Article

Estimating the Tour Length for the Close Enough Traveling Salesman Problem

Debdatta Sinha Roy [1,*], Bruce Golden [2], Xingyin Wang [3] and Edward Wasil [4]

1   Staples Inc., Framingham, MA 01702, USA
2   Robert H. Smith School of Business, University of Maryland, College Park, MD 20742, USA; bgolden@umd.edu
3   Engineering Systems and Design, Singapore University of Technology and Design, Singapore 487372, Singapore; xingyin_wang@sutd.edu.sg
4   Kogod School of Business, American University, Washington, DC 20016, USA; ewasil@american.edu
*   Correspondence: debsroy@terpmail.umd.edu

Abstract: We construct empirically based regression models for estimating the tour length in the Close Enough Traveling Salesman Problem (CETSP). In the CETSP, a customer is considered visited when the salesman visits any point in the customer's service region. We build our models using as many as 14 independent variables on a set of 780 benchmark instances of the CETSP and compare the estimated tour lengths to the results from a Steiner zone heuristic. We validate our results on a new set of 234 instances that are similar to the 780 benchmark instances. We also generate results for a new set of 72 larger instances. Overall, our models fit the data well and do a very good job of estimating the tour length. In addition, we show that our modeling approach can be used to accurately estimate the optimal tour lengths for the CETSP.

Keywords: close enough traveling salesman problem; tour-length estimation; regression models

1. Introduction

Operations researchers have long been interested in estimating the length of tours and routes in the traveling salesman problem (TSP) and the vehicle routing problem (VRP). One of the earliest papers by Beardwood et al. [1] developed analytically derived formulas for the TSP. Christofides and Eilon [2], Hindle and Worthington [3], and Cavdar and Sokol [4] improved these formulas by using empirically estimated parameters. Golden and Alt [5] constructed interval estimates of the optimal solution value. Chien [6] and Kwon et al. [7] used parameters for the shape of the area covering the customers and the depot, the distance between customers, and the coordinates of the customers. Basel and Willemain [8] estimated the optimal tour length for 17 TSP instances using the square root of the number of cities and the variability of tour lengths. Nicola et al. [9] developed empirically based regression models for estimating the travel distance in the TSP, in the capacitated VRP with time windows, and in the multi-region, multi-depot pickup and delivery problem.

As pointed out by Nicola et al. [9], carriers and logistics companies may need to solve a large number of routing problems which might require a significant computational effort. Approaches that generate fast, accurate estimates for the travel distance are highly desirable in practice for a wide range of real-world routing problems. These estimates can be used by companies to make strategic, tactical, and operational decisions quickly [9].

In this paper, we construct empirically based regression models for estimating the tour length in the Close Enough Traveling Salesman Problem (CETSP). In the TSP, a salesman must visit the exact customer location. In the CETSP, a customer has a service region and is considered visited when the salesman visits any point in the customer's service region. The CETSP arises in practical applications including utility meter reading at residential locations using radio frequency identification and using drones for aerial surveillance. In

another application, during multi-visit drone routing [10] it becomes important to quickly assess whether the duration to cover a region of interest would exceed a drone's battery life.

In the CETSP, a service region is assumed to be a circular disk centered at the customer location with a specified radius. The objective is to visit all customers in the shortest distance traveled starting and ending at the depot. The TSP is a special case of the CETSP when all radii of the circular disks are zero, making the CETSP at least as difficult to solve as the TSP. In order to solve an instance of the CETSP, it is not enough to determine the sequence in which the customers are visited. We must also determine the locations at which these customers are visited within their respective service regions. In Figure 1, we show an instance of a CETSP with 12 customers denoted by $C_1, \ldots, C_{12}$ and a depot denoted by C_0. The service region is specified by a circle centered at the customer's location. A feasible CETSP tour is shown by the solid lines with arrows. The tour passes through at least one point in the service region of each customer. If a tour passes through an overlap of several disks, all customers that define those disks are served. A Steiner zone is an overlap of disks. If a Steiner zone is contained in at most k disks, it has degree k. For example, the location of each of customers C_1, C_6, C_8, and C_{11} is within a degree 1 Steiner zone. Similarly, the location of each of customers C_3, C_4, C_5, C_9, and C_{12} is within a degree 2 Steiner zone. Whereas, the location of each of customers C_2, C_7, and C_{10} is within a degree 3 Steiner zone. For more details on Steiner zones, see Wang et al. [11].
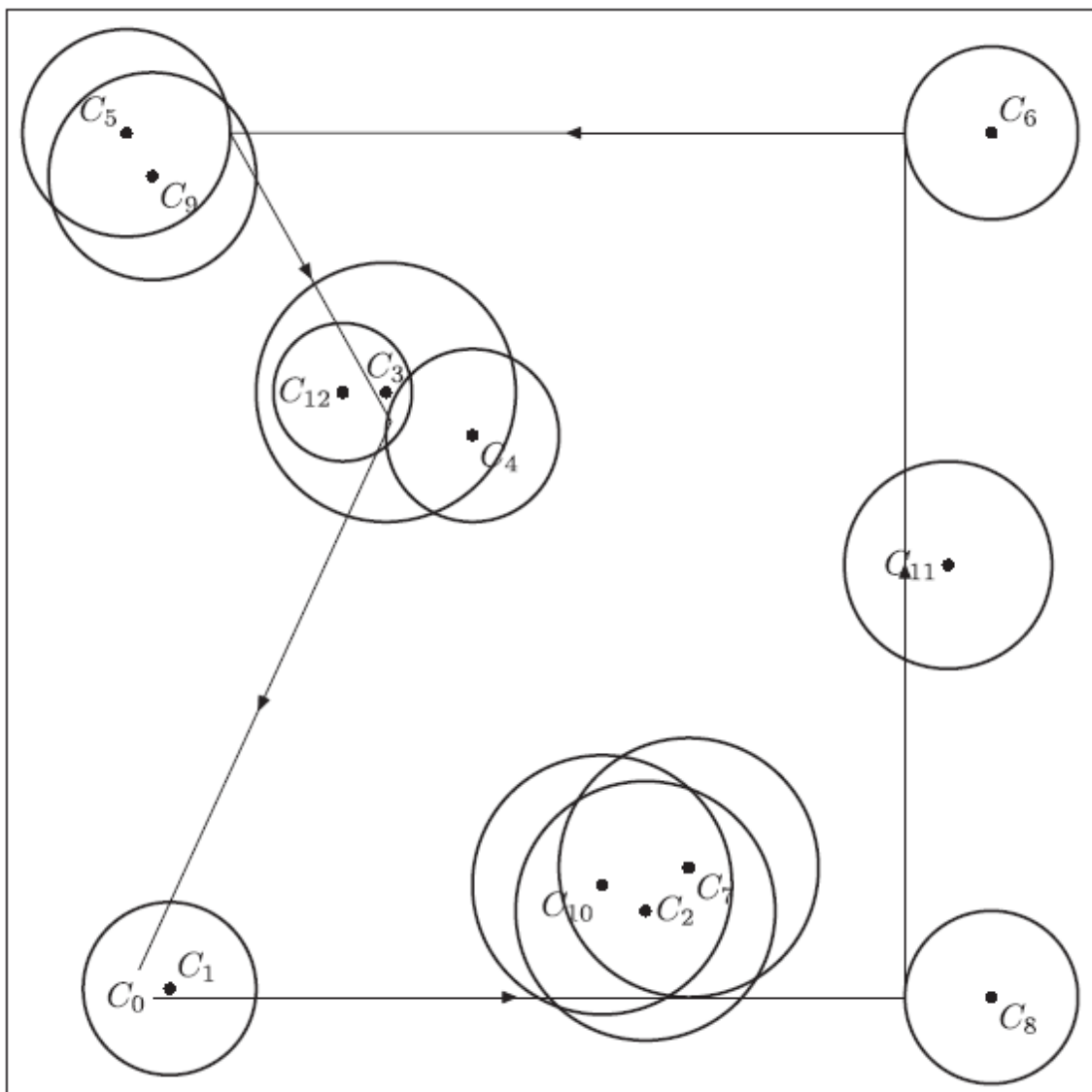

Figure 1. An instance of a CETSP with 12 customers [11].

Progress in computational integer programming has been remarkable over the past 30 years; see Bertsimas et al. [12] for details. In particular, exact solvers for the TSP (e.g., Concorde) can now generate solutions quickly to most large instances with hundreds or thousands of customers. However, there are very few exact approaches for the CETSP. Exact approaches have been developed by Behdani and Smith [13] and Coutinho et al. [14]. Tight bounds have been developed by Carrabs et al. [15]. Coutinho et al. [14] proposed a branch-and-bound algorithm and applied it to instances from the literature. Computation times ranged from seconds for small instances to four hours for large instances with hundreds of customer locations.

In order to generate high-quality solutions quickly, heuristics have been developed and tested for the CETSP. These include the use of supernodes by Gulczynski et al. [16] and Dong et al. [17], Steiner zones by Mennell [18], Mennell et al. [19], and Wang et al. [11], and genetic algorithms by Silberholz and Golden [20], Yuan et al. [21], and Yang et al. [22].

The remainder of the paper is organized as follows. In Section 2, we give the regression models and discuss the fitness measures to test the performance of the models. In Section 3, we show the results of the regression models, best subset model selection, and model validation. In Section 4, we present our conclusions and future directions.

2. Regression Models and Fitness Measures

The Steiner zone variable neighborhood search (SZVNS) heuristic developed by Wang et al. [11] finds high-quality solutions to instances of the CETSP. We use SZVNS tour lengths in our regression models to estimate CETSP tour lengths.

Our regression model can be represented by $y_i = \hat{\beta}_1 + \hat{\beta}_2 \times \mathrm{n}_i + \hat{\beta}_3 \times \mathrm{A}_i + \hat{\beta}_4 \times \mathrm{MinP}_i + \hat{\beta}_5 \times \mathrm{MaxP}_i + \hat{\beta}_6 \times \mathrm{VarP}_i + \hat{\beta}_7 \times \mathrm{SumMinP}_i + \hat{\beta}_8 \times \mathrm{SumMaxP}_i + \hat{\beta}_9 \times \mathrm{MinM}_i + \hat{\beta}_{10} \times \mathrm{MaxM}_i + \hat{\beta}_{11} \times \mathrm{SumM}_i + \hat{\beta}_{12} \times \mathrm{VarM}_i + \hat{\beta}_{13} \times (\mathrm{VarX} \times \mathrm{VarY})_i + \hat{\beta}_{14} \times \mathrm{AvgR}_i + \hat{\beta}_{15} \times \mathrm{SZ}_i$, where $y_i = \mathbb{E}(Y_i)$, $\hat{\beta}_k = \mathbb{E}(\beta_k)$, $k \in \{1, \dots, 15\}$, i denotes a CETSP instance, and $\mathbb{E}()$ denotes the expected value. The dependent variable Y_i is the tour length generated by the SZVNS heuristic. In Table 1, we give the definitions of the independent variables for the regression model. Nodes represent customers and the depot. The size of an instance is captured by n and A. MinP, MaxP, VarP, SumMinP, and SumMaxP capture the distances between nodes. MinM, MaxM, SumM, and VarM capture the distances to the average node represented by the mean x-coordinate and the mean y-coordinate. VarX×VarY captures the spread of the instance across the two axes. AvgR captures the mean of the radii of the customer service regions. The service region radius of a depot is always zero. SZ captures the feature of the instance that is exploited by the SZVNS heuristic. AvgR and SZ are used to capture the geometric features unique to the CETSP. These two independent variables would not be used in a regression model that estimates TSP tour lengths.

We use the 780 CETSP benchmark instances and their tour lengths produced by the SZVNS heuristic given in Wang et al. [11]. The node locations (depot and the customers) are generated randomly and all customers in an instance have the same radius for the service regions. The instances have 6, 8, 10, 12, 14, 16, 18, 20, 25, or 30 customers. The radii for the customer service regions are 0.25, 0.50, or 1.00. There are 30 instances for each combination of the radius of the customer service regions and the number of customers up to 20. There are 10 instances for each combination of the radius of the customer service regions and the number of customers greater than 20.

We use mean percentage error (MPE) and mean absolute percentage error (MAPE) to assess the quality of the approximation of the CETSP tour lengths from the regression model (y_i) with respect to the tour lengths from the SZVNS heuristic (Y_i). MPE and MAPE are defined by $100 \times (\sum_{i=1}^{N} (Y_i - y_i) / Y_i) / N$ and $100 \times (\sum_{i=1}^{N} |Y_i - y_i| / Y_i) / N$, respectively, where N denotes the number of instances. A value of MPE close to zero indicates that there is almost an equal distribution of instances with tour lengths being overestimated ($Y_i < y_i$) and underestimated ($Y_i > y_i$). The value of MAPE is always non-negative, and a small value indicates that the tour-length estimates are close to the SZVNS tour lengths for most of the instances.

Table 1. Definitions of the independent variables for the regression model.

Independent Variable	Definition
n	Number of nodes
A	Area of the smallest rectangle covering all nodes
MinP	Minimum distance across all pairs of nodes
MaxP	Maximum distance across all pairs of nodes
VarP	Variance of distances across all pairs of nodes
SumMinP	Sum of distances to the nearest neighbor of each node
SumMaxP	Sum of distances to the farthest neighbor of each node
MinM	Minimum distance to the average node
MaxM	Maximum distance to the average node
SumM	Sum of distances to the average node
VarM	Variance of distances to the average node
VarX×VarY	Product of variances of the nodes across two axes
AvgR	Average radius of the customer service regions
SZ	Number of Steiner zones of degree three and less that are not dominated by other Steiner zones of degree three and less

We use adjusted R^2, Studentized residuals, Mallows's C_p, and Bayesian information criterion (BIC) to assess the quality of the model fits. Residuals $(Y_i - y_i)$ have a mean of zero. Studentized residuals are scaled residuals with unit variance. Mallows's C_p and BIC are used in the context of model selection where the goal is to find the best model involving a subset of the independent variables. Mallows's C_p addresses the issue of model overfitting by penalizing for adding extra variables. The value of Mallows's C_p should be close to the number of independent variables in the model (p) to indicate the absence of overfitting. BIC penalizes a model for having more independent variables, and the penalty increases as the number of instances (size of the data set) increases. The lower the value of BIC, the better is the model fit.

3. Regression Results

In Table 2, we present the regression results for the 780 instances using all 14 independent variables in our model. The regression model has an adjusted R^2 value of 0.921 which indicates a very good model fit. The variables n and SZ are not significant at the 10% level. The remaining 12 variables are significant at various levels. In Figure 2a, we give the histogram of Studentized residuals for the model with 14 variables. This histogram shows that there is almost an equal distribution of instances with positive and negative residuals. This is also indicated by the MPE value of -0.192% which is close to zero. The MAPE value indicates that the tour-length estimates from the regression model differ by an average of 3.984% from the SZVNS tour lengths.
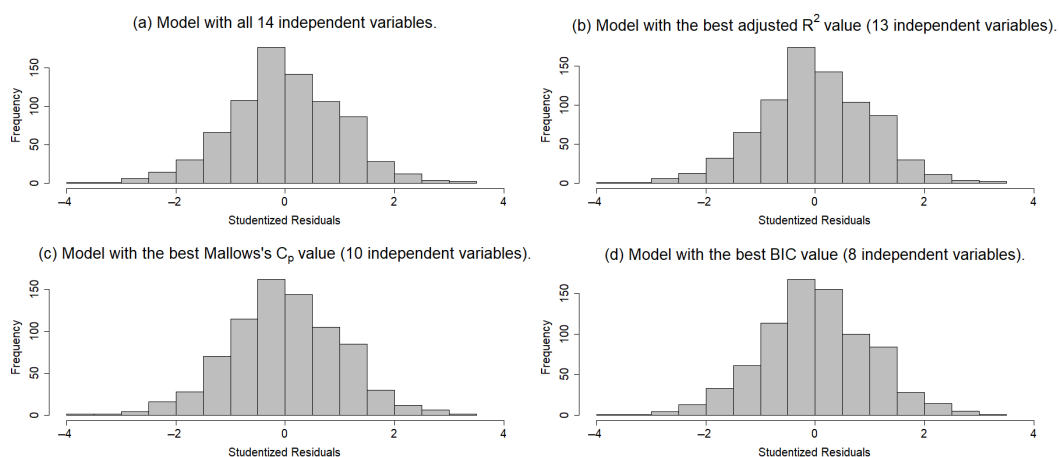
Figure 2. Histograms of Studentized residuals for four models.

Table 2. Regression results.

Coefficient	Mean Values
Intercept (β_1)	15.231 ****
n (β_2)	0.225
A (β_3)	0.048 ****
MinP (β_4)	−0.654 ****
MaxP (β_5)	0.224 **
VarP (β_6)	0.106 *
SumMinP (β_7)	0.361 ****
SumMaxP (β_8)	0.036 **
MinM (β_9)	0.459 ***
MaxM (β_{10})	−0.418 **
SumM (β_{11})	−0.067 **
VarM (β_{12})	0.683 ****
VarX×VarY (β_{13})	0.014 ****
AvgR (β_{14})	−9.092 ****
SZ (β_{15})	−0.026
Adjusted R^2	0.921
MPE	−0.192%
MAPE	3.984%

* $p < 0.1$; ** $p < 0.05$; *** $p < 0.01$; **** $p < 0.001$.

3.1. Best Subset Model Selection

Although the regression model shown with all 14 variables performed well in estimating the SZVNS tour lengths, two of the variables were not significant at the 10% level and one variable was significant only at the 10% level. We now examine whether we can remove some variables to create a parsimonious model without compromising much on the model fit.

We create models with 1 to 14 independent variables in the following way. We start by constructing models with only one independent variable and identifying the 1-variable model that produced the largest adjusted R^2 value. Then we find the 2-variable model that produced the largest adjusted R^2 value, and so on until we consider the 14-variable model. In Table 3, we show the best subset models based on the adjusted R^2 value. For example, out of all possible models with two independent variables, the model containing the variables SumMaxP and AvgR produced the largest adjusted R^2 value. Following this process, we create 14 models that contain 1 to 14 variables. In Table 4, we show the adjusted R^2, Mallows's C_p, and BIC values of the best subset models from Table 3.

Over these 14 models, we select the model that has the largest adjusted R^2 value. This is the model with 13 variables and an adjusted R^2 value of 0.921. There are other models with the same adjusted R^2 value when rounded to three decimal places.

We show the coefficients of the first model in Table 5 under the column labeled "Best adjusted R^2". We note that the adjusted R^2 value of 0.921 indicates a very good model fit. The variable SZ is not in this model and the variable n is not significant at the 10% level. The remaining 12 variables are significant at various levels. In Figure 2b, we give the histogram of Studentized residuals for this model. This histogram shows that there is almost an equal distribution of instances with positive and negative residuals. This is also indicated by the MPE value of −0.192% which is close to zero. The MAPE value indicates that the tour-length estimates from this regression model differ by an average of 3.983% from the SZVNS tour lengths.

Table 3. Best subset models based on adjusted R^2.

Variable	Number of Variables													
	1	2	3	4	5	6	7	8	9	10	11	12	13	14
n									*				*	*
A			*	*	*	*	*	*	*	*	*	*	*	*
MinP								*	*	*	*	*	*	*
MaxP									*	*	*	*	*	*
VarP												*	*	*
SumMinP			*	*	*	*	*	*	*	*	*	*	*	*
SumMaxP	*	*		*		*	*	*		*	*	*	*	*
MinM						*	*	*	*	*	*	*	*	*
MaxM										*	*	*	*	*
SumM					*						*	*	*	*
VarM					*	*	*	*	*	*	*	*	*	*
VarX×VarY							*	*	*	*	*	*	*	*
AvgR		*	*	*	*	*	*	*	*	*	*	*	*	*
SZ														*

Table 4. Adjusted R^2, Mallows's C_p, and BIC values of the best subset models from Table 3.

Best Subset Model	Adjusted R^2	Mallows's C_p	BIC
1-variable	0.598	3189.7	−698
2-variable	0.787	1320.5	−1189
3-variable	0.876	447.1	−1605
4-variable	0.899	218.1	−1762
5-variable	0.911	107.7	−1850
6-variable	0.918	42.1	−1906
7-variable	0.919	30.2	−1912
8-variable	0.920	16.9	−1921
9-variable	0.921	15.0	−1918
10-variable	0.921	13.0	−1916
11-variable	0.921	14.5	−1911
12-variable	0.921	15.7	−1906
13-variable	0.921	16.9	−1901
14-variable	0.921	18.0	−1895

Next, from the 14 models in Table 3, we select the model that has its Mallows's C_p value closest to the number of independent variables in the model. This is the model with 10 independent variables and a Mallows's C_p value of 13.0.

We show the coefficients of the second model in Table 5 under the column labeled "Best Mallows's C_p". We note that the adjusted R^2 value of 0.921 indicates a very good model fit. The four variables n, VarP, SumM, and SZ are not in this model. The remaining 10 variables are significant at various levels. In Figure 2c, we give the histogram of Studentized residuals for this model. This histogram shows that there is almost an equal distribution of instances with positive and negative residuals. This is also indicated by the MPE value of −0.194% which is close to zero. The MAPE value indicates that the tour-length estimates from this regression model differ by an average of 3.995% from the SZVNS tour lengths.

Finally, from the 14 models in Table 3, we select the model that has the smallest BIC value. This is the model with eight independent variables and a BIC value of −1921.

Table 5. Regression results on the three selected best subset models.

Coefficient	Best Adjusted R^2	Best Mallows's C_p	Best BIC
Intercept (β_1)	15.320 ****	15.485 ****	16.334 ****
n (β_2)	0.188		
A (β_3)	0.048 ****	0.046 ****	0.044 ****
MinP (β_4)	−0.668 ****	−0.734 ****	−0.674 ****
MaxP (β_5)	0.224 **	0.230 **	
VarP (β_6)	0.101 *		
SumMinP (β_7)	0.359 ****	0.362 ****	0.362 ****
SumMaxP (β_8)	0.036 **	0.025 ****	0.027 ****
MinM (β_9)	0.455 ***	0.508 ****	0.498 ****
MaxM (β_{10})	−0.410 **	−0.273 **	
SumM (β_{11})	−0.064 **		
VarM (β_{12})	0.685 ****	0.805 ****	0.797 ****
VarX×VarY (β_{13})	0.014 ****	0.011 ****	0.012 ****
AvgR (β_{14})	−9.059 ****	−9.059 ****	−9.059 ****
SZ (β_{15})			
Number of variables	13	10	8
Adjusted R^2	0.921	0.921	0.920
Mallows's C_p		13.0	
BIC			−1921
MPE	−0.192%	−0.194%	−0.202%
MAPE	3.983%	3.995%	4.008%

* $p < 0.1$; ** $p < 0.05$; *** $p < 0.01$; **** $p < 0.001$.

We show the coefficients of the third model in Table 5 under the column labeled "Best BIC". We note that the adjusted R^2 value of 0.920 indicates a very good model fit. The six variables n, MaxP, VarP, MaxM, SumM, and SZ are not in this model. The remaining eight variables are significant at the 0.1% level. In Figure 2d, we give the histogram of Studentized residuals for this model. This histogram shows that there is almost an equal distribution of instances with positive and negative residuals. This is also indicated by the MPE value of −0.202% which is close to zero. The MAPE value indicates that the tour-length estimates from this regression model differ by an average of 4.008% from the SZVNS tour lengths.

All three best subset models based on adjusted R^2, Mallows's C_p, and BIC performed well. In fact, their adjusted R^2 values are nearly the same (about 0.921), as are their MAPE values (about 4%). We recommend the model with the least number of independent variables (Best BIC model with eight variables) to estimate the SZVNS tour lengths for CETSP instances with random node locations. The coefficients of all eight variables in the Best BIC model are highly significant ($p < 0.001$).

3.2. Model Validation

We generated 234 new CETSP instances, similar to the 780 instances we used to develop the regression models, to test the Best BIC model with eight variables. The node locations (depot and the customers) are generated randomly and all customers in an instance have the same radius for the service regions. The instances have 6, 8, 10, 12, 14, 16, 18, 20, 25, or 30 customers. The radii for the customer service regions are 0.25, 0.50, or 1.00. There are nine new instances for each combination of the radius of the customer service regions and the number of customers up to 20. There are three new instances for each combination of the radius of the customer service regions and the number of customers greater than 20. We estimated the SZVNS tour lengths for the 234 new instances using the Best BIC model. The out-of-sample MPE and MAPE values are −0.344% and 4.233%, respectively, which indicate that the Best BIC model performs well in estimating the SZVNS tour lengths on the new CETSP instances.

We generated the optimal tour lengths for the 234 new instances using a branch-and-bound algorithm [14] in order to test the usefulness of the eight independent variables selected in the Best BIC model in estimating the optimal tour lengths. SZ is the only independent variable specific to the SZVNS algorithm and is not a part of any of the three models shown in Table 5. Therefore, the eight independent variables in the Best BIC model should have more general usage in estimating tour lengths. The regression model built using the eight independent variables in the Best BIC model and trained on the optimal tour lengths of the 234 new instances is given by: $optimal_i = 13.394 + 0.054 \times A_i - 0.226 \times MinP_i + 0.316 \times SumMinP_i + 0.033 \times SumMaxP_i + 0.640 \times MinM_i + 0.852 \times VarM_i + 0.018 \times (VarX \times VarY)_i - 8.774 \times AvgR_i$, where $optimal_i$ denotes the optimal tour length for instance i. The signs and the orders of magnitude of the coefficients are the same as the coefficients in the Best BIC model. The regression model for estimating the optimal tour lengths with eight independent variables has an adjusted R^2 value of 0.937, indicating a very good model fit. The MPE value of -0.161%, which is close to zero, indicates that there is almost an equal distribution of instances with positive and negative residuals. The MAPE value of 3.735% also indicates the usefulness of this model in estimating optimal tour lengths.

The eight independent variables in the Best BIC model adequately capture the geometric properties of the CETSP instances where the node locations are generated randomly. The model can be trained on the tour lengths generated from any heuristic or optimal algorithm to accurately estimate the tour lengths for that specific heuristic or optimal algorithm on other similar CETSP instances.

We generated a new set of 72 larger CETSP instances to further assess the Best BIC model with eight variables. The node locations (depot and the customers) are generated randomly and all customers in an instance have the same radius for the service regions. The instances have 35, 40, 45, or 50 customers. The radii for the customer service regions are 0.25, 0.50, or 1.00. There are six new instances for each combination of the radius of the customer service regions and the number of customers. We estimated the SZVNS tour lengths for the 72 new larger instances using the Best BIC model. The out-of-sample MPE and MAPE values are -0.356% and 4.389%, respectively, which indicate that the Best BIC model performs well in estimating the SZVNS tour lengths on the larger CETSP instances.

We tried generating the optimal tour lengths for the new set of 72 larger instances using the branch-and-bound algorithm from Coutinho et al. [14] with a maximum run time of one hour. We were unable to find the optimal solution to seven instances, so we do not report regression results based on optimal tour lengths. This experiment reinforces the need for a regression-based model to quickly estimate the CETSP tour lengths of larger and more difficult instances for a specific heuristic or optimal algorithm. All CETSP instances used in our experiments are given in Sinha Roy et al. [23].

4. Conclusions and Future Directions

We applied regression models to an important problem in the routing literature–the CETSP. We demonstrated that it is possible to quickly and accurately estimate tour lengths using a regression model without generating the actual tours. We showed that the tour lengths generated by the SZVNS heuristic or by an optimal algorithm could be estimated with an average error of about 4% by a regression model with eight independent variables where the node locations are generated randomly. In the future, we would like to develop models for instances with node locations that are generated in different structured ways.

Author Contributions: Conceptualization, B.G.; Data curation, D.S.R. and X.W.; Formal analysis, D.S.R. and X.W.; Supervision, B.G. and E.W.; Writing—original draft, D.S.R.; Writing—review & editing, B.G. and E.W. All authors have read and agreed to the published version of the manuscript.

Conflicts of Interest: The authors declare no conflict of interest.

References

1. Beardwood, J.; Halton, J.H.; Hammersley, J.M. The shortest path through many points. *Math. Proc. Camb. Philos. Soc.* **1959**, *55*, 299–327. [CrossRef]
2. Christofides, N.; Eilon, S. Expected distances in distribution problems. *J. Oper. Res. Soc.* **1969**, *20*, 437–443. [CrossRef]
3. Hindle, A.; Worthington, D. Models to estimate average route lengths in different geographical environments. *J. Oper. Res. Soc.* **2004**, *55*, 662–666. [CrossRef]
4. Cavdar, B.; Sokol, J. A distribution-free TSP tour length estimation model for random graphs. *Eur. J. Oper. Res.* **2015**, *243*, 588–598. [CrossRef]
5. Golden, B.; Alt, F. Interval estimation of a global optimum for large combinatorial problems. *Nav. Res. Logist. Q.* **1979**, *26*, 69–77. [CrossRef]
6. Chien, T.W. Operational estimators for the length of a traveling salesman tour. *Comput. Oper. Res.* **1992**, *19*, 469–478. [CrossRef]
7. Kwon, O.; Golden, B.; Wasil, E. Estimating the length of the optimal TSP tour: An empirical study using regression and neural networks. *Comput. Oper. Res.* **1995**, *22*, 1039–1046. [CrossRef]
8. Basel, J.; Willemain, T.R. Random tours in the traveling salesman problem: Analysis and application. *Comput. Optim. Appl.* **2001**, *20*, 211–217. [CrossRef]
9. Nicola, D.; Vetschera, R.; Dragomir, A. Total distance approximations for routing solutions. *Comput. Oper. Res.* **2019**, *102*, 67–74. [CrossRef]
10. Poikonen, S.; Golden, B. Multi-visit drone routing problem. *Comput. Oper. Res.* **2020**, *113*, 104802. [CrossRef]
11. Wang, X.; Golden, B.; Wasil, E. A Steiner zone variable neighborhood search heuristic for the close-enough traveling salesman problem. *Comput. Oper. Res.* **2019**, *101*, 200–219. [CrossRef]
12. Bertsimas, D.; King, A.; Mazumder, R. Best subset selection via a modern optimization lens. *Ann. Stat.* **2016**, *44*, 813–852. [CrossRef]
13. Behdani, B.; Smith, J.C. An integer-programming-based approach to the close-enough traveling salesman problem. *INFORMS J. Comput.* **2014**, *26*, 415–432. [CrossRef]
14. Coutinho, W.P.; do Nascimento, R.Q.; Pessoa, A.A.; Subramanian, A. A branch-and-bound algorithm for the close-enough traveling salesman problem. *INFORMS J. Comput.* **2016**, *28*, 752–765. [CrossRef]
15. Carrabs, F.; Cerrone, C.; Cerulli, R.; Gaudioso, M. A novel discretization scheme for the close-enough traveling salesman problem. *Comput. Oper. Res.* **2017**, *78*, 163–171. [CrossRef]
16. Gulczynski, D.; Heath, J.; Price, C. The close enough traveling salesman problem: A discussion of several heuristics. In *Perspectives in Operations Research: Papers in Honor of Saul Gass' 80th Birthday*; Springer: New York, NY, USA, 2006; pp. 271–283.
17. Dong, J.; Yang, N.; Chen, M. Heuristic approaches for a TSP variant: The automatic meter reading shortest tour problem. In *Extending the Horizons: Advances in Computing, Optimization, and Decision Technologies*; Springer: New York, NY, USA, 2007; pp. 145–163.
18. Mennell, W.K. Heuristics for Solving Three Routing Problems: Close-Enough Traveling Salesman Problem, Close-Enough Vehicle Routing Problem, Sequence-Dependent Team Orienteering Problem. Ph.D. Thesis, Decision, Operations & Information Technologies, University of Maryland, College Park, MD, USA, 2009.
19. Mennell, W.K.; Golden, B.; Wasil, E. A Steiner-zone heuristic for solving the close-enough traveling salesman problem. In *Operations Research, Computing, and Homeland Defense*; INFORMS: Catonsville, MD, USA, 2011; pp. 162–183.
20. Silberholz, J.; Golden, B. The generalized traveling salesman problem: A new genetic algorithm approach. In *Extending the Horizons: Advances in Computing, Optimization, and Decision Technologies*; Springer: New York, NY, USA, 2007; pp. 165–181.
21. Yuan, B.; Orlowska, M.; Sadiq, S. On the optimal robot routing problem in wireless sensor networks. *IEEE Trans. Knowl. Data Eng.* **2007**, *19*, 1252–1261. [CrossRef]
22. Yang, Z.; Xiao, M.-Q.; Ge, Y.-W.; Feng, D.-L.; Zhang, L.; Song, H.-F.; Tang, X.-L. A double-loop hybrid algorithm for the traveling salesman problem with arbitrary neighbourhoods. *Eur. J. Oper. Res.* **2018**, *265*, 65–80. [CrossRef]
23. Sinha Roy, D.; Golden, B.; Wang, X.; Wasil, E. Instances for the Close Enough Traveling Salesman Problem. Data Set. 2021. Available online: http://doi.org/10.5281/zenodo.4632436 (accessed on 8 April 2021).