

---

## **A logit-based model for facility placement planning in supply chain management**

---

### **Saeed Asadi Bagloee**

Department of Civil and Environmental Engineering,  
Sharif University of Technology,  
Azadi Street, Tehran 11365-11155, Iran  
Email: Saeed.Asadi@parsons.com

### **Matan Shnaiderman**

Department of Civil and Environmental Engineering,  
University of Auckland,  
Auckland 1142, New Zealand  
Email: Matan.Shnaiderman@biu.ac.il

### **Madjid Tavana\***

Business Systems and Analytics Department,  
Lindback Distinguished Chair of  
Information Systems and Decision Sciences,  
La Salle University,  
Philadelphia, PA 19141, USA  
and  
Business Information Systems Department,  
Faculty of Business Administration and Economics,  
University of Paderborn,  
D-33098 Paderborn, Germany  
Fax: +1-267-295-2854  
Email: tavana@lasalle.edu  
\*Corresponding author

### **Avishai Ceder**

Department of Civil and Environmental Engineering,  
University of Auckland,  
Auckland 1142, New Zealand  
Email: a.ceder@auckland.ac.nz

**Abstract:** The facility placement in supply chain management entails suppliers and consumers along with the terminals in between for distributing commodities. This study seeks to find the best terminal placement by taking into consideration the costs for both transportation and terminal construction. We call this a supplier-terminal-consumer (STC) problem and show that the STC is an NP-hard quadratic assignment problem. The NP-hard problems in real-size are proven to be intractable; hence, we develop a two-fold heuristic

method for solving the STC problems. First, we identify the commodity flow by using a logit-based mathematical programming (Logit-MP) methodology based on the demand for the commodity and the locations of the candidate-terminals. We apply Logit-MP in an iterative process and specify the maximum utilisation of the candidate-terminals. Second, the best possible locations for the terminals are identified by analysing the utilisation rates in a geographic information system interface and using an interpolation method for converting the point-based utilisation rates into spatial data. We present numerical results of a large-size transportation case study for the city of Chicago where the commodity, terminals and consumers are interpreted as wheat, silos and bakeries, respectively.

**Keywords:** facility placement problem; supplier-terminal-consumer problem; quadratic-assignment problem; QAP; supply chain management; logit; mathematical programming; MP; geographic information system.

**Reference** to this paper should be made as follows: Bagloee, S.A., Shnaiderman, M., Tavana, M. and Ceder, A. (2015) 'A logit-based model for facility placement planning in supply chain management', *Int. J. Logistics Systems and Management*, Vol. 20, No. 1, pp.122–147.

**Biographical notes:** Saeed Asadi Bagloee is a Senior Transportation Planner and Lead Modeller at Parsons in Dubai, UAE. He studied engineering and holds MS in Transportation Planning and Engineering from Sharif University of Technology in Tehran, Iran. His research interests are in financial engineering, transportation planning, demand forecasting, network modelling, transit planning, mathematical programming, econometric analysis, taxation, budget and monetary planning. He has published his research in several scholarly journals.

Matan Shnaiderman is a member of the Department of Management in Bar-Ilan University. He received his BA, MA and PhD in Mathematics from Bar-Ilan University in Ramat-Gan, Israel. He has been a Post-Doctoral Fellow at the Washington University in Saint-Louis and University of Auckland (Department of Civil and Environmental Engineering) in New Zealand. His main research interests include applied optimisation and mathematical models for supply chain management and logistics, as well as planning of public transportation. He has published papers in *Journal of Optimization Theory and Applications*, *Transportation Research Part B*, *IEEE Transactions on Automatic Control* and more.

Madjid Tavana is a Professor of Business Systems and Analytics and the Lindback Distinguished Chair of Information Systems and Decision Sciences at La Salle University, where he served as Chairman of the Management Department and Director of the Center for Technology and Management. He is a Distinguished Research Fellow at Kennedy Space Center, Johnson Space Center, Naval Research Laboratory at Stennis Space Center, and Air Force Research Laboratory. He was recently honoured with the prestigious Space Act Award by NASA. He holds a MBA, PMIS, and PhD in Management Information Systems and received his Post-Doctoral Diploma in Strategic Information Systems from the Wharton School at the University of Pennsylvania. He is the Editor-in-Chief of *Decision Analytics*, *International Journal of Applied Decision Sciences*, *International Journal of Management and Decision Making*, *International Journal of Strategic Decision Sciences*, and *International Journal of Enterprise Information Systems*. He has published several books and over one hundred research papers in academic journals such as *Information Sciences*, *Decision Sciences*, *Information Systems*, *Interfaces*,

*Annals of Operations Research, Advances in Space Research, Omega, Information and Management, Knowledge-Based Systems, Expert Systems with Applications, European Journal of Operational Research, Journal of the Operational Research Society, Computers and Operations Research, Energy Economics, Applied Soft Computing, and Energy Policy.*

Avishai (Avi) Ceder has arrived in 2007 to the University of Auckland to take the position of Professor-Chair in Transportation, within the Department of Civil and Environmental Engineering where he established, and the Director of, the Transportation Research Centre (TRC). He was the Head of the Transportation Engineering and Geo-Information Department at the Technion, a Visiting Professor at MIT, the University of California at Berkeley, and more. In 2007 he released the book *Public Transit Planning and Operation: Theory, Modelling and Practice*, by Elsevier, UK, which was translated to Chinese by Tsinghua Press, Beijing, 2010; its 2nd edition – early 2014.

---

## 1 Introduction

Efficient facility placement associated with supply chain management is a challenge to researchers and practitioners. Suppose that there is a general chain of supplier-terminal-consumer (STC) of a commodity (e.g., oil, grain, ammunition, groceries) which is produced/manufactured/supplied at one end and is consumed at the other end. Along this chain, intermediate terminals must be introduced to link the suppliers with the consumers. The terminals could be refineries, silos or arsenals. The placement of the terminals is impacted by their construction or set-up costs, and the cost to transport goods from suppliers to terminals and from terminals to consumers. We seek to identify the best terminal placement layout that minimises costs. The commodity flow is from a supplier to a consumer via a terminal. Let this problem be termed STC.

For the sake of clarity assume a large city to be supplied with a generic commodity ‘X’ from outside the city. The commodity is brought to the terminals located around the periphery of the city by heavy commercial trucks, and distributed to consumers by smaller delivery vehicles. We analysed this problem using data from the City of Chicago as a case study, assuming wheat as the commodity. Wheat is supplied and transported from farms to the silo terminals where it is stored and milled. The commodity transformed into the form of flour and is then transported to the consumers, which are bakeries.<sup>1</sup> Thus, the best arrangement of silos will be the locations which minimise the costs involved.

Wheat is transported in a network consists of nodes and links representing junctions and road segments. In addition, the wheat demand is available in an aggregated level represented by zones. For instance, the Chicago model consists of 1778 zones. Zones also are coded as specific nodes (known as centroid) and are virtually connected to the real network by dummy links. We assume that the set of existing zones or centroids are the candidate location of the prospective terminals. The wheat demand is assumed as a quantity matrix between suppliers and the bakeries describing how much of bakery  $q$ 's demand will be provided by a supplier  $p$ . Intuitively, in order to reach the bakery, the supplier must choose the appropriate terminals based on the concept of least perceived costs. The demands for the wheat and all cost components are given, and we shall explain how we set up such parameters in Section 4.

The STC problem introduced here is simple and appears frequently in practice, but it is very hard to solve real-world situations. This problem is a challenge in the areas of computer science and operational research because of the combinatorial nature of its computation. If we drop one end of the chain, the problem becomes supplier-terminal or terminal-consumer and is defined as a quadratic-assignment problem (QAP). The QAP is one of a number of traditional combinatorial problems and is widely recognised as one of the most difficult problems to solve (Burkard et al., 1998).

In QAP, consider the problem of allocating a set of terminals to a set of either suppliers or consumers, with the cost being a function of the distance and flow between the facilities, plus costs associated with a terminal being placed at a certain location. The objective is to assign each terminal to a location such that the total cost is minimised. Specifically, we are given three  $n \times n$  input matrices:  $F = (f_{ij})$ ,  $D = (d_{kl})$  and  $B = (b_{ik})$ , where  $f_{ij}$  is the flow between terminal  $i$  and supplier/consumer  $j$ ,  $d_{kl}$  is the distance between location  $k$  and location  $l$ , and  $b_{ik}$  is the cost of placing terminal  $i$  at location  $k$ . The QAP can be formulated as follows: Let  $n$  be the number of terminals and locations and denote by  $N$  the set  $N = \{1, 2, \dots, n\}$ .

$$\min_{\phi \in S_n} \sum_{i=1}^n \sum_j f_{ij} \cdot d_{\phi(i)j} + \sum_{i=1}^n b_{i\phi(i)} \quad (1)$$

where  $S_n$  is the set of all permutations of terminals' locations denoted by  $\phi$  (Burkard et al., 1998).

Despite more than three decades worth of extensive research, the QAP remains one of the most difficult optimisation problems, and no exact algorithm can solve problems of size  $n > 20$  in a reasonable computational time. In fact, Sahni and Gonzalez (1976) have shown that the QAP is NP-hard.

The STC seems to be a dual QAP consisting of supplier-terminal QAP and terminal-customer QAP. With reference to the case-study of Chicago, any zones (centroids) of the network can be a terminal location candidate and, at 1778 zones, this is far greater than the 20 locations threshold for reasonable computational time. Hence the STC is much more difficult than QAP. The dual nature of the problem along with the number of candidate locations extremely increases the difficulty of the problem compared to the QAP. Also our problem further considers the terminals' capacity, an additional aspect that does not exist in the QAP. It would be interesting to specify the hardness of the introduced problem, but that is left for further study.

Our problem shares some common ground with general problems of 'service facilities placement' in computer science (Laoutaris et al., 2007), in which the number of terminals can be a variable. In addition, the STC is one of the core topics in supply chain management (Huang, 2004). The supply chain is a set of organisations directly linked by flows of products, services, finances and information from a source to a customer. Supply chain management is the integration of key business processes across the supply chain for the purpose of creating value for customers and stakeholders (Lambert, 2008). In other words, supply chain management develops and controls the supply chain (Mentzer et al., 2001). SCM addresses several integral problems, such as distribution network configuration (Aboolian et al., 2009), information sharing (Lee et al., 2000; Ragunathan, 2001; Aviv, 2007), inventory management (Zipkin, 2000; Kogan and Shnaiderman, 2010) and pricing (Fang and Whinston, 2007). Hence the STC can be included in the

SCM category. A thorough review on facility location in the context of supply chain management can be found in the works of Klose and Drexl (2005), Reville et al. (2008) and Melo et al. (2009).

Our target was tackling the STC problem for real life cases and due to the inherent complexity of the problem we developed a heuristic methodology. The heuristic methodology is a hybrid of two components:

- First we develop a behavioural and exact mathematical methodology to understand the commodity (wheat) flow between the suppliers, terminals (silos) and consumers (bakeries). The input data are as follows: silo placement layout, the demand for the commodity (wheat) and the transportation and construction costs. This part of methodology models the behaviour of the suppliers when they choose the appropriate silo based on their own utilities. Utility is the perceived value based on a variety of factors such as transportation costs, storage charge, quality of storage, safety of the silos, and etc. The suppliers' choice behaviour is encoded as a logit model adopted from consumer theory to take the uncertainty of the behaviours into consideration. The literature integrating uncertainty in the supply chain and location decision is still scarce. In order to consider terminal capacity explicitly, we used the logit model(s) in the context of a mathematical programming (MP) model, in which the suppliers compete with each other to reach to the best appropriate silos according to their utilities. In this paper we refer to this component of the methodology as the logit-based mathematical programming (Logit-MP) model. This part of the methodology is inspired by Spiess' work (1996). Spiess (1996) developed a robust model to specify the load factor of the parking lots, paralleling terminals in our terminology, between two origins and destinations.
- Second, we will not seek the exact solution for STC in the quest of identifying locations of the terminals. Instead, we consider all the zones as candidate silo locations having infinite capacity. We run the above described Logit-MP model iteratively to evaluate the load of the wheat in each silo, to specify the silos' utilisations. After each iteration, the weakest node with the lowest utilisation is removed for the next iterations, and its utilisation is saved into a separate 'utilisation' database. When this process is completed and all nodes have been eliminated, the utilisation database is exported to a GIS-based application to analyse the silos' utilisation and their topographic status. From this information, a proposed silo location layout is identified.

The remainder of this paper is organised as follows. Section 2 is devoted to the literature review. Section 3 introduces the mathematical underpinnings of the Logit-MP model along with some additional details provided in appendices A, B, C and D. In Section 4 we present a case study for the city of Chicago to demonstrate the applicability of the proposed methodology. In Section 5 we conclude with our conclusions and future research directions.

## **2 Literature review**

The current paper extends the QAP model commonly used in the literature. Most of the algorithms are heuristic. The first algorithm was proposed by Armour and Buffa (1963)

who determined suboptimum relative location patterns for physical facilities. The authors examined their model using a manufacturing plant layout example, and mentioned that their methodology is general and may be used for a variety of applications. Later research used tabu search (Taillard, 1991; Battiti and Tecchiolli, 1994), genetic algorithms (Fleurent and Ferland, 1994) or both where a set of 29 test problems having between 30 and 100 facilities was investigated (Drezner, 2003). Cordeau et al. (2006) suggested a heuristic method for generalised QAP where the number of facilities must not be equal to the number of sites, and either zero or more than one facility may be assigned to each facility. The authors mention that their problem has applications in container yard management and in the assignment of equipment to manufacturing sites. In order to test their results, they compared them on small examples to results generated by exact formulations, and observed similar optimal values.

Because of the complexity of finding exact solutions to general QAP, research is commonly focused on either lower bounds for the QAP (Assad and Xu, 1985; Carraraesi and Malucelli, 1992; Hahn and Grant, 1998), or exact solutions for special cases. In Christofides and Benavent (1989), the special case of tree QAP, where the non-zero flows resemble a tree, is considered and is exactly solved by a branch-and-bound algorithm based on an integer programming problem formulation. The authors state that their problem has practical applications such as assembly line layout and pipeline design, since in these cases the flow graph is a tree. The algorithm has been tested on 14 examples with sizes up to 25 vertices (machines). It was found that the percents of lower bounds below the optimum values were very small, 1.07% on average. The authors conclude that in general, very short and bushy trees, or very long and thin ones lead to easier problems than in the other cases. Other research which combines lower bounds based on heuristic techniques and an exact approach based on branch-and-bound was undertaken by Maniezzo (1999). In Demidenko et al. (2006), special well solvable cases of QAP with monotone and bimonotone matrices, are considered. Those cases are solvable in polynomial time, and the authors mention that their results “lead to a wide application area concerned with allocation of industrial objects and computer components, optimal management of supply chains and computer networks.” Additional special cases of QAP, which are polynomially (or even less) solvable, can be found in Nagarajan and Sviridenko (2009), and in Zabudskii and Lagzdin (2010).

The superiority of exact solutions upon heuristic ones is discussed by Drezner et al. (2005). The authors state that “solutions at a fraction of one percent from the best known solution values are rapidly found by most heuristic methods.” The paper presents several QAP instances, up to a size of 75, that are difficult using heuristic methods but which are solved well by exact methods such as branch-and-bound and benchmarks. The authors mention that based on previous papers applications of QAP include the placement of electronic modules, the assignment of specialised rooms in a building and other applications in imagery and turbine runner balancing.

In addition, there are studies presenting generalisations and improvements to the QAP. The quartic assignment problem (QrAP) is an extension of QAP with an objective function of higher degree (Chiang et al., 2002, 2006). Nagi (2006) deals with the full assignment problem (FAP), and Zhang et al. (2011) extends it to the full assignment problem with congestion (FAPC), which simultaneously optimises the layout and flow routing. The latter research indicates the following: “FAPC is even more difficult than QAP because it needs to tackle a variant of multicommodity-flow problem for every

layout in the solution space that may contain  $(n+1)!$  permutations as in the QAP case.” The branch-and-price and the column-generalisation algorithms are used to solve the FAPC. The authors compare the QAP, QrAP and FAPC models and conclude that the FAPC is more sophisticated than both the former two, as it requires more data input and more computational effort. Minhas and Sait (2005) consider the VLSI cell placement problem, another generalisation of QAP, which consists of finding suitable physical locations for given cells on the layout. A parallel tabu search algorithm is used for efficient optimisation of a constrained multiobjective problem. In Jaramillo (2009), the multicommodity-flow problem (MFP) and the QAP are integrated to the MFQAP problem dealing with assigning nodes to locations and commodity flows to arcs to minimise the total cost or maximise the profit. Jaramillo (2009) wrote: “the MFQAP has applications in facilities design and logistic systems design... some applications of the MFQAP are the integrated machine allocation and layout problem; an extended distance based facility layout problem; and the dynamic extended facility layout problem.” The author concludes with “it is recommended to develop approximation algorithms for the MFQAP, such as tabu search and simulating annealing that are able to assign nodes to locations and update commodity flows in one step.”

The above QAP related literature clearly indicates that the size of problem is a prohibitive factor impacting the efficiency of the solution. It is worth mentioning that STC problem is a two-fold QAP that adds one more dimension to the difficulty of solving it.

Utilisation of GIS-applications in various fields including transportation science accelerated during the last three decades (Fotheringham and Rogerson, 1993; Viegas et al., 2008; Martinez et al., 2009). GIS-applications provide a set of powerful tools to better understand mostly spatial and convoluted phenomena. Hence, with respect to the spatial characteristic of the STC, we seek spatial analysis and inference modules provided by GIS-applications. Such analysis and inferences would contribute to the final proposed terminal layout.

### 3 Mathematical underpinnings of the Logit-MP model

In this section we first introduce the concept of utility and the logit model. More details can be found in McFadden (1977), Ben-Akiva and Lerman (1985) and Greene (2000). We then briefly outline the framework of the MP problem which has accommodated the logit models. More details can be found in Spiess (1996), Bagloee and Reddrick (2011), and Bagloee et al. (2012).

Suppliers and consumers (bakeries) are denoted by  $P$  and  $Q$  respectively. In between the two is the set of terminals (silos), represented by  $K$ . Supplier  $p \in P$  has to supply the amount of wheat requested by bakery  $q \in Q$  which is denoted by  $G_{pq}$ . This is an entry in the wheat demand matrix. Supplier  $p \in P$  stores and mills wheat at a silo ( $k \in K$ ). The flour is then transported to bakery  $q \in Q$ . The silo is selected by the supplier according to his utility. ‘Utility’, used with discrete choice models, is a critical concept underlying consumer theory (further discussed in Appendix A). There is a unique utility associated with each supplier  $p \in P$ , using silo  $k \in K$  to provide wheat to consumer  $q \in Q$ , denoted by  $U_{pkq}$ . Utility functions are linear functions whose variables are contributing factors to the utility perceived by supplier  $p$  of selecting silos  $k$  to supply bakery  $q$ . These factors may include: transportation costs from  $p$  to  $k$  and from  $k$  to  $q$ , storage charge at silo  $k$ ,

quality of storage and climate condition at silo  $k$ , safety of the silo  $k$ , proximity of silo  $k$  to end users (bakery  $q$ ) etc. For easier formulation, the linear utility function is split into two utility functions:  $-u_{pk}$ ,  $-u_{kg}$  which represent costs in the perceived utility (the minus sign represents a disutility).

Given the logit (utility) functions and the wheat demand, as well as the silos' capacities, a MP model was developed to determine the optimal flow of wheat. Hence the Logit-MP model requires three sets of inputs:

- 1 wheat demand matrix between suppliers and bakeries ( $G_{pq}$ )
- 2 vector of the capacity of the silos ( $C_k$ )
- 3 conductivity matrices.

The linkage between the suppliers, silos and bakeries are provided by the (dis)utility rates. With respect to the logit model expressed in equation (6) in order to facilitate the computation process, we use exponential forms of the utility function:  $\Omega_{pk} = \exp(-u_{pk})$  and  $\Phi_{kg} = \exp(-u_{kg})$ . Smaller  $u_{pk}$  or  $u_{kg}$  indicates lower disutility which implies increased  $\Omega_{pk}$  and  $\Phi_{kg}$ . Thus,  $\Omega_{pk}$ ,  $\Phi_{kg}$  are desirable inputs and convey a positive message. We label these ( $\Omega_{pk}$ ,  $\Phi_{kg}$ ) as conductivity matrices. The output is the wheat flow  $g_{pkq}$ , and consequently the amount of wheat stored at silos is ( $\sum_{pq} g_{pkq}$ ) or the silos' utilisation.

In Appendix B, the mathematical underpinnings and the proof of the Logit-MP model is provided. In Appendix C, a solution algorithm for the Logit-MP is discussed. The problem under study is analogous to a well-known problem in transportation, known as the Hitchcock transportation problem (HTP). Appendix D is devoted to a discussion of the similarity between Logit-MP and HTP. The application of the Logit-MP to the case of the Chicago network is elaborated in the next section.

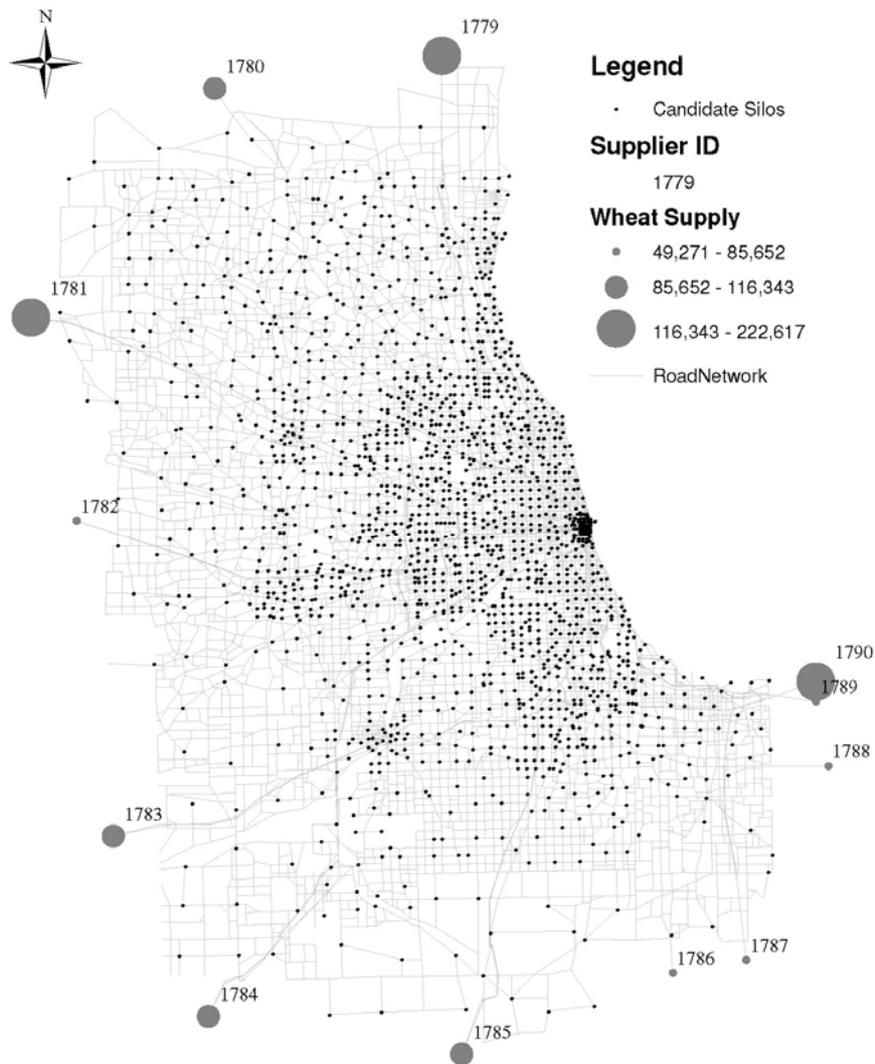
#### 4 Numerical results and the GIS application

We used the large-scale transportation model of the city of Chicago with 1,778 regular zones, 12,982 nodes, and 39,018 road links as the case study. In addition, 12 external zones placed in the peripheral areas represent the city's gates to the outside world. This network has been used by several researchers in the literature (see Bar-Gera and Boyce, 2003). This regional model dataset, like other regional models, also includes travel demand in the form of trip matrices. We interpreted and then permuted the trip demand matrix to a wheat demand matrix between the suppliers and the bakeries as follows:

- All the regular zones in the city are considered as the bakeries (zone number 1 to 1778).
- Zone numbers from 1779 to 1790, 12 external-zones, located on the peripheral area of the city are assumed as the suppliers. Figure 1 depicts Chicago's transportation network, locations of the suppliers (external zones) and the regular zones (candidate silos and bakeries).
- The trip matrix contains 1,360,427 trips per hour for the peak hour time.<sup>2</sup> We need to create the wheat demand from the trip demand which starts as a two-dimensional matrix of  $1790 \times 1790$ . The wheat demand matrix shall be in the form of

(1779...1790) × (1...1778). For each cell in the wheat demand matrix, we sum up cells in the corresponding column of the trip matrix equally, according to the order of the suppliers. For the first supplier (zone 1779) the corresponding cells of 1 to 148, for the second supplier (zone 1780) the corresponding cells of 149 to 2 \* (148) = 296 and likewise for the remaining suppliers are summed. It is worth noting that number of silos is 12 and if we divide the zones (1778 zone) to the silos, the share of each silos is almost 148 zones;  $148 \approx 1778/12$ . Thus the last silo (1790) has to take 2 more zones that is 1629...1778. Figure 1 also shows the magnitude of the wheat supply at the external-zones. These calculations result in a complete and unbiased dataset.

**Figure 1** Transportation network and wheat supply of the city of Chicago



#### 4.1 Encoding and executing the algorithm

Since the host application of the Chicago dataset is the transportation planning software known as EMME/3, we shall be using this software too. Using EMME/3 in this study provides some advantages:

- 1 The way we have translated the STC problem, utilising the transportation terminologies such as nodes, links, and zones support using transportation software
- 2 EMME/3 is designed to handle large scale matrix operations upon which this methodology is based.

Therefore the algorithm was prepared as a ‘macro’ feature of EMME/3 which calls the matrix calculation modules. The Logit-MP model is encoded as a macro called supply-silo-bakery (SSB) macro (‘macro’ simply is a set of EMME/3 commands). We used a PC with a 2.33 GHz CPU and 3.25 GB of RAM for all computations.

#### 4.2 Logit models and conductivity matrices

Each supplier has his own preference when selecting the appropriate silos, depending on the cost, charges and other considerations such as paper work, type of the silo and so on. The logit models may be calibrated given a panel of data pertaining to past supplier behaviour. Due to the widespread usage of logit models, all the econometric applications such as Limdep, ALOGIT and even most of the mathematical or statistical software such as Gauss and SPSS are capable of handling the calibration. The mathematic basis of the logit model calibration can be found in the econometric literature (see McFadden, 1977; Ben-Akiva and Lerman, 1985; Greene, 2000). Since the logit model calibration is not the focus of this study, and the undertaken case-study already is based on a trip data case, we correlate the logit models only to the ‘cost’. The Chicago transportation models yield the congested travel time from assigning the trip demand to the network. We used this travel time as a proxy for the ‘cost’. The following formulations specify the conductivity matrices:

$$\begin{cases} \Omega_{pk} = b.e^{-a.t(p,k)} \\ \Phi_{kq} = b.e^{-t(k,q)} \end{cases} \quad (2)$$

where  $t(m, n)$  is the travel time (in hour) between pair zones (or nodes)  $m$  and  $n$ ;  $a > 1$  is meant to imply that the first leg of the chain is governed by a higher cost compared to the second leg. This would force the suppliers to seek silos before connecting with the bakeries. In other words, the first leg is undertaken by heavy truck vehicles and the second leg is undertaken by lighter delivery vehicles since the mobility of heavy truck is restricted in the inner city network. Let us call this parameter  $a$  truck-charge. We tested our algorithm two times using two rates for truck-charges,  $a = 5$  and  $a = 2$ . Note that parameter  $b$  is used just to provide decent conductivity rates in terms of the number of integer digits. We tested the algorithm for  $b = 1,000$  and we obtained 332.12 as the average of the conductivity matrix of the second leg.

Thus far the input data are as follows: wheat demand ( $G_{pq}$ ); the conductivity matrices ( $\Omega_{pk}, \Phi_{kg}$ ); the infinity assumption for the capacity of the silos ( $C_k = \infty, \forall k$ ). Therefore, we are now prepared to run any scenario of the silos’ status as follows.

### 4.3 A heuristic approach to specify silos locations

This is an iterative process, and the primary output of the mathematical algorithm is the silos' utilisation. At the outset, all the regular zones are considered as candidate silo locations. After each iteration, the zone(s) (candidate-silo) with lowest utilisation is removed, and its utilisation is saved in a dataset called 'utilisation database'. This zone is no longer a candidate in the next iteration. The main concept behind this approach is that the silos' utilisation is computed while they are competing with each other. At each iteration, the poorest or weakest candidate(s) is (are) discarded. The candidate locations which survive until the very last iteration may be regarded as the strong candidates. Apart from the shear rate of the utilisations, the topographic distribution of the silos is very important too. For instance, two closely spaced silos are not a good layout. Therefore, we save the utilisation rates and present them graphically to allow us an opportunity for visual inspection which can increase our understanding of the results. The following is a pseudo-algorithm used to compute the utilisation database:

Step 0 Preparation and initialisation:

- set  $C_k := \sum_{p,q} G_{pq}$  for  $\forall k \in K$
- set  $U_k := k \in K$ , initialising the utilisation database
- call already computed conductivity matrices  $\Omega_{pk}, \Phi_{kg}$ .

Step 1 Run SSB (Logit-MP model) macro

- execute SSB ( $G_{pq}, \Omega_{pk}, \Phi_{kg}, C_k \mid G_k, G_{pk}, G_{kq}$ ), where the first set of arguments are the inputs followed by the outputs entailing contributing silos' utilisation, the wheat flow from suppliers to silos and the wheat flow from silos to bakeries respectively.

Step 2 Post-run process

- $U_k := \max(U_k, G_k)$ , updating the utilisation table.

Step 3 Termination condition

- if  $|\text{sum}(G_k) - \min(G_k)| < 1.0$  then stop
- $K := K - \{k \mid \min(G_k)\}$
- Go to step 1.

In the above pseudo-algorithm, since the SSB macro is sensitive to the silos' capacity, we set the silos capacity as the maximum required value, which is the total wheat demand. Therefore, they implicitly become infinite capacity. At the end of each iteration, in step 2 the utilisation table is updated. Then in step 3, the poorest silos whose utilisation is the minimum is discarded from the next iterations. The remaining silos (updated set  $K$  in step 3) will carry on into the next iterations until no more silos remain. Once only one silo is left with infinite capacity, the algorithm terminates. This is guaranteed when the summation and minimum of the utilisation set are identical. At the end, the last updated  $U_k$  will constitute the utilisation database.

#### 4.4 *GIS applications in the methodology*

In order to analyse the algorithm's sensitivity to the truck-charge parameter, we run the procedure twice: testing  $a = 5$  and  $a = 2$ . We expect that by lowering the truck-charge from 5 to 2, the best location of the silos becomes closer to the city. Since the number of candidate silos is 1778, basically each run comprises 1778 iterations to complete the utilisation table. With respect to the size of the case-study, the computation time was very efficient. Each run took only 1.7 hours.

The utilisation database is then exported to a GIS-based software to graphically analyse the result. We employed the leading software of ArcGIS 9.3 which includes some mathematical tools to undertake graphical analysis such as interpolation. We utilised ArcGIS' interpolation capabilities, included as part of 'spatial analyst toolbox', to convert the spot data of utilisations to smooth distributed surface data. The surface data is more intuitive than the spot and sporadic data. The interpolation tool developed by ArcGIS is the result of various research studies (see ESRI, 2009; Hutchinson, 1988; Hutchinson, 1996; Wahba, 1990). Accordingly, Figures 2 and 3 depict a so-called thermal-plot of the interpolated utilisation database for  $a = 5$  and  $a = 2$  that gives a good indication of how the best locations of the silos are distributed.

It is clear from the thermal-plots (Figure 2 and 3) that there are certain areas coloured with dark, highlighting extremely desirable silo locations. These locations are clearly distinguished and delineated from other areas. Both truck charge levels show desirable areas. However, in the case of truck-charge of  $a = 5$  (a higher charge with respect to the other case;  $a = 2$ ), the density of dark coloured areas is concentrated far from the city centre and closer to the peripheral areas. Furthermore the choices are more acute and more limited in the case of truck-charge of 5 versus  $a = 2$ .

The utilisation database in the form of thermal-plots can be viewed in conjunction with other important data. For instance a thermal-plot of the land price can be superimposed on the utilisation plot to identify the intersection of economical areas to acquire the land and desirable silo locations. The GIS-based applications provide a variety of such tools which are widely utilised in various fields.

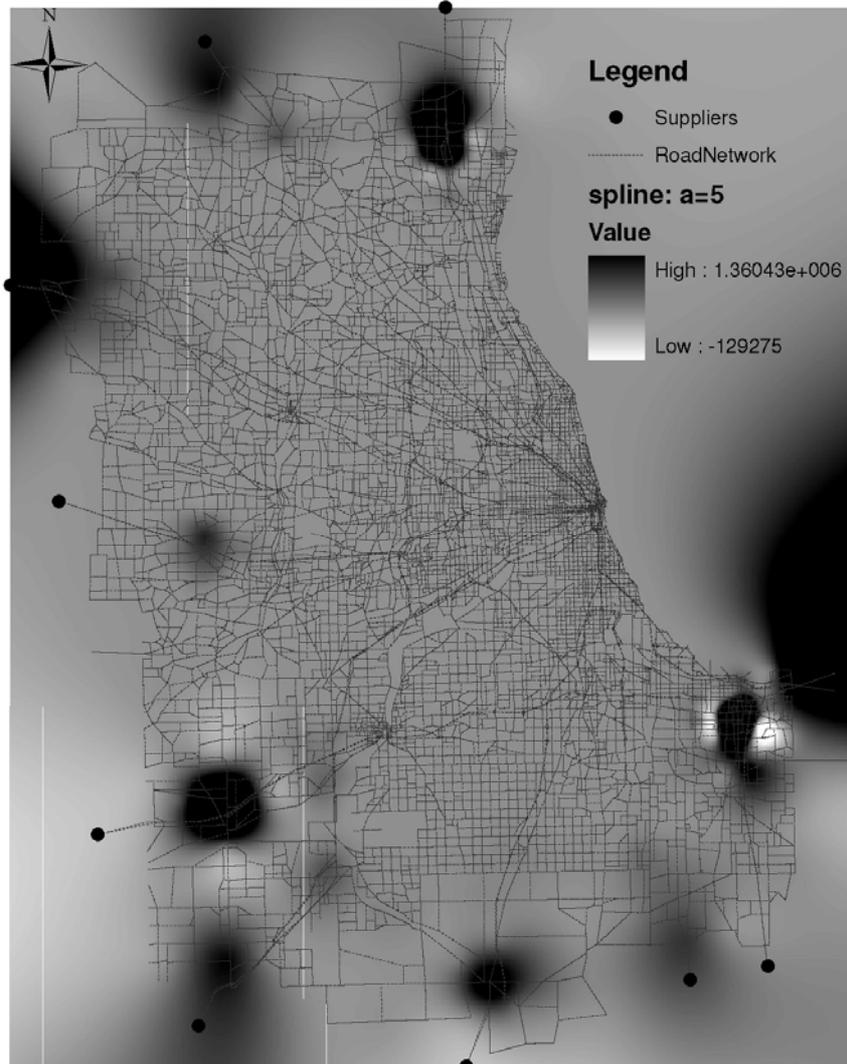
It would be helpful to view the information depicted in the thermal-plots in 3 dimensions. The addition of the third dimension to the interpolated surface thermal plot results in a mountain-like representation. Thus we call such plots mountain schemes. These tools are also provided by ArcGIS 9.3 software in a separate module known as ArcScene.

Accordingly Figures 4 and 5 show the silos' utilisation database as a mountain scheme for the two scenarios of  $a = 5$  and  $a = 2$ . The sharp differences of certain candidates in these figures will make very clear to the planners which locations are the best silo locations.

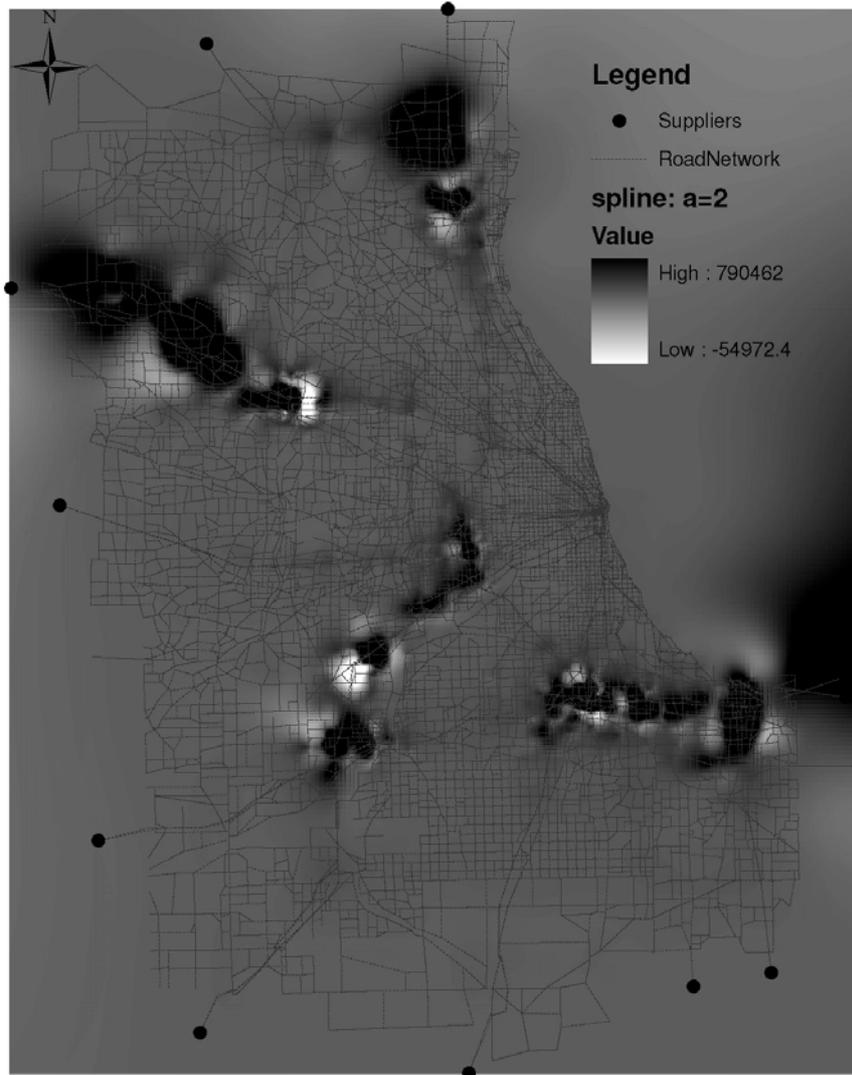
However, selection of locations for significant investments is not a one-sided activity. Efforts by municipalities or counties to actively recruit investment are widely practiced and often include special treatment such as special tax treatment as an inducement to developers. Understanding what kind of industry would find the location especially attractively helps target recruitment efforts. In addition, as a community develops its economic development goals, it is beneficial to understand what advantages the locations can offer, and what industries the community can most effectively compete for. Therefore, it is in the interest of the urban planner, municipalities and counties to incorporate the best locations for large developments such as silos in their master plans.

The thermal plots provide the planner or other authorities with information about the best possible general locations for the silos. This information does not address the number of silos or their precise locations. Addressing these concerns will increase the complexity of the already highly-complex problem and push the problem to the arena of integer programming.

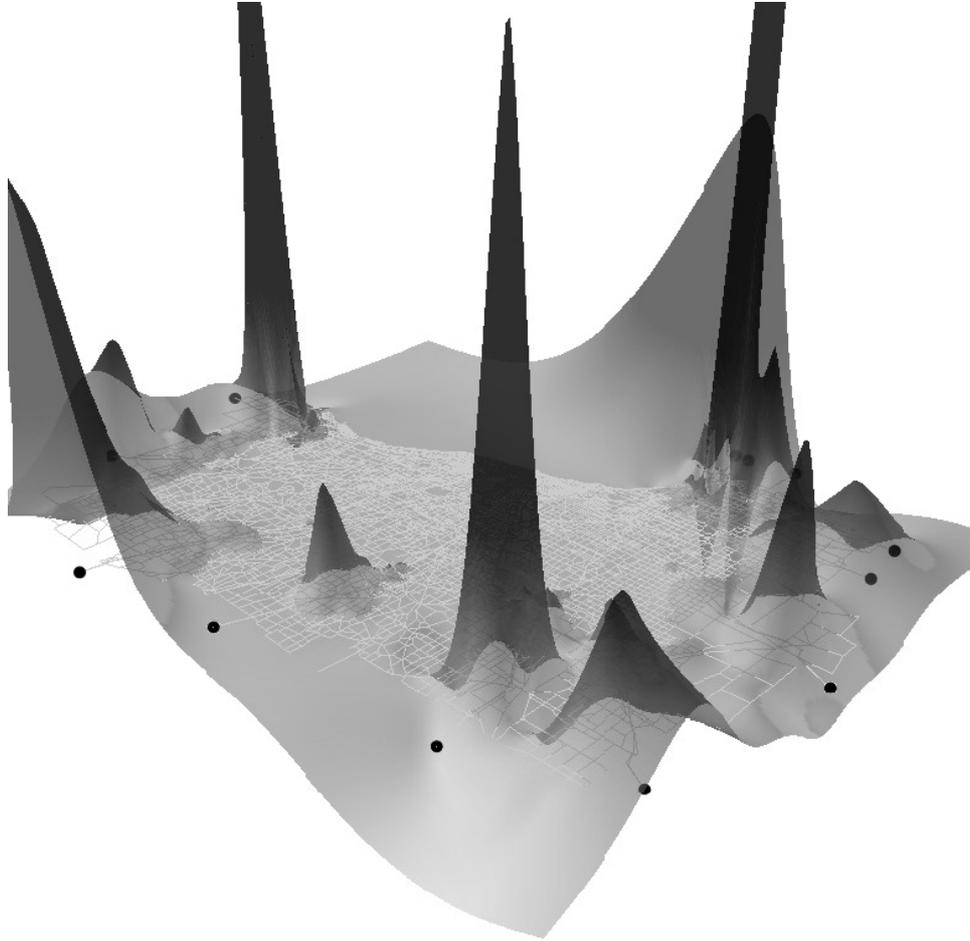
**Figure 2** Thermal-plot of the utilisation table for truck charge of  $a = 5$  in the Chicago network



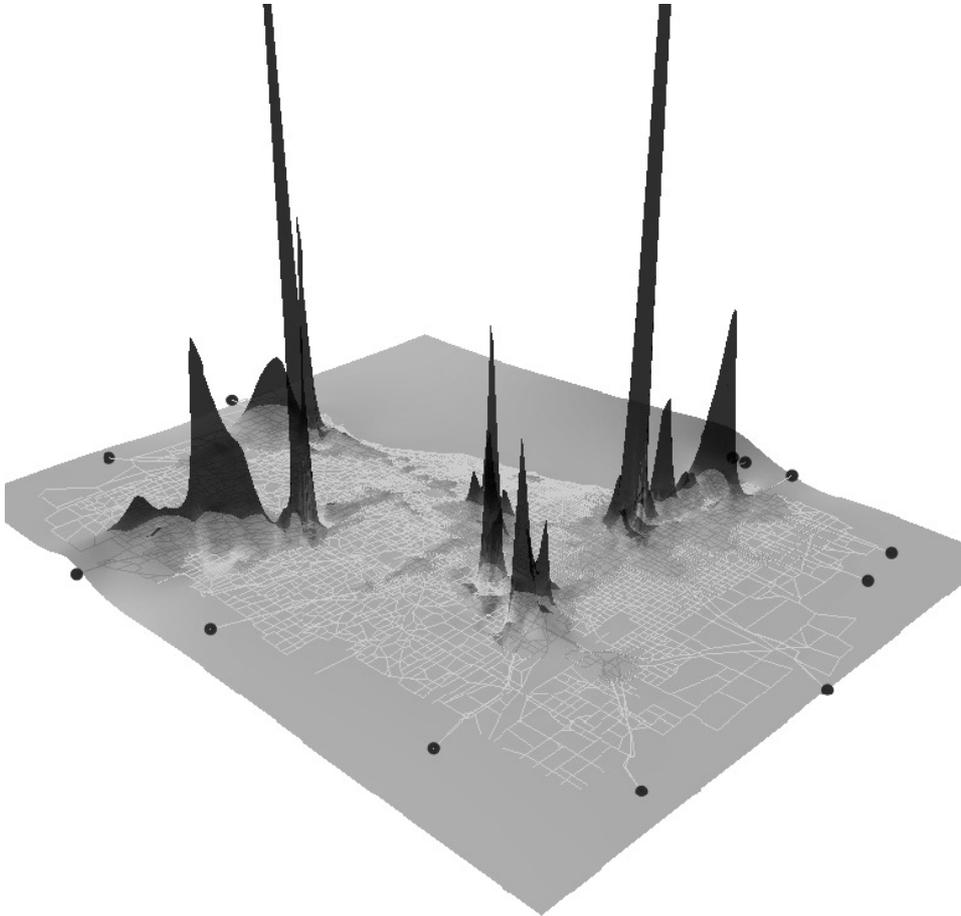
**Figure 3** Thermal-plot of the utilisation table for truck charge of  $a = 2$  in the Chicago network



**Figure 4** Mountain-plot of the silos' utilisation table for truck charge of  $a = 5$  in the Chicago network



**Figure 5** Mountain-plot of the silos' utilisation table for truck charge of  $a = 2$  in the Chicago network



## 5 Conclusions and future research directions

This research studies a common problem in the fields of supply chain management, computer science and operational research. Given a chain of suppliers, terminals and consumers and the flow of a commodity between them, the problem is how to identify the best terminal locations such that an efficient system emerges. We interpreted this as the problem of supplier, silo and bakery assuming that wheat is the commodity. This study addresses actual-size cases. The case study considered 12 suppliers and 1778 candidate locations for the silos and bakeries. The literature review shows that the problem considered is difficult to solve from a computational perspective. As a result, we developed a heuristic methodology that does not guarantee optimality, but can fit real-world problems. The methodology is two-fold. First, a Logit-MP platform is built to identify the commodity flows given a commodity's demand matrix and a scenario of terminal layout. The Logit-MP is sensitive to the capacity of the silos. We used a realistic

behavioural approach by adopting logit models for the suppliers' decision process in choosing appropriate silos. The silos, as other terminals, are limited by a certain capacity while the suppliers are competing to acquire the best silos. Second, the utilisation rates are analysed in a GIS-application module. One of the outputs of the Logit-MP model is estimated silo utilisation. We assumed infinite capacity for the silos and ran the Logit-MP model iteratively to determine its maximum utilisation. The utilisation rates were then exported to ArcGIS 9.3 for a graphical analysis and visual inspection. Two silo location scenarios were developed using two different charges for truck. The resulting two silos' utilisation datasets were presented in the form of a Thermal-plot and a mountain-scheme which showed the location of the silos. Note that consideration of the silos' capacities within the constraints in the Logit-MP model is the contribution of this study to the supply chain literature. Hence, the Logit-MP model paves the way for further studies which can identify number and the size (capacity) of the silos.

### Acknowledgments

The authors would like to thank Claire Bozic from the Chicago Metropolitan Agency for Planning (CMAP) for her insightful comments and suggestions. The authors would like to thank the anonymous reviewers and the editor for their insightful comments and suggestions.

### References

- Aboolian, R., Berman, O. and Krass, D. (2011) 'Profit maximizing distributed service system design with congestion and elastic demand', *Transportation Science*, forthcoming (accepted).
- Armour, G.C., and Buffa, E.S. (1963) 'A heuristic algorithm and simulation approach to relative location of facilities', *Management Science*, Vol. 9, No. 2, pp.294–309.
- Assad, A.A. and Xu, W. (1985) 'On lower bounds for a class of quadratic  $\{0, 1\}$  programs', *Operations Research Letters*, Vol. 4, No. 4, pp.175–180.
- Aviv, Y. (2007) 'On the benefits of collaborative forecasting partnerships between retailers and manufacturers', *Management Science*, Vol. 53, No. 5, pp.777–794.
- Bagloee, S.A. and Reddick, C.G. (2011) 'A logit model for budget allocation subject to multi budget source', *International Journal of Strategic Decision Sciences*, Vol. 2, No. 3, pp.1–17.
- Bagloee, S.A., Asadi, M. and Richardson, L. (2012) *A Methodology for Parking Modeling and Parking Pricing in Traffic Impact Studies*, TRB.
- Bar-Gera H. and Boyce, D. (2003) 'Origin-based algorithms for combined travel forecasting models', *Transportation Research Part B*, Vol. 37, No. 5, pp.405–422.
- Bar-Gera, H. (2012) *Transportation Network Test Problems* [online] <http://www.bgu.ac.il/~bargera/tntp/> (accessed October 2012).
- Battiti, R. and Tecchiolli, G. (1994) 'The reactive tabu search', *ORSA Journal of Computing*, Vol. 6, No. 2, pp.126–140.
- Ben-Akiva M. and Lerman, S. (1985) *Discrete Choice Analysis*, The MIT Press, Cambridge Massachusetts.
- Bradley, S.P., Hax, A.C. and Magnanti, T.L. (1977) *Applied Mathematical Programming*. Reading, Addison-Wesley Publishing Company, MA.
- Burkard, R.E., Çela, E., Pardalos, P.M. and L. Pitsoulis (1998) 'The quadratic assignment problem', in P.P. Pardalos and M.G.C. Resende (Eds.): *Handbook of Combinatorial Optimization*, pp.241–238, Kluwer Academic Publishers, Dordrecht.

- Carraresi, P. and Malucelli, F. (1992) 'A new lower bound for the quadratic assignment problem', *Operations Research*, Vol. 40, Suppl. 1, pp.S22–S27.
- Chiang, W.P., Kouvelis, P. and Urban, T. (2002) 'Incorporating workflow interference in facility layout design: the quartic assignment problem', *Management Science*, Vol. 48, No. 4, pp.584–590.
- Chiang, W.P., Kouvelis, P. and Urban, T. (2006) 'Single and multiobjective facility layout with workflow interference considerations', *European Journal of Operational Research*, Vol. 174, No. 3, pp.1414–1426.
- Christofides, N. and Benavent, E. (1989) 'An exact algorithm for the quadratic assignment problem on a tree', *Operations Research*, Vol. 37, No. 5, pp.760–768.
- Cordeau, J.F., Gaudioso, M., Laporte, G. and Moccia, L. (2006) 'A memetic heuristic for the generalized quadratic assignment problem', *INFORMS Journal on Computing*, Vol. 18, No. 4, pp.433–443.
- Demidenko, V.M., Finke, G. and Gordon, V.S. (2006) 'Well solvable cases of quadratic assignment problem with monotone and bimonotone matrices', *Journal of Mathematical Modelling and Algorithms*, Vol. 5, No. 2, pp.167–187.
- Drezner, Z. (2003) 'A new generic algorithm for the quadratic assignment problem', *INFORMS Journal on Computing*, Vol. 15, No. 3, pp.320–330.
- Drezner, Z., Hahn, P.M. and Taillard, E.D. (2005) 'Recent advances for the quadratic assignment problem with special emphasis on instances that are difficult for meta-heuristic methods', *Annals of Operations Research*, Vol. 139, No. 1, pp.65–94.
- ESRI (2009) *ArcGIS Desktop: Release 9.3*, Redlands, California
- Fang, F. and Whinston, A. (2007) 'Option contracts and capacity management – enabling price discrimination under demand uncertainty', *Production and Operations Management*, Vol. 16, No. 1, pp.125–137.
- FAO (2009) *Wheat Flour: Agribusiness Handbook*, Food and Agriculture Organization of the United Nations.
- Fleurent, C. and Ferland, J.A. (1994) 'Genetic hybrids for the quadratic assignment problem', *Series in Discrete Mathematics and Theoretical Computer Science*, Vol. 16, pp.173–187.
- Fotheringham, A.S. and Rogerson, P.A. (1993) 'GIS and spatial analytical problems', *International Journal of Geographical Information Systems*, Vol. 7, No. 1, pp.3–19.
- Greene, W.H. (2000) *Econometric Analysis*, 4th ed., New York.
- Hahn, P. and Grant, T. (1998) 'Lower bounds for the quadratic assignment problem based upon a dual formulation', *Operations Research*, Vol. 46, No. 6, pp.912–922.
- Huang, W. (2004) *Optimization Models for Sourcing Decisions in Supply Chain Management*, PhD dissertation, University of Florida.
- Hutchinson, M.F. (1988) 'Calculation of hydrologically sound digital elevation models', paper presented at *Third International Symposium on Spatial Data Handling*, Sydney, Australia.
- Hutchinson, M.F. (1996) 'A locally adaptive approach to the interpolation of digital elevation models', in *Proceedings, Third International Conference/Workshop on Integrating GIS and Environmental Modeling*, CA, Santa Barbara, National Center for Geographic Information and Analysis [online] [http://www.ncgia.ucsb.edu/conf/SANTA\\_FE\\_CD-ROM/sf\\_papers/hutchinson\\_michael\\_dem/local.html](http://www.ncgia.ucsb.edu/conf/SANTA_FE_CD-ROM/sf_papers/hutchinson_michael_dem/local.html).
- Jaramillo, J.R. (2009) 'The multicommodity flow quadratic assignment problem', *The 2009 Industrial Engineering Research Conference*.
- Klose, A. and Drexl, A. (2005) 'Facility location models for distribution system design', *European Journal of Operational Research*, Vol. 162, No. 1, pp.4–29.
- Kogan, K. and Shnaiderman, M. (2010) 'Continuous-time replenishment under intermittent observability (technical note)', *IEEE Transaction on Automatic Control*, Vol. 55, No. 6, pp.1460–1465.
- Lambert, D.M. (2008) *Supply Chain Management: Processes, Partnerships, Performance*, 3rd ed.

- Laoutaris, N., Smaragdakis, G., Oikonomou, K., Stavrakakis, I. and Bestavros, A. (2007) 'Distributed placement of service facilities in large-scale networks', in *Proceedings of INFOCOM*, pp.2144–2152.
- Lee, H.L., So, K.C. and Tang, C.S. (2000) 'The value of information sharing in a two-level supply chain', *Management Science*, Vol. 46, No. 5, pp.626–643.
- Luenberger, D.G. (1984) *Linear and Nonlinear Programming*, 2nd ed., Addison-Wesley.
- Maniezzo, V. (1999) 'Exact and approximate nondeterministic tree-search procedures for the quadratic assignment problem', *INFORMS Journal on Computing*, Vol. 11, No. 4, pp.358–369.
- Martinez, L.M., Viegas, J.M. and Silva, E.A. (2009) 'A traffic analysis zone definition: a new methodology and algorithm', *Journal of Transportation*, Vol. 36, No. 5, pp.581–599.
- McFadden, D. (1977) 'Econometric models of probabilistic choice', in Mansky, C.F. and McFadden, D. (Eds.): *Structural Analysis of Discrete Data with Econometric Applications*, MIT Press, Cambridge, Massachusetts.
- Melo, M.T., Nickel, S. and Saldanha-da-Gama, F. (2009) 'Facility location and supply chain management – a review', *European Journal of Operational Research*, Vol. 196, No. 2, pp.401–412.
- Mentzer J.T., DeWitt, W., Keebler, J.S., Min, S., Nix, N.W., Smith, C.D. and Zacharia, Z.G. (2001) 'Defining supply chain management', *Journal of Business Logistics*, Vol. 22, No. 2, pp.1–25.
- Minhas, M.R. and Sait, S.M. (2005) 'A parallel tabu search algorithm for optimizing multiobjective VLSI placement', O. Gervasi et al. (Eds.): *ICCSA 2005*, LNCS, Vol. 3483, pp.587–595.
- Nagarajan, V. and Sviridenko, M. (2009) 'On the maximum quadratic assignment problem', *Mathematics of Operations Research*, Vol. 34, No. 4, pp.859–868.
- Nagi, R. (2006) 'Material handling congestion and the facility layout problem', *9th International Material Handling Research Colloquium (IMHRC 2006)*, Salt Lake City, UH, pp.216–221.
- Ragunathan, S. (2001) 'Information sharing in a supply chain: a note on its value when demand is nonstationary', *Management Science*, Vol. 47, No. 4, pp.605–610.
- ReVelle, C.S., Eiselt, H.A. and Daskin, M.S. (2008) 'A bibliography for some fundamental problem categories in discrete location science', *European Journal of Operational Research*, Vol. 184, No. 3, pp.817–848.
- Sahni S. and Gonzalez, T. (1976) 'P-complete approximation problems', *Journal of the Association of Computing Machinery*, Vol. 23, No. 3, pp.555–565.
- Spiess, H.A (1996) *Logit Parking Choice Model with Explicit Capacities*, EMME/2 Support Center, Aegerten, Switzerland.
- Taillard, E. (1991) 'Robust taboo search for the quadratic assignment problem', *Parallel Computing*, Vol. 17, Nos. 4–5, pp.443–455.
- USDA (1999) Subpart M, sections 810.2201 to 810.2205, in Official United States Standards for Wheat, US Department of Agriculture, Grain Inspection, Packers and Stockyards Administration, Washington, DC.
- Viegas, J.M., Martinez, L.M. and Silva, E.A. (2008) 'Modifiable areal unit problem effects on traffic analysis zones delineation', *Environment and Planning B: Planning and Design*, Vol. 36, No. 4, pp.625–643.
- Wahba, G. (1990) 'Spline models for observational data', paper presented at *CBMS-NSF Regional Conference Series in Applied Mathematics*, Soc. Ind. Appl. Maths, Philadelphia.
- Zabudskii, G.G. and Lagzdin, A.Y. (2010) 'Polynomial algorithms for solving the quadratic assignment problem on networks', *Computational Mathematics and Mathematical Physics*, Vol. 50, No. 11, pp.1948–1955.
- Zhang, M., Batta, R. and Nagi, R. (2011) 'Designing manufacturing facility layouts to mitigate congestion', *IIE Transactions*, Vol. 43, No. 10, pp.689–702.
- Zipkin, P.H. (2000) *Foundations of Inventory Management*, McGraw-Hill Higher Education.

## Notes

- 1 Depending on harvest season, location of the crop, soil, climate, irrigation technique, regional level and other factors, there are various types of wheat. According to ‘Grain Standards of the United States’ there are eight classes of wheat used in inspecting and grading (USDA, 1999). Various available suppliers may also represent various types of wheat. For the Chicago case, to keep a safe side we conservatively considered 12 different types carried to the city from 12 different routes. Silos may also consist of various types of storages facilities to store different types of wheat. Wheat is usually transported by specially designed trucks from the farm to the silos. Since the mobility of such trucks due to size and safety is limited in the inner cities, the silos are built in the peripheral area of the cities. Silos are equipped to mill the wheat to flour. The flour is then transported to the bakeries in the city with a lighter delivery truck (FAO, 2009). Throughout the paper for the sake of brevity we refer to a single commodity (‘wheat’) even when it takes the form of flour when it is transported to the bakeries.
- 2 The main objective of this research is to develop a functioning methodology to address real life cases concern in the supply chain. As shown in the literature review, applications of the methodologies to real life cases are hindered by the size of the case-study. In order to show the merits of the methodology developed in this study we used one of the largest and most challenging network (Chicago). Apart from the size, the data of the Chicago network is easily available on the internet (Bar-Gera, 2012). This makes the Chicago network as a benchmark for the researchers to compare their methodologies. Consequently to make use of the Chicago transportation network, we interpret the trips as the weekly wheat demand in kg. This means that trips are correlated by bread (or wheat) consumption across the city.

## Appendix A

### *Utility concept and logit models*

The perception of supplier  $p \in P$  of the silos’ utilities depends on various factors such as distance or transportation costs and the charges imposed by the silos’ owner and other unknown factors. We can establish a utility function as follows:

$$U_{pkq} = \sum_i w_i . u_{pk}^i + \sum_i w_i . u_{kq}^i + \epsilon_{pkq} \quad \forall p \in P, \forall k \in K, \forall q \in Q \quad (3)$$

where  $U_{pkq}$  is the perceived utility by supplier  $p \in P$  of choosing silo  $k \in K$  to connect with bakery  $q \in Q$ . The utility consists of two arguments: known parameters associated with corresponding impacts ( $\sum_i w_i . u_{pk}^i, \sum_i w_i . u_{kq}^i$ ) and unknown factors ( $\epsilon_{pkq}$ ). The unknown factors are aggregated as a random variable indicating an error term of the utility function.  $u_{pk}^i$  and  $u_{kq}^i$  are the  $i$ -th (dis)utility factor associated with  $w_i$  a weight or impact factor.  $u_{pk}^i$  represents the factors associated with the first leg of the chain, such as the general cost to use silo  $k$  as perceived by supplier  $p$ . The general costs include transportation cost or time and charges of using the silo (the latter can be set such that the silo’s construction cost be considered too). Accordingly,  $u_{kq}^i$  represents factors pertaining to the second leg including transportation costs from silo  $k$  to bakery  $q$ .  $w_i$  determines the role of  $u_{pk}^i$  and  $u_{kq}^i$ :  $w_i > 0$  means that  $u_{pk}^i$  and  $u_{kq}^i$  are desirable characteristics, hence they are ‘utility’ otherwise they are called ‘disutility’.

Given the silos' utility,  $\Pr_{pq}(k)$ , the probability of selecting a specific silo ( $k$ ) on the route of supplier  $p$  to bakery  $q$  may be defined as follows:

$$\Pr_{pq}(k) = \Pr\{U_{pkq} > U_{pk'q} \mid \forall k \neq k', \forall k, k' \in K\} \quad (4)$$

By substituting equation (3) into equation (4) we have:

$$\Pr_{pq}(k) = \Pr\left\{\epsilon_{pk'q} - \epsilon_{pkq} < \sum_i w_i \cdot u_{pk}^i + \sum_i w_i \cdot u_{kq}^i - \sum_i w_i \cdot u_{pk'}^i - \sum_i w_i \cdot u_{k'q}^i\right\} \quad (5)$$

If  $\epsilon_{pk'q}$ ,  $\epsilon_{pkq}$  are assumed to be Weibull distributions, the differences ( $\epsilon_{pk'q} - \epsilon_{pkq}$ ) would be a logistic distribution function and a model known as logit is derived as follows:

$$\Pr_{pq}(k) = \frac{\exp\left(\sum_i w_i \cdot u_{pk}^i + \sum_i w_i \cdot u_{kq}^i\right)}{\sum_{k'} \exp\left(\sum_i w_i \cdot u_{pk'}^i + \sum_i w_i \cdot u_{k'q}^i\right)} \quad (6)$$

The probability of choosing the silos, applied to the wheat demand between the suppliers and bakeries will specify the wheat flow. Knowing the flow of wheat helps us to derive how much wheat will be stored at any specific silo:

$$g_{pkq} = G_{pq} \cdot \Pr_{pq}(k) \quad (7)$$

where  $g_{pkq}$  is the wheat flow from supplier  $p$  to bakery  $q$  by way of silo  $k$ . Since we mostly referred to the utility factors as 'cost' and 'charge', for the sake of simplicity and without loss of generality we define  $w_i = -1$  (costs and charges are not desirable terms, so they are defined as disutility) and then we discard the superscript  $i$  from the utility function implying:  $\sum_i w_i \cdot u_{pk}^i = -u_{pk}$  and  $\sum_i w_i \cdot u_{kq}^i = -u_{kq}$ .

## Appendix B

### Mathematical programming

Spieß (1996) indicated and we shall prove that Eq. (7) is equivalent to solving the following MP:

$$\text{Min} \sum_{pkq} g_{pkq} (\ell n g_{pkq} - 1 + u_{pk} + u_{kq}) \quad (8)$$

$$\text{S.t.} : \sum_k g_{pkq} = G_{pq} \quad p \in P, q \in Q \quad (9)$$

*Lemma 1:* Equation (7) holds at the optimal solution of equation (8) and equation (9)

*Proof 1:* By introducing  $\alpha_{pq}$  as Lagrangian coefficients (or dual variables) for constraints (9), the Lagrangian of the above problem is

$$L(g_{pkq}, \alpha_{pq}) = \sum_{pkq} g_{pkq} (\ln g_{pkq} - 1 + u_{pk} + u_{kq}) + \sum_{pq} \alpha_{pq} \left( \sum_k g_{pkq} - G_{pq} \right) \quad (10)$$

By establishing the first order condition of optimality better known as Kuhn-Tucker, the aforementioned equivalency of the logit form and equation (8) and equation (9) is proved:

$$\nabla_{g_{pkq}} L(\cdot) = 0 \quad \Rightarrow \quad g_{pkq} = \exp(-u_{pk} - u_{kq} - \alpha_{pq}) \quad (11)$$

$$\nabla_{\alpha_{pq}} L(\cdot) = 0 \quad \Rightarrow \quad \exp(-\alpha_{pq}) = G_{pq} / \sum_k \exp(-u_{pk} - u_{kq}) \quad (12)$$

Substituting equation (12) into (11) yields equation (7). End of Proof 1 ■

The proposed MP must adopt a set of new constraints to take the capacity of the silos ( $C_k, \forall k \in K$ ) explicitly into consideration. This can be done by adding the following constraint to equation (8) and equation (9):

$$\sum_{pq} g_{pkq} \leq C_k \quad k \in K \quad (13)$$

This constraint can be easily integrated into the above formulations: By introducing  $\beta_k \geq 0$  as dual variables for Eq. (13) along with  $\alpha_{pq}$  for equation (9) and arranging the Kuhn-Tucker optimality condition for the new problem equation (8), equation (9) and equation (13) we can derive a similar equation to equation (10) (see Spiess, 1996; Bagloee and Reddrick, 2011). The betas have a very important implication: in fact the betas indicates the price of one extra unit of the respective silo's capacity with respect to the objective function equation (8).

## Appendix C

### Solution algorithm

Since the above introduced problem is free of the convexity issue, Spiess (1996) has developed an efficient solution algorithm using the Successive Coordinate Descent (SCD) method (see Luenberger, 1984). The SCD seeks the Kuhn-Tucker conditions through a five-step algorithm. The algorithm is terminated when, for each and every silo, the aggregation of the wheat flows ( $\sum_{pq} g_{pkq}^i$ ) fall in an acceptable proximity of  $\varepsilon$  with respect to the capacity ( $C_k$ ):

$$Err_k = \sum_{pq} g_{pkq}^i - C_k < \varepsilon, \forall k \in K \quad (14)$$

where  $Err_k$  is the total amount of capacity constraint violation at silo  $k$ . In this case-study of Chicago we shall be dealing with rates of million kilograms (kg), hence, assuming  $\varepsilon = 1(kg)$  is more than sufficient. Note that in the above formulation the superscript  $i$ , refers to the iteration number.

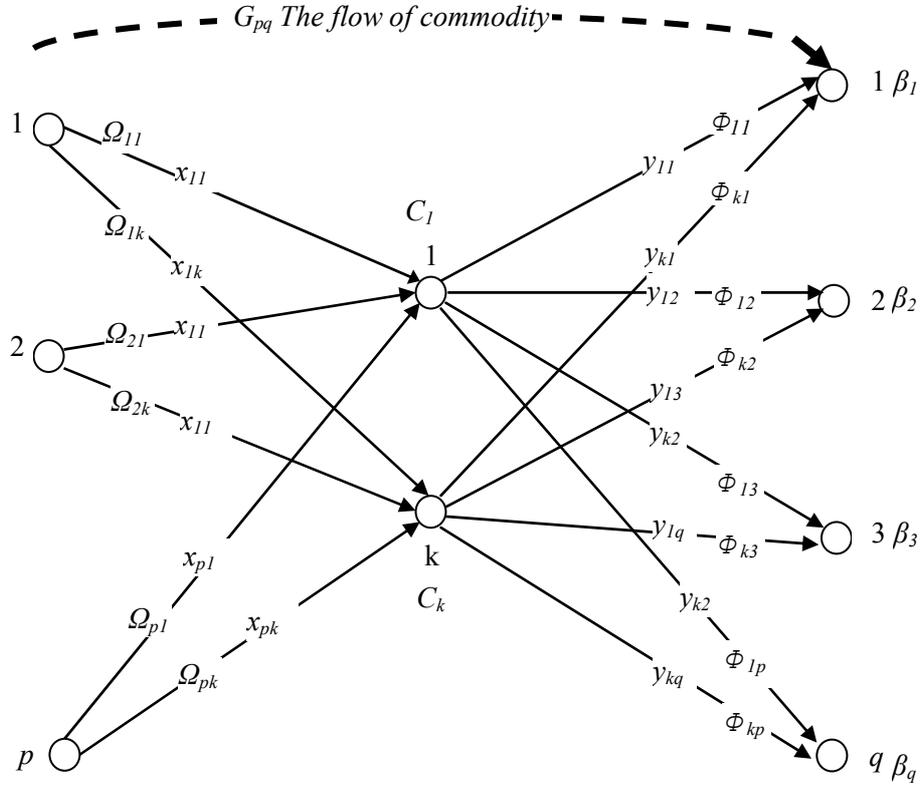
For large size cases, efficient coding and handling implementation of the algorithms may become a serious issue. In this regards, permutation of the implementation process to the format of matrix operations simplifies the process since ordinary commercial optimisation software such as GAMS and MATLAB are able to accommodate the matrix operations required for the STC problems. Fortunately, the solution algorithm can be easily reinterpreted as a series of matrix calculation with simple operations (see Spiess, 1996; Bagloee and Reddrick, 2011). The methodology requires three sets of inputs:

- 1 wheat demand matrix between suppliers and bakeries ( $G_{pq}$ )
- 2 vector of the capacity of the silos ( $C_k$ )
- 3 conductivity matrices.

The linkage between the suppliers, silos and bakeries are provided by the (dis)utility rates. With respect to the logit model expressed in equation (6) in order to facilitate the computation process, we will be using exponential forms of the utility function as:  $\Omega_{pk} = \exp(-u_{pk})$  and  $\Phi_{kq} = \exp(-u_{kq})$ . Smaller  $u_{pk}$  or  $u_{kq}$  means less disutility, which implies more  $\Omega_{pk}$ ,  $\Phi_{kq}$ . Thus  $\Omega_{pk}$ ,  $\Phi_{kq}$  are desirable inputs and convey a positive message. Therefore we label them as conductivity matrices. The output would be the wheat flow  $g_{pkq}$ , and consequently the vector of the silos utilisation ( $\sum_{pq} g_{pkq}$ ) as well as the betas.

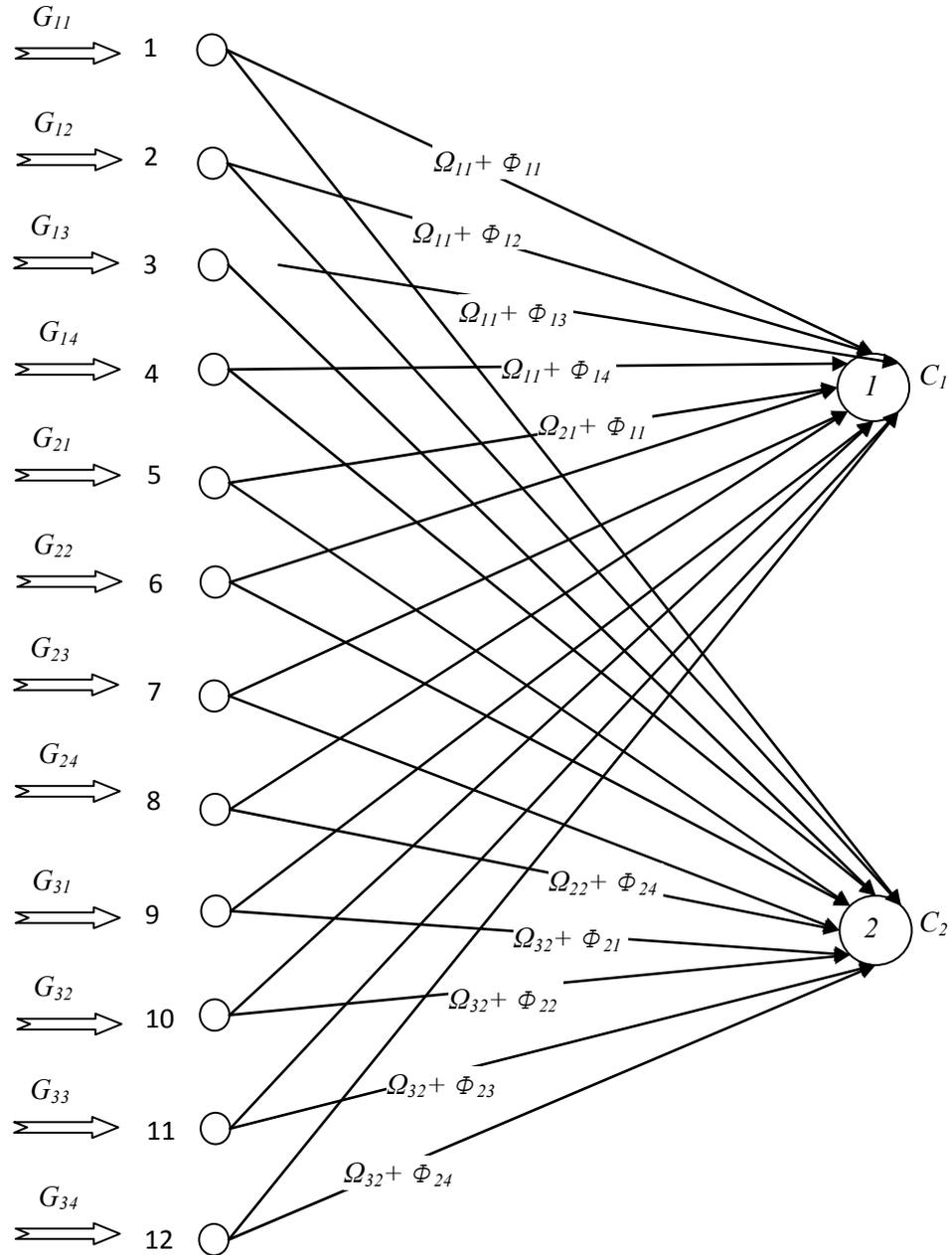
Figure 6 graphically depicts the structure of the mathematical methodology as well as the input and output data. The suppliers are placed on the left hand side. Bakeries are on the right hand side. The commodity flow is from left to right (supplier to bakery). In the middle wheat must be stored in the silos, which are placed in the middle layer. Let us fold the two legs of the supply chain in Figure 6 on the layer of silos. This would lead to a two layer network such that the pair of supplier-bakery demands is placed on one layer and on the other layer the silos are placed where the linkage between supplier-bakery pairs and silos represents the cost applied to get the respective silos. Figure 7 depicts such an arrangement consisting of 3 suppliers, 4 bakeries and only 2 silos in which the combination of supplier-bakery will add up to 12. Figure 6 is reminiscent of the very well-known problem in Operational Research known as the HTP. In the HTP, only the minimum of the total incurred cost or charge (or in this study: disutility) to the (wheat) flow is sought [i.e.,  $Min_{g_{pkq}} \left( \sum_{pk} U_{pk} + \sum_{kq} U_{kq} \right) \cdot g_{pkq}$ ]. The operational research literature yields very efficient solution algorithms to tackle the HTP even for very large size cases (Bradley et al., 1977). In Appendix D we discuss why we did not adopt the Hitchcock problem despite of the aforementioned efficient solution advantage.

**Figure 6** A schematic outlook of the proposed structure for the STC problem



Notes: *Input data:*  $\Omega_{ij}$ ,  $\Phi_{1p}$  conductivity rates for on the first and second leg of the chain;  
 $C_k$  capacity of silo  
*Output data:*  $x_{pk}$ ,  $y_{kq}$  denote the flow on the first and second leg of the chain (i.e.,  
 $g_{pkq} = x_{pk} + y_{kq}$ ),  $\beta_k$  is shadow price for silo's capacity

**Figure 7** Similarity of the STC problem and the HTP



## **Appendix D**

### *Proposed methodology versus HTP*

As discussed, the HTP corresponds to cost minimisation, where the cost (or based on our terminology; disutility) are exactly known. That is not the case in this study. As discussed before, due to the complexity of the problem, many known and unknown factors are influencing the suppliers' decision. Imagine a situation in which a supplier faces two almost identical silos: one with cost (or disutility) of 1.0 and the other with cost (or disutility) of 0.9. The Hitchcock interpretation of the situation suggests that the supplier will approach to the cheapest silos even though the costs difference is tiny. However it is intuitively conceivable that this kind of decisive and absolute situation in a fuzzy real world environment is impossible. Instead we desire a behavioural approach, in which we do not have exact knowledge of the costs and (dis)utilities, rather, the perceived disutilities are statistically distributed. This gives rise to a logit distribution for the behaviour choice which is not only more realistic but also gives solutions that are mathematically much more stable.