ABSTRACT
Transport modelling is a critical step in the design and economic assessment of transport infrastructures. Decisions on large investments must be based on the results of a model that has to cope with the complexities involved in transportation activities under long planning horizons. This paper discusses the methodological issues faced in the development of a model for the economic assessment of the Bolivian Central Bi-oceanic Railway Corridor (CFBC). This large infrastructure will connect the ports in Peru with the ones in Brazil by railway through the Andes. The project faces technical challenges from the harsh geographical and climatic conditions along with the uncertainty on whether the economic benefits will offset the large construction costs. Our goal is to develop a combined freight and passenger transport model that provides traffic absorption forecasts for a long time span of 50 years. The model has to cope with different levels of resolution for regional and international flows and its development is constrained by limited information sources.

Keywords: Modelling and simulation, transportation model, mode choice, multimodality, Railway Corridor

1. INTRODUCTION
The CFBC (Central Bi-Oceanic Railway Corridor) is a railway corridor infrastructure project promoted by the Bolivian Government and funded by the Inter-American Development Bank. It will link the Atlantic and Pacific coasts of the central part of South America. Its construction is expected to save both costs and time of transporting cargo and passengers through the great natural barrier which are the Andes. As of nowadays the only way to cross them is by narrow mountain roads that impede large trucks flow.

The study presented in this paper is part of a wider project for the analysis of prospective trade, market and logistic alternatives. It seeks to determine the levels of passenger and freight demand and the CFBC’s flows absorption among Bolivia, its neighbouring countries and the rest of the world. The goal is to analyse its competitiveness compared to that of other existing alternatives such as the Paraguay-Parana Waterway or the Panama Channel. Import and export trades are regarded as one of the most promising sources of demand and thus accurate international transport modelling is considered a major challenge in the project development.

This study has three main goals. The first one is a prospective analysis of freight and passenger transportation, in different geographic (worldwide, national, regional) and temporal frameworks (long and short term). The second goal is a market study, establishing demand levels, transport service characteristics, competition, commercial strategic etc. The last objective is the study of logistic alternatives in order to obtain multimodal options (where the CFBC is one of the stretches) that generate synergies and increases the competitiveness of the Bolivian Integrated Transport System.

The paper discusses the methodological issues and challenges faced in the development of the project. As this project is still at an early stage of development at the time of this publication, the paper focuses on the discussion of methodological aspects rather than in the analysis of results.

The different approaches used in each step of the model are also explained, remarking the differences with other models employed for transport planning, like the resolution of the model at different levels and the use of optimal order quantity for transforming tonnes of cargo into number of trips.

2. METHODOLOGY OVERVIEW
The models employed for transport planning applications can be divided in those concerning either passengers or freight. The case of Passenger Transport Modelling has been widely studied, generally using the Classical Model of the Four Stages (Ortúzar and Willumsen 2011). In this method, the geographical area under consideration is divided into Traffic Analysis Zones (TAZ), which are the smallest regions in which passenger flows are aggregated. This methodology adopts a successive approach that consists of four main steps:

1. Trip Generation. The trips generated in each TAZ are estimated.
2. Trip Distribution. This step connects each of the trips generated in the previous stage with its destination TAZ. The result is a matrix travel between each pair of origin and destination TAZs (commonly called Origin-Destination OD Matrix).

3. Modal Split. It gives the transport mode that a trip uses (obviously, in the case that more than one transport mode is available for this trip).

4. Traffic Assignment. This step gives the links of the network used for a trip.

This model can and has been adapted to the case of goods carry. However, several challenges are faced for a successful adaptation, mostly related to the difficulty of modelling policy makers’ preferences. Thus, despite of the research effort carried out in the last decades, the freight transport modelling methods are less developed than those applied in passengers modelling (Ortúzar and Willumsen 2011). Freight transport decisions are business management decisions made upon complex criteria. They can be affected by several factors such as those spanning the cost of the transported goods, the transport reliability, the frequency of shipments and the transport time.

In this project the goal is to jointly estimate freight and passenger flows. Thus a multilayer model has been designed in which passengers will be treated as one of the freight classes with specific characteristics to account for passengers’ behaviour. Passengers flow is not expected to be a high fraction of the total demand so estimation inaccuracies are less critical than in freight estimation.

The main decision problem that will determine the competitiveness of the CFBC in relation to other transportation alternatives is a modal choice problem. The key feature of the model must be the ability to determine the allocation of freight flows between CFBC and all transportation alternatives considered.

This modal choice is determined by the interaction between the two main actors who make decisions: cargo shippers and carriers (Samuelson 1977, Holguín-Veras and Willumsen 2011). The shippers determine what freight is transported, to which destinations, in what quantities and how often. The carriers determine the characteristics of the transport system, the level of services and set freight rates. In a competitive market, the interaction between them will determine the service conditions and the transported flows. Shippers have influence on carriers’ decisions about transport settings, because they define the characteristics of shipped goods, and in turn carriers have influence on the shipment sizes and modal choices made by the shippers through their pricing policies or established service levels. This interaction can be modelled as a cooperative game in which previous research results (Jose Holguín-Veras et al. 2009) show that it can be assumed with a high degree of confidence that the decision finally adopted corresponds to the one that minimizes the total shipping costs, including those of both carriers and shippers.

The application of the general four steps analysis framework to this project requires an adaptation to the specific objectives of the study. The extended methodology employed is based on the work of Holguín-Veras (2002), and it is outlined in the Table 1.

OD matrices can be obtained directly from a single initial step in which either by surveying or historical data analysis representative origin-destination matrices can be obtained. In the case where the source destination pairs are clearly defined (as befits a significant percentage of the goods of this study) based on surveying methods, they provide specific information on the types of units transported and thus avoiding more aggregated and less accurate approaches.

<table>
<thead>
<tr>
<th>Table 1: Methodological Schema</th>
</tr>
</thead>
<tbody>
<tr>
<td>Steps of the model</td>
</tr>
<tr>
<td>Freight Generation</td>
</tr>
<tr>
<td>Freight Distribution → OD Matrices</td>
</tr>
<tr>
<td>Trip Generation and Mode Choice</td>
</tr>
<tr>
<td>Network Assignment</td>
</tr>
</tbody>
</table>

2.1. M&S approach as analysis methodology

One of the main sources of uncertainty in any transportation planning model, and given the long planning horizons required in practice, is given by the temporal projection of the variables that feed the model. Growth forecasts of economic activity and commodity flows are subject to great uncertainty that can seriously compromise the reliability of the model results (Flyvbjerg et al. 2006).

The design of logistics and transport services requires collaboration between multiple disciplines such as data mining, statistical analysis, forecasting methods, modelling and simulation and optimization techniques (M. Bielli et al. 2011). These technologies are widely used in various fields of logistics and terminal design (Longo 2010) or transport networks.

The methodology chosen to conduct the evaluation adopts a modelling and simulation approach in order to achieve flexibility and adaptation in model building and experimentation scenarios in terms of regional units, and infrastructural decisional considered.

Main advantages of this kind of methodology are that it is not necessary to build a system to evaluate it and its results are accurate, in general, compared to analytical models. Also it is important to notice that it allows “What if” analysis to be performed so predicting the behaviour of the system under different conditions. This last reason is the most important one, because it is one of the requirements of the Bolivian Government as a result of the project. Simulation is also a tool widely used in the design state (Robinson 2004), as it is a new infrastructure.
Sokolowski and Banks (2010) list a group of advantages, most of them supporting the choice of M&S methodology for this case. As they say, simulation can help explore new policies, operating procedures, etc. in a case like this, in which a new infrastructure is in the planning phase and no real system is yet available for experimentation. Another of these advantages is that new transportation systems can be tested before performing the large investments involved, as can be the use of different train types in Bolivian new infrastructure. This methodology also helps identify the important variables of the system which have more impact on its performance as well as their interactions. Still, the most important advantage is the possibility of assess different scenarios under a “What if” analysis.

Accordingly, two key features of the tool are:

- The scenario generation, which will not only get the desired results of the project to the defined scenarios from market survey but also future updates of the model in the light of developments CFBC demand or its alternatives. This will provide greater flexibility in managing the infrastructure to facilitate decision when there are changes in environmental conditions.

- The sensitivity analysis gives on one hand an element to assess the robustness of the model against the uncertainty in the input data and on the other hand to assess the economic risks associated with the investment. The analysis techniques of the uncertainty in the input data are methodologies developed in recent years and provide a convenient way to validate models in conditions where it is difficult or totally unfeasible have data to conduct an experimental validation of the model (Ankenman and Nelson 2012).

### 2.2. Market Survey

The nature of this study is essentially a market survey which aims to assess the competitiveness of the CFBC to attract freight traffic. Its development by surveying techniques is a fundamental tool for getting information about current commodity flows and criteria affecting transport decisions. Information collected will be used in the calibration of analytical models and the validity of the results depends on this data.

In such studies, usually little information and heterogeneous data sources are available. Getting a set of data of the highest quality possible ensure that the model is valid and has adequate accuracy for the purposes of the study. The difficulties that arise in practice when collecting data sets suggest the adoption of robust models with a level of complexity corresponding to the information availability.

On the other hand, a remarkable set of goods which are potential users of CFBC are iron, soybeans and forest products. This particular set of goods will require a detailed analysis of their origins and destinations, operating restrictions and costs. In this sense, the market survey will provide detailed information about them on which to base the analysis.

The proposed methodology, which is based on principles of the field of logistics, will further evaluate the impact of these particular market segments in the CFBC. Other planning methods with a higher aggregation level would be undesirable for these goods as they would reduce the estimation accuracy of the main sources of demand CFBC.

### 3. MODEL ELEMENTS

#### 3.1. Origin Destination Matrices

The origin-destination matrices are the input model. They collect various goods flows between origins and destination for each group

These source and destination areas, is what are called transportation planning TAZ (Traffic Analysis Zones). These zones are geographic areas able to attract and generate goods or passenger trips. When the goods are clearly identified an element, such as iron ore, these points of origin and destination are easily identifiable. In cases where the production and consumption centres present lower geographic concentration has to pay special attention to the choice of these TAZ, especially with regard to geographical coverage. Excessively large sizes can reduce the accuracy in the calculation of distances, being in close proximity (e.g. two countries), and sizes too small make the characteristics of the area become very heterogeneous, and that increase the complexity of the model and computational cost. Data should be available at the aggregation level sought; therefore choosing the TAZ is also related to the availability of data for the real case.

The accuracy of the model depends on the quality of data of the matrices. It is often necessary to use approximations that transform data into operational data available, i.e. if we aggregated to higher levels of TAZ, these approaches should allow us to disaggregate at the desired level.

#### 3.2. Network

The transportation network contains all information of the infrastructure to be used to perform various transports. It is defined by a set of routes between sources and destinations. In practical applications it is necessary to use some kind of mapping on which to work, such as geographic information systems (GIS) containing all the information and not just geographically but other data such as the characterization of the TAZ population or some economic data from them. It is important to have the tools to add elements to these information systems, because when it comes to evaluating a new infrastructure, they are not usually available

#### 3.3. Mode Choice

The model indicates the choice for a particular user that selects a particular freight under certain conditions. This
is an essential step because this is where the competitiveness of each of the options available is assessed.

This choice model is divided into two distinct but interrelated turns. The first is to obtain the economic order quantities (EOQ) and secondly the model of choice of the individual. This is necessary because the relationship between demand and traffic is not direct (José Holguín-Veras et al. 2011). The need for the definition of an optimal amount of cargo, to minimize the total costs, is either in that an agent may prefer multiple small shipments for short times needs instead of others for which time do not affect, or in accounting for the reduction in costs obtained by taking advantage of economies of scale.

The total cost of logistic operations is obtained from the transport costs of the option chosen and inventory costs associated with consolidating all cargo that will form a shipment. The optimal order quantity depends on the total transport cost and therefore depends on the mode of transport. Thus the modal selection is necessarily affected by the amount of ships and vice versa, requiring the joint consideration of both aspects (Jose Holguín-Veras et al. 2009).

There are multiple mathematical developments that can approximate the shippers mode choice (Ortúzar and Willumsen 2011). Logit models (in any of its various forms) are the most widely used in applied problems in the field of logistics. Logit models are the most widely used Multinomial Logit Models (MNL) and Nested Logit Models (NLM).

- They all start from the same theoretical framework (Ortúzar and Willumsen 2011):
- Decision makers have perfect information to make choices and act rationally.
- There are a set of alternatives (transportation) and there is a set of attributes that quantify the utility of each alternative for the decision maker.
- Each alternative has a net profit for each decision maker, which is how attractive is that option for him.
- The decision maker behaves rationally looking to maximize the utility of your choice.

The variables that are commonly used in utility functions are cost and time; the findings of most studies in this regard (Kreutzberger 2008) indicate that these are the variables on which most decision makers focus their choice.

The time must take into account the total duration of the transport, thus not only the time when the vehicle (for each transport) is moving but respect those rules derivatives (such as resting time) of the waiting times at stations and times intermodal operations of loading and unloading of cargo.

In terms of cost it is important to fully define the costs the options incur because it determines the establishment of pricing policies and the evaluation of the profitability of the infrastructure. There is an important relationship between flows absorbed by a line (Behrens and Picard 2011) and rates as high flows may allow the exploitation of economies of scale.

3.4. Assignment

In the step of assigning traffic the total flow through each section of the network is obtained. There are different methods of allocation depending on the needs of the model. So when the effect of network congestion can be skipped or no material can be used one of the simplest methods is assigning all or nothing (All or Nothing Assignment, AON). In this case, it is assumed that the travel time through each link is not dependent on its charge. Then all traffic flows between pairs of origin and destination can be assigned by the method of shortest path in terms of time, duration, cost or generalized cost function.

With regard to road transport, it is assumed that the congestion effects are fundamentally influenced by its interaction with the transport of passengers. However, for the estimated travel time to analyse the large-scale flow (national or international) the impact of congestion is relatively lower than at smaller scales as is the urban one. Then, simple estimation based on distances and travel times means provides sufficient accuracy.

4. CONCEPTUAL MODEL

4.1. Levels of resolution

The level of detail required in modelling freight worldwide is below than for the regional traffic; besides their choice alternatives are different. We propose the adoption of a method of analysis in two stages. First step is modelling the distribution between ports (entries and exits) of the freight flows. These freight flows are the ones between the countries inside the hinterland of the CFBC and the rest of the world.

On the second step we will model the freight traffic from the zones inside the hinterland of the CFBC and the entry/exits ports.

4.1.1. Zoning international TAZ

We propose the grouping of the different model TAZ in a total of 9 groups that allows pre-filtering of the decision alternatives affecting each source-destination pair to be done.

The TAZ of the hinterland of CFBC define the hinterland of CFBC as the set of regions whose intraregional traffic may be susceptible to recruitment by the CFBC. These TAZ are:

- Western Area: North of Chile, Peru and Bolivia high Andean plateau.
- Central Area: Eastern Bolivian Andes and central regions of Bolivia.
- Eastern Area: western Bolivia, Brazilian states bordering Paraguay and northern Argentina.
The TAZ outside hinterland of the CFBC considered are:

- Asian countries.
- Western strip of North America, Central America and northern South American countries.
- Eastern strip of North America, Central America and northern South American countries.
- Eastern strip of America.
- European and African countries.
- Central and southern Chile.
- Central and southern Argentina and Uruguay.

In Figure 1 a diagram with the above geographical areas is shown. This zoning permits the initial filtering of decision alternatives that will be present between each source-destination pair. For example for the traffic between the hinterland and Western Europe it would not make sense to consider Pacific ports as an alternative outlet.

4.1.2. International Traffic

In this step the Origin and Destination pairs of the matrices are all the countries that have an important commercial relationship with the hinterland of the CFBC and the TAZ of the hinterland. Inland traﬃcs are not considered in both areas.

In this case there are two transportation alternatives, one that uses Atlantic ports as entry/exit ports, and the one that uses the Paciﬁc ports.

The elements of the model are:

- $F_{t,m,i,s,j}$: Flow OD matrix in physical units between TAZ of the hinterland of the CFBC.
- $d_{i,j,k}$: Distance Matrix for each OD pair and each transportation modes.
- $t_{i,j,k}$: Time Matrix for each OD pair and each transportation modes.
- $q_{t,m,i,s,j}$: Economic Order Quantity (EOQ) Matrices for each OD pair and each transportation modes.
- $CT_{t,m,i,s,j}$: Total Cost Matrix for each OD pair.

Figure 2 helps explain the performance of the model. It has two main parts; the first one is related to the international flows and the second one to regional flows. The following notation is employed:

- $t$: Time period (year or trimester).
- $m$: Type freight Index according to Standard Classification of Transported Goods (SCTG).
- $i$: TAZ Origin.
- $j$: TAZ Destination.
- $k$: Each transportation mode.
- $l$: Each type of vehicle, for all transportation modes.

4.1.3. Regional Traffic (CFBC hinterland)

In this step the Origin and Destination pairs of the matrices are between TAZs in the hinterland and the entry/exit ports.

The transportation alternatives are:

- Economic Order Quantity models (for which regression analysis will most likely adopted).
- Long distance tariff model. Represents the shipper cost for all alternatives, in the case of international traffic. It takes into account the haulage costs to CFBC terminals or alternatives passages, Fare of the CFBC or alternatives passages, Port fares, tariffs and customs fees
- Mode Choice Model (1): the entry and exit ports are here chosen
The elements of the model are:

- $F_{t,m,i,j}^{i}$: Flow OD matrix in physical units between TAZ of the hinterland of the CFBC.
- $F_{p,t,m,i,j}^{m}$: OD matrices between the TAZ in the hinterland entry/exit ports
- $q_{t,m,i,j,k}^{m}$: EQ Matrices for OD pair and regional traffic.
- $t_{i,j,k}$: Time Matrix for each OD pair and each transportation modes.
- $CTR_{t,m,i,j,k}$: Total Cost Matrix for each OD of the hinterland and the ports.
- $FT_{t,m,i,j,k}^{m}$: Freight flow matrix for each OD pair and transportation alternative.
- $FT_{t,m,i,j,k,v}^{m}$: Freight flow matrix for each OD pair, each transportation alternative and each vehicle type.
- Network with total freight flows in trips and physical units for each type of vehicle.

The models used in this step are:

- Regional Tariff Model, similar as the international traffic one but in the regional level.
- Mode Choice Model (2): for choosing one of the transportation modes considered.
- Vehicle Choice Model, for each transportation mode, allows the type of vehicle for a shipment to be chosen
- Assignment, this step assigns the freight and passengers flows to the links of the network.

5. MODEL IMPLEMENTATION

Once the conceptual model is defined, a test model is implemented in TransCAD. To do so, a set of macros in GISDK language was developed.

Origin-destination (OD) matrices are the input for the model and the analysis methodology chosen is the modelling and simulation approach, so it is necessary to obtain future forecasts of their values. The first macro implemented uses a gravity model in order to obtain the total flows distribution based on the forecasted production and consumption values of each TAZ. An OD matrix for each commodity and year of the analysis is estimated. After that, a new macro allows the user to choose the set of OD matrices to be employed for the analysis and also the geographical layer that contains the alternative to assess.

Once the scenario has been established in terms of flows and geographical alternative, it is necessary to select the level of resolution. Another macro makes it possible to restrict the analysis to International or Regional chosen flows. Depending on the resolution level, the following macro builds the network. It gives the outline of all the transport modes considered in the analysis, taking into account all the TAZs of the OD matrices. With this last macro the complete scenario of analysis is set, so the main steps of the model - the mode choice and the network assignment - can proceed.

Mode Choice depends on the utility function of each transport mode considered, so there is a macro that introduces these utility functions of the transportation modes considered. As it has been mentioned, the main variables of the utility functions are time and cost, so there are a set of macros that build the matrices of time and cost that feed the mode choice model. These macros take into account the waiting time in the intermodal change nodes and also the cost of these changes. The utility functions and time and cost matrices feed the Mode Choice model, which gives the probability of each transportation mode for each origin-destination pair.

Using these probabilities and the OD matrices, the OD matrices for each transportation mode are built. Once the OD matrices for all the transportation modes are built the next macro provides the assignment to the network. In this case an All or Nothing method is used, where the travel time is the variable to minimize.

![Figure 3: Test Model Implementation.](image-url)
within the planning horizon. It gives a net assignment for every period of time and each freight type or passengers flow. So a new macro calculates the total amount of flows for every period considering all the different freight or passenger flows that are occurring. A last macro represents in the graphical layer the flows of the period chosen. Figure 3 represents and summarizes in a general way the operation of these macros.

6. CONCLUSIONS
The development of an entire and ready to use transport model for the assessment of different transport alternatives within actual constraints is a complex work. This is not only because a minimum set of good quality data is very difficult to obtain, but also due to the intrinsic methodological difficulties to solve.

A simulation based model is adopted since the “what if” analysis naturally fits and expands the options for the assessment of infrastructures and transportation services. It helps reduce uncertainty about what might happen under some economic and services conditions, and their variations, and then make decisions based on it. These decisions could be about the design of a new infrastructure or about the transportation services (fees, frequencies, etc.).

The methodology presented is an adaptation of some different existing models for analysis in two levels. One of them, more aggregated, at international level, where the main entry/exit ports of the country are established. The other level, more disaggregated, is a regional level where the internal traffic flows and the traffic flows to the main ports are obtained. This innovative approach will provide the means for the economic and service assessment of different multimodal alternatives.

REFERENCES

AUTHORS BIOGRAPHY
Rosa Rios Prado, MSc Industrial Engineer. Rosa is a research engineer in the Integrated Group for Engineering Research (GII) of the University of A Coruna (UDC) in Spain. Rosa is mainly devoted to the development of transportation and logistic models for the assessment of multimodal networks and infrastructures.

Diego Crespo Pereira, PhD, MSc Industrial Engineer. Diego is Assistant Professor of the Department of Economic Analysis and Company Management of the UDC. He also works in the GII of the UDC as a research engineer. He is mainly involved in the development and coordination of R&D projects related to industrial and logistic processes optimization.

David del Rio Vilas, MSc Industrial Engineer. He is Adjunct Professor of the Department of Economic Analysis and Company Management. He is mainly involved in R&D projects related to industrial and logistic processes optimization.

Nadia Rego Monteil, MSc Industrial Engineer. Nadia works as a research engineer at the GII of the UDC. Her areas of major interest are in the fields of Ergonomics, Process Optimization and Production Planning.

Oscar de Gregorio Vicente, MSc Telecommunications Engineer and Master in Transportation and Logistics. Oscar is currently Project Manager in the Transport Department of the Spanish company Consultrans S.A.