404823

Sheffi, Y., and Powell, W. B. (1982). An Algorithm for the Equilibrium Assignment Problem with Random Link Times. *Networks* 12, 191–207. doi:10.1002/net.3230120209

Sheffi, Y. (1985). *Urban Transportation Networks: Equilibrium Analysis with Mathematical Programming Methods*. Englewood Cliffs, N.J: Prentice-Hall.

Si, B.-F., Zhong, M., Zhang, H.-Z., and Jin, W.-L. (2010). An Improved Dial's Algorithm for Logit-Based Traffic Assignment within a Directed Acyclic Network. *Transp. Plan. Technol.* 33, 123–137. doi:10.1080/03081061003643705

Vitetta, A. (2016). A Quantum Utility Model for Route Choice in Transport Systems. *Travel Behav. Soc.* 3, 29–37. doi:10.1016/j.tbs.2015.07.003

Wang, J., Peeta, S., and He, X. (2019). Multiclass Traffic Assignment Model for Mixed Traffic Flow of Human-Driven Vehicles and Connected and Autonomous Vehicles. *Transp. Res. Part B Methodol.* 126, 139–168. doi:10.1016/j.trb.2019.05.022

Yai, T., Iwakura, S., and Morichi, S. (1997). Multinomial Probit with Structured Covariance for Route Choice Behavior. *Transp. Res. Part B Methodol.* 31, 195–207. doi:10.1016/S0191-2615(96)00025-2

Conflict of Interest: The authors declare that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

Publisher's Note: All claims expressed in this article are solely those of the authors and do not necessarily represent those of their affiliated organizations or those of the publisher, the editors, and the reviewers. Any product that may be evaluated in this article or claim that may be made by its manufacturer is not guaranteed or endorsed by the publisher.