An integrated model of logistic networks for locating production and distribution points

E. ROMANO, L.C. SANTILLO, P. ZOPPOLI
Department of Materials Engineering and Operations Management
University of Naples "Federico II"
p.le Tecchio – 80125 Napoli
ITALY

www.impianti.unina.it

Abstract: - In the past years we have assisted to a growing number of applications of mathematics for location and routing problems, but just a few studies have been used to support decisions regarding the location of production and/or retail points. The model that we here propose has been used for determining the best location for a pastry shop, and it's retail points, which are dislocated on a local district. Wanting to choose the best solution possible, we must consider the double nature of each kiosk, which is obviously a retail point, but must also be constantly restocked, and for this reason we have to consider the route from the production point. Neither the less, each kiosk must be located where easily reachable by clients. The integrated model determines which locations are characterized with a higher attractiveness for the product, while the second integrated model, supplied with data from the first model, determines which route is optimal, by following a certain number of parameters. When both models are integrated into one we can elaborate valid estimations on location points, but also kiosk and production dimensions.

Key-Words: - logistic network, location routing, vehicle routing, simulation

1. Introduction

The following study has the object to determine the location for production centers and retail centers such as kiosks. For this reason we have realized an integrated model which is able to simulate clients behavior referred to the purchasing of the product, but not only, the integrated model simulates which routes are mainly used by consumers and which routes should be used for distribution.

The integrated model is an instrument that identifies the location of the whole production system, generating also estimations on what size would be best for production and on the dimensions of each kiosk.

The business that we have chosen adapts very well to the necessities of this work: great volume production of baked croissants by a continuous production line. Therefore the business considered is the case of a mono-product line where the plant is completely dedicated to only one product.

These kind of plants, can be simply automated by using specific machinery and by organizing the plants layout in sequentially tasks.

At the end of the process, the products are stocked in refrigerated cells waiting to be distributed to all kiosks.

All kiosk must by equipped with: an appropriate freezer, so to be able to conserve the croissants, preserving its fresh baked characteristics, an oven is necessary for heating the croissants, a cell for thawing and leavening, a coffee espresso machine and finally a fridge for drinks.

2. Models for choosing the location

The models structure enables us to choose the best location for the whole production and distribution system, meaning the production plant but also the retail points, the model is articulated in three levels which constantly interact between themselves for determining the optimized solution, the models are:

- Network model, which simulates the transportation system;
- A model for determining the location of the retail points of the baked goods [7, 9];
- Model for locating the production center.

The first model that we have listed acts as input for the following two models, determining on one side the attractiveness of baked goods for immediate consumption, on the other side, using the transportation networks load, with the relative constraints, we will have a typical routing problem where we can determine the generalized moving cost function of the vehicles.

2.1. Modeling the transportation system

A transportation system are all the social elements (residential areas), economic activities and infrastructures characteristics, such as streets or public transportation, which concur to create a demand for transportation; which if satisfied stimulate a continuous socio-economic growth. The system can be seen as a subsystem of a more vast territorial system, it can be

divided in two principal components: **the demand systems** [5] **and the supply systems** [3, 4].
Building transportation network models is carried out principally in two phases:

– Selection of the infrastructures and/or transportation services that are of importance to the problem that we are considering;
– Building the offer model: which is obtained by structuring the network model and defining all costs functions.

Constructing a network model, normally, proceeds by a sequence of tasks [11]:

– **Delimitation of the area of interest to our study**: defining the physic borders of the problem;
– **Zoning:** all transportation that takes place in a certain territory can begin and end in any point of the territory that we've considered; to simplify the systems modeling it is very useful to distinct the territory in traffic study zones, between which are held vehicle transits of our interest.

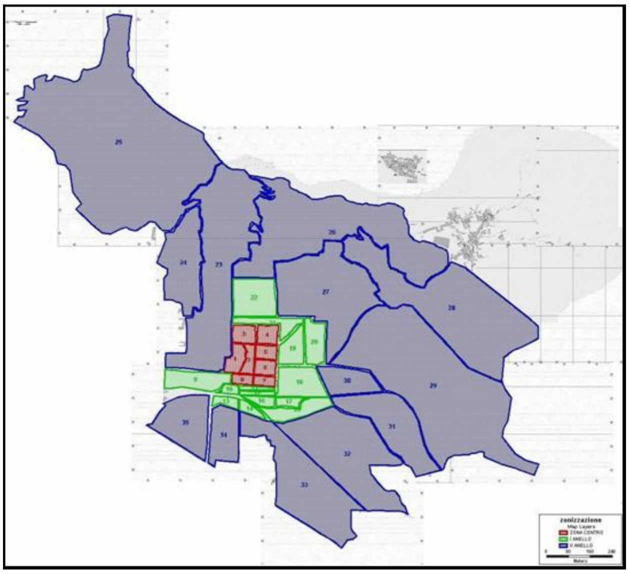

Fig. 1: The macro-zones: ring I, ring II;

– **Extracting the graph**: the graph G consists in a pair of ordinate sets, a set of N elements called nodes, and a set of L elements called arcs: G=(N,L). The nodes can indicate concrete or abstract entities, such as location points on a territory, different physical components or different activities of a component.
– **Defining a network:** the passage from a graph to a network takes place by associating quantity characteristic to the same graph: each graph arc, used to indicate a single transportation route, is characterized by a transferring time and/or by other consuming factors, when a transportation client moves from one starting node to a final node.

Let us define as transportation network T the ordinate set built from a set N of nodes, a set L consisting in pairs of nodes belonging to the set N, and a set of costs indicated with C associated to elements belonging to the set L: T=(N,L,C). To calculate the mean time of transit along any arc, we used a universally known function invented in the States, calibrated even for Italy: the function BPR (Bureau of Public Roads):

$$t_i = t_{0i}\left[1+\alpha\left(\frac{f_i}{Cap_i}\right)^{\beta}\right] \qquad (1)$$

In this expression t_{0i} is the transiting time in total free conditions (which means without other vehicles interfering during transit along the arc); f_i is the flow on arc i; Cap_i is the arcs capacity, that is the maximum number of users that pass the street section in a unit of time; α and β are coefficients which depend on the roads general characteristics, these coefficients must be determined for each road on the basis of experimental data.
The formula used to calculate the transiting time for each arc is the following:

$$t_{0i} = \frac{L}{V_{0i}} + \frac{1}{2}T_c(1-\mu)^2 \quad [h] \qquad (2)$$

where:

$$V_{0i} = 31.1 + 2.8L_{ui} - 12.8T_i^2 - 10{,}4D_i \quad [km/h] \qquad (3)$$

and with:

– Vo: velocity (km/h);
– Lu: width of the roadway, or lane for two-way roads, eliminating the area used for side parking (meters);
– T: grade of curviness of the road, graded on a scale (0,1);
– D: grade of disturbance of circulation, also these graded on a scale (0,1);

The value of D has been fixed to one, where schools with limited parking space, and along narrow streets, zero where the roads are wide and do not obstacles a vehicles circulation. T has been fixed by direct measurements regarding useful width of the roadways.
To study the transportation demand we must define innumerable aspects:

– The classes of user (c) that are mobile;
– The time regarding the beginning and end of a user transit (h);
– The reason for a user movement (s);
– The origin (o) and destination (d) of a movement;
– The kind vehicle used for the movement (m).

These parameters can be organized in origin-destination matrices, each matrix referring to the reason of the movement, time slots of the transit and users class considered (s, h and c).
The estimation of the matrix o/d will then be corrected using mobility models, which uses data regarding the flow of vehicles; these models are based on Generalized Least Squares, which are used to minimize the difference between the counted flow and the flow that would be calculated assigning a starting matrix.
The initial estimation is determined by using behavior models, based on the concept of "casual or aleatority usefulness".
The usefulness that a generic user *i* associates to the alternative *j*, depends on many factors, all factors aren't

even known with certainty and therefore these parameter are represented as aleatory variables. The usefulness perceived, can be considered as the sum of the systematic utility V^i_j, which represents the attended value of usefulness expected and an aleatory residue ε^i_j, indicating the deviation of usefulness perceived by the user *i*, so we can write:

$$U^i_j = V^i_j + \varepsilon^i_j \quad \forall j \in I^i \tag{4}$$

On the basis of the last hypothesis it is possible to determine the probability that a user *i* decides for the alternative *j*, conditioned by his set of past decisions I^i, as the probability that the alternative considered has a perceived usefulness maximum through all the possible alternatives available:

$$p^i\left[j/I^i\right] = \Pr\left[U^i_j > U^i_k, \forall k \neq j, k \in I^i\right] \tag{5}$$

Regarding the functional aspects of the usefulness choosing models, the joined distributions of the residues, most importantly, are two: the Gumbel distribution and the Gaussian multivariate. In the first case the residues have mean null and variance $\pi^2 / 6\cdot\alpha^2$, where α is a distribution parameter, while the covariance is considered zero. This last hypothesis implicates that the residues must be independent, and therefore even the alternatives must be independent. The model that we obtained, is known in literature as **Logit** [3, 8] and has the following functional form:

$$p^i(j) = \frac{\exp(\alpha V^i_j)}{\sum_{k\in I_i} \exp(\alpha V^i_k)} \tag{6}$$

So, choosing a certain alternative has a probability $p^i(j)$ which depends on the parameter α and the systematic usefulness associated to the alternative V^i_j.

So summarizing we can say that a user chooses a certain route depending on his scope (s), his destination (d), the means of transportation (m), the route which shall be used (k) all in a period of time h. Even though the parameters previously listed are interdependent, for analytic and statistical reasons, it is preferable to divide the global demand function in the product of interconnected sub models, each related to a parameter dimension. The sequence of sub models most used is:

$$d_{oh}(s,m,k) = d_o(sh)\left[SE,T\right]\cdot p(d/osh)\left[SE,T\right]\cdot p(m/odsh)\left[SE,T\right]\cdot p(k/modsh)\left[SE,T\right] \tag{7}$$

while we shall use a simplified notation:

$$d_{od}(s,m,k) = d_o(s)\cdot p(d/os)\cdot p(m/ods)\cdot p(k/\mathrm{mod}\,s) \tag{8}$$

The sub models that appear are named **generation model**, **distribution model**, **mean model** (or **mean repartition model)** and **routing model** which, together with a procedure for loading the network, will obtain the vehicle flow on each arc, assumes the name of **assigning model.**

The periods of time that we've considered are two, and precisely: $h_1 = 8 \div 9$ and $h_2 = 18 \div 19$; the same periods of time chosen for modeling the system of offer.

The **generation model** supplies us with the average number of vehicle passages made in the period h by the generic user i belonging to the generic class c, for the reason s, with origin and destination o and $d_o(s)$.

In this work for calculating the number of generic users belonging to a category c which accomplishes a transit for a s reason, by using a generating model based on a longer period of time, on a whole day $d^c_o(s)$, multiplied for the probability which estimates the percentage of alternative mobility during the same time slots, $p^c(h_{i=1,2})$; this probability is obtained by a behavior model deriving from the **Logit** model

$$d^c_o(s,h) = d^c_o(s)\cdot p^c(h_{i=1,2}) \tag{9}$$

where:

$$d^c_o(s) = m(o)\cdot \frac{n^c(o)}{n^{totale}(o)} \qquad \forall c \tag{10}$$

which represents the mean number of daily transitions, and

$$p^c(h_i/s) = \frac{\exp(V^c_{j/s})}{\sum_{j=1,2}\exp(V^c_{j/s})} \tag{11}$$

is the probability of a user, belonging to the category c, moves during the time slot h_i. For calculating $n^c(o)$ we have used data from ISTAT on the stratified population.

For calculating the probability we have used a usefulness aleatory model, where in the analytic formulas 3.13 appear:

- j= the possible alternatives: h_1, h_2;
- V_j= the systematic utilities equal to $V_j = \Sigma_j \beta_j X^c_j$, with X^c_j depending on the alternatives and β_j the relative parameters or homogenizing coefficients The systematic usefulness during h_1 are:

$$V^{c=studenti}_{h1} = \beta_1 X^{studenti}_{h1} + \beta_4 X^{sat}_{h1} \tag{12}$$

$$V^{c=ad\det ti}_{h1} = \beta_2 X^{ad\det ti}_{h1} + \beta_4 X^{sat}_{h1} \tag{13}$$

$$V^{c=altri}_{h1} = \beta_3 X^{altri}_{h1} + \beta_4 X^{sat}_{h1} \tag{14}$$

Analogously, we can deduct the same for the time period h_2. The probabilities of h_1, therefore, must become:

$$p^{c=studenti}_{h1} = \frac{\exp(\beta_1 X^{studenti}_{h1} + \beta_4 X^{sat}_{h1})}{\exp(\beta_1 X^{studenti}_{h1} + \beta_4 X^{sat}_{h1}) + \exp(\beta_5 X^{studenti}_{h2} + \beta_8 X^{sat}_{h2})} \tag{15}$$

$$p^{c=ad\det ti}_{h1} = \frac{\exp(\beta_2 X^{ad\det ti}_{h1} + \beta_4 X^{sat}_{h1})}{\exp(\beta_2 X^{ad\det ti}_{h1} + \beta_4 X^{sat}_{h1}) + \exp(\beta_6 X^{ad\det ti}_{h2} + \beta_8 X^{sat}_{h2})} \tag{16}$$

$$p^{c=altri}_{h1} = \frac{\exp(\beta_3 X^{svago}_{h1} + \beta_4 X^{sat}_{h1})}{\exp(\beta_3 X^{svago}_{h1} + \beta_4 X^{sat}_{h1}) + \exp(\beta_7 X^{svago}_{h2} + \beta_8 X^{sat}_{h2})} \tag{17}$$

For determining the eight coefficients β_1, β_2, β_3, β_4, β_5, β_6, β_7, β_8 we've calibrated the model on the alternative which where chosen by a casual sample of n generic users. For the calibration we've used the method of the Maximum Likelihood.

Table 1: Model calibrated coefficients obtained from experimental data

β_1	β_2	β_3	β_4	β_5	β_6	β_7	β_8
$2\cdot10^{3}$	$1\cdot10$	$7\cdot10^{-1}$	0	0	0	$1\cdot10$	$9\cdot10^{-2}$

Table 2: Probability value obtained from the Logit model

$p_{h1}^{c=studenti}$	$p_{h1}^{c=ad\,det\,ti}$	$p_{h1}^{c=svago}$	$p_{h2}^{c=studenti}$	$p_{h2}^{c=ad\,det\,ti}$	$p_{h2}^{c=svago}$
1	0,841	0,257	0	0,159	0,743

The distribution model gives the percentage $p^c(d/osh)$ of users (belonging to category c) who move from a zone o for the reason s in the slot of time h, arriving at destination d. The model that we have decided to use is the gravitational model, which has the following structure:

$$p_{d/os}^{c} = \frac{Add_d^{\beta_1} \cdot D_{od}^{-\beta_2}}{\sum_{d' \neq d} Add_{d'}^{\beta_1} \cdot D_{od'}^{-\beta_2}} \qquad (18)$$

In which:

- Add_d= is the attractiveness of the area d, depending in general on the transitions reason and the category of users.
- d'= total number of zones in the macro-area
- D_{od}= distance expressed in meters, being the air line distance from the center of the origin and the center of the destination zone d.

The model for choosing the **mean** gives the probability, $p^c(m/oshd)$, of the users belonging to the category c, which bring themselves from the origin o to the destination d for the reason s during the time slot h, by using the mean m. We will consider the only the case of a mono-mean network which considers as only option the car as mean of transportation. These decision has been demonstrated to be realistic by the data obtained from experimental measurements conducted in city, deducting that only 4% of the city population moves by public transportation.

Last of sub models considered for evaluating the demand, following the four phase approach, is the **route choosing model** , which gives the share of users that utilize for their movements the route k of the network related with the mean m, when transiting from origin o to destination d for the reason s. Obviously the user will pick the alternative which he perceives as less costly, and therefore:

$$\hat{C}_k = C_k + \varepsilon_k \qquad \forall k \in I_{odm} \qquad (19)$$

being C_k the average perceived cost for route k, equal to the sum of all the arcs that compose the route, and I_{odm} being all the possible routes which connect the pair o/d by mean m.

The model used for choosing the route, together with the network loading process, is called the **assigning model**. This model, by combining the demand and offer of local mobility is able to determine the vehicle flow on all the roads considered, the networks performance (time and cost for travel) and the impact that each flow has on external system factors.

To resolve the problem by dynamic means, we have used a Geographic Information System (GIS), which enables to elaborate and manage of different nature associated to the territory. The software that has been used in the present model is called TransCAD (produced by Caliper Corporation) [1]. This model joins, the typical characteristics of GIS to those of a software transportation system modeler, becoming a transportation modeling software.

The resolution method that has been chosen, that will be done in a TransCAD environment, is the Stochastic User Equilibrium method. In fact, this method has been assumed as the best because the choosing the route depends on a stochastic usefulness of the routes, which in the same time depend on a series of factor such origin and destination pair o/d, which is a continuous function depending on the costs of transiting on alternative route which connect the same o/d pair. Assigning a user stochastic equilibrium we presume that all users of the road network don't know or perceive the exact cost for each arc/road transition. For this reason transiting along an arc doesn't have a single deterministic cost, but a series of values that vary following a certain probabilistic distribution with the "real" cost as its mean value. Using stochastic trip costs, instead of deterministic values on the arcs, enables to consider a series of factors that not only are not mentioned in the model but which may also be unknown, for example different values attribute by the user to the time factor, and so on).

The outputs coming from the dynamic algorithm TranCAD are, for each arc and flow direction:

- Total flow load on the network;
- Trip time, while the network is loaded;
- V_{oc}= flow/capacity.

The results obtained are indicated in the figures below. As we can see in the central zone of the city (indicated in red) the difference between input flow and output flow is never negative, and this positive fact is confirmed also in h_2 , even if the mean values are different: 496 veihecles/h_1 e 377 veihecles/h_2. Such results where attended, in fact, the data confirms what was instinctively known: such a macro-zone is very attractive being an area for work, study and entertainment during both time slots.

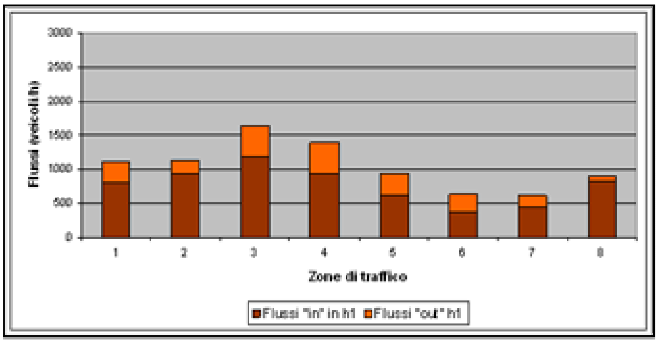

Fig. 2: Incoming flow and outgoing flow in the center traffic zone belonging during h_1

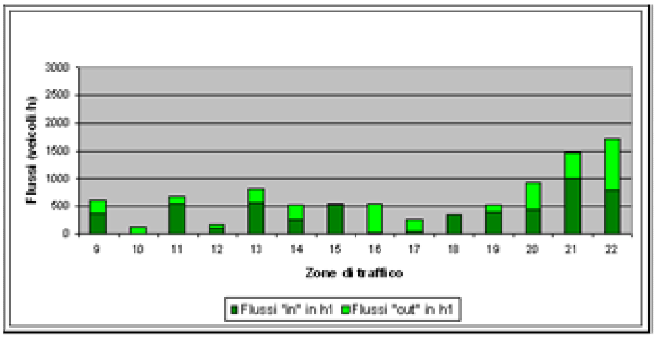

Fig. 3: Incoming flow and outgoing flow in the traffic zone belonging to the I ring in h_1

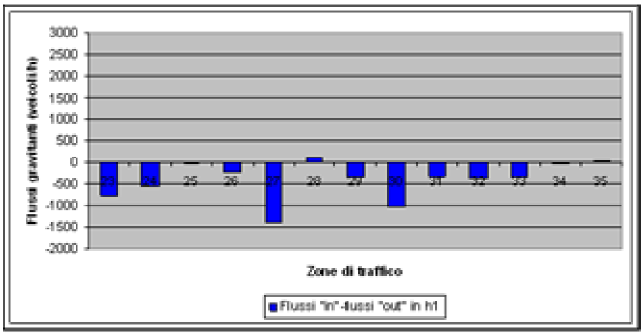

Fig. 4: Incoming flow and outgoing flow in the traffic zone belonging to the II ring in h_1

In the first ring we can see how the situation has changed: in this case we do not have a unique trend, because of the different areas that are here considered that aren't all considered attractive, and this is all more evident during h_2 that in h_1. The traffic zones, belonging to such macro-zone, where the result is negative, are areas with a high residential concentration, therefore the high share of workers that leave this area are not counterbalanced by other flows that are incoming, for this reason these areas are not considered attractive. Confirming such considerations, we can see how in the zone 22 the data on the population is quite high while in zone 14 it is almost absent but with a high attractively.

2.2. Locating the retail points [13]

The criteria used for determining the location points for the kiosks is that to put them where there are potential clients which while allow to maximize profits. A behavior model, for this reason, has been developed which would indicate where to place such kiosks in the urban textile in relation to how a client chooses a certain route between all the alternatives available. In this case the alternatives can consider eventual competitors who offer similar product (in our case bars). To be able to determine where to locate the kiosks, different quality and quantity analysis have been conducted on local data, trying to find which of the last could be used for finding a correlation on business activities and social-economic-territorial parameters. The first data analyzed was the one regarding the concentration of bars on the local territory, 129 in all. We have considered this data for single traffic zones and the whole macro-zone. In the central zone the concentration is quite high, 59 of the 129 bars are located in the heart of the city, in an area that extends for only 2% of the total territory. Immediately adjacent to the center of the city the number and related density of the bars that are present decreases: there are 40 bars on a territory that extend a bit more then the precedent, around 5%of the total area. The decreasing trend is also confirmed on the outskirts of the city where the 23% of the total bars are present on an area that is much larger, precisely 92% of the whole macro-area considered in this study.

An important aspect, observed during the data gathering phase, regarding the counting and locating of the bar that are already present on the territory, is that farther is the bar from the center of the city and more are the services and spaces which they offer, such as newspapers, lottery tickets, cigarettes, piano bar, happy hour and so on, becoming helpful for social encounters. The second information analyzed regards the distance which each bar has respect the related zones' center. What can be clearly seen is that exists a relation between two variables; the relation is that father the distance from the center of the macro-zone the number of bars decrease (figure 3.7), and therefore we can consider the center as a fictitious node of the model where the model considers all zone activities concentrated in the center. So the maximum concentration of bars where there's the most concentration of activities, while these same densities decrease together further the distance from the center of the macro-zone.

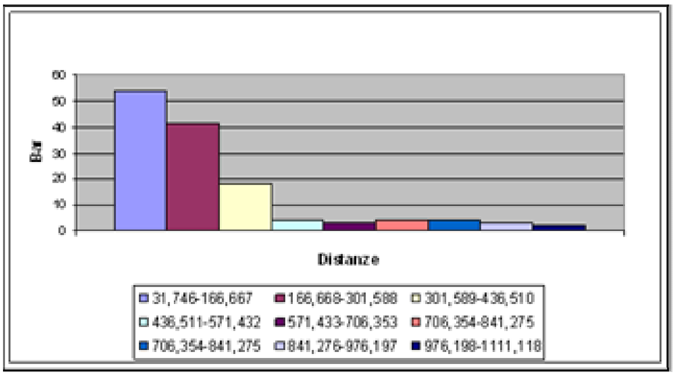

Fig. 5: Functional relationship between number of bars and number of "productive" activities on the territory.

At this point, given the positive results of the analysis previously done on the single variables (number of activities and the distance of the bars from the zone center), it is rational to think of the existence of a relation which ties together the explicative variables, making it able to calculate the variables that interest us: the number of bars. The analytic instrument that lets us calculate such variable is the linear multiply regression model, where the term "regression" indicates the dependence of a variable "dependent" to another variable "independent" [14].
The linear multiply regression model with p independent variables assumes the following expression:

$$Y_i = \beta_0 + \beta_1 X_{1i} + \beta_2 X_{2i} + \ldots\ldots + \beta_p X_{pi} + \varepsilon_i \quad (20)$$

where:

- $\beta_0 =$ intercept
- $\beta_1 =$ angle of Y respect the variable X_1 maintaining all the other variable constant ($X_2, X_3, \ldots, X_p$);
- $\beta_2 =$ angle of Y respect the variable X_2 maintaining all the other variable constant ($X_1, X_3, \ldots, X_p$);
- $\beta_p =$ angle of Y respect the variable X_p maintaining all the other variable constant ($X_1, X_2, X_3, \ldots, X_{p-1}$);
- $\varepsilon_i =$ error that corresponds to the i-th observation .

To reduce as much as possible the errors, so to express at the best the dependency between the Y and the X's, we must determine the values of the constants β_0, β_1, $\beta_2, \ldots, \beta_p$ so to minimize the squared sum of the residuals. This is obtained by using the minimum squared method, which in the case of a simple linear regression is expressed in this way:

$$\sum_{i=1}^{n} e_i^2 = \sum_{i=1}^{n} (y_i - \hat{y}_i)^2 \quad (21)$$

Where:

- y_i is the observed value for the dependent value x_i;
- $\hat{y}_i$ is the corresponding point to the line which better adapts to the equation $\hat{y}_i = a + bx_i$;
- ε_i is the residual

If all the errors, and therefore the residuals, would be nil, all the points would be placed on the regression line.
The independent variables considered to estimate the expected value of the dependent variable, are:

- X_1 = number of the activities;
- X_2 = distance of the bars from the center of each zone;
- X_3 = dummy variable, which value assumes 1 in the heart of the zone and in ring 1 and 0 in the 2 ring;
- Y= number of bars.

Once built the model, we must evaluate the goodness of such model, by using the opportune estimators. The multiply determination coefficients r^2 and r^2 corrected, which assume the respective values of 0.651 and 0.611, demonstrate that the model adapts well to the data that has been observed.
The results relative to the statistic test F indicate the "significance" between the dependent variables and the explicated variables. It's value which is 16,194 results to be superior to the value corresponding to the 95 percent, indicated as the threshold for acceptance of the statistic test. In such a case the nil hypothesis is rejected and therefore the significance of the statistic model is accepted. Verified the validity of the model, we have then made a partial F test so to determine which explicated variables should be inserted in the model because really useful for predicting Y. The first variable that has been tested was "distance of the bars from the center of the related zone": the results of the test, using the statistic T of Student, demonstrated that such a variable is irrelevant on the model respect other variables ($t_{Xi}.>t_{\square}$). Therefore such a result brought to this variables abrogation. The new model, now being constituted by only two independent variables, has ulteriorly verified: the time we investigated on the dummy variable and even this time the variable elimination has registered an improvement of the value of F. Therefore, the models expression has become:

$$Y_{bar,i} = \beta_0 + \beta_1 X_{1i} + \varepsilon_i \quad (22)$$

with X_{1i} being the number of activities present in i.
The value of F, has increased considerably passing to a 16,194, assumed with the regression model of three independent variables to the value 45,695 in the new model having only one independent variable.
The linear regression model determines the number of bars that each zone can have in function of the direct proportionality with other activities classified as "entertainment". It is necessary to correct the number of activities by introducing a model that considers the effective behavior on choosing a determine bar instead of another. It is proper to ask "which characteristics must a bar have so that an eventual client will pick it instead of another bar?", for surely the distance that a client must cover getting to the bar, the same for the competition in services offered by different bars. The model that simulates such a behavior can use information such as "gravitational/potential flow" which indicates the number of potential clients that a bar can have, and also the probability of a client choosing a bar instead of another (Logit model) function of the following attributes:
$X_1 = 1/d_{int}$; this attribute considers the interference between adjacent zones, it is calculated as the inverse of the distance d_{int}, calculated as the mean of the distance between the center of the zone interested and all the other centers of the adjacent zones.

$X_2 = d_m/d_{max}$; this attribute is given as the fraction between the average distance between the center of the entertainment activities and the bar considered and the maximum distance from the center to the perimeter of the zone.
In this case the Logit model assumes the connotations seen in the previous paragraphs, where:
- j= the choosing alternative: 29 traffic zones that we've also considered even in the regression model;
- V_j= systematic utility, which in our case becomes:

$$V = \frac{1}{d_{\text{int}}}\beta_1 + \frac{d_m}{d_{\max}}\beta_2 \quad (23)$$

For determining the two β values we have used the values related to the density of bars in each zone

Table 3: Homogenizing coefficients

β_1	β_2
$9 \cdot 10^{-3}$	$-3 \cdot 10^{-3}$

Knowing the homogenizing coefficients, we have calculated the probability in each traffic zone. In other words the implemented model reproduces a clients choosing behavior, which has moved so to satisfy a need such as: work, entertainment, study, etc…. These same people while moving decide if to stop in a bar for consuming or some other service offered by the bar.

$$p(bar\_choice) = y_{bar} \cdot p\left(\frac{bar\_choice}{\#bar}\right) \quad (24)$$

$$\#\,pers./every\_bar\_zone\_i = \frac{f_i - f_{out}}{\#bar} \cdot p(bar\_choice) \quad (25)$$

The function (3.28) has been calibrated using the experimental data deriving from the local territory of Caserta, determining the average number of clients in the local bars. The reproducing capability of the model also it to be used successively during the provisional phase, where we would determine the probability of success that a newly built kiosk would have. In other words it is necessary to decide in which zones should a kiosk be installed. The pick has fallen on the zones: 1,2, 8 in downtown; 11,13, 15, 18, 21 in the I ring area, and the only area 34 for the outskirts of the town. In particular in the zones 1 and 8 it is counseled to put two selling points, while in all the other zones the model gives as sufficient one retail point.
At this point we have conducted an analysis regarding the topologic structure of each zone that has been chosen as a retail point, so to be sure that each zone could really have the opening of a new kiosk, from this research has emerged that the zone 15 cannot have a kiosk because of the total absence of space. The retail points that have been chosen are reported in the figure 3.8 standing below, a kiosk is evidenced with a yellow point [15, 16].

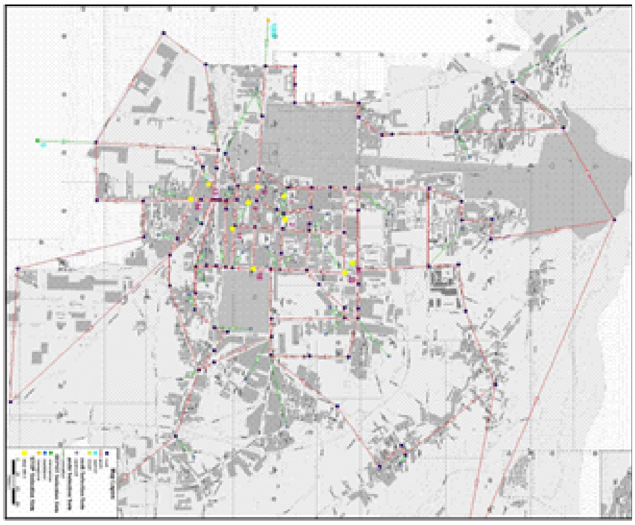

Fig. 6: Placement of the kiosks

2.3 Locating the production plant

Once localized the kiosks, we must decide where are the best points for the production plant, where for the term "best" must be considered under a time and economic view. The pick will be function of the minimum time that the fleet of vehicles consume for furnishing all the kiosks. The choice must be between the only two options which the local territory offers, which are:

1. Marcianise;
2. Casagiove.

For optimizing the route of a fleet of vehicles that have a series of constraints, known as *Vehicle Routing Problems* (*VRPs*) [2, 10, 11, 12, 17], the core of such a problem is determining the route on which are located all the kiosk that must be served, so to minimize the cost of routing and the number of vehicles assigned for completing the task.
One of the variants offered by the VRP, and which adapt very well to time constraints, is the *Vehicle Routing Problem with Time Windows* simply indicated with *VRPTW*. This method extends the VRP problem by adding a constraint: all clients must be served in a certain time slots. More precisely, the VRPTW uses a fleet of vehicles, which all departure from a deposit and must reach a certain number of clients dislocated in different geographic positions, each one having specific time slots that must be respected.
Even though introducing the time constraints makes the number of possible solutions more limited, such an extension of the problem makes any possible solution more realistic, rather then restricting ourselves to a simple, but non realistic problem that regards only the minimizing of the route of our fleet. In fact, other then the last mentioned or the minimizing of the number of vehicles that compose the fleet, we can also minimize the total journey time or even the complete cost of routing and scheduling, which consist in a series of fixed costs regarding the use of vehicles and variable costs, given by the time necessary to reach determined retail points or/and the time necessary for loading and unloading the vehicles [6, 18].

The Geographic Information System of the program TransCAD has an internal module of Routing and Logistics for the resolution of similar problems using an approach of operations research. The Vehicle Routing of TransCAD is able to determine a set of routes so to visit all the clients (represented by the nodes of the graph), from one ore more deposits, so to minimize the global costs or the journey time.

To find a realistic solution we had to put a series of data regarding the kiosks exact location, the quantities that must by delivered at each kiosk and the kind of vehicles that can by used.

For this reason we have positioned the 11 kiosks on the network by building a file .map overlapping the network of the total area of interest.

In the core macro-zone, the 5 kiosks have been placed as following:

- Traffic zone 1= kiosk number 1 and kiosk number 2 both in the center named 1;
- Traffic zone 2= kiosk number 3 in the center named 2
- Traffic zone 8= kiosk number 4 in the center named 8; kiosk number 5 in the node 8 on the boundaries with zone 9 named 9

In the I Ring the 4 kiosks have been placed in this way:

- Traffic zone 11= kiosk number 6 in the center named 11;
- Traffic zone 13= kiosk number 7 in the center named 13;
- Traffic zone 18= kiosk number 8 in the node 13 named 18;
- Traffic zone 21= kiosk number 9 in the node 21 named 22 and the kiosk number 10 in the center named 21;

And, at last, the II Ring the only kiosk has been placed:

- Traffic zone 34= kiosk number 11 in the node 75 named 34;

Analogous procedure has been used to determine the location of the production plant: we created a file .map overlapping it with the network created at first with all the nodes positioned on the network, representing the plant, corresponding with outer centers.

The first location hypothesis sees the deposit in the outer center 151 name 0, the second is in the center 150 named 1000.

Built the file .map we have determined the quantity that must be delivered to each kiosk. It's been necessary, for this reason, to realize a series of interviews in some local bars of the city, during which we noted the quality of the products offered during the day, and the quantity of products that are sold in the two periods of time of our interest.

The data that has been collected has then been divided in macro-zones, while the results are indicated in the following table.

Table 4:Average consumption of croissants

Macro-zone	Average consumption in h_1	Average consumption in h_2	Average day consumption
Center zone	90	50	300
I Ring	55	18	150
II Ring	40	15	100

The data related to the time necessary for delivering has been estimated in:

- Fixed time= 4 minutes for the central zone and 2 minutes for the outskirts. The difference in the values of time depends on the difficultly in parking in the city.
- Unit time= we have hypothesized that the load would be divided in basic load units BLU, organized so the each delivery would be in multiplies of BLU and each BLU would need one minute for unloading.

So to be able to individuate the best solution, in terms of cost and time, we have supposed that it could be utile to investigate on more alternatives, where in each alternative the different location of the production plant is sided by a series of different time slots for delivering the products. Supporting the previous time slots that have been fixed, we added other two time slots so the time slots considered were:

1. $h_1=8\div 9$
2. $h_2=18\div 19$
3. $h_3=11\div 12$
4. $h_4=15\div 16$

The results, considering all the hypotheses, are schematically indicated underneath.

First hypothesis:

- production plant located at Marcianise;
- 1 vehicle with a load capacity of 2350 croissants;
- delivery once a day made between 8:00 and 9:00: the vehicle must arrive at the first kiosk not before 8:00 and at the last kiosk not after 9:00.

Analyzing the results obtained and helping ourselves with the following figures, we have found that some active stops (6) indicated in yellow represent kiosks which the vehicle, respecting all the constraints, is able to supply; their names are indicated in the simulators report, but there are some kiosks which the vehicle is not capable of supplying (5, these nodes are also called Orphans) and these are evidenced in blue.

In these cases it seem to be evident that, if this solution would be adopted, that the five orphans would not be opened because of the location of the production plant and because of the time constraints these kiosks wouldn't be supplied in time.

Second hypothesis:

- production plant located at Casagiove;
- 1 vehicle with a load capacity of 2350 croissants;
- delivery once a day made between 8:00 and 9:00: the vehicle must arrive at the first kiosk not before 8:00 and at the last kiosk not after 9:00.

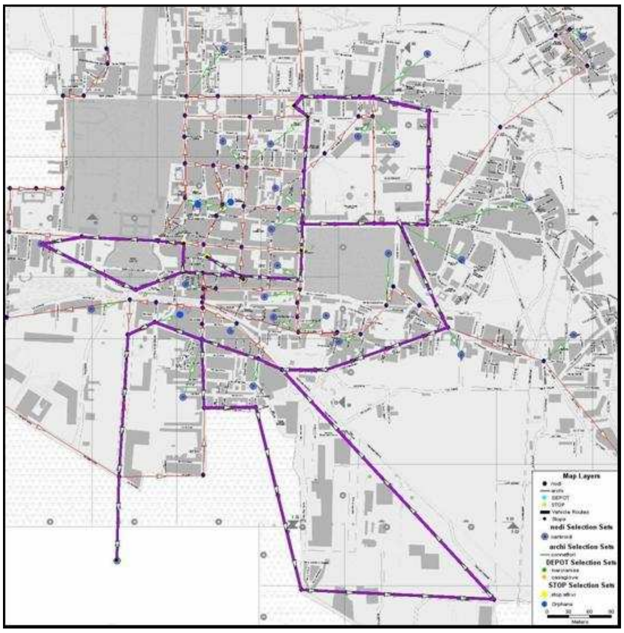

Fig. 7 Hypothesis 1 - Marcianise

Even in this second case (right side of the figure 3.9) there are some orphans, but compared to the first case they have them decreased passing from 5 to 4; and therefore using such alternative would bring to the opening of 7 kiosks on the local territory.

It is evident that having to choose one of the two alternatives, fully respecting all the constraints, we would choose to install the production plant at Casagiove, having that such solution has more kiosks openings, all of which can perfectly be supplied in time.

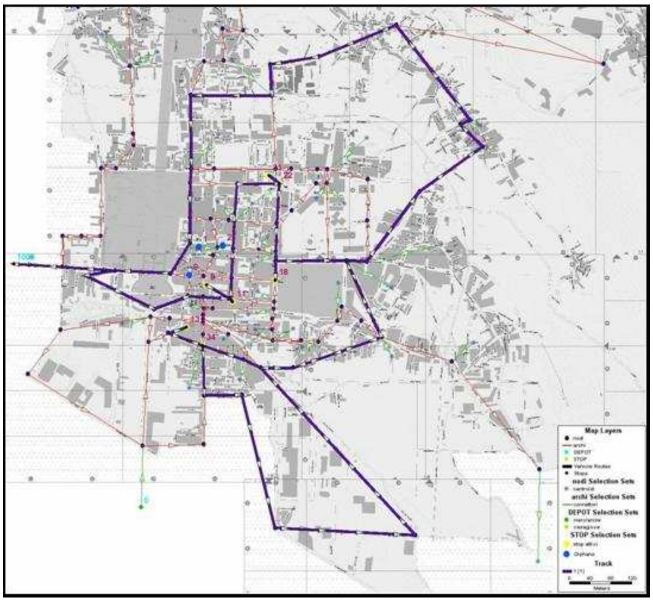

Fig. 8: Hypothesis 2 - Casagiove

The surveys haves been conducted interactively and all results have been proposed in the next table:

Table 5: Hypothesis 1 - Marcianise

Marcianise	Total time (min)	Total distance (m)	First vehicle load	Second vehicle load
h1	269	71096,9	1200	1150
h2	259	69017,2	1200	1150
h3	249	66175,9	1200	1150
h4	267	73803,2	1200	1150

Table 6: Hypothesis 2 - Casagiove

Casagiove	Total time (min)	Total distance (m)	First vehicle load	Second vehicle load
h1	193	46402,9	1200	1150
h2	190	41160,6	1500	850
h3	190	41160,6	1500	850
h4	191	47204,5	1500	850

As we can see, the analysis conducted bring to the conclusion that the best production location is Casagiove because it offers a minor distance that can be transited in minor time.

3. Results analysis and future developments

Each of the model presented in the last paragraphs are characterized by a mutual functional link. The model that schematizes the urban transportation system has the capacity of offering a picture of the traffics flow situation (as a mono-modal network). We have studied the behavior in local transportation on the behaves of students, workers and others (complementary to the first two categories). The load of the network gives somehow an indirect measurement of the attractiveness of certain activities and/or trends, and therefore it constitutes as the input for model that determines all the possible retailing points. The results of this last scheme joined together with the networking model are the platform on which the VRP (Vehicle Routing Problem) has been launched for determining the optimal position for the production system, minimizing, as always, time and delivery costs' (route in our specific case). The transiting time, necessary to reach a retailing point are negatively influenced by the networks load.

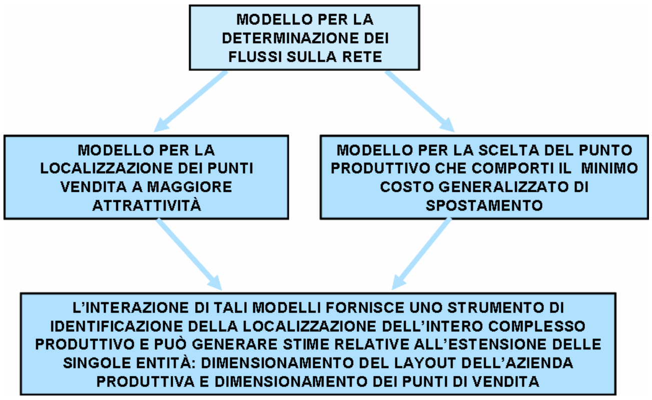

Fig. 9: Logic structure of the models implied.

The proposed structure can usefully by mounted as one model, able to manage all interactions between the sub-models. We'll have in this way, as input all the data regarding socio-economic, territory and infrastructure factors of the system under analysis (manufacturing systems, transportation systems, etc…) but also some parameters regarding the functioning state of the sub-models, which be inserted by the user by a simplified interface. The integrated models output will be, as already declared previously, locating the optimal points for the production and distribution systems, aspects that are independent to the nature of the product considered. The potential of such a model are enormous, in fact the uses of such are truly many. A first immediate use, is related to all cases where the production location isn't unknown: in this case the integrated model chooses the location for each kiosk, eliminating from the set of possible locations all the kiosk that for optimizing reasons, such as time and resources, cannot be supplied while respecting the constraints.
This model becomes a efficient instrument capable of generating evaluations on the dimension that the production plant should have so to satisfy all kiosks without over-dimensioning and also the dimensions that each kiosk should have in harmony with the number of clients that a kiosks has each day. To realize this aspect on the model we will need to make an ulterior integration on the model by interacting with another model that considering all the potential bars, potential flows in each zone and knowing the location of the production plant is able to estimate the consumption of croissant for each single kiosk.

Bibliography

[1] AA.VV. *Transcad – User Manual* Caliper Corporation Newton MA USA
[2] Baptista S., Oliveira R. C.,Zuquete E., *A period vehicle routing case study*, European Journal of Operational Research 139, (2002), pp. 220–229.
[3] Ben-Akiva M.E., 1974. *Structure of passenger travel demand models* Transportation Research Record 526, 26±42.
[4] Cascetta E. *Metodi quantitativi per la pianificazione dei sistemi di trasporto* Padova, CEDAM-casa editrice dott. Antonio Dilani, 1990.
[5] Ferrarese M. *Metodi e Tecniche di Analisi della Domanda (Pax)* Università degli studi di Verona, Facoltà di Economia, Corso di laurea in Economia del Commercio Internazionale, Master in IP of SMEs Internationalization Processes of Small and Medium and Size Enterprises Vicenza 15-16 maggio 2006
[6] Gribkovskaia I., Halskau O., Laporte G, Vlcek M., *General solutions to the single vehicle routing problem with pickups and deliveries*, European Journal of Operational Research 180, (2007), pp. 568–584.
[7] Heragu S., Kusiak A. (1991) *Efficient models for the facility layout problems*. Eur J Oper Res Soc 45:59 – 69
[8] Hensher D.A., 1998. *Understanding travel behaviour: some appealing research directions* Invitational Paper Presented at The University of Oxford RAC/TSU Seminar Series for Future Transport Research. Hertford College, Oxford, 1 July 1998.
[9] Howen S. H., Daskin M. S., *Strategic facility location: A review*, Eurpean Journal of Operational Research 111, (1998), pp. 423-447
[10] Jacobs-Blecha C., Goetschalckx M., *The Vehicle Routing Problem with Backhauls: Properties and Solution Algorithms*, Technical Report 1992-1998, Georgia Tech Research Corporation.
[11] Labbé M., Laporte G., Martín I.R. and González J.J.S., *Locating median cycles in networks* European Journal of Operational Research 160(2) (2005) 457–470.
[12] Laporte G., *Location-routing problems* In Vehicle Routing: Methods and Studies, B.L. Golden and A.A. Assad (eds), North-Holland (1988) 163-197.
[13] Melkote S., Daskin M. S., *Capacitated facility location/network design problems,* European Journal of Operational Research 129, (2001), pp. 481-495.
[14] Orsi R., *Probabilità ed inferenza statistica.* Edizioni Il Mulini 1995.
[15] ReVelle C. S., Eiselt H. A., *Location analysis: A synthesis and survey*, Eurpean Journal of Operational Research 165, (2005) pp.1-19
[16] SahinG., Süral H., *A review of hierarchical facility location models*, Computers & Operations Research 34, (2007), pp. 2310 – 2331
[17] Vigo D. *Il Vehicle Routing Problem: modelli ed algoritmi* DEIS Bologna
[18] Vogt L., Poojari C.A., Beasley J.E. *A Tabu Search Algorithm for the Single Vehicle Routing Allocation Problem* 2005.