

An Instance-Specific Parameter Tuning Approach using Fuzzy Logic for a Post-Processing Topological Map-Matching Algorithm

Carola Alejandra Blazquez, Jana Ries, Roberto Jesus León, and Pablo Andres Miranda

Abstract— Map Matching Algorithms (MMAs) are developed to solve spatial ambiguities that arise in the process of assigning GPS measurements onto a digital roadway network. Scarce systematic parameter tuning approaches exist in the literature for optimizing MMA performance. Thus, a novel framework is proposed for a systematic calibration of the parameters of a post-processing MMA. The calibration approach consists of an Instance-specific Parameter Tuning Strategy (IPTS) that employs Fuzzy Logic principles. The proposed fuzzy IPTS tool determines algorithm-specific parameter values based on instance-specific information a priori to the execution of the MMA. Finally, the proposed IPTS tool is able to adjust to two particular decision maker preferences on algorithm performance, namely solution quality and computational time.

I. INTRODUCTION

Global Positioning Systems (GPS) technology is commonly employed to acquire positioning data. The integration of GPS with Geographic Information Systems (GIS) is part of the innovative advanced technology applied by Intelligent Transportation Systems (ITS) to improve mobility and safety, reduce environmental impacts and energy consumption, and enhance the overall quality of life of individuals. When integrating GPS measurements with a roadway network in GIS, these measurements (represented as data points) are usually projected to the nearest roadway centerline segment to determine the road on which events and incidents occur, point features are located, or a vehicle is traveling.

A map-matching problem (MMP), also known as spatial ambiguity or spatial mismatch, arises when GPS measurements are incorrectly assigned to the roadway segment. The MMP occurs due to roadway network complexities, inadequate GPS data collection, inaccurate digital roadway map and GPS measurements, inadequate sampling rate, or combinations of these situations [1, 2, 3]. Therefore, objects may appear to be located or moving along incorrect roadway centerlines particularly at complex roadway structures such as ramps, divided highways, overpasses and underpasses, parallel roads, and intersections.

As a result, this problem may influence negatively any subsequent evaluation, analysis, planning, or decision-making. The distinct need for reducing incidents (accidents), travel time, congestion, and transportation costs, while increasing customer service, productivity, efficiency, and safety in

numerous ITS applications such as vehicle navigation and route guidance, fleet management, and accident prevention are only some of the major motivations for solving the MMP by determining correct vehicle locations.

A variety of Map Matching Algorithms (MMAs) have been developed to solve the MMP. These MMAs employ control or weight parameters that determine algorithmic performance. Unsuitable parameter values of these algorithms, along with other sources of errors that may arise from the digital roadway map or the GPS measurements, may strongly impact the MMA performance. Basic calibration methods (e.g., trial-and-error) are commonly employed to identify well-performing parameter values prior to executing any type of MMA. The algorithms need to be executed usually for a wide range of parameter values, resulting in extensive experimentation and resource consumption. Traditionally, the parameter values are calibrated for data collected in a specific environmental setting such as dense urban or rural areas [4]. Hence, the reported performance and parameters values of MMAs may be conditioned and biased according to the employed test dataset. In addition, scarce literature exists on calibration methodologies to obtain the best parameter values that optimize the performance of MMAs. For example, Velaga et al [5] determined optimal weight coefficients for data collected in rural areas that minimize the map-matching error. In another study, Oran and Jaillet [6] developed a precise proximity weight formulation for any type of MMAs.

The objective and main contribution of this research is to develop a systematic parameter tuning approach based on fuzzy logic principles for a post-processing topological MMA developed by Blazquez et al. [7]. This approach addresses the lack of systematic tuning techniques in the MMAs literature, and comprehends the performance behavior of MMAs in response to instance-specific information.

II. BACKGROUND AND LITERATURE REVIEW

A. MMAs and Transportation Applications

MMAs have been widely studied for decades by employing different levels of complexity ranging from simple geometrical techniques to complex, advanced inference, filtering, and mathematical modeling methodologies such as Kalman filters, Fuzzy Logic, particle filtering, hidden Markov models, and Bayesian statistics [2, 7, 8, 9, 10, 11, 12]. Topological network-based MMAs have been used

*Research supported by project funds at Universidad Andres Bello Project N° DI-10-18/REG, and the Fondecyt Project N°1140811.

C. Blazquez is a professor at Universidad Andres Bello, Viña del Mar, Chile (corresponding author to provide phone: +56 32 284 5579); e-mail: cblazquez@unab.cl.

J. Ries is a Senior Lecturer at the University of Portsmouth, Portsmouth, UK, e-mail: Jana.Ries@port.ac.uk

R. León is an assistant professor at Universidad Andres Bello, Viña del Mar, Chile, e-mail: roberto.leon@unab.cl

P. Miranda is a full time faculty member at the School of Industrial Engineering at the Pontificia Universidad Católica de Valparaíso, Valparaíso, Chile, e-mail: pablo.miranda@pucv.cl

particularly in ITS applications due to their simplicity, efficiency, rapid, and easy implementation [7, 8, 9, 12]. These MMAs rely on the computation of shortest paths and/or shortest travel time between pairs of matched GPS points using roadway network topology (i.e., link connectivity and contiguity) and turn restrictions. An advantage of topological MMAs is that they provide positions as accurate as possible without the need of large memories [7, 8, 13, 14, 15, 16]. Different types of parameters such as buffer size, traffic flow or congestion information, and road characteristics (e.g., speed limit, number of lanes, etc.) may be employed by these MMAs.

Post-processing MMAs determine the correct road on which a user or vehicle is traveling once the complete or part of the GPS dataset is available. In this study, a post-processing MMA is applied to an instance that is comprised of a set of GPS data points that represents a complete vehicle route [4, 9]. For example, post-processing MMAs are needed to obtain correct paths when collecting multi-modal trip data with GPS receivers for route choice modeling in a GIS environment [1, 14, 17, 18, 19]. Blazquez and Vonderohe [2] employed a topological MMA in post-processing mode to determine the correct roadway centerline on which a fleet of winter vehicles has traveled to compute performance measures for decision-makers and planners. Automatic road toll systems based on the total mileage driven per vehicle equipped with GPS technology are another example that requires post-processing MMAs to prevent errors in road use charging [20]. Hence, the execution of MMAs in post-processing mode has been widely studied and implemented in transportation and logistics applications [21].

B. Topological Map Matching Algorithm

The post-processing MMA employed in this study resolves the MMP by determining the correct roadway centerline on which a vehicle is traveling. First, this MMA selects all roadway centerlines within a buffer around each analyzed GPS data point and orthogonally projects it to the closest roadway centerline (also known as “snapping”). Subsequently, Dijkstra’s algorithm is employed to solve the shortest path problem between previously snapped and current snapped points using network topology and turn restrictions. A path is considered viable and snapped data point locations are accepted if the difference between the computed travel speed and the average recorded vehicle speed for the data points is within a speed range tolerance. If a path is rejected, data points are snapped to alternative roadway centerlines contained within their buffers, shortest paths are recalculated, and speeds are compared once again. If no other roadway centerline candidates exist within the buffers or no viable paths are obtained, then the algorithm tests for feasible paths between preceding and subsequent data points [13]. This study utilizes the main algorithm-specific parameters of the post-processing MMA: buffer size (BZ) and speed range tolerance (ST). BZ is employed to select road segments as candidates for snapping locations, and ST provides a range for accepting or rejecting paths between a pair of snapped points when comparing recorded and computed vehicle speeds.

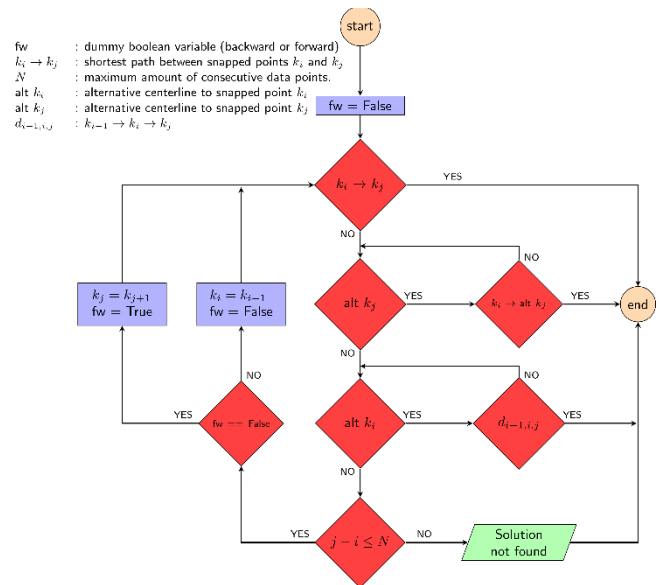


Figure 1. Flow Diagram of the post-processing MMA

Fig. 1 shows the flow diagram of the MMA with the step sequence design. First, the MMA computes the shortest path between pairs of snapped GPS data points k_i and k_j . If this path is not feasible, then the algorithm obtains alternative roadway centerline candidates within the buffer around data point k_j . Consequently, feasible paths are tested between data point k_i and the newly snapped point $alt k_j$. If no viable paths are found, then alternative roadway centerline candidates within the buffer around data point k_i are obtained, and shortest paths are recalculated and speeds are compared between the immediate neighboring points k_{i-1} and k_j . The dummy Boolean variable fw helps to advance either backward or forward to test paths with consecutive data points until the MMP is solved or N consecutive data points have been tested.

C. Parameter Setting Strategies

Heuristic algorithms are known to perform according to the calibration of a set of algorithm-specific parameters ϕ . This so-called parameter setting problem for an algorithm A searches to determine parameter values ϕ^* that optimizes the algorithmic performance $f_A(i, \phi)$, where i is the problem instance under consideration [22]. Strategies for addressing the parameter setting problem can be classified into *parameter control* and *parameter tuning*, in which the setting of parameter values is integrated into the algorithm or determined a priori to executing the algorithm, respectively. Parameter control strategies (PCS) commonly take the form of an adaptive strategy that allows to re-evaluate parameters ϕ based on the current performance of the algorithm. Whereas, parameter tuning strategies (PTS) are mainly based on a Design of Experiments (DOE), as in Coy et al. [23], Adenso-Diaz and Laguna [24], and Dobslaw [25] to find the best parameter setting prior to the execution of an algorithm.

PCS require modification of the algorithm itself to integrate a reactive strategy that usually depends on the performance behavior of the algorithm. These approaches are predominantly justified if some of the structural information of the problem instance is not known prior to executing the algorithm. This concept may be applied dynamically to adjust MMA parameters at different points of algorithmic steps.

In contrast, PTS aim to provide a robust parameter values that can be applied to a broad variety of instances [25]. In the context of combinatorial optimization, recent studies have proposed Instance-specific Parameter Tuning Strategies (IPTS) designs using instance-specific information within parameter tuning strategies [26, 27, 28]. In doing so, IPTS assess instance-specific information a priori to the execution of an algorithm, and subsequently, determine the best parameter values φ^* for the algorithm. An advantage of these strategies is an explicit understanding of relevant instance-characteristics of the problem. This study proposes a framework for calibrating algorithm parameters of a topological MMA using IPTS.

III. IPTS FRAMEWORK FOR A POST-PROCESSING MMA

A. Framework Structure

This study presents an IPTS framework to determine a set of algorithm-specific parameter values based on instance-specific information. The IPTS tool provides parameter values φ_i^* for one particular instance i [27]. Subsequently, this study aims to determine the best parameter values φ_i^* for a given set of GPS measurements. Fig. 2 shows the proposed parameter calibration framework. Its structure is based on finding a parameter setting φ_i^* for a certain instance i prior to executing a MMA. Instance i is defined as a set of GPS data points that is used by the post-processing MMA to determine the underlying vehicle route. The IPTS tool is conducted for any instance i , allowing to adjust the MMA parameters based on the characteristics of the instance.

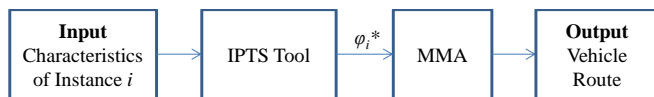


Figure 2. Diagram of the IPTS framework

The instance is characterized by two sets of information, namely scenario-specific and data-specific information. The former may include relevant information about the surrounding area and characteristics of the vehicle. Some of this information remains constant along the vehicle route such as type of vehicle or weather conditions, while other information may vary as the vehicle is traveling such as maximum road speed, type of surrounding built environment (e.g., land use), type of road, etc. In contrast, data-specific information is related to GPS measurements (e.g., number of detected satellites, vehicle speed) and sampling frequency (e.g., 2 seconds or 5 minutes).

B. Design of the Fuzzy Logic IPTS

This study proposes an IPTS design that uses a fuzzy inference system to determine a set of algorithm-specific parameter values based on instance-specific information. Zadeh [29] introduced fuzzy logic as an alternative concept to classic set theory to allow dealing with non-crisp information. Fuzzy logic allows a partial membership of information into multiple sets that are described by linguistic terms, and has been applied in parameter tuning strategies in the related literature (e.g., in the form of a fuzzy inference system [27]). This fuzzy inference system provides a high flexibility in such that relevant specific elements that define an algorithm (e.g., control parameters, design concepts) can easily be adjusted, impacting the algorithm performance. Particularly, in the

MMP context, instance-specific information may be employed to adjust MMA parameters.

The basis of fuzzy logic is the ability to provide a linguistic label for any crisp parameter value, by assigning any value to one or more fuzzy sets with a certain membership degree. This membership degree is assessed by a membership function that is created for each fuzzy set such as ‘low’, ‘medium’, and ‘high’. The proposed fuzzy inference system, which includes a set of membership functions for each variable and a rule base, is based on a manual design. This design is obtained from a statistical analysis. Hence, the fuzzy sets are a subjective interpretation of the individual parameter levels.

The proposed framework consists of the following two design stages:

(I) Investigation of the impact of instance- and algorithm-specific information on the MMA performance

An initial stage analyzes the impact of changes of instance-specific information and algorithm-specific parameters, and possible interaction effects on the MMA performance. In other words, a statistical analysis indicates whether the variation of one parameter, with all others being equal, leads to a statistically significant change in the performance (main effect), and whether the impact of one independent variable varies depending on the level of a second independent variable (2-way interaction).

A factorial design is proposed in combination with a regression analysis based on a set of training problem instances. This statistical analysis provides an understanding of the significance of instance- and algorithm-specific parameters with respect to MMA performance.

(II) Design of the fuzzy IPTS tool

In order to design the fuzzy IPTS tool, every instance characteristic is considered as an input variable and the best parameter values φ_i^* is referred to as an output variable. In the *fuzzification* process of the fuzzy IPTS tool, instance-specific information is mapped to a set of membership functions that are associated with linguistic terms by identifying a membership degree μ that is characterized by (1) with n being the number of membership functions.

$$\mu_j \rightarrow [0, 1], \quad j = 1, \dots, n \quad (1)$$

Subsequently, the fuzzy rules are designed based on the insights drawn from the previous statistical analysis. Each rule comprises information on the input and output variables. Each rule considers a unique combination of fuzzy sets that each variable is assigned to. A set of IF-THEN rules is activated based on the identified membership degrees as follows:

$$\text{If } X_1 \text{ is } A_1 \text{ and } X_2 \text{ is } A_2, \text{ Then } Y_1 \text{ is } B_1 \text{ and } Y_2 \text{ is } B_2$$

Where, X_i and Y_i are the input and output variables, respectively, and A_i and B_i are the linguistic terms associated with the fuzzy sets. For example, “IF sampling frequency (SF) is high, THEN buffer size (BZ) is medium”.

Finally, the fuzzy IPTS tool *defuzzifies* this information to derive crisp values for each output parameter (i.e., each parameter of the algorithm) by using the construct of a Mamdani-type fuzzy inference system. The *defuzzification* process may apply the centroid method described by (2), in

which centers of gravity for the fuzzy sets are employed to determine crisp values [30].

$$z = \frac{\int \mu(z) z dz}{\int \mu(z)} \quad (2)$$

The outcome of the IPTS tool described above is a set of best parameter values ϕ_i^* . This set is used by the MMA as input information to generate complete vehicle routes, as shown in Fig. 2. Finally, the performance of the fuzzy IPTS tool is evaluated with set of test instances using the best parameter values for each instance in the execution of the post-processing MMA.

IV. RESULTS

A. Data

The data employed in this study were collected by winter maintenance vehicles (also known as snowplows) in Portage County, Wisconsin in the United States. Sixteen different instances (i.e., vehicle routes) were collected during a winter season. Fig. 3 presents a data sample of the collected GPS data points at a main highway interchange, and the roadway centerline map with a nominal scale of 1:12,000. The measured data captured coordinates, timestamps, and vehicle speeds. Twenty-four instances are used as the training dataset to investigate the main and interaction effects of any parameters involved. Eighteen instances are used as test datasets when executing the MMA with the output parameter values provided by the IPTS fuzzy tool. Two instance-specific parameters, namely sampling frequency (SF), and share of high-speed roads (HS), are identified initially as relevant instance characteristics in this study. The MMA was executed in Python 2.7.8 with an arcpy library from ArcGIS software.

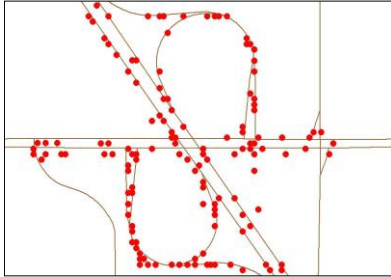


Figure 3. Data sample for Portage County

B. IPTS Framework Design

Stage 1: Design of Experiments (DOE)

The set of training instances were solved for different parameter combinations, as shown in Table I. SF and HS are input variables (i.e., instance-specific information) of the proposed IPTS framework, whereas BZ and ST are output variables (i.e., algorithm-specific parameters) that are determined by the IPTS tool. The tested SF values are 10, 20, and 30 seconds, and the percentage of HS along highways varies, clustering around 30%, 55%, and 70% for the training instances. The latter is computed as a ratio between the number of GPS points collected along highways and all road types (highways, local roads, and ramps). The selection of BZ and ST levels was guided by an experimental study of parameters of the algorithm performed in [31].

Due to the post-processing nature of the MMA, the correct or incorrect snapping result of each data point is known in

advance. The performance of the MMA was assessed by comparing the snapping result of each data point to its true route with ‘1’ representing a correct snapping, and ‘0’ otherwise. Solution quality $SQ(i)$ of the algorithm for instance i is determined by the share of correct snaps using (3) with m_i being the number of correctly snapped points, and n_i being the total number of data points of instance i .

$$SQ(i) = \frac{m_i}{n_i} \quad (3)$$

A linear regression was used to investigate the statistical significance of main and 2-way interaction effects with respect to the MMA performance. The analysis has been conducted for two different decision maker preferences regarding the algorithmic performance, i.e., best solution quality (SQ) and computational time ($TIME$).

Table II shows the regression analysis results stating the coefficients of the significant factors and the corresponding significance level. The R^2 values for SQ and $TIME$ are 0.66 and 0.53, respectively. This table highlights that ST is not statistically significant for any of the two performance preferences (i.e., SQ and $TIME$) with the main and interaction effects. Therefore, ST was disregarded from further analysis. In contrast, BZ shows to be a key parameter in the results. Improved solution quality and longer computational times are obtained when BZ increases.

TABLE I. FACTORIAL DESIGN LEVELS

Parameter	Level 1	Level 2	Level 3
<i>Instance-Specific Parameters</i>			
Sampling Frequency (SF) [ft]	10	20	30
Share of High Speed Roads (HS)	0.3	0.55	0.7
<i>Algorithm-Specific Parameters</i>			
Buffer Size (BZ) [ft]	20	40	60
Speed Tolerance (ST) [mi/h]	15	25	35

TABLE II. RESULTS OF THE LINEAR REGRESSION ANALYSIS

Independent Variables	Performance preference	
	Solution Quality (SQ)	Computational Time ($TIME$)
	Standardized Coefficients Beta	
<i>Main effects</i>		
Sampling Frequency (SF)	0.119	-0.761***
Share High Speed Roads (HS)	0.454***	-0.379**
Buffer Size (BZ)	1.407***	0.487**
Speed Tolerance (ST)	0.020	-0.169
<i>Interaction effects</i>		
HS x SF	-0.111	0.288*
HS x ST	-0.013	0.105
HS x BZ	-0.623***	0.011
SF x ST	-0.020	0.135
SF x BZ	-0.179	-0.370**
ST x BZ	-0.004	-0.004

***p < 0.01, **p < 0.05, *p < 0.1

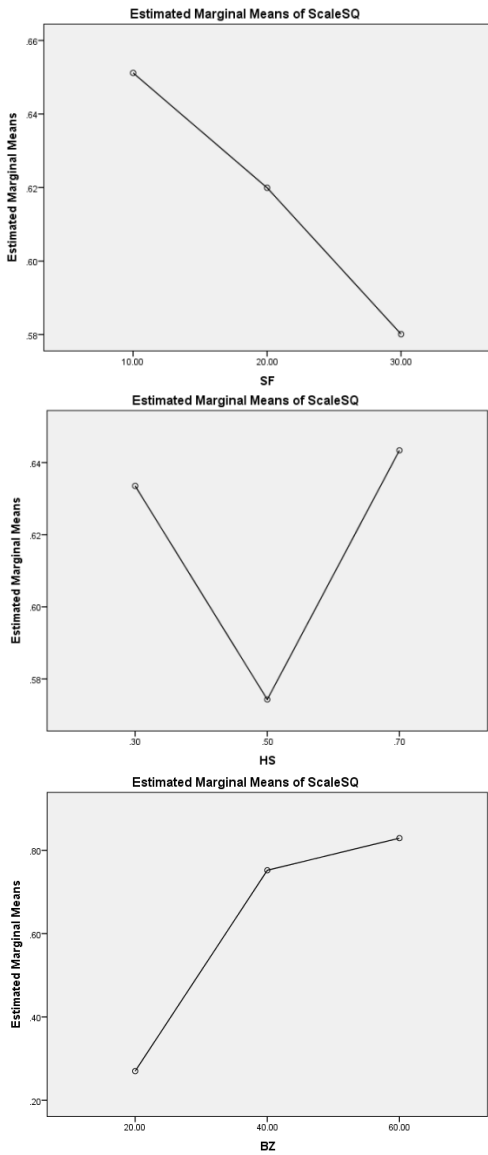


Figure 4. Significant Main Effects of Sampling Frequency (SF), Share of High Speed Roads (HS), and Buffer Size (BZ) on Solution Quality (SQ)

With respect to instance-specific information, Fig. 4 and 5 show that instances with larger HS values results in better SQ performance and shorter computational times when compared to instances with smaller HS values. More importantly, the algorithm results in lower SQ for instances that have an equal share of high and low speed roads, i.e., HS=0.5. In particular winter maintenance applications, the difficulty of solving instances with HS=0.5 may be related to entering and exiting high speed roads at interchanges or critical nodes in the infrastructure. This finding underpins the need for the algorithm to be fine-tuned according to this type of road characteristics.

The interaction effect of HS and BZ on SQ is observed in Fig. 6a) due to the similarity in algorithmic behavior when BZ takes values of 40 and 60. More precisely, the performance converges for high HS values. In contrast, Fig. 6a) and b) indicates that a smaller values of BZ lead to substantially worse performance in SQ and shorter computational times, respectively. This result may be related to not finding any roadway candidates within the smaller buffer sizes.

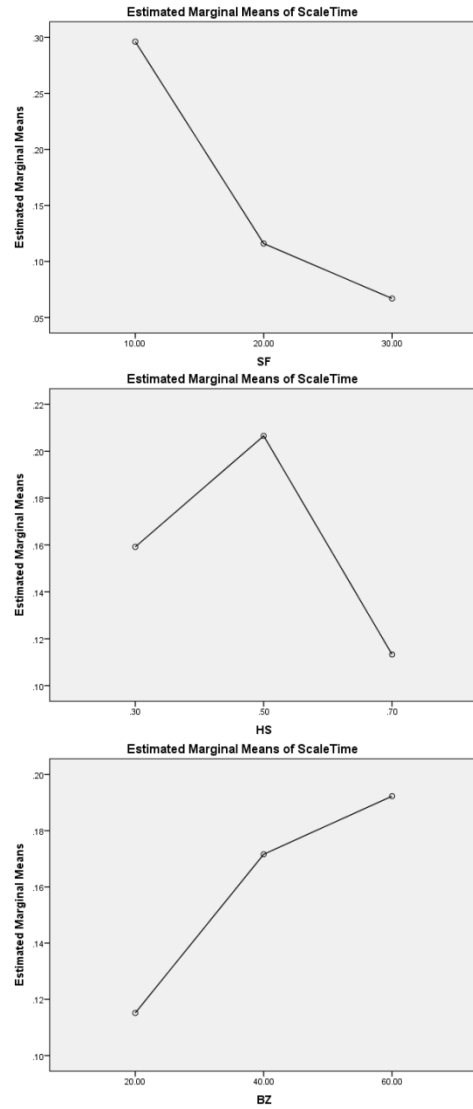


Figure 5. Significant Main Effects of Sampling Frequency (SF), Share of High Speed Roads (HS), and Buffer Size (BZ) on Computational Time (TIME)

Stage 2: Design of the Fuzzy IPTS tool

Table III shows the design of the triangular-shape membership functions for the input (instance-specific parameters) and output (algorithm-specific parameters) of the IPTS tool. The regression analysis was utilized to inform the manual design of a rule base for the fuzzy IPTS tool. Table IV shows the corresponding rule base of the proposed fuzzy IPTS system. Each rule consists of the information on two input variables (i.e., SF and HS) and the information on one output variable (i.e., BZ). For example, if SF is *small* and HS is *high*, then BZ is set to *large* when seeking the best solution quality (SQ) and is set to *very small* when searching for low computational time (TIME).

Following the evaluation of HS and SF for any instance i , the IPTS tool assigns a certain BZ value for obtaining the preference of SQ and TIME, as shown in the surface view in Fig. 7a) and b), respectively. Fig. 7a) suggest that buffer sizes are set particularly high for difficult instances which have a somewhat balance between high speed roads and others

(HS=0.5). For preference SQ , the impact of HS dominates and SF does not interact with this effect. Conversely, the results in Fig. 7b) indicate that BZ is set substantially lower for such difficult instances due to the impact on computational time. Due to the interaction effect with SF, BZ values are set particularly low for small sampling frequencies, as the impact on computational time is stronger than for larger SF values.

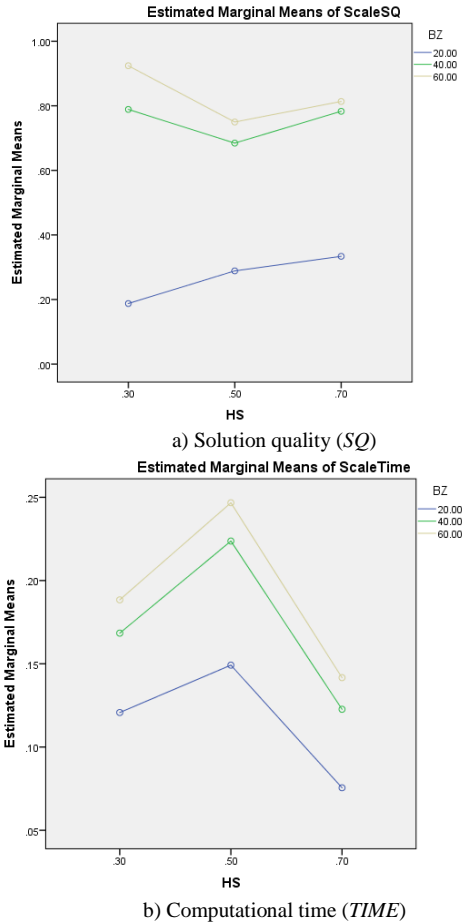


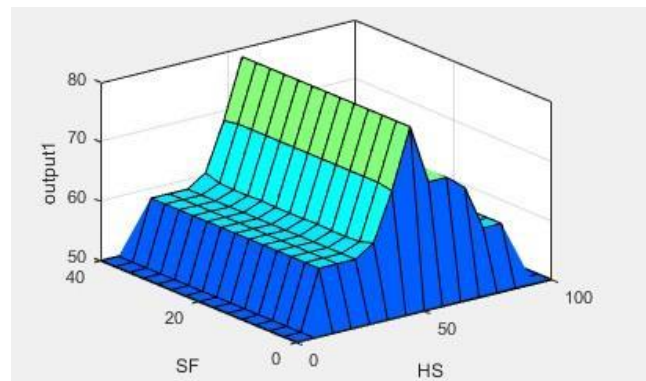
Figure 6. Interaction Effects of High Speed Roads (HS) and Buffer Size (BZ) on a) solution quality (SQ) and b) computational time ($TIME$)

TABLE III. MEMBERSHIP FUNCTIONS FOR INPUT AND OUTPUT

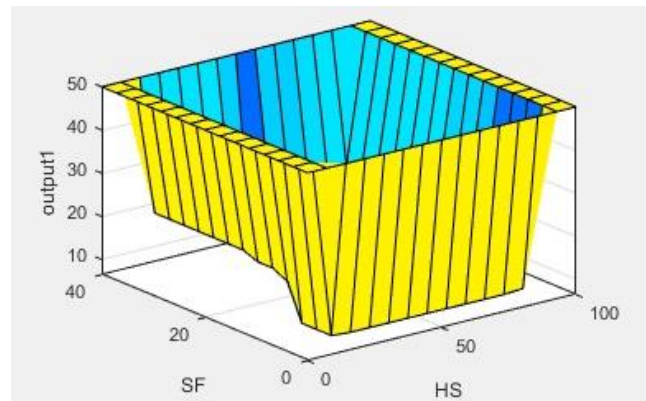
INPUT			
	Min	Medium	Max
Sampling Frequency (SF) [s]			
Small	0	10	20
Medium	10	20	30
Large	20	30	40
Share of High Speed Roads (HS) [%]			
Low	10	30	55
Medium	35	55	75
High	50	70	90
OUTPUT			
	Min	Medium	Max
Buffer Size (BZ) [ft]			
Very Small	0	0	20
Small	0	20	40
Medium	20	40	60
Large	40	60	80
Very Large	60	80	100

TABLE IV. RULE BASE FOR THE FUZZY IPTS

INPUT		OUTPUT	
Instance-Specific Parameters		Algorithm-Specific Parameters	
Sampling Frequency (SF)	Share of High Speed Roads (HS)	Buffer Size (BZ)	
		Solution Quality (SQ)	Computational Time (TIME)
Small	Low	Large	Very small
Medium	Low	Large	Small
Large	Low	Large	Small
Small	Medium	Very Large	Very Small
Medium	Medium	Very Large	Very Small
Large	Medium	Very Large	Very Small
Small	High	Large	Very Small
Medium	High	Large	Small
Large	High	Large	Small



a) Solution quality (SQ)



b) Computational time ($TIME$)

Figure 7. Surface view for sampling frequency (SF) and share of high speed roads (HS) on buffer size (BZ) for a) solution quality (SQ) and b) computational time ($TIME$)

C. Execution of Fuzzy IPTS tool prior to MMA

The designed Fuzzy IPTS was tested using 18 instances. Table V shows the total number of GPS data points, and the SF and HS values for each of these instances. This table demonstrates the diversity across the test dataset. The fuzzy IPTS tool was used to identify the algorithm-specific parameter BZ for each test instance a priori to the execution of the MMA given the preference of SQ and $TIME$, as shown in Table VI. For comparison purposes, Table VII shows the results for each instance for the fixed BZ value of 60 [ft] suggested in [32].

TABLE V. TEST INSTANCE CHARACTERISTICS

Instance	Number of data points	Sampling Frequency (SF) [s]	Share of High Speed Roads (HS)
1	358	10	0.37
2	179	20	0.37
3	179	20	0.37
4	120	30	0.38
5	119	30	0.36
6	119	30	0.39
7	232	10	0.89
8	117	20	0.89
9	117	20	0.89
10	78	30	0.87
11	78	30	0.90
12	77	30	0.90
13	176	10	0.56
14	88	20	0.57
15	88	20	0.56
16	58	30	0.59
17	59	30	0.54
18	59	30	0.56

TABLE VI. ALGORITHM-SPECIFIC PARAMETER VALUES BZ FOR THE TEST INSTANCES FROM THE IPTS TOOL RESULTS

Instance	Fuzzy Buffer Size (BZ) [ft]	
	Solution Quality (SQ)	Computational Time (TIME)
1	63	7
2	63	20
3	63	20
4	64	20
5	62	20
6	66	19
7	60	10
8	60	20
9	60	20
10	60	20
11	50	50
12	50	50
13	74	6
14	73	15
15	74	14
16	72	16
17	75	12
18	74	14

TABLE VII. RESULTS WITH FIX BUFFER SIZE OF 60 [FT] FOR THE TEST INSTANCES

Instance	Solution Quality (SQ)	Computational Time (TIME)[s]
1	91.90%	15.454
2	92.18%	6.813
3	92.18%	7.268
4	88.33%	4.548
5	89.92%	5.039
6	91.60%	5.595
7	91.81%	9.223
8	91.45%	5.302
9	90.60%	4.809
10	87.18%	2.892
11	91.03%	3.308
12	85.71%	3.161
13	94.89%	7.634
14	94.32%	3.529
15	89.77%	4.165
16	91.38%	2.446
17	93.22%	2.600
18	86.44%	2.629

TABLE VIII. COMPARISON OF PERFORMANCE RESULTS

	Fix Buffer Size (BZ) of 60 [ft]	Fuzzy Buffer Size (BZ) (Preference = SQ)	Fuzzy Buffer Size (BZ) (Preference = Time)
	Solution Quality (SQ) [%]		
Median	91.42%	91.53%	55.34%
Mean	90.77%	90.51%	52.95%
Standard Deviation	0.025%	0.035%	0.174%
Computational Time (TIME) [s]			
Median	4.68	4.67	3.29
Mean	5.36	5.34	3.49
Standard Deviation	3.18	3.17	1.51

Table VIII presents the comparison results for the performance of the MMA using the BZ values proposed by the Fuzzy IPTS and the fixed BZ value of 60 [ft]. This table shows subtle improvements with respect to *SQ*. The results confirm that the execution of the IPTS tool prior to the MMA helps to strategically select the best algorithm parameter values for each analyzed instance rather than employing a trial-and-error approach for calibrating algorithm parameters.

V. CONCLUSIONS

There is limited research in the area of parameter calibration methods for MMAs, such as a trial-and-error approach or descend algorithms. Moreover, scarce insights are available about the impact of instance-specific characteristics and any potential impact on MMA performance. In order to fill this gap, this study proposes a systematic approach that relies on an instance-specific parameter tuning strategy (IPTS) prior to running a post-processing topological MMA. Fuzzy logic is used in the IPTS design, enabling the consideration of imprecise information to underpin its suitability in the context of map matching.

In particular, the IPTS provides the best algorithm parameter values tailored to instance-specific information, in which additional experimental efforts are prevented while yielding parameter values that optimizes algorithmic performance. This manifests one main contribution of this research. In addition, this study allows the parameter tuning to adjust to a particular performance preference, i.e., maximizing solution quality or minimizing computational time. Future research foresees to provide a combination of both performance preferences, as in [22], which is relevant in the implementation of real-time MMAs.

The proposed framework integrates instance-specific information in the determination of algorithm-specific MMA parameters using fuzzy logic principles. The main advantage of the proposed fuzzy IPTS tool is the level of flexibility with which it can be redesigned or adjusted to other types of datasets collected from different transportation applications (e.g., fleet management). Additional algorithm and instance-specific parameters (e.g., environment settings) may also be used by this tool, revealing another contribution of this research. Future research should study the interaction between the size, diversity, and characteristics of the training data and the performance of the IPTS tool.

In order to design the IPTS tool, a manual expert-based approach is presented that uses insights from a statistical analysis, which feeds the set-up of the IPTS tool. This approach may be modified using a semi-automated or automated approach allowing for less dependency on expert

knowledge with respect to the algorithm [33], and thus may be applicable for tuning parameters of any type of MMA.

Further research includes implementing the IPTS tool to calibrate the MMA parameters for each individual GPS data point. Subsequently, the MMA is executed for a complete instance by considering different parameter values for each data point. Furthermore, the IPTS tool may be integrated into the MMA, in order to be executed in real-time applications. Additionally, further algorithmic advances in the MMA may be incorporated, including the usage of the shortest travel time criterion for solving the MMP, particularly under congested traffic environments, as in [34, 35, 36].

ACKNOWLEDGMENT

The authors are grateful for the contributions of Dr. Darren Scott from McMaster University, and Portage County Highway Department for providing the data.

REFERENCES

- [1] M. Bierlaire, J. Chen, and J. Newman, "A probabilistic map matching method for smartphone GPS data", *Transp. Res. Part C*, vol. 26, pp. 78-98, 2013.
- [2] C. Blazquez, and A. Vonderohe, "Simple Map-Matching Algorithm Applied to Intelligent Winter Maintenance Vehicle Data", *Transportation Research Record: J. of the Transportation Research Board*, vol. 1935, pp. 68-76, 2005.
- [3] W. Chen, Z. Li, M. Yu, and Y. Chen, "Effects of sensor errors on the performance of map-matching", *J. of Navigation*, vol. 58, pp. 273-282, 2005.
- [4] R. Dalumpines, and D. Scott, "GIS-Based map-matching: development and demonstration of a post-processing map-matching algorithm for transportation research", in *Advancing geoinformation science for a changing world. Lecture Notes in Geoinformation and Cartography*, S. Geertman, W. Reinhardt, and F. Toppen, Eds. Berlin: Springer-Verlag, 2011.
- [5] N. Velaga, M. Qudus, and A. Bristow, "Developing an Enhanced Weight-Based Topological Map-Matching Algorithm for Intelligent Transport Systems", *Transp. Res. Part C*, vol.17, no. 6, pp. 672-683, 2009.
- [6] A. Oran and P. Jaillet, "A Precise Proximity-Weight Formulation for Map Matching Algorithms", 10th Workshop on Positioning, Navigation and Communication (WPNC), 2013, pp. 1-6.
- [7] C. Blazquez, P. Miranda, and A. Ponce, "Performance of a New Enhanced Topological Decision Rule Map Matching Algorithm for Transportation Applications", *J. of App. Res. & Tech.*, vol. 10, pp. 929-940, 2012.
- [8] C. Blazquez, "A Decision-Rule Topological Map-Matching Algorithm with Multiple Spatial Data", In *Global Navigation Satellite Systems: Signal, Theory and Applications*, J. Shuanggen, Ed. Croatia: Intechweb, 2012, pp. 215-240.
- [9] M. Hashemi, and H. Karimi, "A Critical Review of Real-Time Map-Matching Algorithms: Current Issues and Future Directions", *Comput. Environ. Urban*, vol. 48, pp. 153-165, 2014.
- [10] G. Taylor, and G. Blewitt, *Intelligent Positioning GIS-GPS Unification* West Sussex: John Wiley & Sons, Ltd., 2006.
- [11] M. Ren, and H. Karimi, "A hidden Markov model-based map-matching algorithm for wheelchair navigation", *J. of Navigation*, vol. 62, no. 3, pp. 383-395, 2009.
- [12] Y. Zhao, *Vehicle location and navigation systems*, Norwood, MA: Artech House, Inc., 1997.
- [13] C. Blazquez, and A. Vonderohe, "Effects of Controlling Parameters on Performance of a Decision-Rule Map-Matching Algorithm", *ASCE J. of Transportation Engineering*, vol. 135, no. 12, pp. 966-973, 2009.
- [14] F. Marchal, J. Hackney, and K. Axhausen, "Efficient map matching of large global positioning system data sets", *J. of the Transportation Research Board*, vol. 1935, no. 1, pp. 93-100, 2005.
- [15] M. Qudus, W. Ochieng, and R. Noland, "Current Map-Matching Algorithms for Transport Applications: State-Of-The Art and Future Research Directions", *Transp. Res. Part C*, vol. 15, pp. 312-328, 2007.
- [16] Z. Wang, and Z. Yang, Z. "Research on the Map Matching of Typical Region Based on the Topological Analysis" In *Proc. 2nd Int. Conf. on Intelligent Computation Technology and Automation*, Zhangjiajie, China, 2009.
- [17] D. Papinski, D. Scott, and S. Doherty, "Exploring the Route Choice Decision-Making Process: A Comparison of Planned and Observed Routes Obtained Using Person-Based GPS", *Transp. Res. Part F*, vol. 12, pp. 347-358, 2009.
- [18] T. Rasmussen, J. Ingvarsson, K. Halldorsdottir, and O. Nielsen, "Improved methods to deduct trip legs and mode from travel surveys using wearable GPS devices: A case study from the Greater Copenhagen area", *Comput. Environ. Urban*, vol. 54, pp. 301-313, 2015.
- [19] L. Shen, and R. Stopher, "Review of GPS Travel Survey and GPS data-processing methods", *Transp. Reviews*, vol. 34, no. 3, pp. 316-334, 2014.
- [20] K. Sheridan, T. Dyjas, C. Botteron, J. Leclere, F. Dominic, and G. Marucco, "Demands of Roads. An Assisted-GNSS Solution Uses the EGNOS Data Access Service", *GPS World*, vol. 22, no. 3, pp. 28, 2011.
- [21] C. Blazquez, and P. Miranda, "A real-time topological map matching methodology for GPS/GIS based travel behavior studies", In *Mobile technologies for activity-travel data collection and analysis*, S. Rasouli, and H. Timmermans, Ed. New York: IGI, 2014, pp.155-173.
- [22] J. Ries, P. Beullens, and D. Salt, "Instance-specific multi-objective parameter tuning based on fuzzy logic", *European J. of Operational Research*, vol. 218, pp. 305-315, 2012.
- [23] S. P. Coy, B. L. Golden, G.C. Runger, and E.C. Wasil, "Using experimental design to find effective parameter settings for heuristics", *J. of Heuristics*, vol.7, no. 1, pp. 77-97, 2001.
- [24] B. Adenso-Diaz, and M. Laguna, "Fine-tuning of algorithms using fractional experimental designs and local search", *Operations Research*, vol. 54, no. 1, pp. 99-114, 2006.
- [25] F. Dobsław, "A parameter tuning framework for metaheuristics based on design of experiments and artificial neural networks" In *Proc. Int. Conf. Computer Mathematics and Natural Computing*, 2010, pp. 1-4.
- [26] M. Kern, "Parameter Adaption in Heuristic Search – A Population-Based Approach", Ph.D. Thesis, University of Essex, 2006.
- [27] J. Ries, and P. Beullens, "A semi-automated design of instance-based fuzzy parameter tuning for metaheuristics based on decision tree induction", *J. of the Operational Research Society*, vol. 66, no. 5, pp. 782-793, 2015.
- [28] R. Pavon, F. Diaz, R. Laza, and V. Luzon, "Automatic parameter tuning with a Bayesian case-based reasoning system", *Expert Systems with Application*, vol. 36, no. 2, pp. 3407-3420, 2009.
- [29] L.A. Zadeh, "Fuzzy sets", *Information and Control*, vol. 8, pp. 338-353, 1965.
- [30] O. Cordon, "A historical review of evolutionary learning methods for Mamdani-type fuzzy rule-based systems: Designing interpretable genetic fuzzy systems", *Int. J. of App. Reasoning*, vol. 52, pp. 894-913, 2011.
- [31] C. A. Blazquez, *Decision-Rule Algorithm for Map-Matching in Transportation*, PhD Thesis, University of Wisconsin-Madison, 2005.
- [32] C. A. Blazquez, J. Ries, and P. A. Miranda, "Towards a parameter tuning approach for a map-matching algorithm" In *Proc. IEEE Int. Conf. Vehicular Electronics and Safety*, Vienna, 2017, pp. 85-90.
- [33] J. Ries, P. Beullens, and Y. Wang, "Instance-specific parameter tuning for meta-heuristics", *Meta-heuristics optimization algorithms in engineering, business, economics, and finance*, T. Vasant, Ed. Hershey: IGI Global, 2012, pp. 136-170.
- [34] J. Tang, Y. Song, H. Miller, and X. Zhou, "Estimating the most likely space-time paths, dwell times and path uncertainties from vehicle trajectory data: A time geographic method", *Transp. Res. Part C*, vol. 66, pp. 176-194, 2016.
- [35] T. Hunter, *Large-scale, low-latency state estimation of cyberphysical systems with an application to traffic estimation*, Ph.D. Thesis, University of California, 2014.
- [36] W. Li, D. Nie, D. Wilkie, and M. C. Lin, "Citywide estimation of traffic dynamics via sparse GPS traces," *IEEE Intell. Transp. Syst. Magazine*, vol. 9, no. 3, pp. 100-113, 2017.



Dr. Carola Blazquez received her Ph.D. degree in GeoSpatial Information Engineering from the University of Wisconsin-Madison, United States in 2005. In 2006, Dr. Blazquez joined the Engineering Sciences Department faculty at the Universidad Andres Bello in Chile, where she presently holds a position as a professor. Her research interests include Intelligent Transportation Systems (ITS), GIS for Transportation (GIS-T), and transportation safety.



Jana Ries received her PhD in Operational Research from the Department of Mathematics at the University of Portsmouth, UK. She is a Senior Lecturer in the Operations and Systems Management Group at the University of Portsmouth. Her main research is in algorithmic design and parameter calibration for heuristic algorithms in combinatorial optimization with a particular interest in the development of decision support systems in the areas of transport and manufacturing.



Dr. Roberto Leon received his Ph.D. in Computer Engineering from Universidad Tecnica Federico Santa María in Valparaiso, Chile in 2016. He is currently working as an assistant professor at the Universidad Andres Bello in Viña del Mar, Chile. His research interests include scientific computing, numerical simulation, and computational geometry.



Dr. Pablo Miranda received his Ph.D. in Logistics and Operation Research from Pontificia Universidad Catolica de Chile in 2004. Currently, he works as an associate professor at Pontificia Universidad Católica de Valparaiso, Chile. His research interests are models and algorithms for addressing real world problems in industry, focusing on operation, transport, logistics and supply chain management.