

## Hotel Location when Competitors May React:

### A Game-Theoretic Gravitational Model

Pedro Godinho<sup>a,\*</sup>, Paul Phillips<sup>b</sup>, Luiz Moutinho<sup>c</sup>

<sup>a</sup> CeBER and Faculty of Economics, University of Coimbra, Av. Dias da Silva, 165, 3004-512 Coimbra, Portugal; e-mail address: [pgodinho@fe.uc.pt](mailto:pgodinho@fe.uc.pt).

<sup>b</sup> Kent Business School, University of Kent, Canterbury, UK; e-mail address: [P.A.Phillips@kent.ac.uk](mailto:P.A.Phillips@kent.ac.uk).

<sup>c</sup> University of Suffolk, UK and University of South Pacific, Fiji; e-mail address: [luizammoutinho@gmail.com](mailto:luizammoutinho@gmail.com).

This paper presents a hotel location model that incorporates concepts from both game theory and gravitational site location models. We consider a hotel chain intending to build new hotels in a given region. Customers travel to the region to visit some specific points, termed “attractions”, and they choose a hotel according to room price, location and hotel attractiveness. Competitor hotels react to the new hotels by changing prices, in order to maximize their own profits, so the final set of prices will be a Nash equilibrium. We propose an iterative procedure for finding the equilibrium prices and a genetic algorithm-based procedure for finding the optimal strategy, in terms of new hotels to be built and respective typologies. Using a mini case, we illustrate and analyse the influence of several parameters. Then, we present computational experiments, concluding that the proposed procedures are effective in finding good solutions for the model.

**Keywords:** Tourism site location; Game theory; Genetic algorithms; Spatial interaction models

\* Corresponding author

**Acknowledgements:** The authors would like to thank António Mexia Santos, Miguel Emauz and Ivana Kolar for some very useful insights on the hotel business. The authors would also like to thank two anonymous reviewers for their useful comments and suggestions.

© 2018. This manuscript version is made available under the CC-BY-NC-ND 4.0 license <http://creativecommons.org/licenses/by-nc-nd/4.0/>.

This is a pre-copyedit version of an article published in the Tourism Management journal. The final version of the article is available online at: <https://doi.org/10.1016/j.tourman.2018.06.014>

## **Highlights**

- A game theoretic/spatial interaction model is proposed for hotel site location
- Customers choose according to price and non-price attributes
- The specific sites that customers intend to visit are taken into account
- Game theory is used to determine the equilibrium prices, considering the competitors' reactions
- A genetic algorithm is applied to determine the optimal hotel location and typology decisions

## 1. INTRODUCTION

Location remains a potential source for competitive advantage for the accommodation sector (Adam and Amuquandoh, 2014; Yang, Luo and Law, 2014) with location decision-making gaining increasing attention from academic and business community in the past two decades (Chou, Hsu and Chen, 2008). Moreover, the financial crisis of 2008 has led to a subsequent global economic downturn marking the beginning of the “new normal”, characterised by fundamental changes in the appetite for risk taking (Phillips and Moutinho, 2014). Hotels remain a key element of the tourism industry, so new approaches to enhance strategic decision-making of hotel investors will benefit the growth and development of the tourism industry.

Geographic location is important to a diverse range of retailers. For example, Wal-Mart operations in rural markets generate on average higher returns than its operations in more competitive urban markets (Ghemawat, 1986). Hotel location decisions may be quite problematic, especially in regions in which the market is mature and a significant supply already exists. In such cases, hotel chains must assume that competitors already in place will react to new hotels, which in the past led to room rate reduction in order to avoid losing customers. Given the scale and level of investments and maturity of the global hotel sector, opportunities exist for a more diverse range of methodological, philosophical and theoretical approaches. Williams and Baláž, (2015) contend that there is a need for stronger theoretical understanding of the different concepts of tourism risks.

This study considers a game-theoretic approach to address the hotel location problem. In this paper we consider that hotels will set prices in order to maximize their profits. We also assume that demand takes into account not only the hotel price but also other attributes of the hotel.

A paucity of prior game theory research in the context of hotel location (see Yang, Huang, Song and Liang, 2008) provides a stimulant for this hotel location study. We aim to define an operational model that can be used both to find the optimal decisions in realistic and complex situations, and to analyse the outcome of such situations, particularly the impact of specific parameters on the outcome. To achieve this, we moved away from simplified models that would lead to closed-form

mathematical solutions and define a more complex model that can be used in regions with different characteristics.

This study has some links to the operational component of the work of Arenoe, van der Rest and Kattuman (2015), by extending it in several areas. First, similarly to Arenoe et al. (2015), we assume that customers may choose not to be lodged in any of the alternatives explicitly included in the model. But, while Arenoe et al. consider a utility threshold, we assume that there is an alternative that consists of not choosing any of the modelled hotels and which has an attractiveness to the customer. Second, we explicitly model the places that the hotel guests wish to visit, allowing different customer segments to have different visitation patterns. Other authors (e.g., Hung, Shang and Wang, 2010) use proxies like the city centre to identify the places that attract visitors, but those will be just rough approximations in polycentric cities or when attractions are very far apart in a city. Another difference from the work of Arenoe et al. (2015) is the goal of the model. Determining the equilibrium prices is the final goal of Arenoe et al., but it is only the intermediate goal for this study. We assume that the hotel chain we are considering intends to build some new hotels, keeping the investment expenditure within a predefined budget. We want to find the optimal strategy (sites and hotel typology) for opening new hotels, in order to maximize the total profit of the chain. After the new hotel or hotels are built, new competition makes all hotels rethink their prices, and the set of prices becomes a game-theoretic Nash equilibrium.

## **2. RELATED WORKS**

Yang et al. (2014) classify prior hotel location research into theoretical, empirical and operational models. Their analysis covers a diverse mix of academic disciplines from hospitality and tourism, geography, economics, marketing, finance and urban planning. The authors delve into the literature and identify four theoretical categories, six empirical categories, and three operational categories. The authors also recognise that some of the models do not fit these categories, since there are some diverse models concerning hotel location.

Most models in the theoretical category and several models in the empirical category try to explain the spatial location or room pricing choices of the hotels. Since we take the perspective of defining the best locations based on the ability to attract demand, it is of interest to us to analyse models aiming to explain how customers make their choices. Masiero, Heo and Pan (2015) notice that there are limited studies focusing

on the relationship between hotel attributes and room pricing from a customer perspective. The authors propose a new discrete choice model for determining the customer's willingness to pay based on a set of room attributes. Lee, Kim, Kim and Lee (2010) evaluate the importance of different factors in the satisfaction of frequent individual traveller / foreign independent traveller guests of five-star hotels in Korea. The authors consider six factors – tourism attraction, convenience, safety, surrounding environment, traffic and accessibility – and they consider several attributes for each factor. They conclude that tourism attraction is the most important factor in explaining the satisfaction of hotel guests.

Such empirical studies might provide a solid foundation for operational models, but in fact there are very few operational models based on detailed quantitative definition of customer behaviour. For example, the operational categories defined in Yang et al. (2014) – checklists, statistical prediction models and Geographic Information Systems-based models – do not consider a direct model of customer behaviour.

Among the authors considering a model of customer behaviour as the foundation for operational decisions, we can find Moutinho and Paton (1991). Moutinho and Paton propose a spatial interaction model for tourism site selection and analysis, the LOCAT model, based on the probability that tourists will patronise a given site location. It attempts to measure the total attractiveness of a particular site location taking into account the impact of the degree of accessibility, total catchment population and level of product uniqueness.

Arenoe et al. (2015) consider a model of customer behaviour based on conjoint analysis. The authors assume that buyers respond both to price and non-price differences, so the price charged by a hotel manager must take into account the prices charged by other hotels, as well as the characteristics of the different hotels. Any realistic operational model must consider that competitors will not remain indifferent to decisions that may affect them – in the case of Arenoe et al., they will be affected by the price decisions made by other hotels, so they will react to them. To incorporate these reactions, Arenoe et al. define a game-theoretic model of hotel pricing.

From an economics' perspective competitiveness within a sector can be viewed through various lenses. Since the pioneering work of Von Neumann and Morgenstern (1945), which provided greater insights into game theory and economics, economists and mathematicians use the approach to assess decision-making in uncertain situations.

Game theory remains an active area of research in economics and is particularly useful for studying interactions among large numbers of participants (Cheung, 2014). Rubinstein (1990) posits that game theory is a key tool for the construction of the modern theory of industrial organisations. Behaviour of firms can be modelled in either continuous strategy (e.g., Awaya and Krishna, 2016; Laraki, Solan and Vieille, 2005) or discrete strategy sets (e.g. Ciliberto and Tamer, 2009; Godinho and Dias, 2013; Seim, 2006). A behaviour model of the firm will incorporate informational and computational assumptions and relate to more complex phenomenon (Prietula and Watson, 2008). Knowing how rational players behave in a strategic context has appeal to business and management scholars. Niou and Ordeshook (2015) stress the importance of game theoretic decision-making, which provides insight for the managers as one assumes that other decision-makers are not fixed targets, and that they take into account their knowledge of the manager, and that the manager knows that they know. More recently, the ability to identify and predict behaviours to capture value in intense competitive markets is an emerging theme in the value capture theory (Gans and Ryall, 2017). Research that enables scholars to rethink fundamental ideas in competitive environments will provide academic as well as practical benefits which can flow through to the bottom line. Kuechle (2014) does highlight that entrepreneurial activity varies across regions and the phenomenon persists over time, and this impacts choice of hotel location.

Game theory can amplify the interaction of competition from a behavioural modelling perspective (Clarke-Hill, Li and Davies, 2003). Moreover, the field of game theory provides a lens through which hotel location decision-making can be analysed. Game theory-based models have been used in the context of hotel competition, but mostly within very simplified models of price or quantity competition (e.g., Baum and Mudambi, 1995; Chung, 2000; Gu, 1997; Guo, Ling, Dong and Lian, 2013; Song, Yang and Huang, 2009; Yang et al., 2008). The goal of such studies is often to determine the equilibrium prices, and revenue maximization strategies, which are primary objectives of hoteliers. This is in part due to the cost structure of hotels. If fixed costs are high and variable or operating costs are low, revenue maximization may be a sensible objective (Friesz, Mookherjee and Rigdon, 2005). However, profit maximization is a more general objective since it does not rely on such an assumption and is more attractive to financially motivated stakeholders.

A factor that many studies consider to determine customer patronage is the closeness to the destinations that customers intend to visit. For example, Kimes and Fitzsimmons (1990) state that eighty percent of the customers of La Quinta Motor Inns visited destinations within four miles of the inn. In this paper we consider the special importance of this factor, which we model explicitly. In order to do it, we incorporate a spatial interaction perspective in the model.

There are few spatial interaction models applied to tourism. Hurley, Moutinho and Witt (1998) propose a spatial interaction model for tourism site location. The model follows the site location approach proposed by Penny and Broom (1988), considering two attributes: distance and a subjective measure of attraction. The authors show that Genetic Algorithms (GAs) perform quite well in obtaining solutions for the model. Godinho, Silva and Moutinho (2015) also propose a hotel location model based on Penny and Broom (1988) and Hurley et al. (1998), including a cost structure and a budget constraint. In both cases, the spatial component is based on the origin of the visitors, and not on the closeness of the hotel to the intended destinations.

### **3. A MODEL FOR TOURISM SITE LOCATION UNDER COMPETITION**

In this paper we consider that hotels will define prices in order to maximize their profits. We assume that demand takes into account not only the hotel price but also other attributes of the hotel. Like Arenoe et al. (2015), we assume that buyers respond both to price and non-price differences, and so the price charged by a hotel manager must take into account the prices charged by other hotels together with the characteristics of the different hotels. For modelling customer choice, Arenoe et al. use a utility-based multinomial logit model, while we choose a spatial interaction model, in the line of Moutinho and Patton (1991), Hurley et al. (1998) and Godinho et al. (2015). However, our spatial interaction model incorporates an attractiveness parameter, which encompasses all the relevant attributes found on the previous empirical studies (number of stars, hotel amenities, safety, surrounding environment, etc.)

Our work was developed after contacts with hotel managers, who helped us identifying additional relevant factors that should be incorporated in the model. Among these factors, which we included in the model, are capacity constraints, segmented demand (heterogeneous customers, with different preferences and different average length of stay), annual periods of different demand (and different prices) and the possibility of different customer segments wanting to visit different venues within the

considered region, which we term “attractions”. This is done by considering that, depending on the hotel location, a customer will have a generalized cost for visiting the attractions to which he/she wants to go. This expands the use of spatial interaction models to the case in which the proximity to specific attractions within a region is taken into account. According to the hotel managers we contacted, a single hotel or chain is not able to influence the aggregate demand in a region that is already served by a reasonable number of hotels, but the distance from the most important attractions within a region may significantly influence the hotel demand. The parameters of the model, and respective notation, are summarized in Table 1. Table 2 shows the decision variables and calculated values, and the respective notation.

We take the perspective of a profit-maximizing hotel chain that faces competition that already operates some hotels and is considering opening new hotels. We consider that decisions concern a touristic region, which may be a city, a seaside region, etc. In this region there are already  $N_e$  hotels belonging to the hotel chain and  $N_c$  hotels run by competitors. Each hotel has a specific attractiveness level  $W_j$ , which may be the result of multiple factors, like the hotel rating (number of stars), surrounding environment or hotel amenities like the existence of a swimming pool. Each hotel has a given structure of costs and revenues, which includes the room price, additional revenue for each occupied room (e.g., average meal sales), variable cost per occupied room, fixed costs (independent of the occupation level) and capacity.

To simplify the expressions, we define a “variable cost net of additional revenue” that incorporates the variable cost and the additional revenue per occupied room:  $c_{jt} = vc_{jt} - ar_{jt}$ . The annual profit of the hotel is then defined as:

$$P_j = \sum_{t=1}^T \left[ (p_{jt} - c_{jt}) \cdot n_{jt} - F_{jt} \right] \cdot d_t \quad (1)$$

It is assumed that the general touristic demand of the region will not be affected by the decisions made by the company. The region possesses attractions that are the main reason leading visitors to come to the region. In each period, customers visit the region attractions according to a pattern that may change for different customer segments. This pattern of visits is used to define a perceived generalized cost of the local trips for visitors lodged in a given site,  $g_{jkt}$ , calculated as:

$$g_{jkt} = \sum_{a=1}^A \alpha_{ja} \cdot v_{kta} \quad (2)$$



Table 1 – Model parameters and respective notation

Existing hotels:	
• Hotels from the considered chain	$N_e$ hotels (identified by $j = 1, \dots, N_e$ )
• Competitor hotels	$N_c$ hotels (identified by $j = N_e + 1, \dots, N_e + N_c$ )
Sites for potential new hotels	
	$N_p$ sites (identified by $j = N_e + N_c + 1, \dots, N_e + N_c + N_p$ )
Different periods per year with different demand characteristics (high season, low season, etc.)	
• Number of days in period $t$	$d_t$
Hotel characteristics (hotel $j$ ):	
• Additional revenue for each booked room in period $t$	$ar_{jt}$
• Variable cost, for the hotel, of having a room occupied for a day in period $t$	$vc_{jt}$
• Daily fixed costs of the hotel in period $t$	$F_{jt}$
• Hotel capacity (per day) in period $t$	$C_{jt}$
• Hotel attractiveness	$W_j$
Attractions (sites visited by the customers)	
	$A$ attractions (identified by $a = 1, \dots, A$ )
• Perceived generalized cost of a trip from hotel site $j$ to the attraction site $a$	$\alpha_{ja}$
Demand segments	
	$K$ types of customers (identified by $k = 1, \dots, K$ )
• Daily demand (number of rooms in all establishments) by customers of segment $k$ in period $t$	$R_{kt}$
• Cost-sensitivity of customer segment $k$	$\beta_k$
• Joint total attraction exerted by other establishments (not explicitly modelled) over visitors of type $k$	$MO_{kt}$
• Average number of daily visits of customers of segment $k$ to attraction $a$ in period $t$	$v_{kta}$
Set of typologies for new hotels (number of hotel stars, whether to build a swimming pool, etc.)	
	$q = 1, \dots, Q$
Construction cost for building a hotel of typology $q$ at site $j$ ( $j \in \{N_e + N_c + 1, \dots, N_e + N_c + N_p\}$ )	
	$CC_j(q)$
Budget for building new hotels	
	$B$

Table 2 – Decision variables and calculated values, and respective notation

Decision variables	
Daily room price for hotel $j$ (or for a new hotel located at site $j$ ) in period $t$	$p_{jt}$
<ul style="list-style-type: none"> <li>Vector of room prices for the different periods</li> </ul>	$\mathbf{p}_j = [p_{j1} \ \dots \ p_{jT}]$
Typology chosen for a new hotel located at site $j$ , $j \in \{N_e + N_c + 1, \dots, N_e + N_c + N_p\}$ (the value zero is used when no hotel is built)	$q(j)$
Calculated values	
For hotel $j$ (or for a new hotel located at site $j$ ):	
<ul style="list-style-type: none"> <li>Variable cost net of additional revenue</li> </ul>	$c_{jt} = vc_{jt} - ar_{jt}$
<ul style="list-style-type: none"> <li>Average number of rooms sold daily in period <math>t</math></li> </ul>	$n_{jt}$
<ul style="list-style-type: none"> <li>Annual profit of the hotel</li> </ul>	$P_j$
<ul style="list-style-type: none"> <li>Average daily demand, if there were no capacity constraints</li> </ul>	$D'_{jt}$
<ul style="list-style-type: none"> <li>Attraction exerted to visitors of segment <math>k</math> in period <math>t</math></li> </ul>	$I_{jkt}$
<ul style="list-style-type: none"> <li>Perceived generalized cost of the local trips for visitors from segment <math>k</math> lodged at the hotel in period <math>t</math></li> </ul>	$g_{jkt}$
Average daily demand for other establishments (not explicitly modelled) in period $t$	$DO_t$
Total profit	$TP$

We assume that customers choose a hotel considering both hotel attractiveness, room price (other hotel expenses are optional and assumed not to be an integral part of the decision) and the perceived generalized cost of the local trips to the attractions. Following the logic of gravitational models, the attraction exerted by hotel  $j$  to visitors of segment  $k$  in period  $t$  is defined as:

$$I_{jkt} = W_j \cdot e^{-\beta_k \cdot p_{jt}}. \quad (3)$$

We also assume that, apart from the considered hotels, there are other establishments that are not explicitly modelled, possibly smaller and assumed to be less adaptable, which also exert some attraction over visitors. The spatial interaction model assumes that total demand is split among the hotels and other establishments proportionally to the attraction exerted by them. In the absence of capacity constraints, daily demand for hotel  $j$  in period  $t$  would be:

$$D'_{jt} = \sum_{k=1}^K \frac{I_{jkt}}{\sum_{j'=1}^{N_e+N_c} I_{j'kt} + MO_{kt}} \cdot R_{kt} \quad (4)$$

For the other establishments:

$$DO_t = \sum_{k=1}^K \frac{MO_{kt}}{\sum_{j=1}^{N_e+N_c} I_{j'kt} + MO_{kt}} \cdot R_{kt} \quad (5)$$

In the presence of capacity constraints,  $n_{jt} = \min\{D'_{jt}; C_{jt}\}$ , and when  $D'_{jt} > C_{jt}$  the remaining demand,  $D'_{jt} - C_{jt}$ , is split among the other hotels and establishments, proportionally to the respective attraction. Using (1) and (4) we can say that, for a given hotel  $j$ , the profit is given by:

$$P_j = \sum_{t=1}^T \left[ (p_{jt} - c_{jt}) \cdot \min \left\{ \sum_{k=1}^K \frac{I_{jkt}}{\sum_{j'=1}^{N_e+N_c} I_{j'kt} + MO_{kt}} \cdot R_{kt}; C_{jt} \right\} - F_{jt} \right] \cdot d_t \quad (6),$$

with  $I_{jkt}$  defined by (3).

In this model we assume that all hotels will define prices that maximize their annual profit – the set of prices is a game-theoretic Nash equilibrium. This means that if we define  $p_j$  as the vector of prices of hotel  $j$  in the different periods, then these prices should be defined in a way as to maximize the annual profit. If we look at profit  $P_j$  as a function of the price vector of hotel  $j$ , we can write:

$$\mathbf{p}_j = \arg \max_{\mathbf{p} \in \mathbb{R}_+^T} \{P_j(\mathbf{p})\} \quad (7)$$

This assumption of profit maximization by each hotel may be useful in calibrating the model in applications with real world data. In fact, it is reasonable to assume that, for a hotel chain that already possesses some hotels in a given region, some parameters will be available, or can be estimated with some ease – this is, for example, the case for  $p_{jt}, g_{jkt}, R_{kt}, \beta_k$ . However, other parameters may be more difficult to determine, such as the values of  $W_j$  for competitor hotels. Assuming that hotels are defining the prices that maximize their profit, we may indirectly estimate such parameters.

The problem we address considers that a hotel chain intends to build some new hotels and has an investment budget for that. There is a set of sites for potential new hotels and, in each of these sites, a hotel may be built according to a set of pre-defined typologies (a typology may encompass the number of hotel stars, what amenities to build, etc.). So, in this stage there is a discrete set of potential decisions for the

company, comprising the choice of sites and typologies for the new hotels (in the literature, location decisions are often modelled as discrete – see, e.g., Godinho and Dias, 2013; Seim, 2006).

The typology, along with the specific characteristics of the potential location, define the hotel attractiveness, fixed and variable costs and capacity, as well as the construction cost. As an example of the impact of location on attractiveness we can notice that, for the same typology, a more pleasant surrounding environment will increase the attractiveness. So, the company wants to determine the subset of sites in which to build new hotels, as well as the typologies of the hotels to be built, in order to maximize the total annual profit. Assuming that  $q(j)$  denotes the typology chosen for potential site  $j$ , with  $q(j) = 0$  representing the cases in which no hotel is built in site  $j$  and higher values of  $q(j)$  representing costlier and more attractive hotels, we can define the objective of the hotel chain as maximizing the following total profit function:

$$TP = \max_{q(j), j=N_e+N_c+1, \dots, N_e+N_c+N_p} \left\{ \sum_{j=1}^{N_e} P_j(\mathbf{p}_j) + \sum_{\substack{j=N_e+N_c+1 \\ q(j)>0}}^{N_e+N_c+N_p} P_j(\mathbf{p}_j) \right\} \quad (8)$$

In order to avoid cluttering notation, we omitted the dependence of prices and profits on the typologies of the new hotels. We will additionally consider a budget constraint. Denoting by  $B$  the available budget and defining that  $CC_j(0) = 0$ , that is, that the construction cost is zero when no hotel is built at a given site  $j$ :

$$\sum_{j=N_e+N_c+1}^{N_e+N_c+N_p} CC_j(q(j)) \leq B \quad (9)$$

The prices and profits considered in (8) will be the ones resulting from each hotel being a profit maximizer. This means that we assume that competitors will react rationally to the company strategy, so, when there are new hotels in the region, they redefine the room prices in a way that maximizes their profits. This leads to a game between the hotels, with a continuous strategy set for each hotel (price competition is often modelled with continuous strategy sets – for an example, see Awaya and Krishna, 2016). Models with a first stage that consists on choosing location and a second stage that is a game in which all competitors make decisions that will define how demand will be split among them can also be found in several other works (e.g., Saidani, Chu and Chen, 2012, in a retail context). We will assume that no competitor hotels will be closed

so we do not have to take into account competitor fixed costs in order to define competitor strategy. Each hotel from the chain we are considering is also assumed to be a profit maximizer, meaning that they will also define room prices that maximize their own profits. So, expression (7) must hold for each hotel. In fact, we can consider that there is a simultaneous decision about which hotels to build, which typology to choose for each one, and what prices to define. The final set of prices will constitute a Nash equilibrium, as usually considered in game theory, in which each hotel chooses a price that maximizes its own profit.

#### **4. SOLVING THE MODEL**

Solving this model entails two challenges: determining the Nash equilibrium prices for a given configuration of hotels and finding the best strategy for opening new hotels. The second problem depends on the first – in order to assess the company profit with a given configuration, we must find the Nash equilibrium prices for such configuration.

It is possible to determine the Nash equilibrium prices by using a simple algorithm that iterates over all hotels and, for each hotel, applies a numerical algorithm for determining the optimal price. Detailing the proposed procedure, we start by considering the first hotel, and determine the optimal room price, given a set of initial prices for all other hotels. In order to determine the optimal price, a numerical algorithm like the Broyden–Fletcher–Goldfarb–Shanno algorithm (e.g., Broyden, 1970) or Brent’s method (Brent, 1973) may be used. We then go to the second hotel, and determine the optimal price in a similar manner. We go on, until we reach the last one, and complete the cycle over all hotels. After that, we return to the first one, and repeat the cycle. The algorithm stops when, after such a cycle, the maximum price change is below a given threshold (that is, when all price changes are very small). When this happens, all hotels are very close to the optimal price, and we have a good approximation to the Nash equilibrium.

If the number of potential new sites is low, then the problem of finding the best strategy for opening new hotels can be tackled by enumeration: it is possible to explicitly evaluate each possible configuration and to choose the configuration that maximizes the annual profit while meeting the budget. This is the procedure we follow in the analysis of the mini case presented in Section 5. However, in most cases that will

not be possible, so we propose using a Genetic Algorithm (GA) for determining the optimal strategy for opening new hotels.

GAs are meta-heuristics that employ random choice as a guide to a search for the optimal solution, providing a population-based stochastic search procedure based on principles of natural genetics and survival of the fittest (a classical reference is Holland, 1975). They operate through a simulated evolution process on a population of string structures, named chromosomes, which correspond to candidate solutions on the search space. The first step in using them is to represent the possible solutions of the problem we are solving by a string of genes that can assume some values from a specified finite range or alphabet. For the problem we are considering, a solution will be a strategy for building new hotels, so the string of genes can consist on a string of integer values, one for each potential new site, with each value representing the typology of the hotel to be built or zero in case no hotel is built at that site. Alternatively, these integer values can be converted into binary values, which are more adequate to some GA implementations.

When starting the utilization of a GA, an initial population of chromosomes is randomly or heuristically generated, and this population will then evolve across several generations. At each generation, the performance of each chromosome is evaluated by computing its fitness value. In the case of the problem we are considering, the fitness will be the annual profit corresponding to the considered strategy, that is, the fitness function is given by expression (8), with the prices resulting from a Nash equilibrium in which each hotel is a profit maximizer (as defined in (7)) and the profits calculated according to (6). So, differently from other applications of genetic algorithms, the calculation of the fitness value of a chromosome is not straightforward, but it resorts to the previous determination of a Nash equilibrium for the set of hotels.

The population evolves through successive generations by using some operators that intend to search for the best solutions. Selection, crossover and mutation are the most common operators. The selection operator defines which chromosomes will be combined in order to generate new individuals. It usually consists of a randomized procedure that gives priority to the individuals with better values of the fitness function. There are several different methods for defining this operator, with two popular methods being roulette wheel selection (the probability of the selection of a chromosome is proportional to its fitness) and linear rank selection (the probability of the selection of a chromosome is proportional to the rank of its fitness among the population, with larger values of fitness being assigned the larger ranks).

After the two individuals are selected, there is a probability of the chromosomes being combined through the crossover operator – this probability is termed the “crossover probability”. If that does not happen, the individuals are passed to the next generation unchanged. The crossover operator implements a mating scheme between pairs of parents to create two new individuals that carry out the characteristics of both parents. Due to its importance, several crossover techniques were developed. For a binary codification, single-point crossover is often used: this operator randomly chooses a locus and exchanges the sub-sequences before and after that locus between two parent chromosomes to create two descendants.

The mutation operator is used to guarantee the genetic diversity of the population, changing randomly the values of one or more genes. This operator prevents the premature convergence of the method towards a local optimum, considering new points in the search area. Different crossover schemes may be defined – usually, each gene may be changed with a pre-defined probability.

In passing from a generation to the next one, it is possible that the chromosome corresponding to the best solution is lost and all the chromosomes in the new generation are worse than that one. One way to avoid this is by using an elitist selection strategy, where a certain fraction of the best performing individuals is kept intact into the next generation.

A GA starts by generating and evaluating the initial population. Selection, crossover and mutation are then performed, resulting in a new population. The suitability of each chromosome in the new population is evaluated and the entire cycle – selection, crossover and mutation – repeated. It is necessary to define when to stop this cycle. Usually the GA is stopped after a predefined number of generations, or if no improvements are perceived in the fitness function for a specified number of generations.

Finally, we notice that, in problems like the one we are considering, there may be unfeasible solutions – that is, the GA may generate solutions in which the budget constraint does not hold. Disregarding unfeasible solutions is a simple and popular way to handle such solutions, but it is usually not the best option (see, e.g., Michalewicz, 1995). Other approach that is often used is resorting to repair algorithms, which receive an unfeasible solution generated by the GA and “repair” it, transforming it into a feasible one. This approach has been reported as performing particularly well in several applications, at least when an unfeasible solution can be easily transformed into a

feasible one (Coello Coello, 2002). Notice that, usually, the repaired solution does not replace the unfeasible one in the population, but it is only used to calculate the fitness of the unfeasible solution.

For this model, the repair procedure we recommend consists of identifying the typology reductions that seem less likely to cause significant reductions in the fitness function and performing successive reductions until the budget constraint holds. Notice that a reduction from the lower considered typology effectively consists of not building any hotel at a given site. Detailing this procedure, we start by ordering the typologies by the order of global attractiveness (that is, by increasing order of the associated  $W_j$ ). Then, for each potential new site  $j$  and for each typology  $q$ , we calculate the value of the annual profit for a solution that consists of building just one hotel, of typology  $q$ , located at site  $j$ : let us call these values  $PS_{j,q}$ . Afterwards, for each potential new site and for each typology, we calculate the ratio  $ra_{jq}$  between the reduction in  $PS_{j,q}$  and the reduction in construction cost when we reduce the typology of the hotel from  $q$  to  $q-1$ :

$$ra_{jq} = \frac{PS_{j,q} - PS_{j,q-1}}{CC_j(q) - CC_j(q-1)}. \quad (10)$$

Small values of  $ra_{jq}$  indicate that there is a small expected reduction in annual profit for each unit of saving in construction costs. So, when a solution generated by the GA exceeds the budget, we successively reduce the construction costs by reducing the typology of new hotels, starting with the lowest values of  $ra_{jq}$ , until the total construction costs fall within the budget.

## 5. USING THE MODEL IN A MINI CASE

As a first approach for applying the model, we define a mini case illustration that can be analysed by enumeration of the alternatives. This mini case intends to show the type of analysis of alternatives that the model allows and also to determine whether the model provides sensible results for straightforward situations. We define a setting and analyse how several parameters influence the choice between a large hotel, two small hotels, and a medium-sized, high quality hotel situated in a premium location. The parameters we analyse are the average number of visitors per day, distribution of days by the high and low season, percentage of visitors in the most cost-sensitive segment,



distribution of visitors by the high and low season, attractiveness of existing hotels and cost-sensitivity of the most affluent customers. Concerning each of these parameters, we aim to answer the following questions:

- How does the profit of the alternatives change when the parameter value is changed?
- How does the best alternative change when the parameter value is changed?

For some parameters, it is easy to define expectations for the answers to these questions. For example, a higher number of visitors is expected to increase the profit for all alternatives and particularly favour the construction of a large hotel; a higher concentration of visitors in the high season is also expected to favour the construction of a large hotel. However, for others, the impact on the outcome is not so clear at the outset. We will now define the setting and then analyse the results.

We consider a small town with three main attractions ( $A=3$ ): a historic core (A1), a traditional commercial downtown area (A2) and an area where there is an old convent and a museum (A3). In this town there are already four hotels (H1-H4). There are three sites (S1-S3) where a company – which owns none of the hotels H1-H4 – is considering building new hotels. In site S1, it is possible to build either a large, high quality hotel or a smaller, medium quality hotel; in site S2, it is possible to build a small, medium quality, hotel; in site S3 it is possible to build a medium-sized, very high quality hotel. Figure 1 presents a simplified representation of the town.

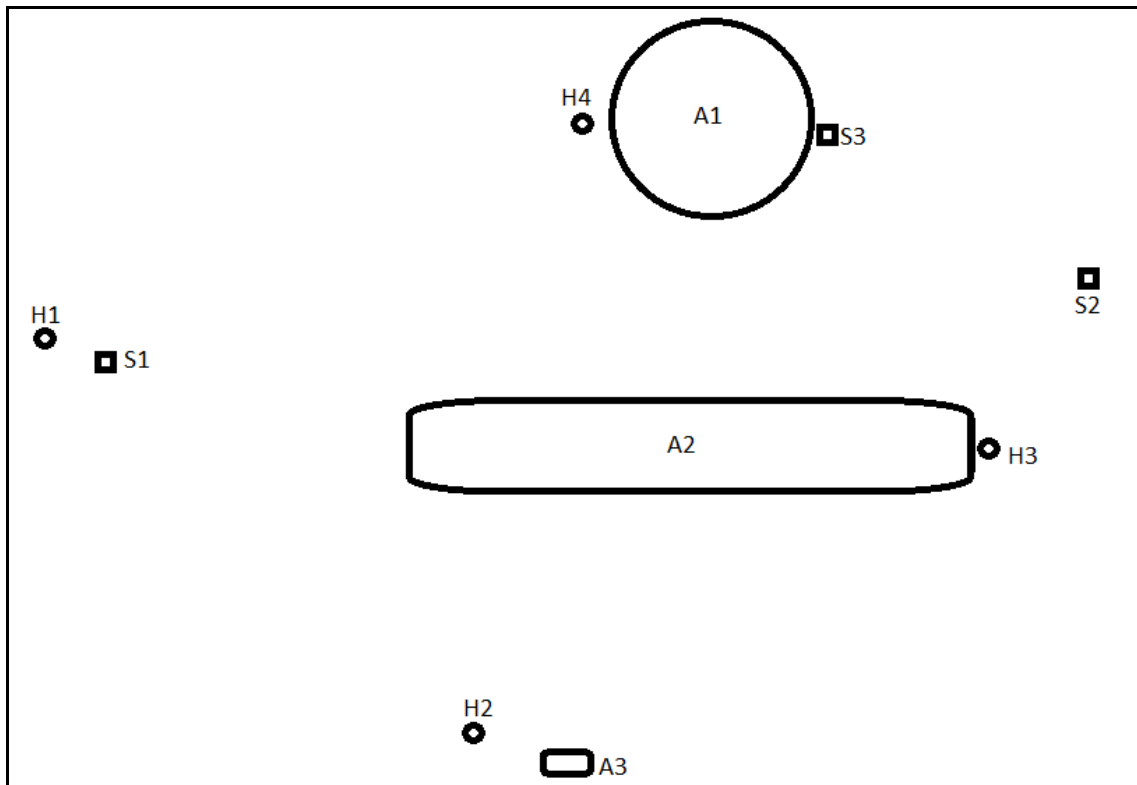
The existing hotels have different quality, H3 being the least attractive hotel, followed by H1 and H4, and H2 being the most attractive hotel. Hotels H1 and H3 are large hotels, while H2 and H4 are smaller hotels. The parameters used to define both the existing hotels and the potential new hotels are shown in Table 3. Table 4 presents the generalized cost of a trip from each hotel site to each attraction.

We consider three different periods per year ( $T=3$ ), with a duration of 122 days each. The number of visitors per day is 1200 in the high season, 800 in the middle season and 400 in the low season.

For the visitors, we consider two segments ( $K=2$ ), each with 50% of the visitors: segment 1 of cost-sensitive visitors ( $\beta_1 = 0.050$ ) and segment 2 of more affluent and less cost-sensitive visitors ( $\beta_2 = 0.015$ ). The pattern of visits to attractions is assumed

to be identical for both segments and for all periods: an average of 0.7 visits/day to A1 (historic core), 0.5 visits/day to A2 (commercial downtown) and 0.2 visits/day to A3 (old convent and museum). The joint attraction exerted by other establishments which are not explicitly modelled ( $MO_{kt}$ ) was defined as twice the average attraction exerted by existing hotels, for each customer segment and for each period.

Figure 1 – Representation of the town considered in the analysis



A1, A2 and A3: attractions (A1: Historic core; A2: Commercial downtown; A3: Old convent and museum); Small circles (H1, H2, H3 and H4): existing hotels; Small squares (S1, S2 and S3): potential locations for new hotels.

Table 3 – Parameters of the hotels used in the mini case (in this case, the parameters do not change across periods)

Hotel ( $j$ )	Attractiveness ( $W_j$ )	Capacity ( $C_j$ )	Variable cost ( $c_j$ )	Fixed cost ( $F_j$ )
H1	1.5	250	3	-
H2	3.0	100	6	-
H3	1.0	250	3	-
H4	2.0	80	5	-
New, large hotel at S1	2.0	260	7	1200
New, small hotel at S1	1.2	100	6	1000
New hotel at S2	1.2	100	6	1000
New hotel at S3	2.5	140	10	2000

-: not relevant for the model.

Table 4 – Generalized cost of a trip from each hotel site to each attraction ( $\alpha_{ja}$ )

	A1	A2	A3
H1	25	8	25
H2	30	15	3
H3	17	0	15
H4	1	12	30
S1	23	6	23
S2	17	5	20
S3	0	10	28

It was assumed that the company budget would allow it to choose one of the following alternatives:

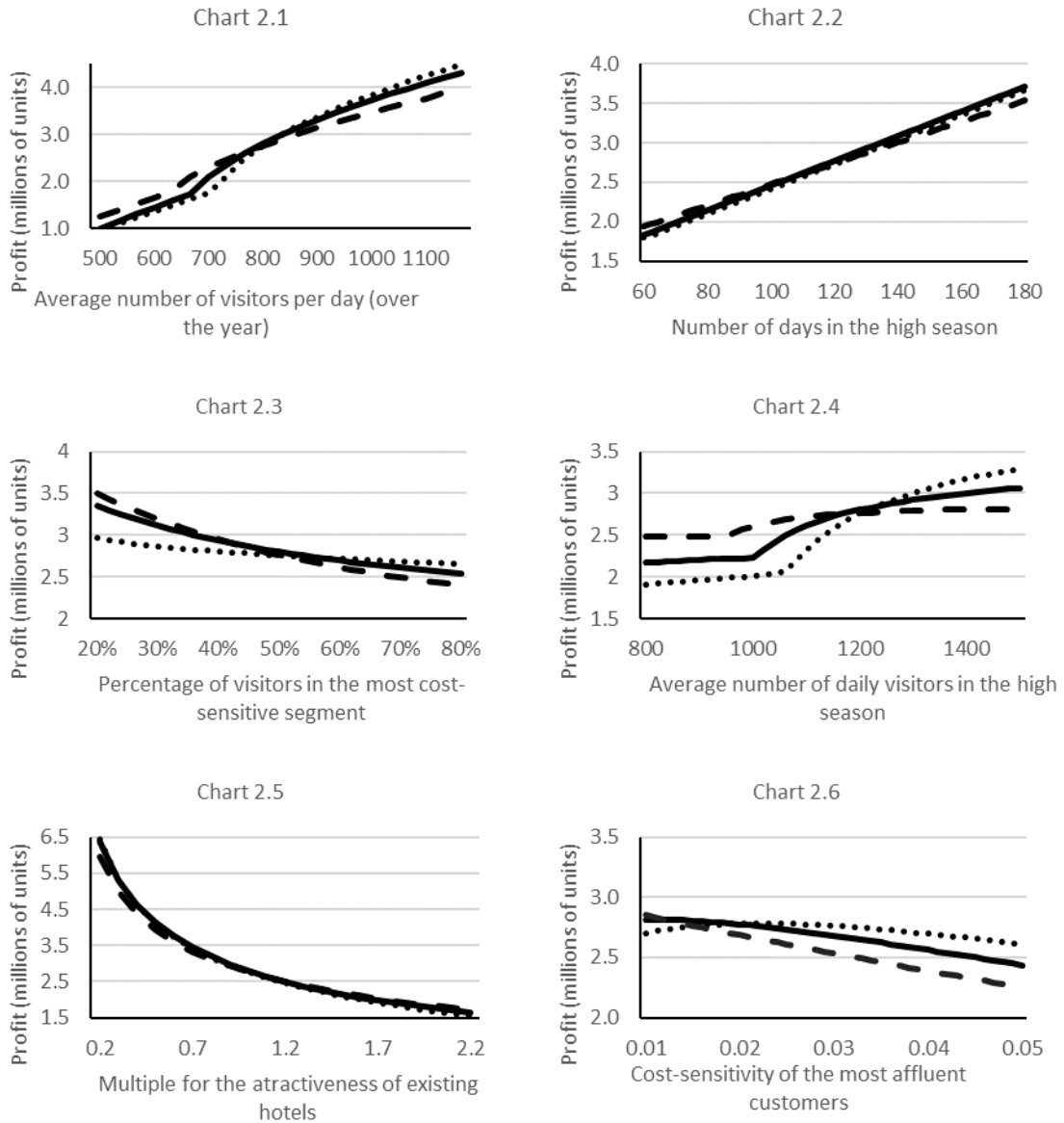
- Alternative 1: building two small hotels at sites S1 and S2;
- Alternative 2: building a large hotel at site S1;
- Alternative 3: building a luxury, medium sized hotel at site S3.

In the central scenario described here, the best alternative is Alternative 1, with an annual profit of 2 804 090 monetary units. Alternatives 2 and 3 are not far behind, with profits of 2 757 675 and 2 764 483, respectively.

The close profit values for the three alternatives make this a good case to analyse the influence of different parameters in the attractiveness of the three types of alternatives. The results we obtained are represented in Charts 2.1-2.6, in Figure 2.

In Chart 2.1 of Figure 2 we can see that, as expected, an increase in the number of visitors benefits all alternatives while a decrease harms all alternatives. We can also see that the alternatives show different sensitivity to changes in the number of visitors. The profit of Alternative 2 (large hotel) displays the wider variations, while the profit of Alternative 3 (luxury hotel) presents smaller changes. So, as the number of visitors increases, Alternative 2 becomes the most profitable as Alternative 3 is the most profitable for smaller numbers of visitors.

Figure 2 – Charts representing the impact of the changes in some parameters on the profit of the alternatives



Solid line: profit of Alternative 1 (building two small hotels at sites S1 and S2); dotted line: profit of Alternative 2 (building a large hotel at site S1); dashed line: profit of Alternative 3 (building a luxury, medium sized hotel at site S3).

Chart 2.2 shows the effects of changing the distribution of days by the high and low season (with the middle season unchanged). The X-axis of the chart shows the number of days in the high season, with a change in the number of days being accompanied by a symmetrical change in the number of days in the low season (the sum of days in the high and low seasons is always 244). In the chart we can see that the profit always increases when the number of days in the high season increases, which

was expected, since these are the days with the greatest number of visitors. We can also see that Alternative 1 (two hotels) is always more profitable than Alternative 2 (large hotel), and Alternative 3 (luxury hotel) is the most profitable when the number of days in the high season is small.

Chart 2.3 shows the impact of changing the percentage of visitors belonging to the most cost-sensitive segment. Alternative 3 is the best when few visitors are highly cost-sensitive and the worst in the opposite situation. This was to be expected, as Alternative 3 consists of building a luxury hotel. Alternative 2, which consists of building a large hotel, is the most profitable when the percentage of visitors in the most cost-sensitive segment is large.

Chart 2.4 shows the effect of changing the distribution of customers between high and low season. The X-axis of the chart shows the average number of daily visitors in the high season, with a change in this number of visitors being accompanied by a symmetrical change in the number of daily visitors in the low season (the sum of these two values is kept constant at 1600). In the chart we can see that the profit tends to increase when visitors are more concentrated in the high season, but the rate of increase is higher for intermediate values (in the 1000-1200 range) than for more extreme values. We can see that Alternative 3 (luxury hotel) is the most profitable when customers are more evenly distributed between high and low season, and Alternative 2 (large hotel) is the best when customers are more concentrated in the high season.

In Chart 2.5 we can see that profits are reduced by an increase in the attractiveness of competitor hotels, as was to be expected. In this chart, the X-axis shows the value by which the attractiveness of competitor hotels was multiplied. We can see that the differences among the profit of the alternatives is always small, with Alternative 2 (large hotel) becoming the most profitable when the other hotels are less attractive and Alternative 3 (luxury hotel) being the best when they are more attractive.

Finally, Chart 2.6 shows the impact of changing the cost-sensitivity of the most affluent (less cost-sensitive) visitors. For Alternatives 1 and 3, profit decreases when the cost-sensitivity of these customers increases, as was expected. For Alternative 2 (large hotel), profit starts by increasing and then decreases when cost-sensitivity increases. The initial increase is somewhat unexpected, but a detailed analysis showed that, for low cost-sensitivity, an increase in sensitivity makes it easier for a large hotel to compete with hotel H2 (the highest quality competitor hotel) for the most affluent visitors, even allowing it to increase prices in the high season and still fill all the rooms.

So, when the cost-sensitivity of the most affluent visitors is lowest, an increase in this sensitivity is beneficial for Alternative 2. Comparing the three alternatives, we can see that Alternative 2 is the most profitable when the most affluent visitors show higher cost-sensitivity, and Alternative 3 (luxury hotel) is the best in the opposite situation.

In summary we can say that, when we consider changes in the cost-sensitivity or distribution of visitors among segments, a luxury hotel (Alternative 3) shows the worst performance in the situations in which the profit of all alternatives is lower, and the best performance in the opposite situations. However, when we consider changes in all other parameters, a luxury hotel is the safest alternative, in the sense that it ensures the higher profit in the worst situations, and a large hotel is usually the alternative that achieves a higher profit in the best situations.

## **6. TESTING THE MODEL WITH SIMULATED DATA**

We then analysed the model with simulated data. In this instance, the results may be dependent upon the method of simulation. However, simulations over a wide range of possible situations will define trends that can be used to obtain generalizable consequences of the decisions. More importantly, in the case of a new model, it allows us to determine whether the model provides sensible answers for straightforward situations (e.g., does the profit increase when the number of visitors increase? Does the profit decrease when competition increases?) and also if the model is able to provide answers when the analysed cases are not so straightforward. These tests with simulated data were also used to gather indications of whether the GA easily converges to an optimal solution (or a solution close to the optimal).

We started by defining 32 types of problems, based on five characteristics: number of existing hotels, dispersion of the cost-sensitivity of different customer segments, differences of demand in different periods, distribution of customers over the different segments and budget size. For each type of problem, five problems were generated randomly and each one was solved seven times. By comparing the results obtained in successive applications of the GA, we can get an indication of whether it seems to be converging: if the dispersion of the results is small, this means that all runs are leading to solutions of similar quality, and that is an indication that the GA is converging.

The model was implemented in R language, resorting to the “rootSolve” package for solving the problems that allow the determination of equilibrium prices and

the “GA” package for the GA (Scrucca, 2013). The codification of solutions was made considering that an integer value from zero to  $Q$  corresponded to each potential site, initially assuming that zero corresponded to the case in which no hotel is built at the site and a positive value corresponded to the typology of the hotel to be built. These integers were converted to binary values using Gray code and resorting to the minimum of digits that were necessary to represent all the values. Gray code was used to avoid the situation in which similar typologies could only be reached with very different chromosomes (the so called “Hamming cliff” – see, e.g., Schaffer, Caruana, Eshelman and Das, 1989). Usually a given number of bits will correspond to the representation of more integer numbers than necessary (in these applications we considered four typologies plus the case of no construction, so we needed three bits, but three bits allow the representation of eight numbers – three more than we wanted). To take this issue into account, we considered that the  $Q$  higher numbers represented the different typologies, and the remaining (lower) values corresponded to the case in which no hotel is built at the site.

We followed the methodology presented in Section 4 and, based in some preliminary experiments, we used the following parameters for the GA: population size: 70; number of generations: 75; crossover probability: 80%; probability of mutation (for each parent chromosome): 15%. The 5% best elements of the population were automatically included in the following generation (elitism).

The selection method was linear rank selection. In this method, each individual in the population is given a rank, according to its fitness value. If there are  $N$  individuals in the population (in this application,  $N=70$ ), the best individual is given rank  $N$ , the second best is given rank  $N-1$ , and so on until the worst is given rank 1. Then, the probability of an individual being selected is proportional to its rank in the population. For crossover, we used single-point crossover (this method is briefly described in Section 4). For mutation, we used uniform random mutation, with the random selection of a gene from the individual to be mutated, and the change of that gene (if the gene has the value zero, it changes to one and, if it has the value one, it changes to zero).

### **6.1. Data used in the tests**

We considered a region with eight attractions ( $A=8$ ), represented by a 20x20 square. We considered three periods ( $T=3$ ) with different characteristics: a “high season” (period  $t=1$ ) with 65 days/year; a “middle season” (period  $t=2$ ) with 100

days/year and a “low season” (period  $t=2$ ) with 200 days/year. We considered there were four relevant hotel typologies ( $Q=4$ ; although other attributes can be incorporated in the definition of a typology, in this example we considered that the typologies corresponded to two star, three star, four star and five star hotels), numbered from 1 to 4 from the least attractive typology to the most attractive one.

Customers are divided into four segments ( $K=4$ ), numbered from 1 to 4 from the more cost-sensitive (these can be seen as the less “wealthy” customers) to the less cost-sensitive (the more “wealthy” ones). Each customer segment has a pattern of visits to attractions with a random number between one and a half and three average visits per day. This pattern was assumed to be the same on all time periods, and was generated randomly, and independently for all customer segments.

Both attractions, existing hotels and potential sites for new hotels were distributed randomly over the bi-dimensional 20x20 square that represents the region. We considered that customers incur a generalized travel cost of one monetary unit per unit of Euclidean distance, to travel from the hotels to the attractions. We always assumed there were  $N_p=16$  potential new locations, but the number of existing hotels varied for different problem types.

Hotel attractiveness and costs depend on the hotel typology. Hotel typology of existing hotels is generated randomly, using a discrete uniform distribution. Hotels of the lowest typology have a base attractiveness of  $W_j=1$ , and base attractiveness increases by 0.5 for each increase in hotel typology. Since location-specific factors may have an impact on hotel attractiveness, the attractiveness is calculated by perturbing the base attractiveness by a random factor in the interval  $[-20\%,+20\%]$ . In the case of new hotels, the same random perturbation factor is applied to all typologies.

Net variable costs incurred by the hotels was assumed to be  $c_{jt}=5$  monetary units per day and per occupied room, for the lowest typology, and it was assumed to increase by 1.5 monetary units for each increase in typology. Hotel capacity was assumed to be a random value drawn from a uniform distribution, between 100 and 200, for existing hotels and for all typologies. For new hotels, the capacity corresponding to the lowest typology was drawn from a random distribution between 137 and 200. We assumed that better typologies would require more space per room, effectively decreasing the hotel capacity. Hotel capacity was assumed to decrease by 10% for each increase in typology (so, for the highest typology, the minimum possible capacity is 100). Fixed costs depend both on the hotel typology and on its capacity. For the lowest typology, there is



a base fixed cost of five monetary units per day per room, and this base cost increases by one and a half monetary units for each increase in typology. It is assumed that specific factors may have an impact on fixed costs, so base fixed costs are perturbed by a random factor in the interval  $[-15\%,+15\%]$  (in the case of new hotels, the same random perturbation factor is applied to all typologies). The joint attraction exerted by other establishments which are not explicitly modelled ( $MO_{k,t}$ ) was defined as a random value, between 50% and 100% of the average attraction exerted by existing hotels, for each customer segment and for each period.

In the initial situation, we consider that hotels define the room prices as the equilibrium prices, and demand is divided among the alternatives according to the model – that is, demand is split according to (4) and (5).

Construction costs for new hotels are defined as the sum of two terms: one dependent on the typology and the other dependent on both typology and number of rooms. The first term has a base value of 2 500 000 monetary units for the lowest typology and increases by 50% with each increase in typology. In order to account for location-specific costs, this component is multiplied by a random factor in the interval  $[-10\%,+10\%]$  (for each site, the same random perturbation factor is applied to all typologies). The second component has a base value of 20 000 monetary units per room for the lowest typology, and increases by 50% with each increase in typology. In order to account for the impact of location-specific costs, this component is multiplied by a random factor in the interval  $[-2\%,+2\%]$  (once again, for each site, the same random perturbation factor is applied to all typologies). A budget of 60 000 000 monetary units was defined for the construction costs of new hotels.

## **6.2. Problem types**

We defined 32 types of problems, based on five characteristics: number of existing hotels, dispersion of the cost-sensitivity of different customer segments, differences of demand across periods, distribution of customers over the different segments and budget size. For each of these characteristics we considered 2 scenarios, and we defined a problem type for every possible combination of scenarios. For each type of problem, seven problems were generated randomly, leading to a total of 224 problems. We will now explain the values considered for each one of these characteristics.

### ***A. Number of existing hotels***

We considered a large number of existing hotels (AL), with a total of 48 hotels (40 competitor hotels and eight hotels belonging to the chain) and a small number (AS), with a total of 24 hotels (20 competitor hotels and four hotels belonging to the chain).

### ***B. Cost-sensitivity of customer segments***

We considered two scenarios of cost-sensitivity: cost-sensitivity very different for different segments (BD) and cost-sensitivity doesn't change much across different customer segments (BS). In the first scenario we have  $\beta_k = 0.032, 0.026, 0.020$  and  $0.014$  for customer segments 1, 2, 3 and 4, respectively. In scenario BS we have  $\beta_k = 0.026, 0.024, 0.022$  and  $0.020$  for customer segments 1, 2, 3 and 4, respectively.

### ***C. Differences of demand between periods***

For this characteristic, we considered that the differences might be either large (CL) or small (CS). In the first scenario we have a daily demand of 4500, 2500 and 1000 rooms per day in periods 1, 2 and 3, respectively. In scenario CS we have a daily demand of 3500, 2500 and 2000 rooms per day in periods 1, 2 and 3, respectively.

### ***D. Distribution of customers over different segments***

For this characteristic, we considered that customers may be evenly divided by all segments (DE), with 25% of customers belonging to each segment, or concentrated on the segments with higher cost-sensitivity (that is, on the least wealthy segments) (DL). In the latter scenario, we have 35% of the customers in segments 1 and 2, and 15% of the customers in segments 3 and 4.

### ***E. Budget***

For this characteristic, the levels for the budget are a large budget (EL) of 60 000 000 monetary units or a small budget (ES) of 30 000 000 monetary units.

Each of these 224 problems was solved five times, to check whether the genetic algorithm seemed to be converging with the defined parametrization. A small dispersion of values of the fitness function (total profit) is a sign that the genetic algorithm seems to be converging, while a large dispersion indicates that it is not converging.

### **6.3. Results and discussion**

Tables 5 and 6 present the results of the tests. For each problem type (defined by the respective characteristics), we present the profit of the chain after the new hotels are opened and the increase in profit due to these new hotels (from the point of view of the chain, and in relation to the situation before new hotels are opened), the average number of new hotels opened according to the optimal strategy and the percentage of hotels belonging to each typology. The presented values are averages for the seven problems of each type, and for each individual problem we considered the solution that led to the highest value of the fitness function (that is, the highest profit). In the last column of these tables we present the average value of the standard deviation of the profit (as a percentage of the maximum profit), calculated over the five solutions obtained with the genetic algorithm. We can see that this standard deviation is always less than 1.5% (and, in all but one case, less than 1%), providing evidence that the values of the fitness function are very close for all runs of the genetic algorithm, when it is applied to the same problem. So, we have reasons to believe that the parametrization used for the genetic algorithm is allowing it to converge to the optimal solution, or a solution very close to it, for all types of problems.

Table 5 – Results of the tests for instances with a large budget (EL)

Problem type				Avg. profit with optimal strategy	Increase in profit due to new hotels	Average number of new hotels	Average percentage of new hotels in each typology				Std. dev. profit
A	B	C	D				Typ. 1	Typ. 2	Typ. 3	Typ. 4	
AS	BS	CS	DE	10 151 316	3 979 484	5.9	2%	49%	39%	10%	0.26%
AS	BS	CS	DL	10 037 015	4 788 430	5.9	2%	32%	66%	0%	0.18%
AS	BS	CL	DE	17 318 951	7 128 783	4.7	0%	9%	67%	24%	0.63%
AS	BS	CL	DL	17 176 313	7 432 190	4.9	0%	18%	56%	26%	0.56%
AS	BD	CS	DE	10 298 887	3 856 547	6.1	2%	49%	49%	0%	0.17%
AS	BD	CS	DL	8 900 183	2 909 774	5.7	5%	28%	65%	3%	0.13%
AS	BD	CL	DE	16 730 340	6 973 364	4.7	0%	9%	64%	27%	0.72%
AS	BD	CL	DL	15 272 252	6 087 234	4.6	0%	9%	56%	34%	0.55%
AL	BS	CS	DE	5 195 250	1 869 009	5.0	0%	17%	57%	26%	0.24%
AL	BS	CS	DL	5 001 779	1 688 392	4.9	0%	9%	68%	24%	0.38%
AL	BS	CL	DE	3 712 622	1 288 195	4.7	0%	12%	45%	42%	0.51%
AL	BS	CL	DL	3 373 461	1 322 191	4.1	0%	3%	28%	69%	0.54%
AL	BD	CS	DE	5 472 396	1 983 524	5.1	0%	19%	61%	19%	0.26%
AL	BD	CS	DL	4 131 809	1 615 695	4.7	3%	12%	45%	39%	0.28%
AL	BD	CL	DE	3 442 427	1 361 926	4.3	0%	7%	33%	60%	0.79%
AL	BD	CL	DL	3 430 524	1 200 163	4.0	0%	0%	21%	79%	0.84%

Characteristics of the problem type: AL – large number of existing hotels; AS – small number of existing hotels; BS – cost-sensitivity doesn't change much across different customer segments; BD – very different cost-sensitivity for different customer segments; CS – small differences in demand for different periods; CL – large differences in demand for different periods; DE – Demand equally distributed over all customer segments; DL – Demand mostly belonging to the most cost-sensitive customer segments. Details on these characteristics are presented in Subsection 6.2. Increase in profit due to new hotels calculated as the difference between the annual profit after and before the opening of new hotels. Increasing typology number indicates more attractive and costlier hotels. The last column shows the average standard deviation of the profit corresponding to the best strategy identified in successive runs of the genetic algorithm. The standard deviation is calculated over five runs of the genetic algorithm, and the average is calculated over seven problems of each type.

Table 6 – Results of the tests for instances with a small budget (ES)

Problem type				Avg. profit with optimal strategy	Increase in profit due to new hotels	Average number of new hotels	Average percentage of new hotels in each typology				Std. dev. profit
A	B	C	D				Typ. 1	Typ. 2	Typ. 3	Typ. 4	
AS	BS	CS	DE	8 565 403	2 442 837	3.7	42%	31%	27%	0%	0.57%
AS	BS	CS	DL	8 318 377	2 723 092	3.6	16%	60%	24%	0%	1.27%
AS	BS	CL	DE	15 142 073	5 049 967	3.6	12%	64%	24%	0%	0.80%
AS	BS	CL	DL	14 732 613	5 258 903	3.7	31%	54%	15%	0%	0.70%
AS	BD	CS	DE	8 024 801	2 795 102	4.0	54%	32%	14%	0%	0.68%
AS	BD	CS	DL	7 200 621	2 211 094	3.7	54%	15%	31%	0%	0.63%
AS	BD	CL	DE	15 524 850	6 082 476	4.0	32%	64%	4%	0%	0.67%
AS	BD	CL	DL	14 541 163	5 363 188	3.9	22%	74%	4%	0%	0.57%
AL	BS	CS	DE	4 689 805	1 211 061	3.0	0%	43%	57%	0%	0.67%
AL	BS	CS	DL	4 636 982	1 148 206	3.1	5%	50%	45%	0%	0.65%
AL	BS	CL	DE	3 361 517	846 218	2.9	0%	35%	60%	5%	0.89%
AL	BS	CL	DL	3 074 997	866 741	2.9	0%	35%	60%	5%	0.68%
AL	BD	CS	DE	4 728 541	1 193 495	3.3	13%	48%	39%	0%	0.70%
AL	BD	CS	DL	3 634 749	918 913	3.1	9%	41%	50%	0%	0.67%
AL	BD	CL	DE	3 269 875	822 748	2.4	0%	18%	59%	24%	0.49%
AL	BD	CL	DL	2 689 396	769 678	2.7	0%	32%	58%	11%	0.93%

Characteristics of the problem type: AL – large number of existing hotels; AS – small number of existing hotels; BS – cost-sensitivity doesn't change much across different customer segments; BD – very different cost-sensitivity for different customer segments; CS – small differences in demand for different periods; CL – large differences in demand for different periods; DE – Demand equally distributed over all customer segments; DL – Demand mostly belonging to the most cost-sensitive customer segments. Details on these characteristics are presented in Subsection 6.2. Increase in profit due to new hotels calculated as the difference between the annual profit after and before the opening of new hotels. Increasing typology number indicates more attractive and costlier hotels. The last column shows the average standard deviation of the profit corresponding to the best strategy identified in successive runs of the genetic algorithm. The standard deviation is calculated over five runs of the genetic algorithm, and the average is calculated over seven problems of each type.

As expected, the existence of a large number of hotels has a very negative influence both on the total annual profit and on the additional profit due to the new hotels, since in this case there is more competition. It is interesting to analyse the profit in the cases of large and small differences of demand between periods. If the number of existing hotels is small, then the largest profit is obtained when demand is concentrated in the high season – in this case, hotels are able to charge very high room prices in the

high season. On the other hand, if the number of existing hotels is large, the highest profits are obtained when there are small differences in demand, since in this case prices during the high season are not so high, and the demand is higher during the low season (which is longer than the high season). As expected, when customers are evenly distributed by all segments, the profit is larger than in the case that most clients belong to the more cost-sensitive segments, and when the budget for building new hotels is smaller, profits are also decreased (since the investment must be decreased). Finally, we notice a slight decrease in the profits when the differences in cost-sensitivity between customer segments increase.

It is also interesting to analyse the average number of hotels belonging to each typology, for each problem type. It is very clear that the typology of new hotels is related to the number of existing hotels – higher typology hotels tend to be built especially when there are already many hotels. An explanation for this is that, when there is intense competition, highly attractive hotels are necessary to entice customers. More hotels of the highest typology (typology 4) tend to be built when there are large differences in demand among the periods, especially when these are accompanied by big differences in cost-sensitivity across customer segments or large budgets – big differences in cost-sensitivity make it more worthwhile to attract the wealthiest customers to more attractive (higher typology) hotels, with higher room prices, and a larger budget makes it easier to accommodate building higher typology hotels. On the other hand, hotels of the lowest typology (typology 1) tend to be built more often when there are big differences in cost-sensitivity across customer segments. When this happens, it is important to attract the less wealthy customers with hotels that provide significant capacity and may charge low room prices. Hotels of the lowest typology also tend to be built when the budget is smaller, which is explained by the lack of funds to build higher typology hotels.

Finally, we may notice that the number of hotels to be built decreases when the number of existing hotels is larger, when the differences in demand among periods increase and when the budget is small. In the first two cases it is advantageous to build higher typology hotels, leaving money for a smaller number of hotels. In the latter case there is less money to invest so it is possible to build less hotels.

A summary of the main trends found on the empirical tests is presented in Tables 7 and 8. Table 7 shows the main impacts found when the analysed characteristics are changed, and Table 8 shows some impacts found only in specific scenarios.

Table 7 – Summary of the main impacts found when the analysed characteristics are changed

Characteristic changed	Profit achieved with optimal strategy	Additional profit due to new hotels	Number of new hotels	Typology of new hotels
Increase in the number of existing hotels (AS→AL)	Decrease	Decrease	Decrease	Increase in the percentage of hotels of highest typologies
Increase in differences in cost-sensitivity among different customer segments (BS→BD)	Slight decrease	Slight decrease	No clear trend	Slight increase in extreme typologies (1 and 4)
Increase in differences of demand among periods (CS→CL)	No clear trend	No clear trend	Decrease	Increase in the highest typology (4)
Increase in number of customers belonging to the least wealthy segments (DE→DL)	Decrease	Slight decrease	No clear trend	No clear trend
Increase in budget (ES→EL)	Increase	Increase	Increase	Increase in the percentage of hotels of highest typologies

Table 8 – Summary of some impacts found only in specific scenarios

Scenario	Characteristic changed	Impact
Small number of hotels (AS)	Increase in differences of demand among periods (CS→CL)	Increase both in profit and in additional profit due to new hotels
Large number of hotels (AL)	Increase in differences of demand among periods (CS→CL)	Decrease both in profit and in additional profit due to new hotels

## 7. CONCLUSIONS AND MANAGERIAL IMPLICATIONS

In this paper we proposed a model for hotel pricing under competition and for determining the optimal hotel-building decisions. The main goal was to develop a model that could be applied to operational decisions in realistic and complex situations. We also intended to develop a model that can be used to analyse how the optimal decisions are altered by changes in specific aspects – that is, by changes in the model parameters. To reach these goals, we could not rely on simplified analytical approximations that might lead to attractive closed-form mathematical solutions – instead we set out to define a model intended for computational application.

The model assumes that customers make their choices based on the price, attractiveness and proximity to the attractions they intend to visit. The approach followed in this model, incorporating game theory within a spatial interaction model that may integrate inputs from studies that analyse the determinants of customers' choices, is quite novel. In fact, it goes beyond the operational categories proposed by Yang et al. (2014), and the closest work we could find is Arenoe et al. (2015). However, with respect to operational aspects, the proposed model is more general than the one presented by Arenoe et al.: our model adds a spatial interaction component and it goes beyond the simple calculation of equilibrium prices to the definition of hotel building strategies. Given the complexity of the model, we proposed the use of a GA for obtaining the best solutions.

We undertook two applications of the model. First, we used a simple setting and then we resorted to simulated data. These applications led us to conclude that the model is able to reach sensible solutions for straightforward situations and also interesting answers for not so straightforward cases.

The application with simulated data was also able to provide evidence that the model and the GA work with reasonably sized problems. Not only do the results make sense, but successive runs of the GA in the same problem provide very similar values of the profit. This is a strong indication that the GA is converging to the optimal, or a near-optimal, solution.

As we said, the goal of the paper is to propose a model and the applications were designed to provide evidence that the model works sensibly. However, despite the limitations of the analysis, it produced some pleasing ramifications. In the simulation analysis we could see that if the number of existing hotels is already large, hotel chains



achieve the highest level of profitability when market demand is evenly distributed throughout the year; however, if the number of hotels is small, profits are larger when demand changes widely among different periods, thus allowing the hotels to raise prices in the high season.

The issue of cost-sensitivity triggers multiple niches. When different customer segments have very different cost-sensitivity, hospitality organisations tend to increase their profits by building of new hotels either in the upper or lower ends of the market. The mini case that was analysed (Section 5) also produced an interesting result regarding cost-sensitivity. An increase in the cost-sensitivity of the wealthiest customers might indeed benefit a large high quality hotel that, this way, might be able to compete with luxury hotels for those customers, even being able to increase prices without losing demand.

In the mini case we were able to find out that, among the alternatives of building two small hotels, one large hotel and one medium-sized luxury hotel, this latter option seems to be the one that is less sensitive to changes in parameters unrelated to cost-sensitivity or distribution of customers across segments. This means that, if there is high uncertainty concerning, e.g., future number of visitors in the region or the attractiveness of existing hotels, building to a high-end hotel may be the safest option. On the other side, if uncertainty concerns the cost-sensitivity of customers or the distribution of customers across segments, this becomes the riskier option, with the construction of a large hotel becoming the safest option.

All things considered, the use of both game theory and a spatial interaction model seems to have a significant impact on the results. The best strategy and the hotel profits are clearly influenced by the anticipation of the competitors reaction, as expected by the use of game theory (e.g., much lower profits are expected if there are already many hotels), and the choices prescribed show that the types of hotels to be built are clearly influenced by the situation at the outset, as would be expected when we are using a spatial interaction model (e.g., highest typology hotels being mostly built when there are big differences in cost-sensitivity across customer segments).

We believe this model is a step forward in the definition of realistic models for hotel pricing and hotel site location. However, there are several ways in which this model may be extended. First, generalizing the model in order to consider that hotel attractiveness may be different for different client segments would be useful, although this may lead to additional complexity. Using an explicit model to determine how the

hotel attractiveness may depend on different attributes may be another interesting way to extend the model. Finally, we stress the importance of making a real-life application of the model for assessing the usefulness of its practical usage by hotel chains.

## REFERENCES

- Adam, I., & Amuquandoh, F. E. (2014). Hotel characteristics and location decisions in Kumasi Metropolis, Ghana. *Tourism Geographies*, 16(4), 653-668.
- Arenoe, B., van der Rest, J. P. I., & Kattuman, P. (2015). Game theoretic pricing models in hotel revenue management: An equilibrium choice-based conjoint analysis approach. *Tourism Management*, 51, 96-102.
- Awaya, Y., & Krishna, V. (2016). On communication and collusion. *American Economic Review*, 106(2), 285-315.
- Baum, T., & Mudambi, R. (1995). An empirical analysis of oligopolistic hotel pricing. *Annals of Tourism Research*, 22(3), 501-516.
- Brent, R. (1973). *Algorithms for Minimization without Derivatives*. Englewood Cliffs N.J.: Prentice-Hall.
- Broyden, C.G. (1970). The convergence of a class of double-rank minimization algorithms. *Journal of the Institute of Mathematics and Its Applications*, 6, 76-90.
- Cheung, M.W. (2014). Pairwise comparison dynamics for games with continuous strategy space. *Journal of Economic Theory*, 153, 344-375.
- Chou, T. Y., Hsu, C. L., & Chen, M. C. (2008). A fuzzy multi-criteria decision model for international tourist hotels location selection. *International Journal of Hospitality Management*, 27(2), 293-301.
- Chung, K. Y. (2000). Hotel room rate pricing strategy for market share in oligopolistic competition - eight year longitudinal study of super deluxe hotels in Seoul. *Tourism Management*, 21(1), 135-145.
- Ciliberto, F., & Tamer, E. (2009). Market structure and multiple equilibria in airline markets. *Econometrica*, 77(6), 1791-1828.
- Clarke-Hill, C., Li, H., & Davies, B. (2003). The paradox of co-operation and competition in strategic alliances: towards a multi-paradigm approach. *Management Research News*, 26(1), 1-20.
- Coello Coello, C. (2002). Theoretical and numerical constraint-handling techniques used with evolutionary algorithms: a survey of the state of the art. *Computer Methods in Applied Mechanics and Engineering*, 191(11-12), 1245-1287.

- Friesz, T. L., Mookherjee, R., & Rigdon, M. A. (2005). An evolutionary game-theoretic model of network revenue management in oligopolistic competition. *Journal of Revenue & Pricing Management*, 4(2), 156-173.
- Gans, J., & Ryall, M.D. (2017). Value capture theory: A strategic management review. *Strategic Management Journal*, 38(1), 17-41.
- Ghemawat, P. (1986). Wal-Mart stores' discount operations. Harvard Business School case No 9-387-018 on Wal-mart).
- Godinho, P., & Dias, J. (2013). Two-player simultaneous location game: Preferential rights and overbidding. *European Journal of Operational Research*, 229(3), 663-672.
- Godinho, P., Silva, M., & Moutinho, L. (2015). Tourism Site Location Based on a Genetic Algorithm. *Tourism Analysis*, 20(2), 159-172.
- Gu, Z. (1997). Proposing a room pricing model for optimizing profitability. *International Journal of Hospitality Management*, 16(3), 273-277.
- Guo, X., Ling, L., Dong, Y., & Lian, L. (2013). Cooperation contract in tourism supply chains: the optimal pricing strategy of hotels for cooperative third party strategic websites. *Annals of Tourism Research*, 41(2), 20-41.
- Holland, J. H. (1975). *Adaptation in Natural and Artificial Systems*. University of Michigan Press, Ann Arbor, USA.
- Hung, W. T., Shang, J. K., & Wang, F. C. (2010). Pricing determinants in the hotel industry: Quantile regression analysis. *International Journal of Hospitality Management*, 29(3), 378-384.
- Hurley, S., Moutinho, L., & Witt, S. (1998). Genetic algorithms for tourism marketing. *Annals of Tourism Research*, 25(2), 498-514.
- Kimes, S. E., & Fitzsimmons, J. A. (1990). Selecting profitable hotel sites at La Quinta motor inns. *Interfaces*, 20(2), 12-20.
- Kuechle, G. (2014). Regional concentration of entrepreneurial activities. *Journal of Economic Behavior & Organization*, 102, 59-73.
- Laraki, R., Solan, E., & Vieille, N. (2005). Continuous-time games of timing. *Journal of Economic Theory*, 120(2), 206-238.
- Lee, K. W., Kim, H. B., Kim, H. S., & Lee, D. S. (2010). The determinants of factors in FIT guests' perception of hotel location. *Journal of Hospitality and Tourism Management*, 17(1), 167-174.

- Masiero, L., Heo, C. Y., & Pan, B. (2015). Determining guests' willingness to pay for hotel room attributes with a discrete choice model. *International Journal of Hospitality Management*, 49, 117-124.
- Michalewicz, Z. (1995). Heuristic methods for evolutionary computation techniques. *Journal of Heuristics*, 1(2), 177-206.
- Moutinho, L., & Paton, R. (1991). Site Selection Analysis in Tourism: The LOCAT Model. *Service Industries Journal*, 11(1), 1-10.
- Niou, E., & Ordeshook, P. C. (2015). *Strategy and Politics: An Introduction to Game Theory*. Routledge.
- Penny, N.J. and Broom, D. (1988), "The Tesco Approach to Store Location", in Wrigley, N. (Ed.), *Store Choice, Store Location and Market Analysis* (pp. 106-119). London: Routledge Publishers.
- Phillips, P., & Moutinho, L. (2014). Critical review of strategic planning research in hospitality and tourism. *Annals of Tourism Research*, 48, 96-120.
- Prietula, M.J., & Watson, H.S. (2008). When behavior matters: Games and computation in A Behavioral Theory of the Firm. *Journal of Economic Behavior & Organization*, 66(1), 74-94.
- Rubinstein, A. (1990). *Game theory in economics*. Edward Elgar Publishing.
- Saidani, N., Chu, F., & Chen, H. (2012). Competitive facility location and design with reactions of competitors already in the market. *European Journal of Operational Research*, 219(1), 9-17.
- Schaffer, J. D., Caruana, R. A., Eshelman, L. J., & Das, R. (1989). A study of control parameters affecting online performance of genetic algorithms for function optimization. In Schaffer, J. D. (Ed.), *Proceedings of the Third International Conference on Genetic Algorithms* (pp. 51–60). San Mateo, CA: Morgan Kaufmann.
- Scrucca, L. (2013). GA: A Package for Genetic Algorithms in R. *Journal of Statistical Software*, 53(4), 1-37.
- Seim, K. (2006). An empirical model of firm entry with endogenous product-type choices. *The RAND Journal of Economics*, 37(3), 619-640.
- Song, H., Yang, S., & Huang, G. Q. (2009). Price interactions between theme park and tour operator. *Tourism Economics*, 15(4), 813-824.
- Von Neumann, J., & Morgenstern, O. (1945). *Theory of games and economic behavior*. Princeton, NJ: Princeton University Press.

Williams, A. M., & Baláž, V. (2015). Tourism Risk and Uncertainty Theoretical Reflections. *Journal of Travel Research*, 54(3), 271-287.

Yang, Y., Luo, H., & Law, R. (2014). Theoretical, empirical, and operational models in hotel location research. *International Journal of Hospitality Management*, 36, 209-220.

Yang, S., Huang, G. Q., Song, H., & Liang, L. (2008). A game-theoretic approach to choice of profit and revenue maximization strategies in tourism supply chains for package holidays. *Journal of China Tourism Research*, 4(1), 45-60.