Article

Using Genetic Algorithms for Real Estate Appraisals

Vincenzo Del Giudice [1], Pierfrancesco De Paola [1,*] and Fabiana Forte [2]

[1] Department of Industrial Engineering, University of Naples "Federico II", 80125 Naples, Italy; vincenzo.delgiudice@unina.it

[2] Department of Architecture and Industrial Design, University of Campania "Luigi Vanvitelli", 81031 Aversa, Italy; fabiana.forte@unina2.it

* Correspondence: pierfrancesco.depaola@unina.it; Tel.: +39-320-66-11-888

Abstract: The main aim of this paper is the interpretation of the existing relationship between real estate rental prices and geographical location of housing units in a central urban area of Naples (Santa Lucia and Riviera of Chiaia neighborhoods). Genetic algorithms (GA) are used for this purpose. Also, to verify the reliability of genetic algorithms for real estate appraisals and, at the same time, to show the forecasting potentialities of these techniques in the analysis of housing markets, a multiple regression analysis (MRA) was applied comparing results of GA and MRA.

---

1. Introduction

The real estate sector requests accurate appraisals for property values, this is because various stakeholders pursue several objectives as keep under control real estate values or update their profitability returns [1]. For these reasons, the interpretation of the housing market becomes a critical issue together to its evolution and dynamism, and at the same time appropriate techniques in able to adequately analyze the characteristics of real estate market are strongly needed [2].

Land and building are the two main components that are involved in the building production process. They are economic factors that are distinguishable even when the urban and building transformation process are completed. The scarcity of buildings in an urban area is then closely related to the available lands for constructions. This is a technical constraint that mainly depends by the geographical location of lands and by the urban planning policies, the latter are aspects that cannot be overcome by intrinsic factors of the real estate market.

Urban land is a limited factor, however it has an unlimited duration over the time. Due to these its features, the urban land is subject to a specific and intrinsic income. In fact, geographical location shapes the real estate markets and makes them sensitive to economic and social variations. Therefore, the nature of each urban site is unique, and this uniqueness strongly conditions the value and usefulness of the same urban land.

This said, any real estate market analysis cannot be conducted without a necessary identification of sub-markets, or areas where the mechanisms of price formation and the factors that affect real estate demands and supplies are homogeneous [3–17].

There is no doubt that more accurate is the analysis of real estate market if it is performed for each single submarket. To identify homogeneous real estate submarkets in an urban area, a key parameter is geographical location.

2. Targets and Research Design

Genetic algorithms (GA) have wide potential in the real estate field and present many advantages, especially from a practical point of view, with great benefits for real estate market actors, being able to highlight the variation of rental values for several sub areas belonging to the same market segment, even in small local real estate markets.

Moreover, the use of GA allows to predict, quantify, and locate where and how the rental values vary in a specific urban context, with a possibility to correlate these variations with any phenomenon or economic effect (i.e., modeling of locational variables, delimitation of areas with homogeneous values, etc.).

For this reason, GA were used in this work with the aim of identifying the effect on real estate rental values derived by geographical location of housing units. Again, in order to verify the reliability of GA for real estate appraisals and to show the forecasting potentialities of GA in the analysis of housing markets, a multiple regression analysis (MRA) was applied comparing, at the end of the work, results of GA and MRA applications.

MRA has been chosen for comparison purposes since it is a well-established methodology over time that provides robust and reliable results in real estate appraisals.

The analysis is carried out through the elaboration of rental prices of housing units located in a central urban area of Naples (Santa Lucia and Riviera of Chiaia neighborhoods).

Due to the opacity of Italian real estate markets, GA can aid to better interpret different segments of local real estate markets, or even help in the prediction and interpretation of the phenomena related to the genesis of rewards of position, with particular reference to problems of transformation and investments for urban areas affected from particular projects or plans, and in order to optimize the choices of use of goods and resources.

3. A Brief Literature Review

In the last several years, problems concerning real estate market analysis has grown up very fast, as this knowledge is very relevant for real estate predictions, investments, and taxation issues.

The most relevant difficulty in the forecast of real estate prices has always been the subjectivity of judgments to be processed for to obtain reliable real estate values [18]. For this reason, multiple regression analysis (MRA) has been considered for long time as the most flexible technique able to provide reliable predictions and information on real estate values and market analysis. As possible alternative, mostly in economic or financial fields, artificial neural networks (ANNs) have been tested by researchers for forecasting purposes [19–25].

In literature, however, the comparison between MRA and ANNs often presents controversial aspects and does not clearly show the superiority of one or the other method [26–29].

As alternative to above approaches genetic algorithms have been experimented over the time.

GA are stochastic search techniques that may explore wide and complex spaces based on Charles Darwin's evolutionary principle [30–33]. Multi-parameter optimization problems, characterized by objective functions subject to constraints, are the main operating areas suitable for GA. Initialization, selection, crossover, and mutation are the four stages in which the search process is developed by GA [34].

The history of GA began with the first computer simulations on evolution, the latter began to spread in the 1950s. In the 1960s, Rechenberg [35] introduced the "evolutionary strategy", and the topic was further expanded by Schwefel [36]. Instead, Fogel, Owen, and Walsh [37] developed "evolutionary programming", or a particular technique whose possible solutions were finite state machines. During those years, other experiences were carried on automatic learning and algorithms inspired to evolution and genetic [38].

However, only with Holland [31], and after with Goldberg [32], GA has become wide spread in many scientific fields, mainly thanks to the improvement of power of computational machines, making an ever wider practical application of those methods possible.

Holland's original idea, citing Mitchell [39], "was not to design algorithms to solve specific problems, but rather to formally study the phenomenon of adaptation as it occurs in nature and to develop ways in which the mechanisms of natural adaptation might be imported into computer systems".

Holland's GA, as presented in his book [31], is a technique useful to pass from an initial population of chromosomes to a new one, more fit to the environment, using a natural selection mechanism and the genetic operators of crossover, mutation, and inversion.

GA applications are very limited in real estate field. In international literature, Lertwachara [40] presents a GA method in identifying attractive stocks, for to measure the performance of profitable investment strategy; Ahn et al. [41] use ridge regression with GA to enhance real estate appraisal forecasting about Korean real estate market; Ma et al. [42] propose an approach given by the combination of hierarchical genetic algorithm and least squares method to optimize a particular neural network, with the aim to predict the real estate price index. In the Italian literature, only two works are findable on the use of GA in real estate appraisal: Manganelli et al. [43] test GA to interpret the relationship between real estate prices and property location, De Mare et al. [44] using GA verify the economic and territorial impacts caused by high speed railways.

Nowadays, the diffusion of GA is so large that a majority of big companies use them to solve a wide range of problems, like scheduling, data fitting, trend spotting and budgeting problems.

4. Characteristics of Genetic Algorithms

GA are adaptive techniques, particularly suitable in the resolution of optimization problems, able to seek—for a complex and prefixed fitness function—its maximum or minimum point.

In these circumstances, other traditional analytical techniques are not usable due to the impossibility of exploring the space of possible random solutions.

More generally, GA are adaptive heuristic search techniques based on the ideas of natural selection and genetic evolution. As already mentioned previously, at the basis of GA there are the concepts of simulation processes in natural systems necessary for evolution, the latter initially postulated by Charles Darwin with his theory on survival [39–45].

Then GA represent, in other terms, a clever utilization of a random search within a preset search space in order to solve a prearranged problem.

GA have been applied in a wide number of ways but their effective representation and selection of fitness function are the key parameters in their applications.

The relevance of GA in the applications is attributable to possibility of using them easily, even though they are, at the same time, reliable and adaptable algorithms able to search better solutions quickly in complex multidimensional problems. In particular, there are three specific conditions that, if verified, suggest the use in operational applications of GA:

- the space of solutions is complex, wide or its knowledge is very poor;
- the knowledge about a domain is poor or it is not possible to decode or restrict the space of solutions;
- traditional tools or techniques of mathematical analysis are unavailable.

In analogy with a biological system, it is possible to assume that also the built environment is characterized by a given issue to be solved and it is constituted by subjects of the population that are solvers for the problem to be addressed.

On the individuals of the population, the environment exerts a pressure that is continuously in evolution, pressure identifiable as a specific fitness function able to provide for each person a high accuracy degree about solutions, proposed by the same people, for the problem under investigation.

Derived from biology field, crossover and mutation are main genetic operators considered in study cases. Whereby, referring to the concepts about natural selection, the best individuals of a

given generation, if coupled together, will produce better subsequent and future generations of new individuals that can provide an entire population of possible solutions for the problem of interest.

In this way, there is a continuous improvement of genetic characteristics and in the number of individuals compared with previous generations. Thus, the better features are crossed and exchanged in the course of future generations. Exploring in detail the most suitable research areas, the optimal solution is identified by best individuals generated in the time.

Generally, five domains can be distinguished where GA are usefully applied, even if these domains do not have clearly defined boundaries [43]:

- Control;
- Design;
- Simulation and identification;
- Planning;
- Classification, modelling and machine learning.

For the latter domain, the use of GA is directed to the construction of models able to interpret an underlying phenomenon or for forecasting purposes [9].

5. Case Study

For the implementation of GA, the constitution of an initial population (random or derived by heuristic rules) is the first step; after, an evolutionary cycle has to be generated, and in correspondence of each iteration must be produced a new population using genetic operators to the previous population.

GA focuses on the identification of maximum or minimum point for a default fitness function. Therefore, the best solutions must be selected with subsequent iterations and, then, must recombine between them, so a continuous evolution towards the optimal state occurs. In particular, each individual generated from its previous population is assigned a specific value that depends on quality degree of solutions. Subsequently, more suitable individuals are recombined to produce new generations until there is convergence to the value attributable to the best individual.

The proposed model focuses on the interpretation of rental prices as the result of a multicriteria choice process.

The criteria of selection may be expressed by the characteristics of housing units, which are fundamental in the mechanism of formation of real estate values. For to determine the marginal price of geographic location, which contributes to define real estate rental prices, the original sample (initial population) has been defined with detection of real estate rental prices that refer to the recent past (last two months) and to housing units with similar building type and located in a central urban area of Naples (45 housing units located in Santa Lucia and Riviera of Chiaia neighborhoods).

About this particular issue (real estate sample size), Green [46] used statistical power analysis in order to compare the performance of several rules-of-thumb for how many subjects were required for MRA. Marks et al. [47] specifies a minimum of 200 subjects for any regression analysis. Tabachnick and Fidell [48] suggested that although 20 subjects per variable would be preferable, the minimum required subjects should be five. Harris [49] argued that the number of subjects should exceed the sum of 50 and the number of predictor variables. Schmidt [50] determined that the minimum number of subjects per variables lies in the range of 15 to 20. While Harrell [51] suggested that 10 subjects per variable was the minimum required sample size for linear regression models to ensure accurate prediction in subsequent subjects. Finally, Del Giudice [52] argued that the required sample size minimum is equal to 4–5 subjects per independent variable).

As reported above, also considering the opacity of Italian real estate markets as in the present case, the reduced real estate sample used in this work does not prejudice the operational usability of MRA or GA models.

Only non-homogeneous real estate characteristics, besides the geographic location, were detected for each sampled unit and in particular (see Table 1):

- real estate rental price (monthly) expressed in euro (RERP);
- commercial area expressed in square meters (AREA);
- maintenance status (MAIN) expressed with a scores scale: 2 if the housing unit is in optimum conditions, 1 if maintenance status is good, 0 otherwise (mediocre status);
- number of floor level of housing unit (FLOOR).

Table 1. Statistical description of real estate data.

Variable	Std. Dev.	Median	Mean	Min	Max
RERP	6.000	963.51	1.692	400	6.000
AREA	71.29	100	122.311	30	460
MAIN	0.63	2	1.49	0	2
FLOOR	1.55	2	2.44	0	7

Other real estate characteristics (as panoramic views, noise, etc.), manifesting with the same modalities in all the sampled units were excluded from the analysis.

The geographical position (GEO) was expressed assigning a prefixed sequence of characters to each housing unit, so as to classify every housing unit as falling to a defined area among those already identified in a study focused on the city of Naples [4]. The cited study used a geoadditive model based on penalized spline functions for the housing market segmentation and, specifically, by this work it is possible to subdivide the urban area considered in five subareas with homogeneous values as Figure 1 shows (A, B, C, D, E). More precisely, mentioned sub-areas were so defined (in terms of monthly rental price):

- Area A: mean rental price up to € 1,700.00 (sequence: 1,0,0,0,0);
- Area B: mean rental price from € 1,700.00 to € 1,750.00 (sequence: 0,1,0,0,0);
- Area C: mean rental price from € 1,750.00 to € 1,800.00 (sequence: 0,0,1,0,0);
- Area D: mean rental price from € 1,800.00 to € 1,850.00 (sequence: 0,0,0,1,0);
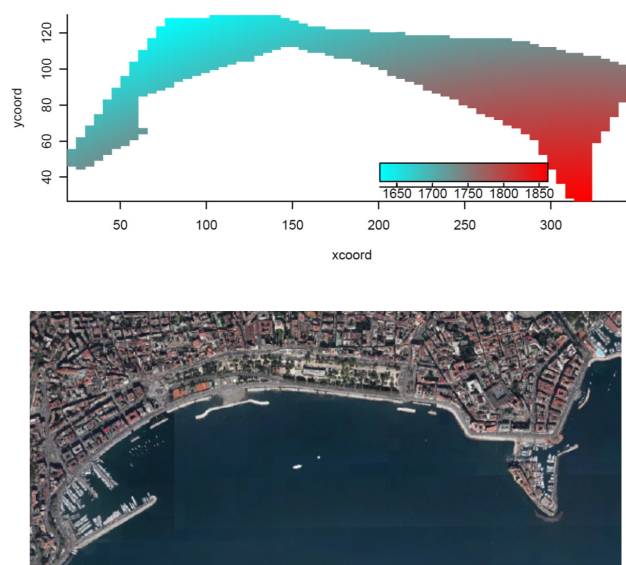- Area E: mean rental price over € 1,850.00 (sequence: 0,0,0,0,1).

Figure 1. Spatial distribution of real estate rental values [4] compared with aerial view of urban area derived from Google Maps.

Obviously, residential units that falling in the mentioned five subzones (A, B, C, D, E) are included into a homogeneous urban area in terms of services and infrastructure qualification, in fact, the differences in value for the different sub-areas are small. Sub-areas with the highest values are characterized by historical and touristic amenities nearby, while sub-areas with the lowest values are almost exclusively residential.

Residual sum of squares (*RSS*) is the fitness function adopted for the proposed model to obtain residuals as the difference between sampled real estate rental prices and fitted values provided by the model [43]:

$$RSS = \sum_i \varepsilon_i^2 = \sum_i (y_i - \bar{y}_i)^2 = \sum_i (y_i - b_0 - b_i x_i)^2, \; for \; i = 1, \, 2, \, \dots, \, n. \tag{1}$$

where y_i are the observed rental prices, $\bar{y}_i$ are the fitted values, x_i is the measure of *i-th* real estate characteristic selected, b_0 is the statistical error (or constant), and b_i is the coefficient of *i-th* real estate characteristic (marginal price with which the *i-th* real estate characteristic contributes to the rent).

"Genetic algorithms" tool was used in MATLAB software for to process the proposed model.

In MATLAB software, numerous iterations were performed, changing in several iterative cycles, one by one, constraints and model key parameters, directing the process to the minimization of the sum of squared residuals. Data processing has allowed determination of genetic code able to generate the optimal solution:

- Solver: GA—Genetic Algorithm;
- Number of variables: 9;
- Population type: double vector;
- Population size: 45;
- Mutation: constraint dependent;
- Mutation: proportional scaling fitness function;
- Selection: selection function as stochastic uniform;
- Reproduction: elite count equal to 1, crossover fraction equal to 0.1;
- Crossover: heuristic function;
- Migration: both;
- Evaluate fitness and constraint functions: in serial;
- Stopping criteria: maximum number of generations equal to 5,000;
- Plot function: best fitness, stopping.

Table 2 shows the values of b_i coefficients.

Table 2. Results of GA application.

Genetic Algoritms		
$b_0 = -20.74$ Characteristics/Area		b_{ij}
1	AREA	13.21
2	MAIN	10.67
3	FLOOR	0.90
4	A	47.45
5	B	23.53
6	C	136.69
7	D	−34.47
8	E	−37.30

The vector of coefficients directly expresses the marginal rental prices for each single real estate characteristic. The rental price function is additive as follows (K_j are the marginal contributions of real estate characteristics):

$$y_i = f(K_j) = K_0 + \sum_{j=1}^{n} K_j = b_0 + \sum_{j=1}^{n} b_{ij} x_{ij} \tag{2}$$

The additive linear form with intercept was chosen because, in this work, the aim is to identify the marginal contribution on real estate rental values derived by geographical location of housing units. Otherwise, without the intercept, the analysis would provide average prices and not the marginal prices of real estate characteristics (which, as noted, are different between them).

Expected values are calculated with Equation (2) and, then, also the error between fitted values and observed rental prices.
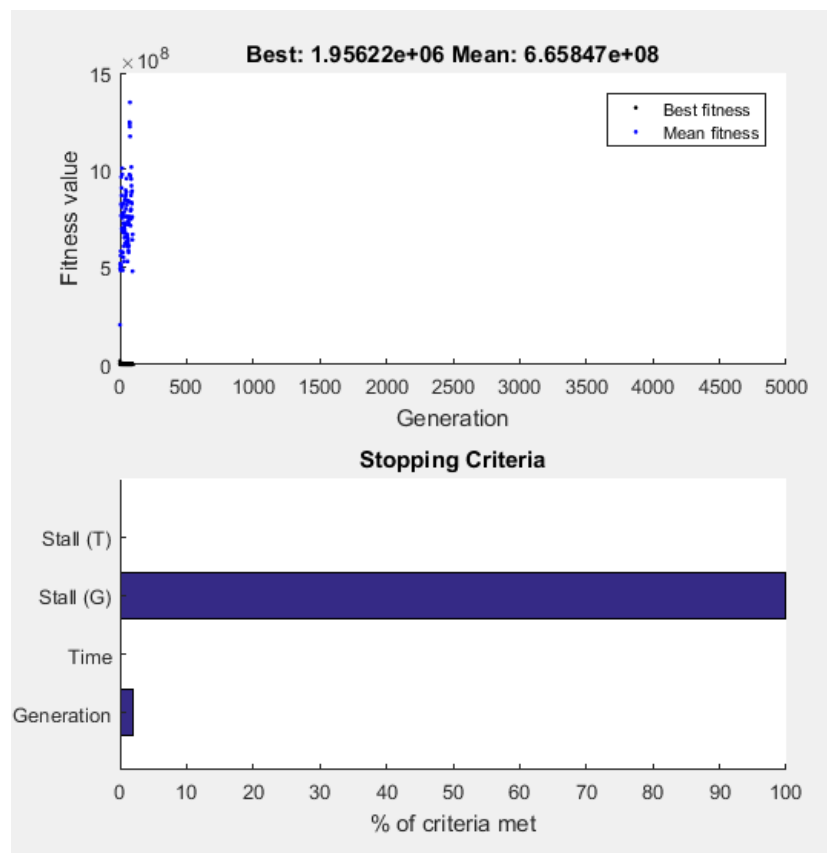
Figure 2 and Tables 2 and 3 show model results.

Figure 2. Fitness function value and stopping criteria generate from MATLAB software.

Table 3. Comparison between GA and MRA.

Description	GA	MRA
Max overestimation (%)	25.67%	21.23%
Max underestimation (%)	24.04%	48.52%
Absolute average percentage error (%)	10.62%	11.50%
Number of overestimation in the sample (>15%)	6	4
Number of underestimation in the sample (<−15%)	4	8

The optimal run has carried out 95 iterations (on a maximum of 5,000 prefixed iterations) and the best objective function value was equal to 1.956×10^6 (see Figure 2).

For variable "AREA", the marginal price is €/sqm/month 13.21; for variable "MAIN", the marginal price is €/month 10.67 for each additive score point; for variable "FLOOR", the marginal price is €/month 0.90 for each additional floor level; for geographical location (GEO), the corresponding marginal prices are differentiate respect to five areas previously selected and, precisely: €/month 47.45 if housing unit is in area "A", €/month 23.53 if housing unit is in area "B", €/month 136.69 if housing unit is in area "C", €/month 34.47 (negative sign) if housing unit is in area "D", €/month 37.30 (negative sign) if housing unit is in area "E".

From these results, some considerations can be expressed:

1. Variables "AREA", "MAIN" and "FLOOR" are all positive, showing an increase in the real estate rental prices in correspondence of increasing amount of the respective real estate characteristics; this circumstance is in line with the normal dynamics of housing market;

2. Housing units that fall in zones A and B are those who, according to the study previously conducted on the segmentation of real estate market for the urban area considered (which, remember, was done with a geoadditive model), have lower average prices; this fact is indirectly confirmed from data processing carried out with GA (housing units in zones A and B have the lower marginal prices);

3. Housing units that fall in zone C are those who should have intermediate prices between the five different areas considered, but this circumstance is not detected in GA implementation; on the contrary, housing units in zone C have the higher marginal price;

4. Housing units that fall in zones D and E are those who should have higher prices between the five different areas considered, but instead the GA implementation shows that housing units in zones D and E have negative marginal prices;

5. The floor level variable (FLOOR) has an impact almost irrelevant on real estate rental prices (marginal price equal to €/month 0.90 for each additional floor level); in fact, compared to the average real estate rental price, this variable is about 0.13% in its average amount.

From an exclusively numerical point of view, some of above considerations (see Points 3 and 4) may seem controversial, but can be explained considering that—even falling in a narrow geographic area—in the D and E zones there is the presence of a great castle (so-called "Castel dell'Ovo"), many luxury hotels and a pedestrian area. With the same environmental characteristics, the housing supply in areas D and E is much more scarce and, consequently, real estate rental prices tend to be higher, although housing units have less fruition ability for parking difficulties, vehicle traffic due to the great presence of tourists, and more distance from urban services mainly located in zone C. The latter conditions, typical of the surrounding environment, tend to justify and clarify because GA implementation provides negative marginal prices for D and E areas. It seems, therefore, that the GA has a greater ability to interpret current and future behavior of potential users of residential units, comparing GA with geoadditive models.

With the purpose of showing the forecasting potentialities of GA techniques in the analysis of housing markets, the results of GA implementation are also compared with those derivable from the application of MRA (with intercept) on the same real estate sample. From a general and theoretical point of view, GA and MRA have the same potentialities and application fields, and they are suitable to be used with the same assumptions and basic principles (known and established for MRA, then this facilitates the experimentation of GA); for these reasons, the comparison between the two techniques can be performed both in terms of residue analysis both comparing the marginal prices of real estate characteristics.

Firstly, comparison between GA and MRA highlights that marginal prices, for each single real estate characteristic, are different to each other. Output data of MRA show a determination index (R^2) equal to 0.953, a determination index corrected ($R^2{}_C$) equal to 0.950, a standard error equal to 214.53, and a positive *F-test* (over 95% significance).

Secondly, although there is a substantial similarity between GA and MRA in terms of results, GA show their superiority in order to interpret the real estate rental market (10.62% for GA versus 11.50% for MRA in terms of absolute average percentage error, equal to a better prediction in favor of GA in measure of 7.65% about), as also it is detectable from the low value of statistical error ($b_0 = -20.74$, respect a intercept of MRA equal to -36.89).

Lastly, the hedonic function obtained with GA implementation, in order to process a direct appraisal of real estate rental prices in Santa Lucia and Riviera of Chiaia neighborhoods, may synthetically be defined as follows (see Table 2 for the amounts of coefficients):

$$RERP \ (monthly) = b_0 + b_{AREA} \cdot AREA + b_{MAIN} \cdot MAIN + b_{FLOOR} \cdot FLOOR + b_{(A,B,C,D,E)} \cdot GEO$$

where $b_{(A,B,C,D,E)}$ represents the marginal price of geographical location (GEO) in correspondence of each sub-areas specified (from A to E).

6. Conclusions

Hedonic models are widespread in urban studies, but for them statistical techniques as genetic algorithms, linear programming, semi-parametric or non-parametric regressions, artificial neural networks are certainly less common [3–13,28,29,53–65].

Currently, only a few works use GA for real estate appraisal purposes [40–44]. The current future research lines in real estate field about GA are the possible integrations with other techniques as artificial neural networks. In fact, ANNs and GA demonstrate powerful problem solving ability as they are based on quite simple principles, but take advantage of their mathematical nature: non-linear iteration. Integrations of GA and ANNs have a wide field of real-world applications (i.e., automotive design, engineering design, robotics, evolvable hardware, biomimetics invention [66]) but also suffer from several operational disadvantages (the choice of basic parameters as network topology, learning rate, initial population, etc. often already determine the success of data processing) and in the future it will be necessary try to work on these limitations.

Genetic algorithms were used in this case study with the aim to determine the marginal price with which the geographic location takes part in the formation of real estate rental prices. This problem has been addressed using a preliminary segmentation of a central urban area of Naples developed with a geoaddictive model on the same real estate data. With respect to this segmentation, genetic algorithms show some difference, in that they are able to better interpret the current and future behavior of potential users of housing units. Even if the results obtained by genetic algorithms are excellent (absolute average percentage error equal to 10.62%), they do not improve, significantly, similar performance obtainable with traditional parametric approaches (better forecasting results for GA in measure of 7.65% about respect to MRA).

The improvement obtainable with genetic algorithms is mostly due to use of non-linear functions. In this sense, the computing capabilities of genetic algorithms and the modelling ability of other and more complex prediction or fitness functions utilizable for appraisal purposes (in order to better consider the non-linearity of typical real estate phenomena), might certainly provide even better results in the future.

Author Contributions: This paper is to be attributed in equal parts to the authors.

Conflicts of Interest: The authors declare no conflict of interest.

References

1. Manganelli, B.; Morano, P. Estimating the market value of the building sites for homogeneous areas. *Adv. Mater. Res.* **2014**, *869–870*, 14–19.
2. Manganelli, B. *Real Estate Investing*; Springer: Cham, Switzerland, 2015.

3. Del Giudice, V.; De Paola, P. Undivided real estate shares: Appraisal and interactions with capital markets. *Appl. Mech. Mater. Trans. Tech. Pubbl.* **2014**, *584–586*, 2522–2527. [CrossRef]

4. Del Giudice, V.; De Paola, P. Spatial analysis of residential real estate rental market. In *Advances in Automated Valuation Modeling*; d'Amato, M., Kauko, T., Eds.; Studies in System, Decision and Control Series 86; Springer: Cham, Switzerland, 2017; pp. 155–162.

5. Del Giudice, V.; De Paola, P.; Forte, F. The appraisal of office towers in bilateral monopoly's market: Evidence from application of Newton's physical laws to the Directional Centre of Naples. *Int. J. Appl. Eng. Res.* **2016**, *11*, 9455–9459.

6. Del Giudice, V.; Evangelista, P.; De Paola, P.; Forte, F. Knowledge Management and Intellectual Capital in the Logistics Service Industry. In Proceedings of the International Conference on Knowledge Science, Engineering and Management (KSEM 2016), Passau, Germany, 5–7 October 2016.

7. Del Giudice, V.; Manganelli, B.; De Paola, P. *Depreciation Methods for Firm'S Assets, ICCSA 2016, Part III, Lecture Notes in Computer Science, 9788 (2016)*; Springer: Cham, Switzerland, 2016; pp. 214–227.

8. Manganelli, B.; Del Giudice, V.; De Paola, P. *Linear Programming in a Multi-Criteria Model for Real Estate Appraisal, ICCSA 2016, Part I, Lecture Notes in Computer Science, 9786 (2016)*; Springer: Cham, Switzerland, 2016; pp. 182–192.

9. Del Giudice, V.; De Paola, P.; Manganelli, B. *Spline Smoothing for Estimating Hedonic Housing Price Models, ICCSA 2015, Part III, Lecture Notes in Computer Science, 9157 (2015)*; Springer: Cham, Switzerland, 2015; pp. 210–219.

10. Del Giudice, V.; De Paola, P. Geoadditive Models for Property Market. *Appl. Mech. Mater. Trans. Tech. Pubbl.* **2014**, *584–586*, 2505–2509. [CrossRef]

11. Del Giudice, V.; De Paola, P. The effects of noise pollution produced by road traffic of Naples Beltway on residential real estate values. *Appl. Mech. Mater.* **2014**, *587–589*, 2176–2182. [CrossRef]

12. Del Giudice, V.; De Paola, P.; Manganelli, B.; Forte, F. The monetary valuation of environmental externalities through the analysis of real estate prices. *Sustainability* **2017**, *9*, 229. [CrossRef]

13. Del Giudice, V.; De Paola, P.; Cantisani, G.B. Rough Set Theory for real estate appraisals: An application to Directional District of Naples. *Buildings* **2017**, *7*, 12. [CrossRef]

14. Tajani, F.; Morano, P.; Locurcio, M.; Torre, C.M. Data-driven techniques for mass appraisals. Applications to the residential market of the city of Bari (Italy). *Int. J. Bus. Intell. Data Min.* **2016**, *11*. [CrossRef]

15. Morano, P.; Tajani, F.; Locurcio, M. GIS application and econometric analysis for the verification of the financial feasibility of roof-top wind turbines in the city of Bari (Italy). *Renew. Sust. Energ. Rev.* **2017**, *70*, 999–1010. [CrossRef]

16. Morano, P.; Tajani, F. Bare ownership of residential properties: Insights on two segments of the Italian market. *Int. J. Hous. Mark. Anal.* **2016**, *9*, 376–399. [CrossRef]

17. Goodman, A.C.; Thibodeau, T.G. Housing Market Segmentation. *J. Hous. Econ.* **1998**, *7*, 121–143. [CrossRef]

18. Wiltshaw, D.G. A comment on methodology and valuation. *J. Prop. Res.* **1995**, *12*, 157–161.

19. Ahn, J.J.; Lee, S.J.; Oh, K.J.; Kim, T.Y. Intelligent forecasting for financial time series subject to structural changes. *Intell. Data Anal.* **2009**, *13*, 151–163.

20. Chen, W.S.; Du, Y.K. Using neural networks and data mining techniques for the financial distress prediction model. *Expert Syst. Appl.* **2009**, *36*, 4075–4086. [CrossRef]

21. Dong, M.; Zhou, X.S. Knowledge discovery in corporate events by neural network rule extraction. *Appl. Intell.* **2008**, *29*, 129–137. [CrossRef]

22. Lee, K.; Booth, D.; Alam, P. A comparison of supervised and unsupervised neural networks in predicting bankruptcy of Korean firms. *Expert Syst. Appl.* **2005**, *29*, 1–16. [CrossRef]

23. Lu, C.J. Integrating independent component analysis-based denoising scheme with neural network for stock price prediction. *Expert Syst. Appl.* **2010**, *37*, 7056–7064. [CrossRef]

24. Oh, K.J.; Han, I. Using change-point detection to support artificial neural networks for interest rates forecasting. *Expert Syst. Appl.* **2000**, *19*, 105–115. [CrossRef]

25. Versace, M.; Bhatt, R.; Hinds, O.; Shiffer, M. Predicting the exchange traded fund DIA with a combination of genetic algorithms and neural networks. *Expert Syst. Appl.* **2004**, *27*, 417–425. [CrossRef]

26. Dehghan, S.; Sattari, G.; Chehreh, C.S.; Aliabadi, M.A. Prediction of uniaxial compressive strength and modulus of elasticity for Travertine samples using regression and artificial neural networks. *Min. Sci. Technol.* **2010**, *20*, 41–46. [CrossRef]

27. Hua, C. Residential construction demand forecasting using economic indicators: A comparative study of artificial neural networks and multiple regression. *Constr. Manag. Econ.* **1996**, *14*, 125–134. [CrossRef]

28. Nguyen, N.; Cripps, A. Predicting Housing Value: A Comparison of Multiple Regression Analysis and Artificial Neural Networks. *J. Real Estate Res.* **2001**, *22*, 313–336.

29. Worzala, E.; Lenk, M.; Silva, A. An Exploration of Neural Networks and Its Applicationto Real Estate Valuation. *J. Real Estate Res.* **1995**, *10*, 185–202.

30. Goldberg, D.E. *Genetic Algorithms in Search, Optimization and Machine Learning*; Addison-Wesley: Boston, MA, USA, 1989.

31. Holland, J.H. *Adaptation in Natural and Artificial Systems: An Introductory Analysis with Applications to Biology, Control and Artificial Intelligence*; The MIT Press: Cambridge, MA, USA, 1975.

32. Oh, K.J.; Kim, T.Y.; Min, S. Using genetic algorithm to support portfolio optimization for index fund management. *Expert Syst. Appl.* **2005**, *28*, 371–379. [CrossRef]

33. Koza, J. *Genetic Programming*; The MIT Press: Cambridge, MA, USA, 1993.

34. Wong, F.; Tan, C. Hybrid neural, genetic, and fuzzy systems. In *Trading on the Edge*; Deboeck, G.J., Ed.; Wiley: New York, NY, USA, 1994; pp. 243–261.

35. Rechenberg, I. *Evolutionsstrategie: Optimierung Technischer Systeme Nach Prinzipien der Biologischen Evolution*; Frommann Holzboog: Stuttgart, Germany, 1973.

36. Schwefel, H.P. *Numerical Optimization of Computer Models*; Wiley: New York, NY, USA, 1981.

37. Fogel, L.J.; Owens, A.J.; Walsh, M.J. *Artificial Intelligence through Simulated Evolution*; Wiley: New York, NY, USA, 1966.

38. Reed, P.; Minsker, B.S.; Goldberg, D.E. Designing a competent simple genetic algorithm for search and optimization. *Water Resour. Res.* **2000**, *3612*, 3757–3761. [CrossRef]

39. Mitchell, M. *An Introduction to Genetic Algorithms*, 5th ed.; MIT Press: Cambridge, MA, USA; London, UK, 1999.

40. Lertwachara, K. Selecting stocks using a Genetic Algorithm: A case of real estate investment trusts. *Int. J. Comput. Internet Manag.* **2007**, *15*, 20–31.

41. Ahn, J.J.; Byun, H.W.; Oh, K.J.; Kim, T.Y. Using ridge regression with genetic algorithm to enhance real estate appraisal forecasting. *Expert Syst. Appl.* **2012**, *39*, 8369–8379. [CrossRef]

42. Ma, H.; Chen, M.; Zhang, J. The prediction of real estate price index based on improved Neural Network Algorithm. *Adv. Sci. Technol. Lett.* **2015**, *81*, 10–15.

43. Manganelli, B.; De Mare, G.; Nesticò, A. Using Genetic Algorithms in the housing market analysis. In Proceedings of the International Conference on Computational Science and Its Applications (ICCSA 2015), Banff, AB, Canada, 22–25 June 2015.

44. De Mare, G.; Lenza, T.L.L.; Conte, R. Economic Evaluations using Genetic Algorithms to Determine the Territorial Impact Caused by High Speed Railways. *Int. J. Soc. Educ. Econ. Manage. Eng.* **2012**, *6*, 3313–3321.

45. Nga, T.; Skitmoreb, M.; Wongc, K.F. Using genetic algorithms and linear regression analysis for private housing demand forecast. *Build. Environ.* **2008**, *43*, 1171–1184.

46. Green, S.B. How many subjects does it take to do a regression analysis. *Multivar. Behav. Res.* **1991**, *26*, 499–510. [CrossRef] [PubMed]

47. Marks, M.R. Two kinds of regression weights that are better than betas in crossed samples. In Presented at the Meeting of the American Psychological Association, September 1966.

48. Tabachnick, B.G.; Fidell, L.S. *Using Multivariate Statistics*, 6th ed.; Pearson: Boston, MA, USA, 2012.

49. Harris, R.J.; Quade, D. The Minimally Important Difference Significant Criterion for Sample Size. *J. Educ. Behav. Stat.* **1992**, *17*, 27–49. [CrossRef]

50. Schmidt, F.L. Statistical significance testing and cumulative knowledge in psychology: Implications for the training of researchers. *Psychol. Methods* **1996**, *1*, 115–129. [CrossRef]

51. Harrell, F.E., Jr. *Regression Modeling Strategies*; Springer-Verlag: New York, NY, USA, 2001.

52. Del Giudice, V. *Estimo e Valutazione Economica dei Progetti*; Paolo Loffredo Editore: Napoli, Italy, 2015.

53. Narulaa, S.C.; Wellingtonb, J.F.; Lewisb, S.A. Valuating residential real estate using parametric programming. *Eur. J. Oper. Res.* **2012**, *217*, 120–128. [CrossRef]

54. Kontrimasa, V.; Verikasb, A. The mass appraisal of the real estate by computational intelligence. *Appl. Soft Comput.* **2011**, *11*, 443–448. [CrossRef]

55. Bourassa, S.C.; Cantoni, E.; Hoesli, M. Predicting house prices with spatial dependence: A comparison of alternative methods. *J. Real Estate Res.* **2010**, *32*, 139–159.

56. Brooks, C.; Tsolacos, S. International evidence on the predictability of returns to securitized real estate assets: Econometric models versus neural networks. *J. Prop. Res.* **2003**, *20*, 133–155. [CrossRef]

57. Chica-Olmo, J. Prediction of housing location price by a multivariate spatial method: Cokriging. *J. Real Estate Res.* **2007**, *29*, 92–114.

58. Juan, Y.K.; Shin, S.G.; Perng, Y.H. Decision support for housing customization: A hybrid approach using case-based reasoning and genetic algorithm. *Expert Syst. Appl.* **2006**, *31*, 83–93. [CrossRef]

59. Peterson, S.; Flanagan, A.B. Neural network hedonic pricing models in mass real estate appraisal. *J. Real Estate Res.* **2009**, *31*, 148–164.

60. Rossini, P. Artificial neural networks versus multiple regression in the valuation of residential property. *Aust. Land Econ. Rev.* **1999**, *3*, 1–12.

61. Wilson, I.D.; Paris, S.D.; Ware, J.A.; Jenkins, D.H. Residential property price time series forecasting with neural networks. *Knowl.-Based Syst.* **2002**, *15*, 335–341. [CrossRef]

62. Del Giudice, V.; Manganelli, B.; De Paola, P. Hedonic Analysis of Housing Sales Prices with Semiparametric Methods. *Int. J. Agric. Environ. Inf. Syst.* **2017**, *8*, 65–77. [CrossRef]

63. Del Giudice, V.; De Paola, P.; Cantisani, G.B. Valuation of Real Estate Investments through Fuzzy Logic. *Buildings* **2017**, *7*, 26. [CrossRef]

64. Simonotti, M.; Salvo, F.; Ciuna, M.; De Ruggiero, M. Measurements of rationality for a scientific approach to the market-oriented methods. *J. Real Estate Lit.* **2016**, *24*, 403–427.

65. Simonotti, M.; Salvo, F.; Ciuna, M. Multilevel methodology approach for the construction of real estate monthly index numbers. *J. Real Estate Lit.* **2014**, *22*, 281–302.

66. Mahajan, R.; Kaur, G. Neural networks using genetic alghoritms. *Int. J. Comput. Appl.* **2013**, *77*, 6–11.