Available online at www.sciencedirect.com
ScienceDirect
Building and Environment 43 (2008) 1171–1184
BUILDING AND ENVIRONMENT
www.elsevier.com/locate/buildenv

Using genetic algorithms and linear regression analysis for private housing demand forecast

S. Thomas Ng[a,*], Martin Skitmore[b], Keung Fai Wong[c]

[a]Department of Civil Engineering, The University of Hong Kong, Pokfulam Road, Hong Kong
[b]School of Urban Development, Queensland University of Technology, GPO Box 2434, Brisbane Q4001, Australia
[c]Department of Civil Engineering, The University of Hong Kong, Pokfulam Road, Hong Kong

Received 18 May 2005; received in revised form 12 November 2006; accepted 23 February 2007

Abstract

An accurate prediction of prospective construction supply and demand, especially the private residential market, is paramount important to policy makers, as it could help formulate strategies to cultivate/stabilize the economy and satisfy the social needs (at macro level). Despite that, a realistic prediction of future private residential demand is never an easy task, as it is governed by a number of social and economic factors. In this paper, four leading indicator models are developed and compared for directly forecasting Hong Kong private sector residential demand. These comprise a (i) Linear Regression Analysis (LRA) model, (ii) Genetic Algorithms (GA) model, (iii) GA-LRA model, where LRA is used to select the indicator variables; and (iv) GA-LRA model with Adaptive Mutation Rate (AMR) to reduce the likelihood of local optima. The findings indicate that the GA-LRA model with AMR provides the most accurate forecasts and over a longer time horizon. In providing a range of possible forecasts, the model also provides an opportunity for the decision-maker to exercise judgment in selecting the most appropriate forecasts.

Keywords: Forecasting; Housing; Demand; Supply; Private sector; Models; Genetic algorithm

1. Introduction

Precise estimation of demand for new residential properties is never a simple task, as it could be influenced by a number of dynamic factors, viz., demographic change, economic pattern, government policy and external environment [1]. While many major cities are confronted with a shortage of public housing and a soaring private property price [2], it is usually the government's responsibility to formulate suitable long-term housing strategies and policies to regulate and accommodate the housing needs of different sectors such that a sufficient amount of land and housing units are available to satisfy the demand. In order to make housing policy decisions, it is first necessary to estimate both short-term and long-term future housing demand.

As the housing stock is relatively inelastic in the short run, an overly conservative prediction in the housing demand could result in a shortage of residential supply. However, no one would ever imagine an overly optimistic housing forecast could also lead to profound effects to the locality especially on the overall economy. Recent example in Hong Kong (HK) has illustrated that a surplus supply of residential units had an inverse relation to the price of real estates (the property price in HK plummeted by almost 60% between 1998 and 2003). Reliable estimation of new residential property not only concerns policy makers, planners and home purchasers/tenants, but could also determine the survival of many companies related to the construction sector [3].

Despite its strategic significance, little research has been carried out to enhance the methods for predicting the residential demand. In some cases, estimations are made according to a projection of flats required for new households (e.g. new marriage, divorce, new immigrant, etc.) and existing families (e.g. those affected by redevelopment

*Corresponding author. Tel.: +852 2857 8556; fax: +852 2559 5337.
E-mail addresses: tstng@hkucc.hku.hk (S. Thomas Ng), rm.skitmore@qut.edu.au (M. Skitmore), kfwkf@graduate.hku.hk (K.F. Wong).

0360-1323/$ - see front matter
doi:10.1016/j.buildenv.2007.02.017

programs). Surely, demographic change would have significant implication to the housing demand, yet one should not ignore the impacts of economic change on the desire of property purchase [4]. According to Hillebrandt [5], the effects of economy on construction occur at all level and in all aspects of economic life, hinting that the economy (e.g. income, interest rate, etc.) may somehow influence the demand for residential properties, especially on private housing.

This paper reports on a comparison of four leading indicator models for forecasting HK private sector housing supply (as a proxy for demand) directly. These comprise a (i) Linear Regression Analysis (LRA) model, (ii) Genetic Algorithms (GA) model, (iii) GA-LRA model, where LRA is used to select the indicator variables; and (iv) GA-LRA model with Adaptive Mutation Rate (AMR) to reduce the possibility of local optima. The findings suggest that the GA-LRA model with AMR provides the most accurate forecasts and over a longer time horizon. In providing a range of possible forecasts, the model also provides opportunity for the decision-maker to exercise some judgment in selecting the most appropriate forecasts.

2. Economic indicators

The findings of previous research studies (e.g. [6,7]) realized that a close relationship exists between the construction and economic cycles, and thereby swings in the economy can be treated as indicators of the prospective movement in the construction industry and *vice versa*. The cyclical indicator technique can be used to exploit this for forecasting purposes. Although not without its shortcomings (i.e. its apparent lack of theoretical basis and inability to explain transmission processes), this technique can be used in any market-oriented economy [8].

Having been used for various aspects of construction forecasting [7,9,10,11], the suitability of economic indicators for forecasting the private residential demand should be explored. In developing a similar system for the private sector, it is clear that some aspects of the public sector model are useful. Indicator variables such as marriage, divorce, etc., are likely to be relevant to both sectors. In addition, public housing demand can also be treated as an indicator for future public sector housing supply (as a proxy for demand) and the government forecasts for these are expected to be reasonably accurate.

By observing the economic indicators used in similar topics of other countries, together with those of the HK government, and considering the availability and consistence of measurement of data in HK, a list of candidate economic leading indicators as shown in Table 1 were selected for building the forecasting model. These economic indicators have been used in comparable studies such as Goh [9,10] and Killingsworth [6] and they should therefore be appropriate for model development. Time series data for the indicators are available from the "*Hong Kong Monthly Digest of Statistics*", which is one of the general statistical digests compiled by the Census and Statistics Department in HK—with historical records dating from the early 1980s to the present time. To ensure a sufficient amount of data is available for model estimation, 20 years of quarterly records were used for all the time-series data relating to construction output and other economic indicators.

Table 1
List of leading indicator variables

Economic indicators	Abbreviation
Newly completed public housing	PUBH
Disposal of government land	LAND
Unemployment rate	UER
Property index	PROIN
Heng Sang index	HIS
Gross domestic product	GDP
Gross domestic product—construction	GCON
Composite consumer price index on housing item	HCPI
Total housing stock	HSTOCK
Government consumption expenditure	GCE

Where the indicators recorded in the digest were not exactly in quarterly form (e.g. the Hang Seng and marriage indices, which are in monthly form, or housing stock, which are in yearly form), it was necessary to estimate the quarterly figures by either aggregation or interpolation of the figures involved.

3. Modeling techniques

3.1. Linear regression analysis

As suggested by Hanke [12], the main statistical forecasting techniques available are LRA and Autoregressive Integrated Moving-average (ARIMA) techniques. In the construction industry, LRA and ARIMA models have often been used to model and forecast construction variables such as demand and price owing to their relative simplicity in both concept and application [2,7,10,13–20], sometimes in conjunction with other techniques such as artificial neural networks [21], decision support systems [22] and geographic information systems [23].

In using LRA for private sector housing forecasts, an obvious starting point is the model:

$$y = \beta_0 + \beta_1 x_1 + \beta_2 x_2 + \beta_3 x_3 \ldots + \beta_n x_n, \quad (1)$$

where the dependent variable y represents the value of private housing supply predicted; and x_1 x_2 … x_n are economic indicators with the coefficients β_0 β_1 … β_n to be estimated from the data.

3.2. Genetic algorithms

GA are stochastic search and optimization algorithms based on the principles of natural evolution [24] and their ease of use has enabled many applications to be identified in solving business, scientific and engineering optimization