Freight distribution performance indicators for service quality planning in large transportation networks

Francesco Paolo Deflorio [1], Guido Perboli[2+3] , Roberto Tadei [2]

1 - POLITECNICO DI TORINO - DITIC (DEPARTMENT OF HYDRAULICS, TRANSPORT AND CIVIL INFRASTRUCTURES) - Transport Engineering Working Group.

corso Duca degli Abruzzi, 24 - 10129 - Torino - Italy - Europe
Tel. +39 011 564 56 01; Fax. +39 011 564 56 99; e-mail: francesco.deflorio@polito.it

2 - POLITECNICO DI TORINO - DAUIN (DEPARTMENT OF CONTROL AND COMPUTER ENGINEERING) - Optimization and Operations Research (ORO) Working Group.

corso Duca degli Abruzzi, 24 - 10129 - Torino - Italy – Europe

3 - Centre Interuniversitaire de Recherche sur les Reseaux d'Entreprise, la Logistique et le Transport, Montréal (QC), Canada.

Abstract

This paper studies the use of performance indicators in routing problems to estimate how transportation cost is affected by the quality of service offered. The quality of service is assumed to be directly dependent on the size of the time windows. Smaller time windows mean better service. Three performance indicators are introduced. These indicators are calculated directly from the data without the need of a solution method. The introduced indicators are based mainly on a “request compatibility”, which describes whether two visits can be scheduled consecutively in a route. Other two indicators are introduced, which get their values from a greedy constructive heuristic. After introducing the indicators, the correlation between indicators and transportation cost is examined. It is concluded that the indicators give a good first estimation on the transportation cost incurred when providing a certain quality of service. These indicators can be calculated easily in one of the first planning steps

without the need of a sophisticated solution tool. The contribution of the paper is the introduction of a simple set of performance indicators that can be used to estimate the transportation cost of a routing problem with time windows.

Keywords

Freight distribution service, time windows, service quality, requests compatibility indicators, experimental analysis.

1 Introduction

The freight transportation sector is constantly changing as a consequence of the growth and transformation of the economic activity. In recent years companies have been reducing their storage areas to save resources, at the same time they tried to offer high quality services to customers in terms of freight availability and the ability to meet the delivery times. In this context, the new technological developments have been a positive factor for the expansion of new markets and new consumer needs. The growth of e-commerce and postal shopping, as well as a hectic life-style, have reinforced the importance of "just in time" policies in freight distribution, also to implement policies of freight consolidation in specific urban centers (Marcucci and Danielis 2008). Moreover, the service quality of a transportation carrier is often related to the travel time, and can vary according to both socio-economics and trip characteristics (Puckett et al. 2008). The total travel time of a vehicle trip depends on several aspects such as actual travel time, waiting and access time, congestion, deadlines or service features, etc. (Wardman 1998).

This study relates to time constrained freight distribution planning, like parcel express distribution. In this type of service, one of the customers' needs is defined by a time interval, known as the *time window* within which the customer wants the freight to be delivered or picked up by the transportation service. The width of the time window can be considered a quality factor of the service, but its nature and configuration also affects the total transportation costs (Cordeau et al. 2002). This issue should be considered at the moment of defining the service settings and characteristics. This phase is a typical phase of the planning process (see table 1 in Gayialis and Tatsiopoulos 2004 for a general framework) and it precedes the route optimization operations, for this reasons there is a possibility that not all the decision-makers have the expertise to use adequate optimization tools to simulate the various alternatives of the service configuration. The aim of this paper is to present a methodology to evaluate how the transportation cost level of a freight distribution service with time windows is affected by the width of time windows themselves. In fact they are to be considered as the expression of the quality of service. This method is based on several indicators and the cost levels are measured in terms of total travel time. Although this evaluation could also be carried out applying suitable Vehicle Routing Problem with Time Windows (VRPTW) algorithms, we remark that these tools are not always available for decision makers who are responsible of the quality setting of the distribution services. Moreover, the service setting phase generally precedes the detailed vehicle routing operation phase, and operators must be able to explore several service hypothesis (e.g. with different time windows settings), before accepting an appropriate solution. Therefore, the use of a fast and reliable tool is crucial for a feasible service planning process, and this is precisely the reason why the proposed a-priori indicators do not need any VRP optimization tool to be computed.

In practical applications, many distribution services operate without time requirements, therefore they do not define a TW. In this case, some useful information on the location of requests can be easily drawn from historical data bases, and the delivery/pick-up time can also be obtained from the delivery notes or from the users' stated preferences. For this type of services, for example, introducing an established TW would represent a possible upgrade to standards that better meet the users' needs. Given these considerations, in the planning of a new service it would be also necessary to assess the impact that the mentioned improvement in quality would have on transportation costs. On the other hand, in case of a newly introduced service where data about the requests are not available, a demand prediction process should be performed. This usually occurs in the transportation planning phase, when demand modelling is applied and data can also be simulated, based on where population, businesses and activities are located, as well as on socio-economic and land use information, road network configuration and any other useful information coming from similar systems.

The paper is organized as follows. The first section covers how the existing studies have treated the time windows issue. The paper then presents the relative planning problems found in literature. After that, the proposed methodology to compute the selected indicators is described. Before reporting the analysis of the service performance indicators, the general characteristics of the experimental design are described and the instances for testing the indicators are presented. Finally, to evaluate the capabilities of the proposed indicators in estimating the transportation cost variations, several test scenarios are shown, where a comparison has been carried out with the results obtained by means of a recent and reliable VRPTW algorithm.

2 An overview on time constrained freight distribution

In time constrained freight distribution, the high number of carriers and the tough competition of different companies make quality and price important aspects. Moreover, price settings are related to the different costs of the distribution service, including transportation costs. For a transportation carrier, variable costs are related to transportation times and the total distance travelled by the vehicles. These costs depend on various factors:

- The road network configuration, which determines the travel time between nodes. This travel time can be translated into costs (working hours, fuel costs, parking fees, etc), and also vehicle emissions, which are one of the main causes of air pollution (which also depends on the vehicle characteristics).
- The nature of the demand, expressed in terms of quantity, location and delivery (or pickup) time.
- The quality of the service offered, given by the width of the time windows for delivery (or pickup) time.

To the authors' knowledge, studies evaluating how time windows width directly affects the total cost of the service cannot be found, though a vast literature on optimization of the VRPTW is available. In fact, this is one of the variants of routing problems most studied and applied. In this context, indeed, the study of indicators is mainly focused to describe the computational effort and therefore to link it to the characteristics of the instances (Cordeau et al. 2002, Toth and Vigo 2002b).

Vehicle Routing has become a central problem in the fields of logistics and freight transportation. In some market sectors, transportation costs constitute a high percentage of the value added of goods. Therefore, the use of computerized methods for transportation can result in savings ranging from 5% to as much as 20% of the total costs (Toth and Vigo 2002b). VRP problems are present in the literature in many variants (see Toth and Vigo (2002), Baldacci et al. (2007)) for the main contributions in the area and Perboli et al. (2008) for a comparison of the main heuristic methods in the Capacitated VRP case.

In the following, we report a brief overview of the basic literature of the VRPTW, which is the context where the proposed indicators have been applied.

In this problem, time constraints are introduced to highlight the importance of a timely arrival of the freight, which is a common characteristic of applications such as express courier carriers, postal services, newspaper distribution, and e-commerce. A Time Window (TW) is defined as the interval of time within which a vehicle has to arrive to a node, and it is usually characterized by an early arrival time (*EAT*) and a late arrival time (*LAT*). Two types of time window constraints can be defined as follows:

- *Hard time windows*, which are constraints forcing each vehicle to reach customers in the interval defined by the TW. Some variants of the problem give a vehicle the possibility to have an idle time at destination in order to reach the lower time limit.
- *Soft time windows*, defined in the objective function, are represented by an increasing cost penalty in case the vehicle arrived at destination outside the time window interval.

A detailed survey of this class of problems has been proposed by Cordeau et al. (2002). Most of these techniques tend to solve these problems with a heuristic approach, as exact solutions are feasible only for small size problems. About these methods, Braysy and Gendreau (2005a, 2005b) show the main heuristic techniques in a detailed survey, they are usually tested on a group of instances (Solomon 1987), each one of them presents up to 100 freight requests. These requests are grouped into sets following the spatial distribution of the request destinations. The test instances have been extended, from 200 to 1000 customers, to study larger problems (Homberger and Gehring 2005) and the main outcomes on the last meta-heuristics show that the computation times have to rise considerably, before a good result can be obtained. For the same problem, Pisinger and Ropke (2007) propose a general heuristic for different VRP variants based on an adaptive large neighborhood search (ALNS) framework. This method is an extension of Shaw's large neighborhood search framework (1998) combined with a layer that adaptively chooses among a number of insertion and removal heuristics to intensify and diversify the search. This algorithm provides extremely high quality solutions, therefore its results have been used in the following for the comparative analysis.

3 Definition of the new service performance indicators

We present here a methodology to evaluate at an initial stage of planning operations, the effectiveness of a time window configuration for a given service quality in relation to a set of freight distribution requests located in specific nodes of the road network.

3.1 General definitions

Let us consider a service that involves a number of freight requests, within a given geographical area, using a fleet of vehicles travelling on the road network where, for each arc, travel time is assumed to be constant. Each request *R* is identified by a location in a node of the network, the quantity of freight to be delivered, and the time interval within which the freight has to be delivered, which is defined by an Early Arrival Time (EAT_R), and a Late Arrival Time (LAT_R). The fleet is homogeneous and consists of NV_{TOT} vehicles with an individual capacity equal to K. To comply with a request the service must deliver the freight and meet the time constraints. The various requests should be combined by the service provider in order to produce feasible routes for the available vehicles. The result of this planning activity or, in other words, how the requests are combined, depends on the requests level of compatibility (which is related to the demand configuration, time windows, road network and vehicle characteristics).

An idea to evaluate the requests compatibility has been proposed by Fischetti et al. (2001). The authors define a compatibility flag of a pair of requests (R_i and R_j) as a binary attribute; if a feasible circuit visiting the destination point of request R_i before serving request R_j exits, its value is equal to 1, else the flag is set to 0. By using this attribute we can determine whether request R_i can be served before request R_j with the same vehicle, either consecutively or not. However, we cannot use this to compare compatible cases, to establish priorities, or to determine how flexible this compatibility or incompatibility is. Therefore, we define this concept as follows.

Let us consider two requests R_A and R_B, defined by their own location, respectively A and B, the quantity of freight to be delivered, respectively d_A and d_B, and the time window, defined by EAT (respectively EAT_A and EAT_B) and LAT time (respectively LAT_A and LAT_B). The distance (expressed as a time measure) between A and B on the road network is noted as t_{AB}. It is also possible to model the service time at the request location, which can be written respectively as t_A and t_B. The pair of requests R_A - R_B is assumed as compatible if a vehicle serving both requests R_A and R_B can visit B after delivering or picking up the freight at A without slack period and without visiting other customers in between *A* and *B*, still complying with the time constraints defined by the time windows.

Let us suppose that we want to serve the requests in *A* and *B* consecutively and we want to calculate the earliest arrival time from *A* to comply with this condition. The vehicle will arrive at *A* at least at EAT_A and will not leave *A* before $EAT_A + t_A$. To ensure that request *B* is met, and according to the definition given for compatibility which guarantees the derivation of a simple and fast set of indicators, the vehicle cannot arrive at *B* before EAT_B. Therefore, the time between arrival at *A* and successive arrival at *B* is t_A+t_{AB}. The early arrival time at *A* of a vehicle serving *A* and *B* consecutively can be written in the following way (see Fig. 1):

$$EAT_{A/B} = \max \{EAT_A , EAT_B - (t_A + t_{AB})\}$$

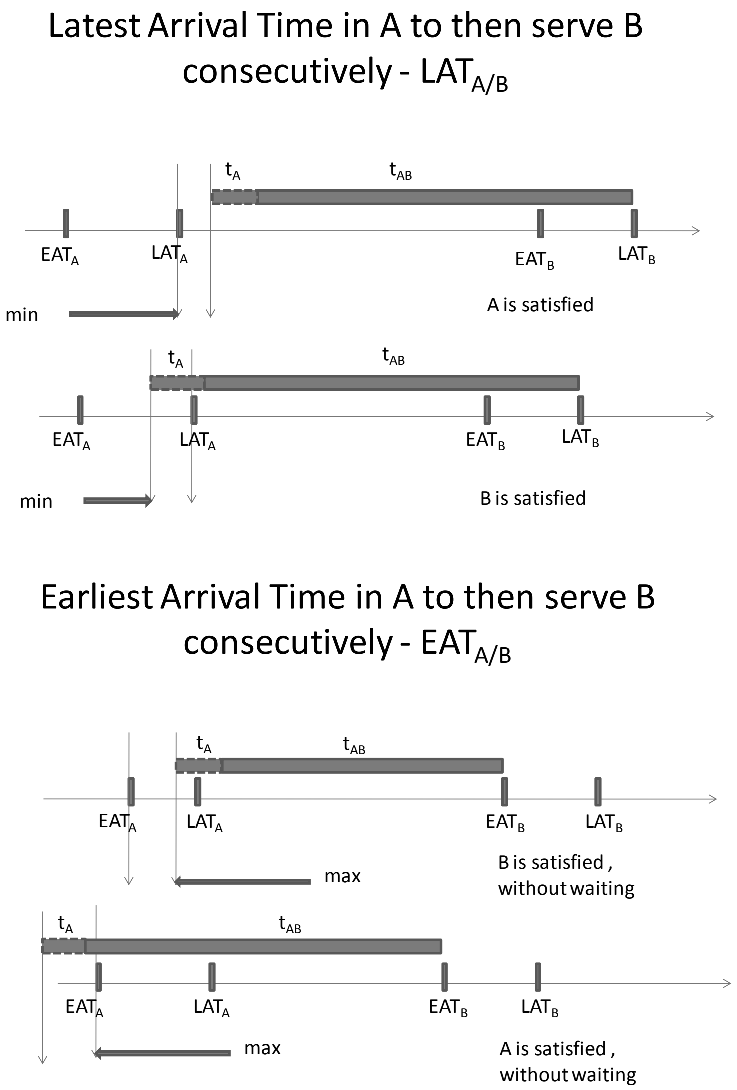


Fig. 1 - Earliest and latest departure time from *A* to satisfy consecutively requests *A* and *B*

In the same case, the vehicle cannot arrive at *A* after LAT_A and has to arrive at *B* before LAT_B, considering the travel time t_{AB} and the service time at *A*, t_A. The latest arrival time at *A,* which will enable a vehicle to serve consecutively the requests in *A* and *B* will be the following (see Fig. 1):

$$LAT_{A/B} = \min \{LAT_A, LAT_B - (t_A + t_{AB})\}$$

The *compatibility time interval* (*CTI*) of the pair of requests R_A - R_B is therefore defined as the interval between the earliest arrival time and the latest arrival time at *A* with a vehicle that needs to deliver to *A* and *B* consecutively, i.e. $LAT_{A/B}$ - $EAT_{A/B}$. We can write this in the following form:

$$CTI_{A/B} = \min\{LAT_A, LAT_B - (t_A + t_{AB})\} - \max\{EAT_A, EAT_B - (t_A + t_{AB})\}$$

This value defines the time interval including a possible actual arrival time at *A* if the subsequent request along the route is *B*. The term $CTI_{A/B}$ can be positive or negative. If $CTI_{A/B}$ is positive, then request R_A can precede request R_B directly. The higher the numeric value, the higher the overlapping time interval of the requests and the easier to serve them with the same vehicle. If $CTI_{A/B}$ is negative, request R_A cannot precede request R_B directly. However, this result can have two alternative meanings (Fig. 2):

- early arrival at *B*: R_A precedes R_B while the relative time windows are separated by a time interval bigger than the service time in A (t_A) plus the time to connect them (t_{AB});, i.e. $LAT_A < EAT_B - (t_A+t_{AB})$;
- late arrival at *B*: if B were served after A, its TW would not be met, since the arrival in B would be after LAT_B, so that it would be impossible to carry out the sequence A-B in the indicated order, i.e. $LAT_B - (t_A+t_{AB}) < EAT_A$.

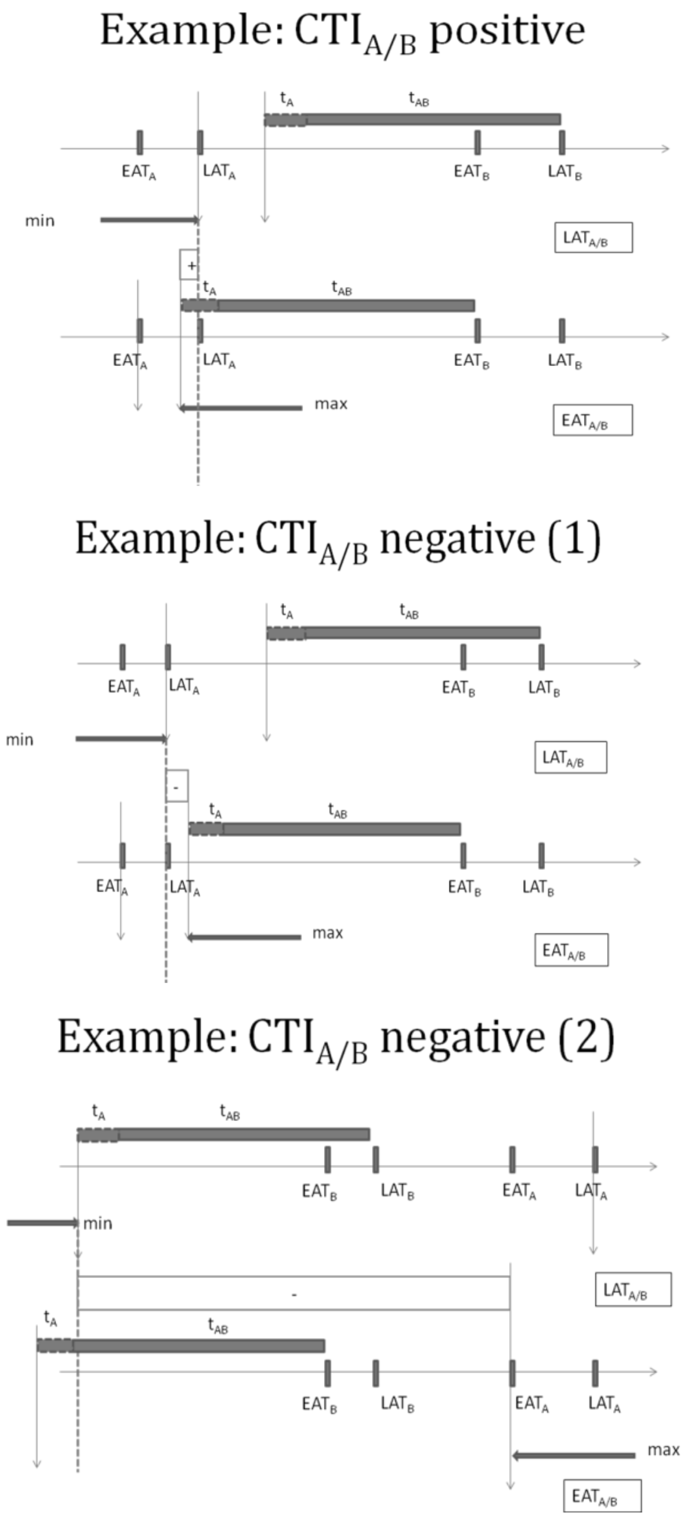


Fig. 2 - Examples of positive and negative pair compatibility time interval cases

In the first case it is still possible to deliver requests in A and B in the same vehicle trip, for example delivering to other customers in between or stopping a vehicle (slack pause) in order to arrive at B before LAT_B. This possibility is quantified by the value of $CTI_{A/B}$: if its absolute value is high, it will be more difficult to serve R_A and R_B using the same vehicle. In the second case it is not possible to serve R_A before R_B in the same vehicle trip.

The compatibility time interval for each pair of requests can be collected into a square matrix of dimension n_R (the total number of requests). This matrix is called Request Compatibility Matrix (*RCM*). To define this matrix, it has been decided to sort the requests by increasing Earliest Arrival Time, in order to rapidly separate most of the negative compatibilities with a late arrival at B. In this way, the negative compatibilities under the main diagonal identify the late arrival incompatibilities. In the upper diagonal, the negative elements may indicate the early or late arrival in B.

The *RCM* includes a first relation between the geographical and service characteristics of the network (links between nodes, maximum high speed allowed, trip times, etc.) and the nature of the demand (*RCM* is affected by node location of requests and time constraints).

3.2 Set of indicators

To give a synthetic evaluation of a freight distribution problem, we need to extract information from the data in the *RCM*. This matrix refers to each pair of requests, but does not give explicit information about the service compatibility (which are the potential requests that can be included in the same vehicle trip). From the *RCM*, we can obtain the number of positive $CTI_{A/B}$, which is represented as n_S^+, the number of early negative $CTI_{A/B,}$ noted n_S^{a-}, and the number of late negative $CTI_{A/B}$, which we will call n_S^{r-}. Therefore, we can define the following simple indicators expressing a priori the difficulty of a set of requests for a given network.

ACTI is the Average Compatibility Time Interval of requests, which represents the average value of all the positive pair compatibilities and it contains also *CTI* from and to the depot:

$$ACTI = \frac{\sum_{CTI_{A/B}>0} CTI_{A/B}}{n_S^+}$$

PPC is the percentage of positive compatibilities in *RCM*:

$$PPC = \frac{number\ of\ positive\ CTI_{A/B}}{number\ of\ elements\ in\ RCM} \quad (\%)$$

For each request R_A, the minimum travel time is calculated considering the travel time of t_{AB} each request R_B compatible with R_A. For those particular requests which are not compatible with any other ones, the minimum time is assumed equal to the travel time from/to the depot. This value is defined as the Minimum Travel Time between R_A and any compatible request R_B, and noted $MT_{comp}(A)$. Then, the average of $MT_{comp}(A)$, assumed as the mean value of all requests, is called Average of the Minimum Time Between each request *A* and any Compatible Request *B*, and noted *AMTBCR*:

$$AMTBCR = \frac{\sum_{A} \min_{CTI_{A/B}>0} \left(t_{AB}\right)}{n_R}$$

The first indicator (*ACTI)* quantifies the average compatibility time intervals between the requests, the second one (*PPC)* shows the proportion between positive and negative compatibilities and the third indicator (*AMTBCR*) gives an estimate of the time required to connect two requests in a plan. It should be noted that these compatibility-based indicators are calculated for pairs of requests and they give a rough idea of how the request configuration fits with the routing problem. However, these indicators cannot represent all the system components and the influence of the vehicle capacity in particular.

Therefore, in order to extract more suitable information for our problem from the *RCM*, we have defined two further indicators NV_I and T_I. These aggregate the information on pair compatibility time interval contained in *RCM*, for a sequence of requests. This makes it possible to extend the request compatibility time interval concept to the whole vehicle trip on the road network. In order to develop these indicators based on vehicle trips, we present a constructive heuristic which allows us to obtain a first simple and fast estimate of travel times and number of vehicles.

Using the *RCM* it is possible to subdivide the set of requests into a number of sub-sets that can be served in a feasible sequence. We should recall that the aim of this study is not to find an optimal solution for the distribution service, but to define a measure for the assessment of the compatibility of different requests, which depends on the demand and road network characteristics. We built a greedy algorithm in order to produce such feasible requests sub-sets. If each sub-set in our problem is viewed as a vehicle with a fixed capacity, then each sequence of requests represents a route for that vehicle.

In the *RCM* to each couple *A-B* we associate $CTI_{A/B}$. Let us consider a request R_A and a request R_B. Request R_B can be consecutively met after R_A if:

- the compatibility time interval $CTI_{A/B}$ is positive;
- the vehicle capacity constraints are complied with.

In the following we present a simple and efficient constructive heuristic that, according to our tests, is among the most efficient ways to solve the specific partitioning problems generated by RCM. The partitioning procedure works as follows: firstly, we initialize all the routes (i.e. we suppose that all vehicles are empty). Secondly, we take an empty vehicle starting from the depot, we evaluate all the requests that have not been already served, and add the request R_i with the lowest transportation time from the depot to the route. Then from this request we evaluate each request R_j with a compatibility time interval $CTI_{A/B} > 0$. The request that meets this condition with the lowest transportation time from i is added to the route.

We repeat these steps until there are no requests left with positive compatibility (with respect to the last request of the route) or if we cannot add a request without violating the capacity constraint. In these cases, the vehicle returns to the depot (to close the circuit and the sequence), we take another vehicle (creating another set), and repeat the process.

To generate a realistic configuration for the sub-set of requests, at each step, among all the possible options, we select the best partial solution (according to the minimum route travel time criterion). However, the result obtained in this way will not be assumed as the solution to the optimization problem, but as a rapid procedure to estimate the further two indicators, namely the number of vehicles needed and the total transportation time.

The algorithm, implemented with Microsoft Excel and Visual Basic, stops when all requests are assigned to a vehicle, that is, when each request belongs to a set. With this procedure also the completion time for each request is computed. Therefore, after obtaining each vehicle trip sequence, the corresponding transportation time is known.

The following definitions have, therefore, been adopted for the two indicators:

- T_i: the total transportation time as the sum of the time taken for each route;
- NV_i: the number of vehicles that are not empty, i.e. the number of sets created.

In this analysis, no attempt has been made to integrate the two indicators because the impact of these two factors on the total cost depends on the policy of the distribution service company. We will, therefore, try to analyze these two indicators separately, in order to propose an overall view of the problem, without referring to specific company objectives or preferences.

Finally, we remark that the capacity constraints do not affect the *ACTI*, *PPC* and *AMTBCR* indicators, while they are taken into account in the NV_i and T_i ones. For these last indicators, in this phase of the study, the fleet has been assumed to be composed by vehicles with the same capacity.

4 Computational results

In the organization of freight distribution services, we can observe two opposing factors. In order to increase profits, the transportation carriers wish to reduce costs, which means increasing service efficiency. However, this can have a negative effect on quality standards and may make it difficult to achieve the level of service expected by the user, who could change provider if the requested quality is not achieved. The level of service, in our study, is defined by the ability to comply with the time window and it is quantified by the width of the time window (the shorter the waiting period, the higher the quality of the service). The more complex and restrictive the time windows, the more difficult it becomes to maintain a good level of efficiency.

In the following, the results obtained show a strong correlation between the transportation cost level of an established distribution service with time windows, estimated by means of the proposed indicators, and the transportation cost estimated for the same service by a reliable optimization tool.

4.1 Experimental design

In order to evaluate the indicators, with regard to their capability in estimating the transportation cost level of a freight distribution service, reliable results obtained by a reference VRPTW optimization tool are needed to demonstrate their validity. We chose those of a recent study (Pisinger and Ropke 2007). This is a quite robust algorithm and it provides extremely high quality solutions. Moreover, it can be applied to many vehicle routing variants (the algorithm improved 183 best known solutions out of 486 benchmark tests, and that corresponds to 5 different vehicle routing optimization variants). The best results were obtained for the VRPTW, which improved the best solutions for 122 out of 300 large scale instances. For the configurations of the requests, they referred to Homberger and Gehring's VRPTW benchmark problems (2005), obtained by extending the well-known Solomon's benchmark problems (1985) for large networks from 200 to 1000 customers (Homberger and Gehring 2005).

In this phase of the study we perform a test on the proposed indicators to check their validity. To ensure a reliable test, we chose the results computed by ALNS, which show a high quality. In the application phase of the indicators, no tools other than the data of the requests are required to proceed with the estimation of the transportation cost level. Thus, ALNS results are only used in a simulation framework described in figure 3 to perform a comparative analysis of the quality of the computed indicators.

These test cases can be grouped into 2 different types of scenarios according to the time horizon used (short= type 1 and large= type 2). Each set of instances can be divided into three sub-sets (R, C, RC) based on the nature of the random generation of the requests at nodes of the network. In group R the customers are generated from a random uniform distribution, in problems belonging to group C, the customers are clustered, while group RC combines requests generated with both types of instances. Each sub-set includes 10 instances with various time windows, spatial and temporal distribution of customers.

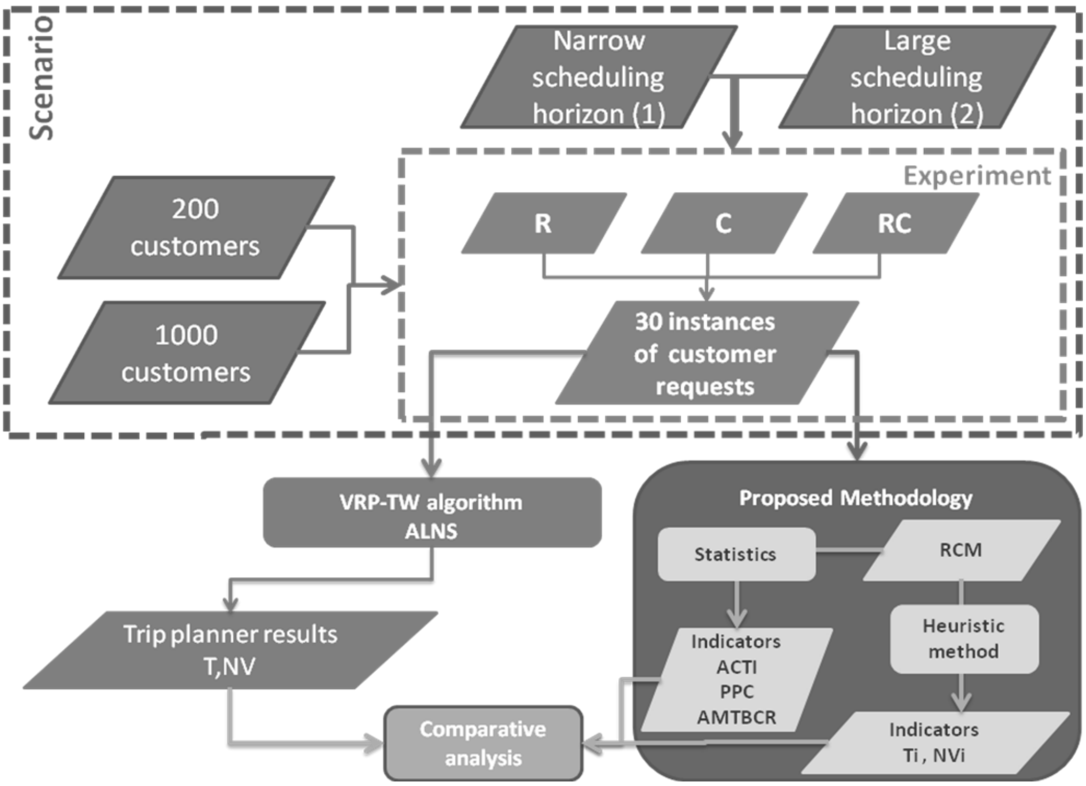


Fig. 3 - Chart representing the experimental procedure

Since one of the main requirements of indicators is to work on large networks, the analysis has been performed on cases of 1000 customers, but a deep preliminary test with the instances composed by 200 customers has also been done.

The chart in Fig. 3 represents the main blocks of the data and procedures used in the experimental process, where the block named “proposed methodology” contains data and procedures related to the computation of the proposed indicators. This block is connected to input data coming from instances of scenario and gives as output all the proposed indicators, computed therefore without any VRP algorithm. Results are finally compared (block “Comparative analysis”) with the trip planning results published for the same instances and obtained by means of the ALNS algorithm.

Travel times separating two customers correspond to their relative Euclidean distance; the service time at the location of the request has been obtained from the given instance data of the scenarios analyzed.

The fleet size, vehicle capacity and customer service times of the instances chosen for our experiments are reported in Tab. 1.

Tab. 1 – Main parameters of selected instances

Customers	Scheduling Horizon type	Density	Fleet Size	Vehicle Capacity	Service Time
200	1	R	50	200	10
200	1	C	50	200	90
200	1	RC	50	200	10
200	2	R	50	1000	10
200	2	C	50	700	90
200	2	RC	50	1000	10
1000	1	R	250	200	10
1000	1	C	250	200	90
1000	1	RC	250	200	10
1000	2	R	250	1000	10
1000	2	C	250	700	90
1000	2	RC	250	1000	10

The objective of this experimental phase is to determine which indicators better define the difficulty degree of an instance and its consequence on the service cost variations, in order to evaluate the demand characteristics related to the service supplied and the network configuration, without having to solve a VRPTW problem.

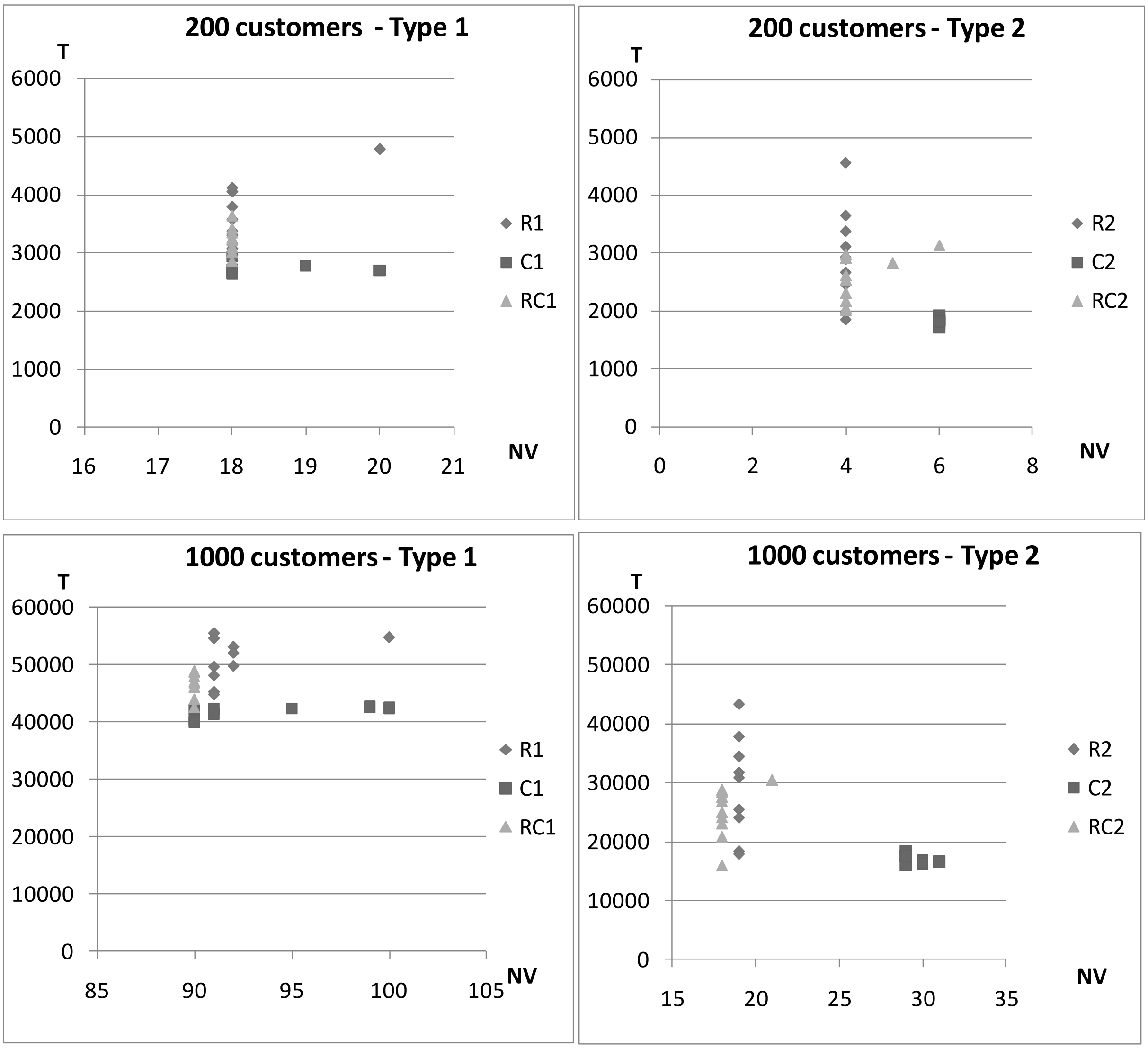


Fig. 4 – Trip planning results (ALNS) for the four scenarios selected and the R,C and RC cases

Then, to confirm by comparison the ability of our indicators to predict the level of difficulty in solving an instance, we used the ALNS results. This was done in order to have a reliable estimate of the service cost, which is generally described by the number of vehicles activated and their total travel time. The data displayed in Fig. 4 are reported to better clarify the use of ALSN results. It is well-known that in practical applications service costs mainly depend on the number of vehicles which must be used to satisfy the requests and their total travel time during the service operations.

These two service cost components are relevant, but often in conflict. On the other hand, in practice the type and number of vehicles required cannot be changed in a short time span. Thus, in our computational tests each instance class is characterized by a similar number of vehicles. More precisely, as shown in the charts in figure 4, in the 200 customer scenario, 18 (type 1) and 4 (type 2) are the most frequent number of vehicles used to satisfy requests. These figures contrast with the 1000-customer scenario, where they reach 90 for type 1 and 18-19 for type 2. Figure 4 also shows that in a very few cases the number of vehicles is greater than the most frequent value and this mainly occurs for the instances in group "C" where customers are clustered.

The comparison has been made by calculating the correlation coefficient ρ (also known as the Pearson coefficient), which shows if there is a linear correlation[1] between the two sets of data (each defined indicator and the total travel time T). As known, the correlation is 1 in the case of an increasing linear relationship, −1 in the case of a decreasing linear relationship, and these values have been assumed as target depending on meaning of the specific indicator (Tab. 2, Tab. 3).

4.2 Performance of the indicators

In this section we report the main results for the case of 200 customers, also used to select the best indicators, and for the case of 1000 customers, which confirms the capabilities of indicators to predict in most cases the variation of the transportation cost level, here estimated by means of the total travel time, for different service quality settings, here defined as the average width of the time windows for the

[1] Other correlation types can also be detected, but for our purpose of identifying a significant trend for such indicators, a linear correlation analysis is considered as adequate.

instance. As an example, a linear model has been calibrated and the trend of the transportation cost predicted based on one of the indicators has been compared with the total travel time estimated by the ALNS algorithm.

Finally, for any homogeneous set of instances also a regression analysis applied to disaggregated data has been applied and all of the diagrams are reported in Annexes A and B, while the values of the coefficient of determination R^2 are collected in two separated tables (Tab. 4 and Tab. 6).

4.2.1 Results for instances with 200 customers

The first selection of the best indicators was carried out on the 200 customer cases and, to comparatively measure the difficulty of an instance, the total travel time (*T*) has been used. It is useful to highlight that there is no certainty that if we have two instances whose solutions require the activation of a different number of vehicles, given an equal total travel time (T), they would actually generate the same transportation costs. Therefore, in order to use the values of T as a measure of the transportation costs, it is important for our purpose that the solutions show an equal, or similar, number of activated vehicles.

Therefore, to be able to work with homogeneous results, we selected high frequency cases with the same number of vehicles (*NV*): respectively, 24 cases with 18 vehicles (Tab. 2), for the narrow scheduling horizon scenarios (R1, C1, RC1) and 18 cases with 4 vehicles (Tab. 3) for R2, C2, RC2. To measure the quality offered by a transportation service on a given instance, we assume the average value of the time windows (*ATW*) for all the customers (excluding the depot), and to better evaluate if the proposed indicators help the evaluation of the service cost variations, we compare their correlation coefficients to those calculated with respect to *ATW*.

The two tables, which refer to the two types of scheduling scenario, then report for each row the computed values of the five indicators defined, the values of ATW and the results of the ALNS used as terms of comparison (NV, T) for any selected instance.

The tables show that the simple measure taken for the instance quality (*ATW*) is also well correlated with the service cost (expressed here as total travel time), particularly for the cases with narrow scheduling time. Therefore, we assume that the reproducibility of the cost variations by means of an indicator is effective only if its correlation coefficient is closer to the target value than to the *ATW* value (these values are indicated in bold in the tables). Observing the values of ρ for both tables we deduce that T_I and *AMTBCR* are the indicators that better describe the variations of travel time between instances for these cases. The tables also show that *PPC* and NV_I have a correlation coefficient higher than the one indicated for ATW, while *ACTI* has a lower coefficient still correlated, though, to the travel time.

Tab. 2 - Results for the narrow scheduling horizon scenarios (type 1)

ID	ACTI	PPC	AMTBCR	Nvi	Ti	ATW	NV	T
C1_2_10	270.8	78.5	2.6	21	3516.8	479.7	18	2644.3
C1_2_2	187.6	47.5	3.5	27	5595.3	347.3	18	2943.8
C1_2_3	389.6	74.5	3.0	26	4517.1	633.1	18	2710.2
C1_2_4	666.7	91.1	2.9	22	3935.5	919.1	18	2644.9
C1_2_9	192.3	61.7	2.7	20	3309.8	360.0	18	2687.8
R1_2_10	66.3	43.5	7.8	22	4519.0	124.5	18	3312.4
R1_2_2	81.9	42.1	8.8	49	8944.1	150.2	18	4059.6
R1_2_3	181.0	70.0	6.2	33	6241.8	290.0	18	3387.6
R1_2_4	314.8	88.7	5.7	21	3986.4	431.0	18	3086.1
R1_2_5	16.7	11.3	14.9	31	7445.4	30.0	18	4125.2
R1_2_6	90.1	46.6	7.6	26	6241.0	165.2	18	3586.8
R1_2_7	186.4	72.9	5.7	26	5158.1	300.0	18	3160.4
R1_2_8	316.8	90.2	5.3	19	3650.5	436.0	18	2971.7
R1_2_9	31.6	21.5	11.0	27	5930.7	60.1	18	3802.6
RC1_2_1	16.3	13.5	10.8	30	6134.5	30.0	18	3647.6
RC1_2_10	80.0	61.5	6.0	20	4002.8	150.0	18	3020.2
RC1_2_2	88.4	47.4	6.0	25	5120.6	165.1	18	3269.9
RC1_2_3	186.3	73.4	5.3	24	4169.4	300.7	18	3034.5
RC1_2_4	315.6	90.6	4.9	19	3409.1	436.1	18	2869.7
RC1_2_5	35.0	27.8	8.7	25	4952.8	64.7	18	3430.0
RC1_2_6	31.8	25.6	8.4	25	4995.9	60.0	18	3357.9
RC1_2_7	48.2	38.7	7.7	22	4365.1	91.4	18	3233.3
RC1_2_8	63.3	49.6	6.5	21	4161.5	119.2	18	3110.5
RC1_2_9	62.9	50.1	6.4	21	4143.5	120.0	18	3114.0
ρ	-0.680	**-0.755**	**0.919**	0.702	**0.869**	**-0.728**		
target	-1	**-1**	**1**	1	**1**	**-1**		

Tab. 3 – Results for the large scheduling horizon scenarios (type 2)

ID	ACTI	PPC	AMTBCR	Nvi	Ti	ATW	NV	T
R2_2_1	65.8	11.7	14.8	14	6410.0	121.2	4	4563.6
R2_2_10	258.0	41.9	7.9	5	2921.3	476.5	4	2666.1
R2_2_2	397.2	50.7	7.5	12	4876.8	708.7	4	3650.5
R2_2_3	847.5	78.2	5.9	9	3672.9	1296.3	4	2892.1
R2_2_4	1469.1	94.6	5.3	5	2272.6	1883.1	4	1981.3
R2_2_5	126.8	22.5	11.0	7	4552.4	240.0	4	3377.2
R2_2_6	453.5	57.0	7.0	6	3705.4	799.2	4	2929.7
R2_2_7	899.9	81.0	5.8	7	2906.1	1357.3	4	2456.7
R2_2_8	1507.8	95.4	5.3	5	2236.8	1914.1	4	1849.9
R2_2_9	199.0	33.3	8.9	8	3932.4	373.1	4	3113.7
RC2_2_10	381.8	62.5	6.1	5	2708.7	600.0	4	2015.6
RC2_2_3	836.6	79.2	5.5	8	2831.2	1296.3	4	2613.1
RC2_2_4	1466.4	95.0	4.8	5	2140.0	1884.6	4	2052.7
RC2_2_5	157.9	36.5	7.8	10	3472.8	279.1	4	2912.1
RC2_2_6	127.9	31.9	8.2	7	3436.7	240.0	4	2975.1
RC2_2_7	205.3	44.6	7.0	6	2962.1	369.8	4	2539.9
RC2_2_8	286.8	54.3	6.5	4	2808.6	485.4	4	2314.6
RC2_2_9	284.7	55.2	6.4	5	2745.3	480.0	4	2176.0
ρ	**-0.595**	**-0.750**	**0.871**	**0.869**	**0.979**	**-0.588**		
target	**-1**	**-1**	**1**	**1**	**1**	**-1**		

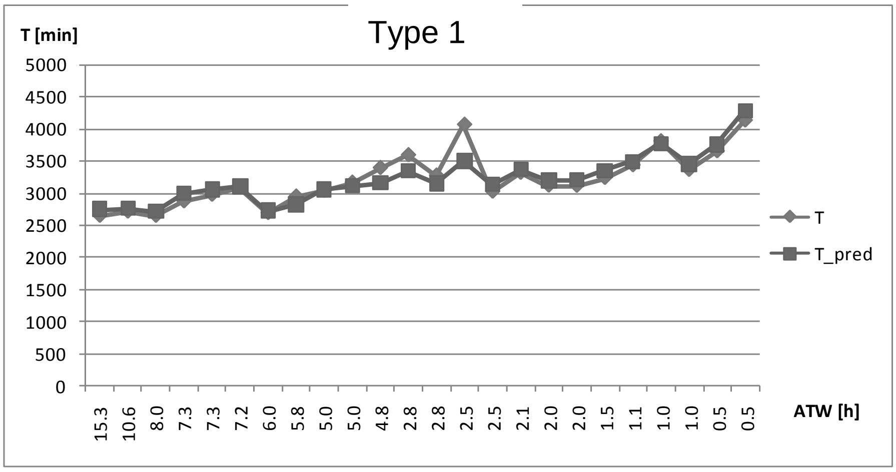
In order to allow a better understanding of the indicators performance, we report here below a more in-depth analysis, by means of a linear regression on the results of the simulation. For editorial matters, we limited this analysis to the AMTBCR indicator which proved to have a good mean behavior.

The linear regression model for the prediction of the total travel time (T_p) for this 200-customers problem based on the best simple indicator (*AMTBCR*) can, then, be calibrated for the two cases (narrow and large scheduling time) as follows:

$T_p = 2371.1 + 128.31 * AMTBCR$ *[type 1]*

$T_p = 930.42 + 245.64 * AMTBCR$ *[type 2]*

The results of the application of these models can be observed in diagrams in Fig. 5, where instances are sorted by decreasing ATW.



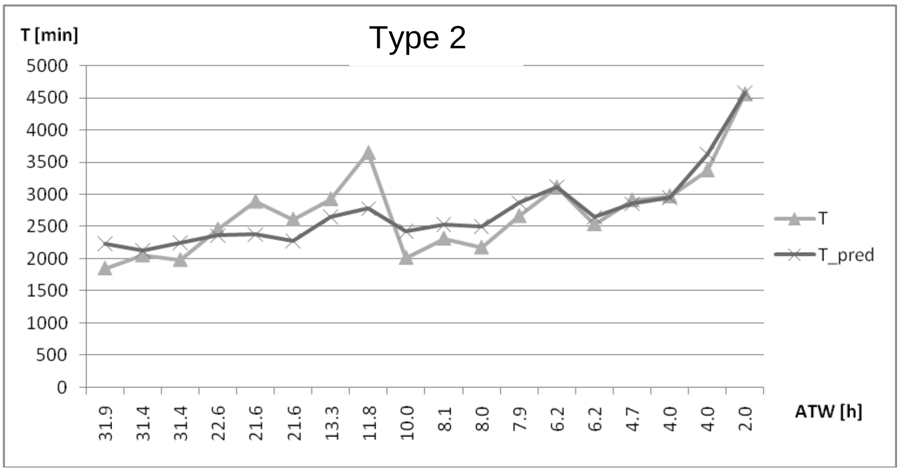


Fig. 5– Comparison between the total travel time by planning (ALNS) and the predicted values for the two types of scheduling horizon

The two diagrams also show that the minimum value of T is different for the two cases, and that occurs for the instances with the widest TW. For these cases we can expect that the solutions can possibly be estimated by a CVRP algorithm (applying only the capacity constraint and not the time windows one), because when the width of the time windows is so large, it has a limited effect on the trip planning solution.

The linear regression models are only made for these examples, with the aim of better showing, from a practical point of view, the punctual difference between the total travel time estimated by ALNS, and the one predicted through a linear model linked to one of the indicators. This analysis basically integrates the one resulting in tables 2 and 3, which show the main results that, for this reason, are computed for all the indicators.

After the selection of the best indicators, it is useful to remark that *AMTBCR* and *PPC* are only related to the demand side of the problem, as well as the time windows setting, while T_I also depends on the capacity of the vehicles used to perform the distribution service. Therefore, whether the vehicle capacity becomes a hard constraint for the transportation problem, the effectiveness of the first two indicators is reduced, since they only quantify the level of difficulty of the requests, without considering the fleet characteristics. In order to explore further properties of the selected indicators, we repeated this analysis for all the 30 instances of the two groups (3D – three distributions joined), since the number of vehicles activated for these results is similar (Fig. 4), also considering the disaggregation for the three types of customers densities (C, R and RC).

Tab. 4 - The coefficient of determination R^2 for the linear regression between total travel time and *indicators (* 200 customers)

Instance Type				
Scheduling Horizon	**Customer Density**	**PPC**	**AMTBCR**	**Ti**
1	C	0.0563	0.0021	0.1404
1	R	0.8013	0.8516	0.8798
1	RC	0.8951	0.8968	0.9264
1	3D	0.2383	0.835	0.6525
2	C	0.8828	0.5398	0.2475
2	R	0.7302	0.7609	0.9842
2	RC	0.5510	0.5942	0.7578
2	3D	0.1372	0.7563	0.5452

It can be observed (see Tab. 4 and for more details Fig. 8 and Fig. 9 depicted in ANNEX A) that a good linear regression can be found, even if the contribution of type-C instances decreases the correlation coefficient in type-1 case. Indeed the first line of the table 4 shows that for these instances the correlation between indicators and travel time is not relevant. This can be explained by observing that, in this case, the number of vehicles used to satisfy the requests is not constant (see figure 4) and therefore the travel time cannot be assumed to be the only factor affecting the transportation cost, as an important role is also played by the number of vehicles. For type-2 case, a different situation can be observed for the indicator PPC, because its trends are related to the type of density (R, C and RC) with a good general correlation, but here a global trend is less relevant. For both types, the slope of the lines is similar, where case RC is between C and R.

4.2.2 Results for instances with 1000 customers

For the 1000-customers problems the ALNS solutions are more heterogeneous with respect to the number of vehicles and the most frequent case (13 times) activates 90 vehicles (Tab. 5)

Tab. 5 – Number of vehicles for ALNS solutions (1000 customers)

NV	**type 1**	**type 2**
18		**9**
19		**10**
21		**1**
29		4
30		5
31		1
90	**13**	
91	**8**	
92	**3**	
95	1	
99	1	
100	4	
Total	30	30

Therefore, to include as many cases as possible, we decided to investigate the behavior of the indicators considering the instances with a similar number of vehicles *NV* (from 90 to 92 for type-1 and from 18 to 21 for type-2) still considering the total travel time when comparing the degree of difficulty.

Tab. 6 - The coefficient of determination R^2 for the linear regression between total travel time and *indicators (*1000 customers)

Instance Type				
Scheduling Horizon	**Customer Density**	**PPC**	**AMTBCR**	**Ti**
1	C	0.6941	0.2723	0.1161
1	R	0.5138	0.2285	0.8821
1	RC	0.9083	0.8154	0.6245
1	3D	0.3076	0.5055	0.7474
2	C	-	-	-
2	R	0.8680	0.7679	0.8668
2	RC	0.7780	0.6267	0.8213
2	3D	0.5842	0.7001	0.8502

We recall that Table 6 only refers to a set of selected instances, because the number of the activated vehicles is much more variable in the 1000-customer scenarios, while Table 4 reports the analysis made for the entire set of instances. Observing the case with 1000 customers summarized in Tab. 6 (for more details see also the diagrams in Fig. 9 and Fig. 11 in ANNEX B) the following points can be outlined:

- In the instances selected for the purpose of comparison, there were only 5 for type-1 and 0 for type-2 in distribution “C” . This is due to the fact that their NV is much more variable than in other instances and travel times are less significant as a service cost measure. It should also be noted that in type-2 instances the vehicle capacity is lower than that used in the remaining instances (see Table 1 for details).
- PPC confirms its ability to describe well the degree of difficulty, while separating the different densities (R, C and RC).

- AMTBCR describes quite well cases RC and it is useful to represent a density-independent behavior (case 3D).
- Ti better describes type R customer density and it confirms its best performance for type-2 horizon scheduling cases.

As for the 200-customers case, in order to show the errors of a possible linear model with regard to *AMTBCR* also for the 1000-cutomers case, we report the diagram in Fig. 6 where the selected instances are sorted by decreasing ATW.

$T_p = 42746 + 313.44 * AMTBCR$ *[type 1]*

$T_p = 16771 + 665.33 * AMTBCR$ *[type 2]*

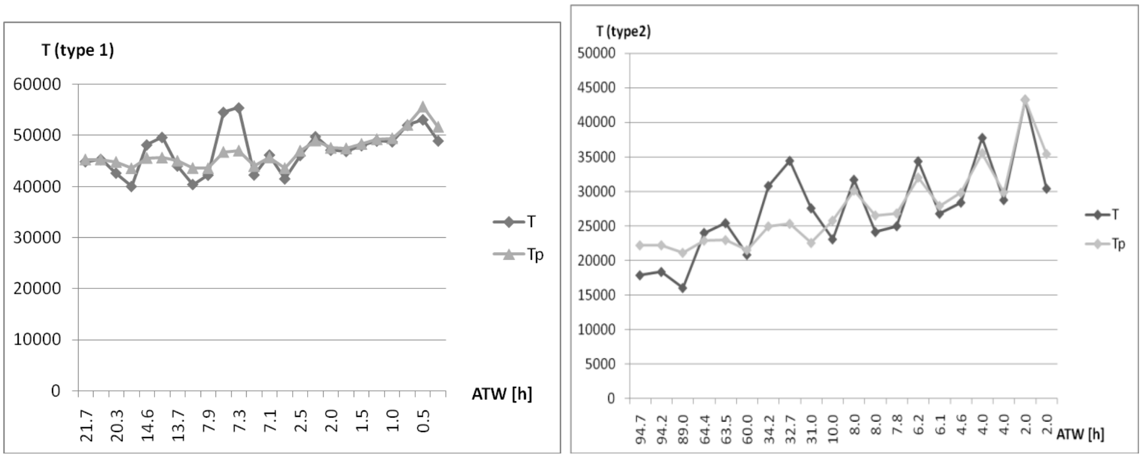


Fig. 6 - Comparison between the total travel time by planning (ALNS) and the predicted values for the 2 cases (1000 customers)

These diagrams show that only few cases do not fit with the model that, therefore, presents a generally acceptable behavior. Finally, we remark that instances do not have only time-window variations but they are also affected by other factors (service time, space and time density of the requests) and, in some cases, they also include a vehicle-capacity variation, since they are not specifically generated for this purpose. However, these instances allow a reliable test of the performance of the indicator because of the stability of their results.

4.3 *Realistic scenarios for the application of the indicators*

4.3.1 Iterative process of service quality setting

From a practical point of view, a possible use of the proposed indicators would be suitable in a context of a simulated variation of the service setting, also for a large number of operating alternatives. Indeed, when the number of scenarios to be analyzed increases, the compatibility with the high computation time of a reliable route planning tool fails, since an iterative process is often needed to support the decision on the time windows setting for the service quality (Fig. 7).

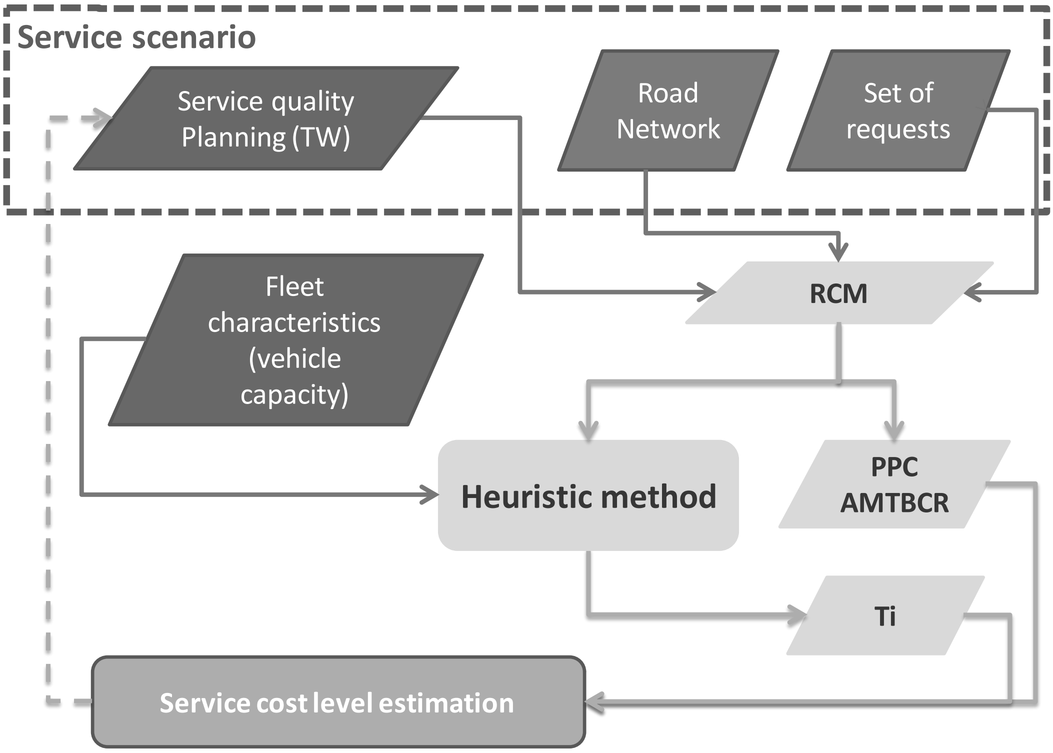


Fig. 7 - Chart representing the iterative process of indicator application for TW setting

Moreover, when the analysis has to be made in the planning process, the analysts are not used to adequately apply sophisticated tools, since these are generally used in the vehicle routing operational phase. On the other hand, it is not recommendable to use data coming from other known experiences to make choices related to the planning of a freight distribution service, since its transportation costs strongly rely on the specific road network configuration and on the nature of the requests (e.g. location and time). Therefore, the proposed indicators, which are quite able to correctly estimate the trend of total travel time of a freight distribution service with time windows, can be profitably used to predict this crucial transportation cost factor, in those cases where the fleet dimension can be assumed almost constant. The chart in Fig. 7 also shows that the T_i indicator should be properly used if vehicle capacity is an important constraint for the problem, while the other two indicators are only based on road network and request data.

4.3.2 Service cost level estimation

The cost trend estimation can be done after a simple calibration of a linear model that requires knowing only a few points, possibly derived from the available previous experiences in similar service conditions (e.g. road networks, fleet dimension). In fact, the first point should be easily estimated when a distribution service is already operating on a given area with an established TW, for the set of the requests and the consequent transportation cost of the service for the given quality level are already known. Another reference point of the line can be estimated with regard to the simple case without TW constraints, which is a baseline for a distribution service. These two points can be considered the necessary references to trace the linear model for a rough estimation of the transportation cost with regard to the performance indicator. This can be computed, as assumed before, if the set of the requests is known, on the base of the real or simulated data.

Moreover, other points can be estimated to improve the calibration of this simple prediction model if more real data scenarios are available for different TW settings or, when accessible, by using an adequate VRP optimization tool employed to solve a selection of relevant instances. Therefore, the indicators in this last context can either have the purpose of extending the transportation cost estimation (carried out by an optimization tool) to a wider range of TW widths, or they can aim to improve the resolution of the estimation between two TW settings, in the scale of the performance indicators.

4.3.3 Final remarks on the role of the indicators

We would like to remark that the role of the indicators is not to substitute the results obtainable with an optimization algorithm, but to easily estimate the level of difficulty of a given set of instances. Indeed, they should be considered as a measure of the service cost related to an instance set. Moreover, our indicators do not need any parameters for their calculation. Thus, AMTBCR and PPC might be considered as linked to the service demand and quality, while T_i is also related to the supply side requirements, such as the vehicle capacity. The analysis carried out has shown that in most cases such indicators are able to reproduce the cost trend, with regard to time windows variation for large network problems. Although the average value of requests TW (ATW) already contains a partial indication on the related transportation cost, since there is experimental evidence of their correlation (see for example Tables 2 and 3), the selected indicators perform better with regard to the estimation of cost levels. This can be explained by considering that the indicators, as has been shown in the experimental analysis, are linked to a greater number of problem data (not only TW). A clear demonstration of this is given for one of the indicators (AMTBCR) in Fig. 5, where the cost trend, which is not linear with regard to ATW, can be clearly represented by means of a linear model.

5 Conclusions

The main contribution of this research is the definition of a methodology to evaluate a-priori how the quality of a freight distribution service with time windows, operating on a given road network to meet a number of requests, affects the service cost. The notion of pair compatibility time interval between two requests has been defined and all the data have been collected into a Request Compatibility Matrix (*RCM)*. From this matrix, a first group of three statistical indicators was defined (*ACTI*, *PPC* and *AMTBCR*) by following simple statistical rules, and, in a second stage, a further group of two indicators (T_I and NV_I), following a planning criterion which takes into account the constraint of the vehicle capacity and, therefore, needing an appropriate computational procedure, was proposed.

The methodology presented has also been evaluated and the ability of the indicators to describe the level of difficulty in planning the requests has been illustrated in a set of experiments based on data used in literature for large network problems. A preliminary analysis allowed us to select the most suitable indicators (*PPC*, *AMTBCR* and T_I). Eventually, an analysis was performed on these indicators with regard to their ability to describe the difficulty to solve instances caused by time windows variations, by comparison with reliable and published results computed by means of a recent VRP algorithm (ALNS). A simple linear model has been built to better show the cost estimation capability with regard to average TW width for one of the indicators (AMTBCR).

The proposed sets of indicators are simple to understand and apply even by non-Operations Research experts and can be computed with a limited computational effort, even on large networks. These indicators can be used to give a first estimation of the transportation cost trends related to the quality of the service.

Finally, we note that in the present analysis time windows were not assumed to be uniform within the same experiment. Further experimental analysis could be carried out to explore requests with homogeneous TW, in order to better control the main factors that affect the quality of the service. In future research, it would also be useful to generate a wider range of instances based on realistic cases and use a specific planning tools to test the indicator accuracy, also in other interesting scenarios, particularly in a dynamic context. Indeed, for their fast computational ability, indicators are suitable to explore, for example, the transportation cost service trends in case of variations of road network travel times. Moreover, they are useful in scenarios where not all the requests are known in advance, and the service operator has to decide if a set of real time requests can be easily served with an acceptable cost increase. Finally, in this study a simple fleet configuration has been assumed, where all vehicles have the same capacity, but other realistic hypothesis can be explored to test the given indicators and, if necessary, to propose some new ones.

Acknowledgments

The authors would like to thank Simone Amico for his fundamental contribution to this work, which has been partially supported by the Ministero dell'Istruzione, Università e Ricerca (MIUR) (Italian Ministry of University and Research), under the Progetto di Ricerca di Interesse Nazionale (PRIN) 2007 "Optimization of Distribution Logistics”. They would also thank Bruno Dalla Chiara and Jesús Gonzalez Feliu for their valuable suggestions on the paper. The authors finally thank the two anonymous reviewers for their useful comments that have significantly improved the quality of the original manuscript.

References

Baldacci, R, Toth P, Vigo D (2007). Recent advances in vehicle routing exact algorithms. 4OR: A Quarterly Journal of Operations Research, 5- 4, pp. 269-298.

Braysy O, Gendreau M (2005) Vehicle Routing Problem with Time Windows, Part I: Construction and Local Search Algorithms. Transportation Science 39-1, pp.104-118.

Braysy O, Gendreau M (2005) Vehicle Routing Problem with Time Windows, Part II: Metaheuristics. Transportation Science 39-1, pp.119-139.

Cordeau JF, Desaulniers G, Desrosiers J, Solomon MM, Soumis F (2002) VRP with Time Windows. In P. Toth and D. Vigo (eds.): The Vehicle Routing Problem, SIAM Monographs on Discrete Mathematics and Applications, vol. 9, Philadelphia, PA, pp. 157-193.

Cordeau JF, Laporte G, Savelsberg M, Vigo D (2007) Vehicle Routing. In: C. Barnhart, G. Laporte (Eds.), *Transportation*, North Holland Elsevier, pp. 367-428.

Fischetti M, Lodi A, Martello S, Toth P (2001) A Polyhedral Approach to Simplified Crew Scheduling and Vehicle Scheduling Problems. Management Science 47 (6), pp. 833-839.

Homberger J, Gehring H, (2005) A two-phase hybrid metaheuristic for the vehicle routing problem with time windows, European Journal of Operational Research 162, pp. 220-238.

Laporte G (1992) The vehicle routing problem: an overview of exact and approximate algorithms, European Journal of Operational Research 59, pp. 345–358.

Marcucci E, Danielis R (2008) The potential demand for a urban freight consolidation centre, Transportation 35, pp. 269–284

Perboli G, Pezzella F, Tadei R (2008) Eve-opt: an hybrid algorithm for the capability vehicle routing problem. Mathematical Methods of Operations Research 68, pp. 361-382

Pisinger D, Ropke S (2007) A general heuristic for vehicle routing problems, Computers & Operations Research 34, pp. 2403-2435.

Puckett SM, Hensher DA (2008) The role of attribute processing strategies in estimating the preferences of road freight stakeholders. Transportation Research Part E 44, 379–395.

Solomon MM (1987) Algorithms for the Vehicle Routing and Scheduling Problem with Time Windows Constraints. Operations Research 35, pp. 254-265.

Gayialis SP, Tatsiopoulos IP (2004) Design of an IT-driven decision support system for vehicle routing and scheduling. European Journal of Operational Research 152, pp. 382–398.

Toth P, Vigo D (eds.) (2002)a The Vehicle Routing Problem, SIAM Monographs on Discrete Mathematics and Applications, vol. 9, Philadelphia, PA, USA.

Toth P, Vigo D (2002)b An overview of vehicle routing problems. In P. Toth and D. Vigo (eds.): *The Vehicle Routing Problem*, SIAM Monographs on Discrete Mathematics and Applications, vol. 9, Philadelphia, PA, pp. 1-26.

Wardman M (1998) The value of travel time: a review of British evidence. Journal of Transport Economics and Policy 32 (3), pp.285-316.

ANNEX A

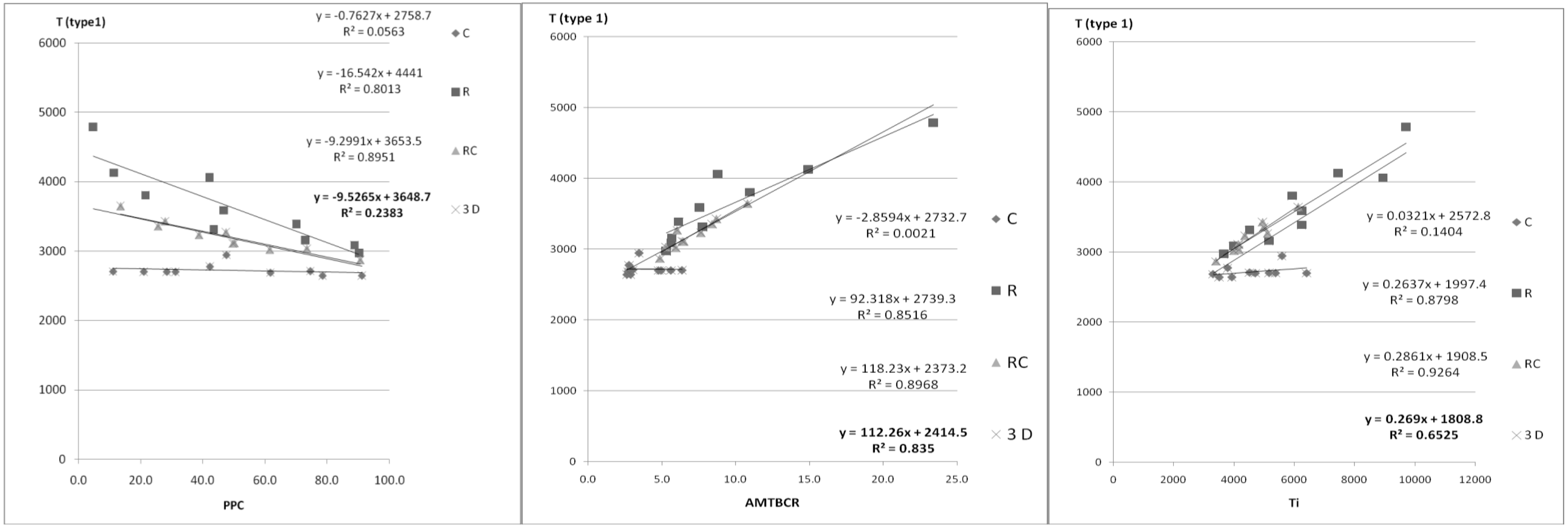


Fig. 8 – Linear regression for total travel time with *PPC*, *AMTBCR* and T_I for type 1 (200 customers)

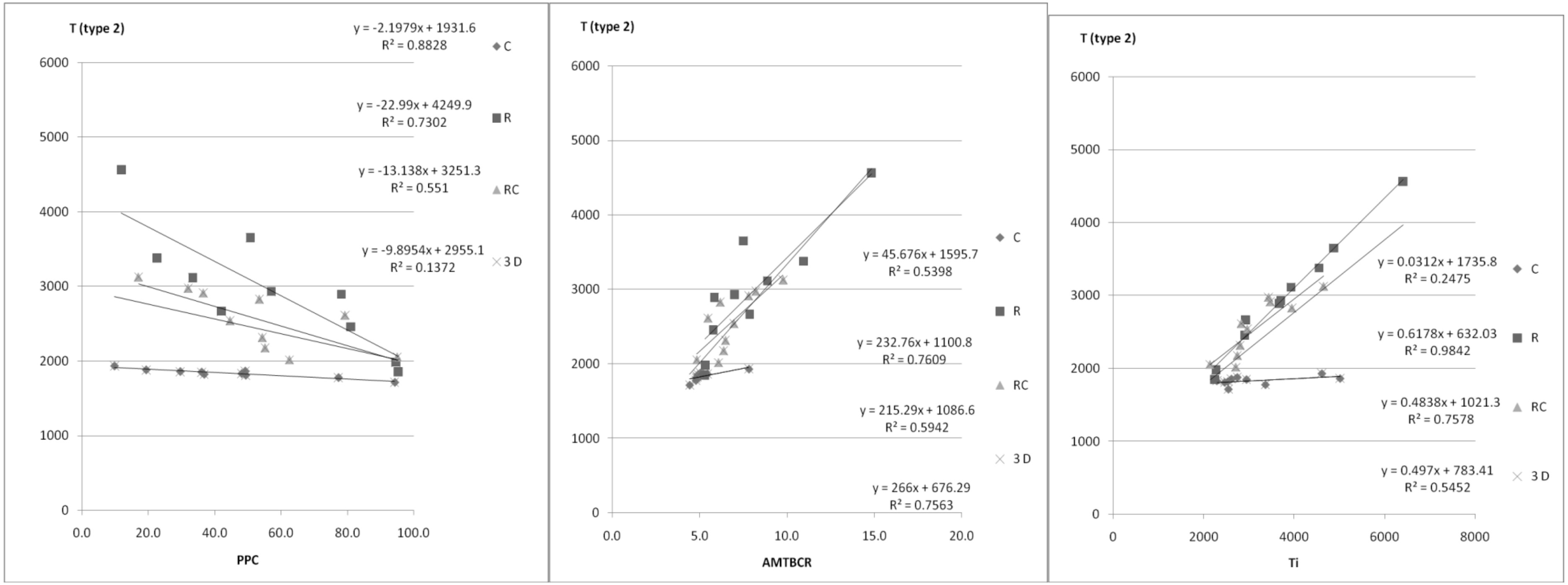


Fig. 9 – Linear regression for total travel time with *PPC*, AMTBCR and T_I for type 2 (200 customers)

ANNEX B

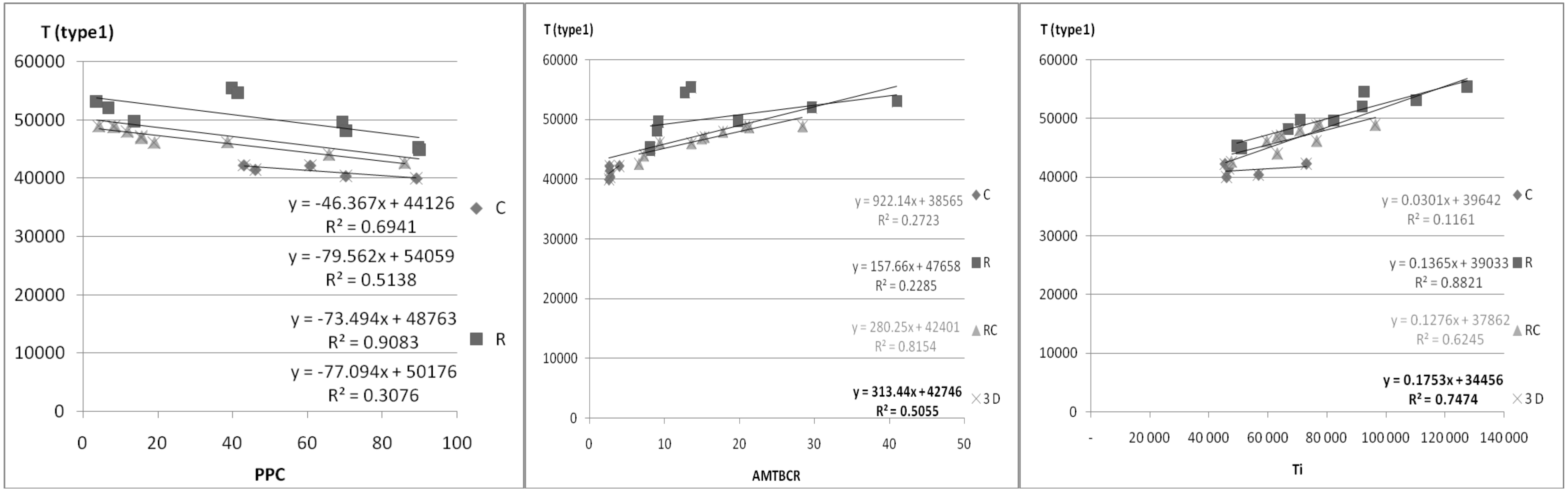


Fig. 10 – Linear regression for total travel time with *PPC*, *AMTBCR* and T_I for type 1 (1000 customers)

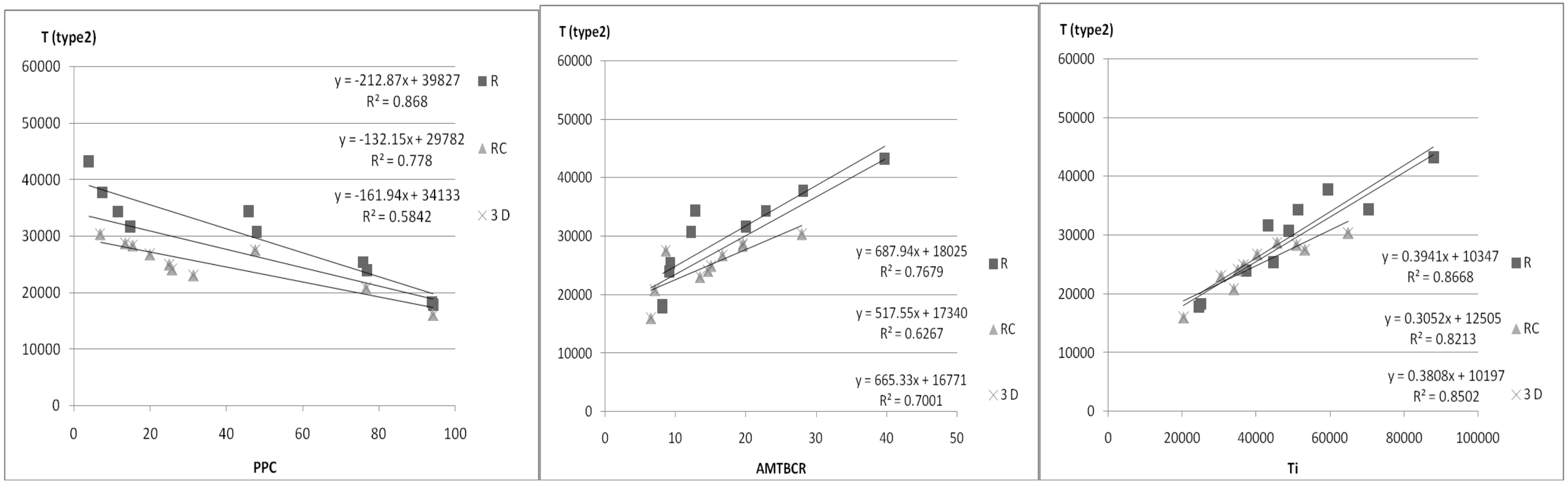


Fig. 11 – Linear regression for total travel time with *PPC*, *AMTBCR* and T_I for type 2 (1000 customers)