Dwelling Projection Model

FINAL REPORT

Prepared by Growth Funding and Analytics Unit:

Yong Srun Vann, Research and Data Analyst

Greg Carstens, Unit Manager

Andy Zhao, Analytics Programme Manager

Nathan Dalgety, Former Analytics Programme Manager

May/2022

Hamilton City Council
Te kaunihera o Kirikiriroa

Abstract

Residential construction has played an increasingly important role in New Zealand's economic development, so forecasting residential construction demand is vital for both policy formulation and implementation, and for designing effective regional infrastructure development plans. However, research on forecasting residential construction demand in New Zealand at a regional level is sparse. Using local data and knowledge of Hamilton City, this study proposes a comprehensive forecasting method that can predict the number of new dwellings consented and completed in the short and medium term. The proposed method combines two consecutive models which have different forecasting methodologies: Sub-model 1 implements multivariant forecasting frameworks using econometric methods to predict the number of new dwellings consented; the output from Sub-model 1 is fed into Sub-model 2 to predict the number of new dwellings completed by applying a machine learning method to council's rich building inspection dataset. This study finds that the use of economic variables and a combination method can enhance forecast accuracy in terms of low mean absolute percentage errors (MAPE) and acceptable root mean square standard error (RMSE). The output produced serves as valuable input for developing local urban and infrastructure planning and growth strategies. Furthermore, the robust method proposed in this study not only contributes to the literature but also provides practical direction to local authorities and researchers in developing new dwelling forecasting methodology at a regional level.

Keywords: modelling, forecasting, dwellings consented, dwellings completed, econometrics, machine learning

Table of Contents

List of Figures and Tables

SECTION 1 - INTRODUCTION

Forecasting dwelling demand is important for policy development and urban planning because the construction market plays a significant role in economic development and provides business and employment opportunities. In particular, understanding future dwelling demand at a local level enables local government to design effective infrastructure investment plans and growth strategies, that will support the sustainability of regional economic growth and the enhancement of healthy development of the construction market.

As New Zealand's largest inland city, Hamilton has experienced rapid economic growth and rising population due to its attraction as a place to live, work and study. Along with this growth, there are changes in household formation and housing demand that impose upward pressure on the local housing market, construction sector and public policy. Keeping housing supply up with rising demand is challenging and becomes one of the local government's commitments to ensure a sustainable, competitive housing market. To accommodate this demand growth, Hamilton City Council (HCC) is looking closely at how it enables housing solutions through ensuring sufficient land and infrastructure capacity are available. Therefore, having a good understanding of future dwelling demand is crucial to making effective policy decisions on urban planning and capital investment.

HCC has assessed the capacity of land for housing in the city and the development capacity needed to sufficiently meet the future demand in accordance with the National Policy Statement for Urban Development 2020[1] (NPS-UD). HCC has been actively involved in research that helps forecast dwelling demand so that effective planning and infrastructure decisions can be made to meet future local housing demand and maintain sufficient housing capacity. An assessment of demand and capacity for housing development is conducted every three years for the Future Proof Area (FPA), which includes Hamilton City, Waikato District, and Waipa District. The report focuses on the projections for urban-dwelling demand in the region and evidence-based implications for planning and sufficient capacity.

In addition, the National Institute of Demographic and Economic Analysis (NIDEA) has developed population and household projections for the Waikato region over the period of 2018-2068 (Cameron & Cochrane, 2021). To date, the NIDEA's household projection has been used to estimate the demand for new dwellings in Hamilton City as it is the best projection per annum available to HCC. However, it is constructed based on demographic factors, such as population size and household formation, but ignores economic factors that could potentially affect the demand for dwellings and construction. Using NIDEA's household

[1] See details in Ministry of Housing and Urban Development at https://www.hud.govt.nz/urban-development/national-policy-statement-on-urban-development-nps-ud/

projection as an indicator of dwelling demand could be less justified in practice because the number of households does not necessarily translate into the number of dwellings. Moreover, the household projection has under-projected compared to the actual number of new dwellings completed in recent years. Although the difference can be acceptable in the sense of the nature of forecasting, it could be that the NIDEA model overlooks the impacts of economic indicators.

The change in the number of new dwellings tends to follow the change in population size in the long term, but it is more closely correlated with economic indicators in the short term. Fundamentally, the development of the construction market is reliant on economic performance. For instance, the change in the demand for new dwellings saw a dip with the global financial crisis and a surprising rise amid the economic impacts of the COVID-19 pandemic regardless of the change in population size. As literature has pointed out that the construction sector has explainable relationships with general macroeconomic factors, the forecast on future demand for new dwellings should reflect the change in the economic environment (Chan, 2001; Hua, 2012; Lewis, 2009; Ofori & Han 2003).

There is no research that has incorporated economic variables into dwelling projections, especially for Hamilton City. Also, the currently available models in the literature only focus on construction demand in terms of the value and number of building consents (hereafter, consents), whereas forecasting the number of dwellings completed has not been studied. While not all dwellings consented will get CCC issued or take the same amount of time to get completed, the forecast on dwelling completed will better indicate when actual demand for council services will occur and what the impacts of economic disruptions such as supply chain issues, have on the completion of new dwellings.

Consequently, HCC's elected members have requested an improved modelling and forecasting method. This has become a research interest and motivation to build a forecasting model that can capture most, if not all, potential factors that are associated with dwelling demand and provide a reliable and robust forecast of both dwellings consented and completed in the short and medium term.

To build a more robust and agile forecasting model, a combination approach is developed by introducing two subsequent models which use various forecasting techniques. Sub-model 1 focuses on developing a forecasting model that can predict the number of dwellings consented on an annual frequency, whereas Sub-model 2 constructs a forecasting model for predicting the number of dwellings completed on a monthly basis. In each model, different techniques have been applied, including econometric and machine learning methods. This study has utilised the most recent data available from both local and national databases, including HCC, Statistics New Zealand (Stats NZ), The Treasury, Reserve Bank of New Zealand (RBNZ) and NIDEA. The combination approach has produced reliable and realistic forecasting outputs that have been validated from both an econometric and practical perspectives. The proposed methodology was built on open-source software that allows researchers to easily update the models and input data when needed.

From a practical perspective, this proposed methodology is highly beneficial. The forecast can be updated frequently to capture up-to-date information once the data are available, especially when Stats NZ releases new data, thus enabling decision-makers to make effective decisions responding to the recent market environment. Historically, elected members relied on NIDEA's household projection which is only updated every five years after Census data are released.

NIDEA's household projection is static and not able to adjust to economic changes such as the recent housing boom. Using a static projection to make long-term city-scale infrastructure investments, adds risk and uncertainty and can lead to inefficient allocation of capital. For example, HCC's current Long-term Plan (LTP) relies on growth projections to underpin its $2 billion capital growth programme. In recent years the NIDEA projection has been well below the actual number of dwellings completed. Having projections that are too low is a significant risk for councils particularly in terms of infrastructure investment and ensuring that there is adequate land ready for development as required by the government. Projections that are too high also pose a risk as infrastructure investment could occur before it is needed. The new forecast will be a valuable input into HCC's Growth Model which estimates the growth in different areas of the city. It will also underpin HCC's growth modelling environment, including the Growth Model, the Three-Waters Models, the Transport Model, the Development Contribution Model and the Asset Management Models, which all rely on dwelling projection to simulate the models.

The remainder of the paper is structured as follows. Section 2 reviews the literature that is relevant to the development of the forecasting model. Section 3 introduces the methodology framework used in this study. Section 4 presents the forecasting output and model evaluation. Section 5 provides a discussion of models and policy implications. Finally, Section 6 concludes the study.

SECTION 2 - LITERATURE REVIEW

This section reviews a collection of literature relevant to dwelling projection in different contexts and countries. The presentation is divided into two sub-sections. Section 2.1 presents the theoretical relationship of construction demand and selecting macroeconomic variables. Section 2.2 discusses the forecasting techniques implemented in the previous studies.

2.1 Theoretical Relationship and Selecting Macroeconomic Variables

A more reliable forecast of construction demand would assist governments and stakeholders to design appropriate policies and strategies to ensure the sustainable development of this industry (Jiang & Liu, 2011). The knowledge of construction demand enables builders, firms, tenderers and other stakeholders to predict the future workload, design efficient planning and strategies and increase profitability (Hua, 2000). However, forecasting construction demand is arduous as there is a complex process that many factors potentially affect the construction demand. Construction demand often fluctuates and is hard to forecast due to its uncertain nature that is associated with the state of the economy and government policies. This also becomes of interest for researchers to seek methods to predict construction demand (Hua, 2012).

Existing literature relies on the general macroeconomic factors that have economic influences on construction demand to construct modelling and forecasting (Hua, 2012). Fluctuation in the construction market can be caused by changes in the local economic conditions (Bon, 1989; Briscoe, 1988; Hillebrandt, 1984; Hindle, 1993). The construction market is closely associated with other sectors and affects the momentum of economic movement (Chan, 2001; Lewis, 2009; Ofori & Han 2003), and apposite planning and strategies to ensure the sustainability of the housing development (Hampson & Brandon, 2004).

Theoretically, the construction activities are related to the economic environment, so economic volatility can impact the construction industry. The performance of the construction sector depends on the state-wide well-being of the economy (Hua, 2012). If the general economic condition is positive, the demand for physical assets is expected to increase, so more investment in construction is favourable. The business cycle can have a direct impact on construction demand. The earlier study by Tan (1989) suggested that the general business cycle is closely associated with construction demand; economic indicators may be used as leading indicators in explaining construction demand (Hua, 1996).

Some important variables are found to have significant relationships with construction demand. The gross domestic product (GDP) has a positive impact on construction demand (Ofori & Han, 2003). House price is another important determinant of the construction market (Hua, 2012). Changes in population size can affect housing demand (Fan et al., 2011). The unemployment rate is used to indicate the macroeconomic stability that can affect the investment in the housing market (Ng et al., 2011). Moreover, the interest rate is a government policy to control money flow in the construction market (Fan et al., 2011).

Literature suggests that a multivariant model is a suitable approach for studying the construction market. Bork & Moller (2018) and Rapach & Strauss (2007) show that using macroeconomic factors outperforms in forecasting house prices in the USA. Although multivariant forecasting models outperform univariate forecasting models in most cases, adding too many unrelated variables, which do not have a direct effect on the dependent variable, may worsen the forecasting performance. Gupta et al. (2011) argued that a model with a small number of explanatory variables performs better than one with many variables. The two significant problems still exist in multivariant forecasting models: how to select the explanatory variables and how to assess whether the quantity of variables used has improved or worsened the forecasting accuracy.

The number of variables selected varies across the existing studies. Akintoye & Skimore (1994) adopted five variables: economic output, construction price, real interest rate, unemployment and profitability to predict the three types of construction demands in the UK. Hua (1996) selected a wide array of variables, which are national income per capita, construction demand, real GDP, building material price index, money supply, consumer price index, property price index, labour force, unemployment rate and homeownership. The explanatory variables included in Jiang & Liu (2011) have only four explanatory variables: population change, national income, interest rate and household expenditure. Hua (2012) included unemployment rate, disposable income, consent, housing stock, national saving, gross fixed capital formation, bank lending, money supply, construction cost, inflation, wage, earing and property price index. Jiang & Liu (2014) argued that the housing demand in the regional market depends on the market and economic condition in the region, in particular, the macroeconomic factors including construction prices, regional GDP, unemployment rates, interest rate and population.

In literature, the statistical selection of variables is important in the modelling process (Hua, 2012). The selection of explanatory variables in macroeconomic modelling follows two main criteria: economic theories and statistical adequacy. Firstly, statistical significance is the procedure of selecting the variables that are found to have a statistically significant relationship with the target variable. The statistical significance is recommended in selecting influencing economic variables. Hua (2000) collected a wide range of economic indicators based on literature reviews and then reduced them by using a Stepwise procedure. Secondly, the theoretical relationship remains a suitable approach for selecting economic variables. This approach is a review of literature and selection of all most possible economic and social factors that can affect construction demand. The authors need to determine the characteristics of the construction demand, identify the economic factors, confirm with economic theory and choose the most appropriate economic variables that can represent those factors (Hua, 2012).

Milunovich (2020) selected the explanatory variables based on the theoretical relationship suggested in the existing literature. The author found that multivariant forecasting frameworks with the inclusion of economic variables can well predict the out-of-sample house price and growth rates in Australia. Drought & McDonald (2011) investigated the ripple effects of the local house price effect in 10 urban areas in New Zealand by using VECM (Vector Error Correction Model). Only the variables that are considered to have some leading information and the key macroeconomic variables are included in the models. The results show that the local house price is constrained within the regions rather than spreading across the regions. This implies that the local economic factors play vital roles in leading local house prices rather than migration and spatial arbitrage. In this regard, it is suggested that the housing market should be studied at a regional or local level rather than a national level because the macroeconomic environment, including population growth, economic development and market structure, may be different across the regional markets (Jiang & Liu, 2014; Meen, 1996).

However, modelling construction demand comes with complexity and efficiency, so the local judgment and knowledge of the data and subject are also important in selecting the variables (Hua, 2012). This author contended that there is no best selection rule or clear-cut evidence to suggest such a case in modelling construction demand. Furthermore, although the selection of economic indicators can be justified by statistical significance, some important variables are theoretically significant and related to construction demand, so even though those variables are not statistically significant, this does not necessarily mean they are not important in modelling construction demand. The statistical significance can be partly influenced by the quality and quantity of the model (Hua, 1996).

Other studies suggest that adding too many explanatory variables can cause a greater problem in terms of overfitting. The forecasting performance of the econometric methods depends on the dataset and the dependent variable while it is advisable to try a wide range of models before choosing the final one (Hallac et al., 2015; Jiang et al., 2017; Kuzin et al., 2013; Stock & Watson, 2012).

Faghih et al. (2021) suggested that the literature review, local judgment and intuition can help researchers to identify the variables used in modelling and forecasting. In practice, the authors predicted the average hourly earnings of construction labour using the VECM framework because VECM is a widely adopted model to estimate the short- and long-run relationships among a set of variables that are cointegrated. The authors obtained a collection of explanatory variables from the literature review to formulate the candidate VECMs where 150 proposed VECMs were created and finally only 25 of them passed the diagnostic tests. Their general framework of model selection is based on the creation of all VECM candidate models with a different set of independent variables and those who pass the diagnostic tests and have the best forecasting performance (Faghih et al., 2021).

2.2 Forecasting Techniques

Statistical forecasting can be generally formulated in two types of models: univariate and multivariate models. The former refers to the model with the inclusion of only one variable, that depends on the past values itself to predict the future values, while the latter uses the multiple variables as explanatory variables (predictor variables) to predict the dependent variable (target variable).

The univariate model has been adopted to predict many macroeconomic variables in forecasting studies, in particular in the construction market. The univariate models include exponential smoothing, auto-regressive (AR) and autoregressive integrated moving average (ARIMA) of the Box-Jenkins (1970), known as the BJ approach. Merkies & Poot (1990) used the exponential smoothing technique to forecast the construction activities in the Netherlands and New Zealand. The AR model predicts the future values based on the historical time series data (Figueiredo et al., 2011) while the ARIMA model adds several econometric features to the data generating process, namely integrated order and moving average, on top of the AR model to be appropriately used with non-stationary data. For instance, Fan et al. (2010) used the AR approach to forecast the construction demand for commercial, industrial and residential types in Hong Kong. Hua & Pin (2000) and Wong et al. (2007) applied ARIMA to forecast construction demand.

However, univariate forecasting methods do not consider potentially exploitable data of the other time series in the same dataset, whereas multivariant models are developed to consider the external variables. The multivariant models become more popular in the forecasting literature due to their high forecasting accuracy power. One of the widely used multivariant forecasting techniques, the Multiple Regression (MR) model has been used in forecasting studies. Akintoye & Skitmore (1994), one of the early studies, used the MR model of five variables (GDP, price level, interest rate, unemployment and manufacturing profitability) to predict private sector construction demand in the UK. Tang et al. (1990) adopted MR to forecast the three construction market types in Thailand. Fan et al. (2010) employed MR to predict the construction volume in Hong Kong.

The most recent and widely adopted multivariant forecasting models are Vector Autoregression (VAR) and VECM (Johansen, 1991), which are an extension of AR models. The principle of these models is that the predicted value of the dependent variable is given by its past values and the explanatory variables included in the model. In econometrics, these models can produce the forecast outputs of each variable in the system based on its lags and lags of all the variables in the model. VECM is more appropriate for forecasting economic time series when the time series are cointegrated; it produces a long-run equilibrium relationship between the dependent and independent variables while a past equilibrium is used as an explanatory variable to explain the dynamic behaviour of the current variable (Fan et al., 2010). In other words, VECM estimates the long-run relationship between dependent and independent variables while the past equilibrium plays as predicting variables to explain the dynamic behaviour of the current variables (Jiang & Liu, 2014). It has been claimed to outperform other forecasting approaches (Fan et al., 2011; Wong & Ng, 2010). Likewise, VECM is found to have a better forecasting performance than the univariate model (Faghih & Kahshani, 2018).

VECM has been widely applied in forecasting various macroeconomic indicators, including economic output (Anderson et al., 2002), unemployment (Bruggemann, 2006), interest rate (Tan & Baharumshah, 1999) and exchange rate (Van et al., 2000). VECM can well predict labour demand in the Hong Kong construction market (Wong et al., 2007). VECM has been adopted to predict construction costs in many studies (Faghih et al., 2021; Moon & Shin, 2018; Shahandashti & Ashuri, 2013; Xu & Moon, 2012). VECM is also used to forecast the house price in New Zealand and is found to produce more accurate forecasts than most of the other models, including AR (Drought & McDonald, 2011).

Fan et al. (2011) used the VECM method to forecast the medium-term construction demand in Hong Kong. VECM has been recognized as a suitable technique to forecast the construction demand as the growth of the construction market is interdependent on the general economy and social development. VECM can capture both short-term dynamic and long-run relationships. The variables included in their study are GDP, interest rate, population growth and unemployment rate. VECM is found to produce MAPE (Mean Absolute Percentage Error) with 3%. The VECM can successfully capture the economic fluctuations caused by economic austerity in the past decade (Fan et al., 2011).

Other studies also support that VECM provides high forecasting performance compared to other techniques. Jiang & Liu (2011) implemented the multivariate forecasting technique of VECM with dummy variables to predict the construction demand in Australia. The authors found that both VECMs with and without dummy variables are suitable for forecasting construction demand. VECM is considered a potential forecasting technique in construction studies (Faghih et al., 2021; Jiang & Liu, 2011, 2014).

Nevertheless, the common limitation of VECM is that it is unable to predict future shocks, including extreme economic disruptions or policy changes, so it is suggested to constantly update the VECM coefficients as the assumption of constant coefficients of the linear model may not always hold constant if the relationships between variables vary to the different economic conditions (Faghih et al., 2021). Thus, the author suggested that future studies should consider this limitation by developing a framework that allows for VECM adjustment in the modelling.

Apart from conventional econometric approaches, machine learning approaches have become an alternative forecasting approach due to their two main functions: being able to capture the non-linear relationship and using cross-validation to select the best model (Milunovich, 2020). Moreover, machine learning has been proved to provide better results if data have non-linear and noisy-type properties (Yu et al., 2009).

The Artificial Neural Network (ANN) model is one of the well-known machine learning methods in the forecasting study due to its powerful, flexible and easy functions (Patterson, 1996). The advantages of the neural network have been known for its ability to perform non-linear modelling without needing to make any functional assumptions about time-series data (Buyukasahin & Ertekin, 2019; Hua, 2000). ANN model has been widely used in machine learning algorithms in many fields, including finance, energy, hydrology and network communications. In its network, the input layers are the input variables. The hidden layers serve as the layers of abstraction, pulling information from the input layers. No rules are setting the hidden layers in the design of ANN architecture. The rule of thumb is to start with one hidden layer and add more as needed. Another adopted rule of thumb is one-fourth of the size of the input layers. The smaller size of hidden layers has a less effective mapping process to pull all information from input layers, but although the larger size of hidden layers increases the processing power of the neural network, it makes the training more complicated and time-consuming. It is suggested that the optimal size of the hidden layer is five (Hua, 1996).

The gradient boosting model (GBM) is another machine learning forecasting method that has earned popularity because of its scalability. It can run 10 times faster than the existing single machine learning algorithms (Truong et al., 2020). It has been broadly implemented in various studies (Nie et al., 2021; Truong et al., 2020; Yoon, 2021). The GBM is an ensemble machine learning model which was developed by Friedman (2001). This model makes a single leaf and builds regression trees, which are decision trees used to estimate a continuous real-valued function. The regression tree is derived from an iterative process that divides the data into smaller sub-data based on the number of nodes and branches. It involved three main elements: (1) the optimization of the loss function; (2) the utilization of weak learners to make predictions; (3) an additive model to add weak learners to minimise the loss function. In general, it uses a base weak learner and boosts the performance of a weak learner by iteratively shifting the focus towards problematic observations that are hard to forecast and thus build a stronger learner (Yoon, 2021).

Yoon (2021) used the Gradient boosting model (GBM) to forecast the real GDP growth in Japan from 2001-2018 and showed that GBM produced more accurate forecasts compared to the benchmarks, suggesting that GBM should be more widely used in macroeconomic forecasting. GBM is also applied to predict house price and energy consumption (Nie et al., 2021; Troung et al., 2019). Their results show that GBM is a decent forecasting method that can produce a lower value of forecasting errors.

SECTION 3 - METHODOLOGY

The objective of this study is to build a forecasting framework that can produce an output more practically useful for urban growth policy implementation and vitally, for HCC's strategic infrastructure planning. Therefore, the development of models in this study would be prioritized to the use case of the output. This also means the developed models should be robust and agile and the outputs should be practically useful for council-wide planning.

This study proposes a combination approach that combines two different sets of forecasting frameworks where several forecasting models were developed to predict the number of new dwellings consented and completed. The following presents details of the methodology framework used in this study.

3.1 Understanding Residential Building Construction Process

Unlike the literature which only focuses on construction demand proxied by consent value or number, the main objective of this study is to forecast two variables: the number of residential dwellings consented (hereafter, DCS) and the number of residential dwellings completed that are issued for Code Compliance Certificate (hereafter, CCC). There are three main reasons the CCC projection is needed. First, although the prediction of DCS remains popular in literature and important policy input, the prediction of CCC which refers to the number of dwellings consented turning into completed dwellings is more relevant and useful for infrastructure investment planning purposes.

Second, the high-quality building inspection and CCC data are available from HCC's regulatory system and can be utilised to develop the forecasting modelling. Lastly and most importantly, CCC projection is the final production of the development process that indicates when people move into their house and the demand for council services materialise. The workflow from DCS-CCC can be illustrated by the building construction process shown in Figure 1.

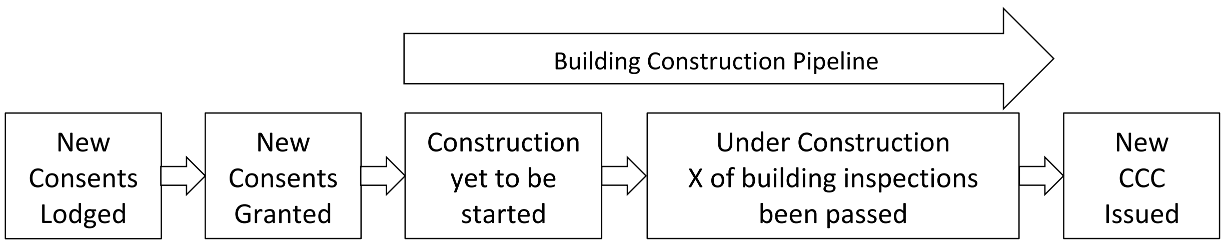

Figure 1 Building Construction Process

Figure 1 illustrates the statutory process for building construction. Initially, the building consent is lodged to get permission for construction. After the consent is granted, it will move to the stage of building construction pipeline where two cases are considered: (1) the consent has been granted but the construction has not yet started; (2) the consent has been granted and the construction has begun and passed *X* number of building inspections by the Council. Once all the required inspections are passed, the CCC application can be lodged and subsequently issued once all necessary documents are submitted. Therefore, this study aims to predict DCS and use them as inputs to predict CCC.

3.2 Combination Approach

As discussed in the previous section, forecasting the number of dwellings completed requires knowing the number of DCS in any specific year. Accordingly, a combination approach is proposed. This approach combines two different forecasting frameworks. In other words, it has two sub-models: the first sub-model focuses on forecasting DCS in the short and medium (hereafter, Sub-model 1); and the second sub-model forecasts the number of CCC in the short, medium and long terms (hereafter, Sub-model 2). The detail of each sub-model is presented in Figure 2:

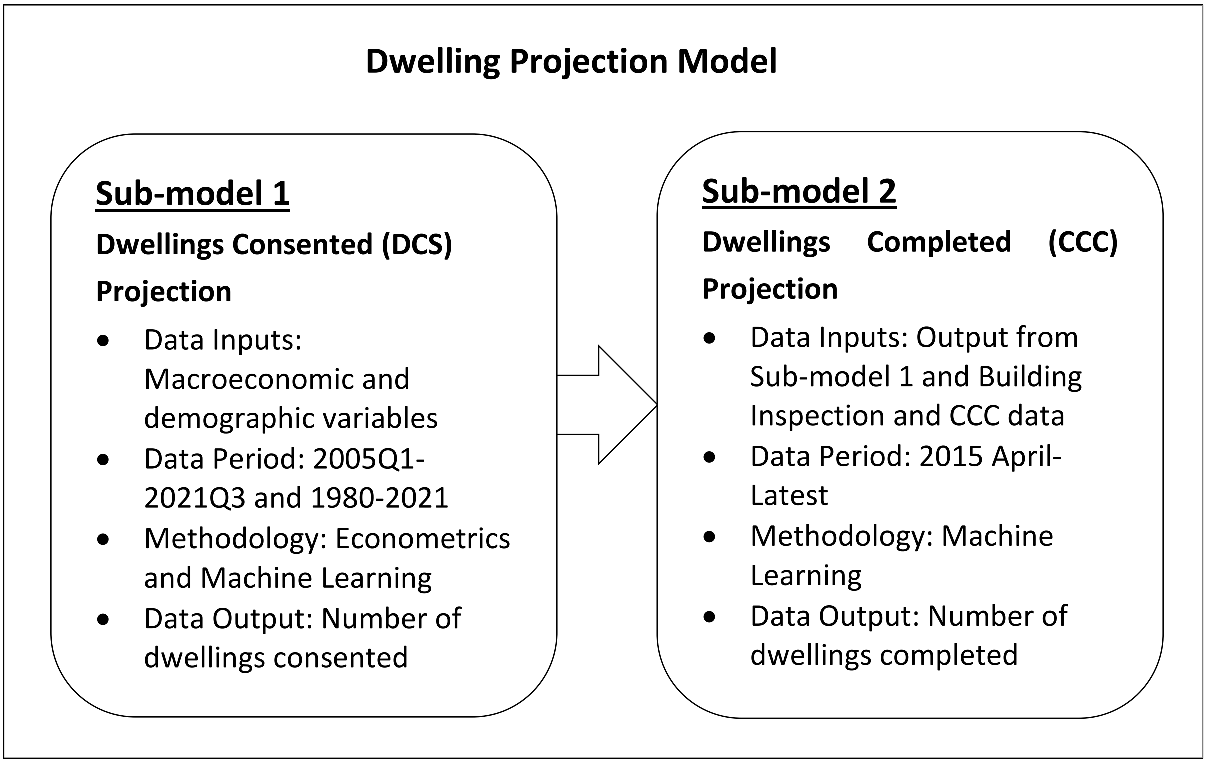

Figure 2 Dwelling Projection Framework

Figure 2 shows a summary of the methodology framework in each sub-model. Specifically, Sub-model 1 relies on the external indicators to predict DCS, i.e., the macroeconomic and demographic variables are used to capture the external impacts on DCS, which have been evidenced in the literature. In particular, the demand for construction (consent volume) depends on the macroeconomic environment (Jiang & Liu, 2014), thus affecting the DCS. The data period mainly depends on the data availability of the variables of interest, DCS. The methodology of Sub-model 1 combines both econometrics and machine learning methods. The output from Sub-model 1 is the predicted number of DCS in the short and medium term and the output is then fed into Sub-model 2.

Sub-model 2 is the next stage of the whole dwelling projection model. The data inputs in this model include the building inspection data, CCC data and DCS output from Sub-model 1. The methodology of Sub-model 2 starts with the estimation of the construction period from DCS to CCC by using a machine learning algorithm. The model obtained is then used to predict the number of CCC based on both current DCS (in the pipeline) and future DCS (the output of Sub-model 1). The details of the methodology of both sub-models are presented in the following sections.

3.3 Sub-model 1: Dwellings Consented Projection

Sub-model 1 uses macroeconomic variables as explanatory variables to explain their impacts on the number of dwellings consented (hereafter, DCS). The proposed methodology is the econometric method of VECM, which has been widely used in the literature to predict the construction demand. VECM uses the past data to study the relationships among the variables and then uses the past values of each variable with the obtained estimates to predict the future values of the variables. This approach allows researchers to conduct the forecast when the future information of other variables is not available. In this regard, VECM is the most suitable approach used to forecast the DCS over the period of 2022-2031 which is the 10-year horizon from the year this study was conducted. The predicted values of DCS produced in the Sub-model 1 are computed by using the predicted values of other economic variables in the VECM system. The output implies the demand for new dwellings based on various economic circumstances.

3.3.1 Sub-model 1: Methodology Framework

Sub-model 1 applies the econometric method of VECM. Figure 3 presents the methodology framework of Sub-model 1, which includes eight steps.

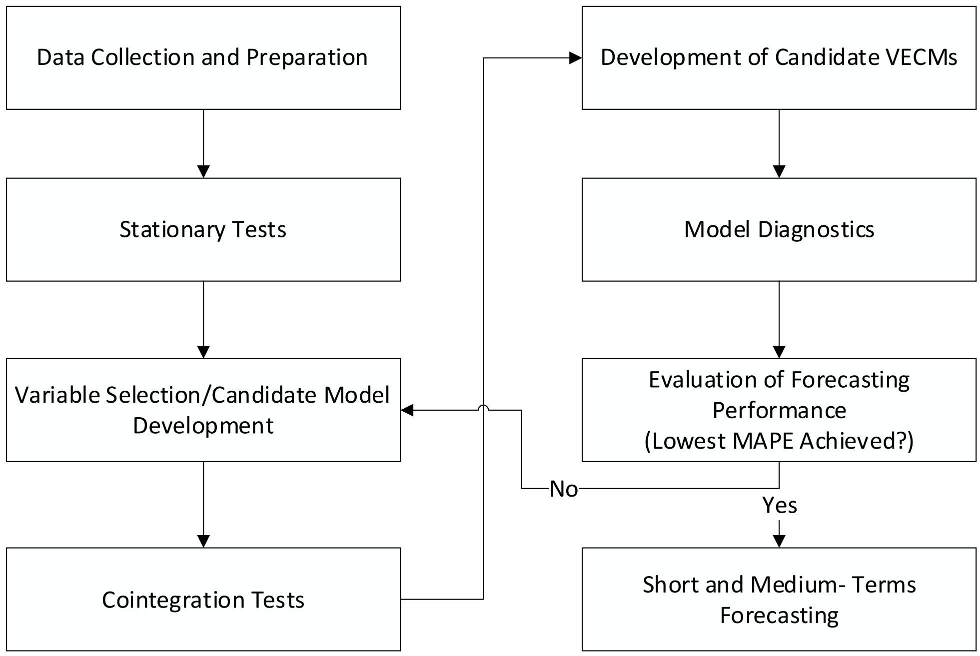

Figure 3 Methodology Framework of Sub-model 1

Firstly, the data on possible variables are collected from various sources and then prepared for analysis. The data division is carried out for training and testing the model. Secondly, the unit root test is applied to each variable to examine the stationary properties in time series. Next, this study implements a similar approach to that of Faghih et al. (2021) to select the explanatory variables, i.e., literature review, local judgment and intuition are used to identify the explanatory variables and build candidate models. The developed candidate model is tested for cointegration to check whether there are any cointegrating ranks and thus to be qualified for VECM estimation. After that, the information derived from the cointegration test is treated in the development of VECM. The estimation result is stored and carried on with the model diagnostic tests of VECM to validate the reliability of VECM.

If the model satisfies all the requirements of diagnostic tests, the out-of-sample forecast is implemented and then compared with the actual data to evaluate the forecasting performance. If the MAPE score is less than 10%, the forecasting power of the model is considered to be 'highly accurate forecasting', and if the MAPE score is between 10-20%, it is 'good forecasting' (Abidin & Jaffar, 2014; Lewis, 1982). Otherwise, the model needs to be modified by adding or removing variables by referring back to Step 3-Variable Selection/Candidate Model Development, and then the same processes are followed iteratively until the lowest MAPE score is achieved. It is also worth mentioning that although MAPE < 20% is the acceptance rule in this study, the return to Step 3 could be necessary if there is a possibility to achieve the model with a lower MAPE score than the current model. Once the lowest MAPE is achieved, the model is employed to produce the ex-ante forecast for both the short- and medium-term. The details of the methodology for Sub-model 1 are described in the followings.

(a) Data Collection and Preparation

Data need to be divided into training and testing. The former is used as an input for modelling and forecasting while the latter is used as an input for validating the forecasting performance. The ratio of data division is 80:20, which is commonly adopted in the literature. Hence, the train data is from 2005Q1-2018Q3, and the last 12 quarters (2018Q4-2021Q3) are retained as the test data. All the time series data used in this study are expressed as the actual level term because transforming original data can lead to overlooking the nonstationary properties and cointegrating relationships in the data and thus causing model misspecification.

(b) Variable Selection and Candidate Development:

The most crucial task in developing forecasting models is to select appropriate and useful explanatory variables, that can potentially improve the forecasting accuracy. There is no universal rule to select the best indicators although there are many techniques suggested in the literature. For instance, Hua (2012) adopted the Stepwise and Statistical Significance methods to decide which variables need to be entered into the model. Some studies rely on theoretical justifications to select the most appropriate variables (Fan et al., 2011; Hua, 1996; Jiang & Liu, 2011; 2014). Likewise, this study relies on the literature review to get a list of possible indicators. This study finds that the variables selected in the literature are vastly different across studies. Notwithstanding this, there are still a number of common variables which have been fundamentally justified to have impacts on dwelling demand by both theoretical and statistical evidence. They include income, house price, interest rate, unemployment and population (Fan et al., 2011; Hua 1996, 2012; Jiang & Liu, 2011,2014).

Despite many approaches available for variable selection, the common purpose of doing that is to achieve the highest forecasting performance. The challenge of adding a large number of explanatory variables is that not all variables will improve forecasting performance but instead create noise in the model. As the main objective of this study is to achieve a higher predictive quality of forecasting models, the subset of explanatory variables that can improve forecasting performance should be selected. In addition, the choice of variables should not be generalised. Different locations will have different responses of macroeconomic variables to the dwelling demand due to different economic conditions and housing markets. Thus, it is important to run an analysis of variable selection specifically for Hamilton City.

To implement this, this study will utilise all the potential variables whose data are available at both local and national levels to be included in the model. This method is beneficial on the ground that all variables will be taken into account equally whereas the statistical selection approach could overlook their actual relationships happening in the real world due to data quality or model misspecification. However, including all variables in the model could also be subject to econometric issues, such as multicollinearity and overfitting. To overcome those issues, the final selected variables are based on the degree of forecasting performance. This means that if the variables added to the model can achieve higher forecasting accuracy or lower forecasting errors in terms of MAPE scores, they will be selected to create a candidate model.

In practice, this study starts with including only the most relevant variables, i.e., the variables with quarterly data available at a Hamilton-City level. They include GDP, house price, house sale, unemployment rate, rent price and land price. Next, more relevant variables which are at the national level are included in the model if the previous model cannot reach reasonably acceptable forecasting accuracy. Those variables are interest rate, inflation, saving rate, migration, population, money supply, household spending, mortgage rate and many more. To reduce subjective judgement, this study proceeds with this approach by adding or removing variables iteratively until reasonably acceptable forecasting errors are achieved. Although it is simply a naïve approach, it is still a reliable approach that can provide the best forecasting models. Also, the models obtained should produce the outcome to be closest to the real world. Finally, the variables found to best predict DCS for Hamilton City are GDP, house price, land value, unemployment and interest rate. These selected variables are similar to those in Akintoye & Skitmore (1994) and Jiang & Liu (2014). The details of the variables selected are presented in Sub-section (h).

(c) Unit Root Test

A stationary time series has a constant mean and time-invariant correlations. Otherwise, the time series is considered to be non-stationary and hence has a unit root (Hamilton, 1994). The number of the order of integration *I(d)* of each time series can be defined by the stationary test, which reports the number of differences to reach the stationary series (Hamilton, 1994). If the time series is non-stationary at the level and becomes stationary after taking the first difference, it is said to have *I(1)*. In some situations, the transformation of non-stationary time series can help covert to be stationary, but it could lead to a loss of information. For instance, if all variables are non-stationary and *I(1)*, there is a potential existence of cointegrating relationships between the variables. Therefore, the stationary test or the determination of the order of integration is necessary to determine whether there is cointegration and to identify the choice of estimation methods (Baltagi, 2011).

One of the most popular stationary tests for time series data is the Augmented Dickey-Fuller (ADF) test developed by Dickey & Fuller (1979) and Said & Dickey (1984). The ADF test identifies the order of integration and existence of unit root with the null hypothesis of having unit root at a specific level of confidence by comparing the t-statistic of the series against the corresponding critical value (Dickey & Fuller, 1979). The ADF test can be formulated as follows:

$$\Delta X_t = \delta t + \beta + \alpha X_{t-1} + \sum\nolimits_{i=1}^{p} \alpha_i \, \Delta X_{t-1} + \varepsilon_t \qquad (1)$$

where ΔX_t denotes the differenced value of time series X_t at time t. δt denotes the time trend. β denotes the drift, α denotes the coefficient value and ε_t denotes the residuals. The null hypothesis is $H_0\colon \alpha = 0$ representing the time series is non-stationary while the alternative hypothesis takes $H_1\colon \alpha < 0$, representing the time series is stationary. The null hypothesis is rejected if the t-statistic of the ADF test is larger than the critical value or the *p-value* is smaller than 0.01, 0.05 or 0.1). Otherwise, the series is non-stationary, and then stationary test of the difference of the data should be taken to check if it is *I(1)* or further.

(d) Cointegration Test

Having more than one non-stationary time series, it is necessary to check if there exists a linear combination of them that is stationary *I(0),* namely the cointegration series (Gujarati, 1995). If there is, variables are said to have a cointegration relationship. In other words, cointegration refers to the long-run relationships between *I(1)* variables. If two variables are cointegrated, they are said to have a similar pattern of long-run fluctuations (Engle & Ganger, 1987).

One of the well-known cointegration tests, the Johansen cointegration test (Johansen, 1988) determines the number of cointegration equations by testing the restrictions imposed by cointegration on an unrestricted vector autoregressive (VAR) model (Johansen, 2000). Johansen cointegration test consists of five different models with different characteristics in terms of trend and intercept of the cointegration equations. Model 1 assumes no trend in a dataset with zero mean. Model 2 represents deterministic data with an intercept but no trend in the cointegration equations. Model 3 assumes data have a linear trend with an intercept but no trend in the cointegration equations. Model 4 has a linear trend with both an intercept and a trend in cointegration equations. Finally, Model 5 assumes a quadratic data trend with an intercept and a trend in the cointegration equations. However, Model 1 and 5 are not practical in the real world and thus implausible for forecasting (Faghih et al., 2021; Fan et al., 2011; Jiang & Liu, 2014).

The cointegration test of Johansen & Juselius (1990) uses the multivariate maximum likelihood approach to identify the number of cointegration ranks without using arbitrary normalization rules. This study uses the cointegration test on the model with data having a linear trend with an intercept but no trend in the cointegration equation, as it is generally used in the literature. The Johansen cointegration test can be performed based on Trace statistics or Max-eigen statistics (Johansen, 1991; Johansen, 2000). In the Trace test, the null hypothesis is that there is the number of cointegration equations is less than or equal to *r* (number of cointegrating ranks), while the alternative hypothesis implies that there are more than *r* cointegration equations existing in the model. On the other hand, the null hypothesis of the Max-eigen test is that the number of cointegration equations is *r*, while the alternative hypothesis is the *r+1* cointegration equation. However, the possible number of cointegration equations should be less than the number of variables included in the model ($r < p$) (Johansen, 1991).

If the cointegration test shows any long-run relationships among the variables exist, a long-run estimation approach should be used to estimate the model, e.g., VECM which can be developed to estimate the relationships among the variables in both the short and long run as well as allows researchers to conduct ex-ante forecasts using their dynamic relationships.

(e) Vector Error Correction Model

Cointegration is a time-series data property in the econometric describing the long-run relationships between non-stationary variables. When two variables are non-stationary at level but become stationary at the first difference, they tend to be cointegrated at order one, implying a linear long-run relationship between the variables.

The VECM is a combination of the VAR model and cointegration restrictions. The long-run equilibrium restriction explains that the relationship of two variables returns to the long-run trend after there is a deviation in the short run. Following Jiang & Liu (2011), the general VECM is represented in Equation 1:

$$\Delta Y_t = C + \Pi Y_{t-1} + \sum_{i=1}^{k-1} \Gamma_i \, \Delta Y_{t-i} + \epsilon_t \qquad (2)$$

where Y_t are the independent variables being *I*(1) and integrated to an *I*(0) vector, C is the intercept, Γ is the matrix representing the short-run dynamic relationship between the independent variables, and ϵ_t is the residuals. In $\Delta = (I - L)$, L is the lag operator. k is the number of lags. The cointegrating rank of Π is the matrix reflecting the long-run equilibrium information, i.e., the number of linear combinations of stationary Y_t and denoted as r. If there are p endogenous variables and cointegrated with rank $(\Pi) = r < p$, the rank of Π can be written as $\Pi = \alpha\beta' = \alpha ecm_{t-1}$ where ecm_{t-1} denotes the error correction term and $\beta' Y_t$ is stationary. This suggests that there exists $r < stationary$ linear combinations of Y_t, β is a vector of cointegrating relationships and α is a loading matrix defining the adjustment speed of the variables in Y to the long-run equilibrium, which is the cointegrating relationship.

The lag length of the VECM can be selected by the Akaike information criteria (AIC). The selection can be done through the test of the lag length selection, and the result is used as input into the Johansen cointegration test to construct the VECM. Specifically, all variables included need to be stationary and cointegrated, so the VECM can be constructed.

(f) Diagnostic Tests

Several diagnostic tests need to be performed to validate the VECM. The tests investigate the presence of serial correlation, heteroscedasticity and normality (Chatfield, 2000; Faghih et al., 2021).

First, the test for serial correlation is necessary to the check the existence of autocorrelation in the residuals of the model which can imply several issues including overlooking the seasonal effect and long-run linear trend in the model and omitting important variables and random noise in the model (Faghih et al., 2021). Ignoring serial correlation where it exists in the model can lead to underestimation of the standard errors and oversized goodness of fit and statistical significance (Asteriou & Stephen, 2007). The Breusch-Godfrey serial correlation LM test can be used to detect the presence of autocorrelation (Breusch, 1978; Godfrey, 1978). This test allows a high-order autoregressive model for the residuals where the null hypothesis is no autocorrelation in the residuals up to a specified order (Faghih et al., 2021; Gujarati, 1995; Jiang et al., 2013).

Second, homoscedasticity refers to having constant variants of the model residuals. Ignoring the homoscedasticity can result in overestimating the goodness of fit and thus inefficient estimates (Asteriou & Stephen, 2007). The autoregressive conditional heteroscedasticity in the autoregressive-typed models can be detected by Engle's (1982) ARCH-LM test. This test is the Lagrange Multiplier test to fit a linear regression model for the squared residuals and determine if the fitted model is significant. The null hypothesis of this test is the residuals are homoscedastic.

Finally, the normality of the model residuals is an important criterion in performing linear regression. Normality refers to having normally distributed residuals; otherwise, violating the normality assumption can lead to the inferential statistics of a regression model not being valid. Jarque-Berra's (1987) JB statistics can be applied to examine the residual normality. JB statistics follows the chi-square distribution, and the values of skewness and kurtosis of the residuals can be used to examine the residual distribution (Faghih et al., 2021). The null hypothesis of JB statistics implies the normal distribution of the residuals.

(g) Selected Candidate VECM

This study proposed three models as follows:

(i) VECM-1

This VECM represents the full model which includes all potential independent variables which are found to be capable of explaining the dependent variable. In this case, the first VECM for the Sub-model 1, namely VECM-1, can be written as:

$$\begin{aligned}\Delta DCS_t = C + \alpha(ecm_{t-1}Y_{t-1} + \rho_0) + \sum_{i=1}^{k}\theta_{1,i}\,\Delta DCS_{t-i} + \sum_{i=1}^{k}\theta_{2,i}\,\Delta GDP_{t-i} \\ + \sum_{i=1}^{k}\theta_{3,i}\,\Delta AHP_{t-i} + \sum_{i=1}^{k}\theta_{4,i}\,\Delta LVL_{t-i} + \sum_{i=1}^{k}\theta_{5,i}\,\Delta UNE_{t-i} \\ + \sum_{i=1}^{k}\theta_{6,i}\,\Delta INT_{t-i} + \sum_{i=1}^{k}\theta_{7,i}\,\Delta POP_{t-i} + \epsilon_t\end{aligned} \tag{3}$$

where α is the adjustment coefficient, ρ_0 is the intercept of cointegrating equations, Y_{t-1} are the *I(1)* vectors at time t-1. $\theta_{j,i}$ represents the short-run dynamics of the relationships between the independent variables and the target variable. At time t, DCS_t denotes the number of dwellings consented (residential dwellings demand), GDP_t denotes the gross domestic products, AHP_t denotes the average house price, LVL_t is the land value, UNE_t is the unemployment rate, INT_t denotes interest rate and POP_t denotes changes in population size. All variables included here are at a city level, except the interest rate variable which is only available at the national level.

(ii) VECM-2

Concerning the problem of interpolating variables, this study introduces another model which excludes the population variable from VECM-1, namely VECM-2. VECM-2 is developed to check if the removal of the interpolated variable can improve forecasting performance. In this regard, VECM-2 has five explanatory variables: GDP, house price, land value, unemployment and interest rate.

(iii) VECM-3

Both VECM-1 and VECM-2 assume that all variables are endogenous, whereas the interest rate variable could be exogenous as it is measured by national data. For this reason, VECM-3 was proposed to consider the exogeneity property of the interest rate variable and to check whether this consideration enhances forecasting performance. VECM-3 will use the estimates of VECM-2 as its coefficients to compute forecasts, but the difference between VECM-2 and VECM-3 is the exogeneity function of interest rate in VECM-3. This means that interest rate can impact other variables in the system but is not impacted by other variables. Therefore, VECM-3 depends on the exogenous predicted values of interest rate to predict the future values of the other variables.

The exogenous predicted values of the interest rate were obtained by using a separate forecast methodology, namely ARIMA. This method can well predict future values using its own past values. The interest rate data consists of the historical data of 2005Q1-2021Q3 and forecast data of 2021Q4-2026Q2 whereas the former is used as train data and the latter is set as test data. The data were collected from Stats NZ and the Treasury of New Zealand. The train data were used to develop an ARIMA model where the order is automatically selected by AIC. MAPE score is used to evaluate the forecasting performance.

(h) Data and Variables

The data used in this study are based on quarterly time series collected from 2005Q1-2021Q3, which are available at the time the analysis is conducted. Most data were collected at the local level from various sources. The main variable of interest is DCS, which is measured by the number of dwellings consented in HCC's consenting database. It is noted that the DCS data excludes the dwellings in retirement village as dwellings in retirement village are frequently volatile and project-based. This exclusion helps reduce the noise in the data and improve forecasting performance. The consent data have been used as a proxy of construction demand by Fan et al. (2011), Gyourko & Saiz (2006), Hua (2012) and Tse et al. (1999).

One of the main explanatory variables is GDP, measured by the real GDP of Hamilton City collected from Stats NZ. It is widely accepted in the literature that GDP is a leading indicator of economic performance that could ultimately cause the fluctuation in housing demand. A rising GDP could translate into higher demand for new dwellings (Ofori & Han, 2003). Similarly, the individual demand for new dwellings tends to go along with income level. House price is also an important determinant of demand. This variable is proxied by the average house sale price collected from HCC's consenting database. Residential housing is generally regarded as household consumption goods, so the rising price could reduce the number of dwellings demanded (Hua, 2012).

In addition, land price is another fundamental determinant in predicting the total dwelling cost. The change in land price implies the change in land availability for development, which could affect the construction decision. This variable is measured by the average land value collected from HCC's consenting database. The unemployment rate is an indicator of macroeconomic stability that can affect investment in the housing market (Ng et al., 2011). The unemployment rate is measured by the rate of unemployed people to the total population and collected from Stats NZ. The interest rate can affect the investment level and the money supply in the market, which can potentially impact the loanable fund in the housing market. Interest rate is a macroeconomic variable that can affect the overall economic stability of an economy (Woodford, 2003). A lower interest rate reduces the lending costs of borrowing and encourages more investments in the construction sector as there are more loanable funds available in the market (Fan et al., 2011). The interest rate is proxied by the bank bill rate. The data are presented in the national level and collected from the RBNZ and Treasury of New Zealand.

It is generally accepted that the change in population size could affect the housing demand (Fan et al., 2011). However, the population data is not available on a quarterly basis for Hamilton City. The annual population estimates for Hamilton City collected from Stats NZ and NIDEA are converted to quarterly frequency by using the same quarterly proportion values (quarterly/ annual) of DCS. This approach assumes that the quarterly change in DCS is affected by the quarterly change in population size in each quarter.

3.3.2 Forecasting Construction

The forecasts ($f_{t+h|t}$) of the vector are computed based on the information derived from the model estimates. The forecast construction in this study is based on the multiple-step forecasts using an iterated approach. The iterated procedure needs only one estimated model to compute the forecasts for any *h* time ahead and make efficient use of data (George, 2019).

The iterated forecast approach developed by Pfaff (2008) is applied and takes the following equations:

$$y_{t+h|t} = A_1 y_{t+h-1|t} + \cdots + A_p y_{t+h-p|t} + CD_{t+h} \tag{11}$$

where $y_{t+h|t}$ is the forecasted variable at h ahead, and p denotes lag number of the vector model. CD_{t+h} is the constant or/and trend terms.

3.4 Sub-model 2: Dwellings Completed Projection

Unlike Sub-model 1, Sub-model 2 predicts the number of CCC in both short and medium term from two different components using the machine learning approach of the Gradient Boosting Machine (GBM) algorithm as the forecasting method. The GBM simulates the construction process from the building consent granted to construction completed with Code Compliance Certificate being issued. The algorithm uses historical construction timeframes as a basis to forecast the total time length. To make it simple, this study calls it the 'construction period' (the time length from consent granted to CCC granted). In short, GBM is used to predict the construction period which will be summed with the date on which consents are progressing to compute the date of CCC being issued.

3.4.1 Methodology Framework

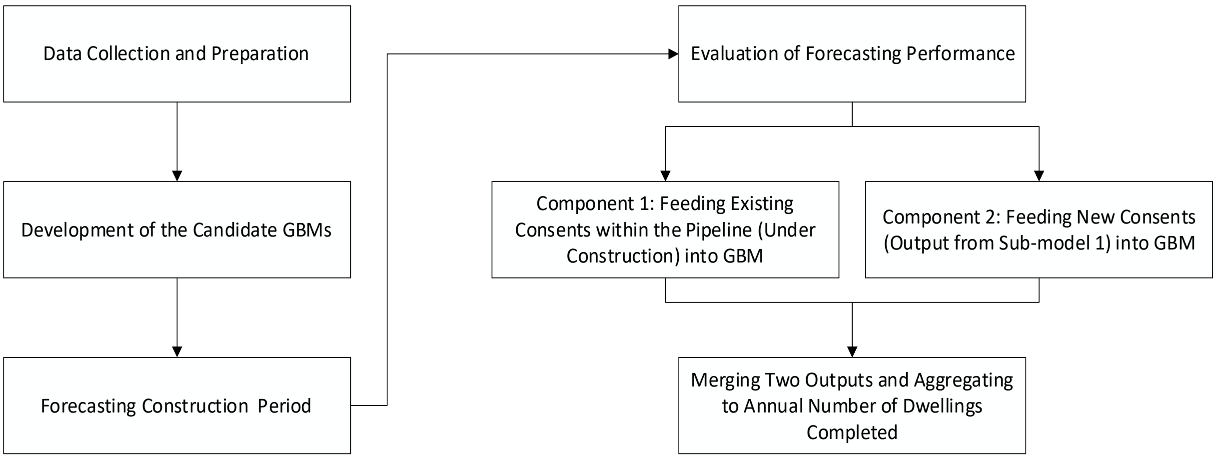

Figure 4 Sub-Model 2: Dwellings Completed Projection Framework

First of all, the data collection and preparation are performed. The input data are collected from April 2015 when the housing boom started to capture the most recent construction pattern. On a high level, the average length of construction time reflects a combination of local labour force capacity, material supply and the council's work volume. The data are divided into train (80%) and test (20%) datasets. The data are on a monthly frequency, which enables monthly updating to reflect the latest activity in the construction market. Next, the train data is fed into the Gradient Boosting algorithm, and then the GBM is developed and implemented to forecast the construction period and validate against the test dataset. The loss function minimises the difference between the predicted construction period and the actual construction period in terms of the number of months, provided by RMSE. The smaller RMSE suggests the lower errors.

After the best candidate GBM is obtained, it is used to predict the total future number of CCC by merging outputs from two components: the existing dwellings consented in the pipeline (under construction) and the new dwellings consented (output from Sub-model 1). First, the existing dwellings consented in the pipeline or under construction refer to the dwellings which are granted for consents either being ready for construction (pipeline) or being in the construction process (under construction). During construction, it is required to pass number of building inspections performed by the council before a CCC can be issued. Therefore, the way to test whether construction has started is to check the number of building inspections that have been completed. Second, the new dwellings consented are the ones that have either been lodged but not been granted, or will be lodged/granted in the future.

The first component is derived from the prediction using the building inspection data in the pipeline (under construction), including all building consents that have not received CCC. The data are collected from HCC's database. The data are fed into the trained GBM to compute the length of the construction period for each existing dwellings consented and derive the completion date. They are then aggregated into the annual frequency to understand the number of dwellings completed in each future year. In the next step, the second component is the predicted CCC number based on the number of new dwellings consented that have either lodged but not granted or not been lodged, in particular, new DCS produced by Sub-model 1.

To implement this, the quarterly or annual data of DCS from Sub-model 1 will be converted into monthly frequency using the phasing technique because GBM's construction period data are in monthly frequency and the conversion can reflect the seasonal difference and enable monthly model updating. The phasing technique assumes that the future monthly proportion of DCS (monthly/quarterly or monthly/annual) in each year will stay constant over time. The used proportion is based on the year 2019 data as this year is the normal year without any major economic disruption or pandemic.

Next, the historical CCC data are used to compute the dwelling proportion by several categories, such as growth cell (location), dwelling typology (detached or attached), median dwelling number per consent and median dwelling value. The new DCS data are then distributed proportionally based on the historic dwelling proportion. This method is beneficial because it considers the heterogeneity effects in data distribution that the different locations and typologies have different construction completion timeframes. After that, the distributed new DCS data are fitted by the GBM algorithm to predict the construction period (the time length from consent granted to CCC issued) for each new dwelling consented.

Finally, both predicted values from the two components are merged and aggregated to annual frequency to compute the total number of CCC in each year.

3.4.2 Gradient Boosting Model (GBM)

The GBM is a decision tree-based machine learning algorithm that is developed by Friedman (2001). The algorithm is a widely used method that can perform both regression-based and classification-based predictions. Boosting means using the algorithm method to improve the weak prediction trees to become stronger prediction trees. It is developed to improve the forecast accuracy and robustness of the model through building an additive model in a forward stage-wise fashion. This model makes a single leaf and builds regression trees, which is a decision trees used to estimate a continuous real-valued function. The regression tree is derived from an iterative process that divides the data into smaller sub-data based on the number of nodes and branches. It involved three main elements: (1) the optimization of the loss function; (2) the utilization of weak learners to make predictions; and (3) an additive model to add weak learners to minimise the loss function. In general, it uses a base weak learner and boosts the performance of a weak learner by iteratively shifting the focus towards problematic observations that are hard to forecast and thus build a stronger learner. According to Nie et al. (2021), the process of minimizing the loss function is calculated by the following equation:

$$\tilde{F}(x) = \underset{F(x)}{\operatorname{argmin}} L_{y,x}(\, y, F(x)) \tag{12}$$

where $\tilde{F}(x)$ is the approximation of function $F(x)$ mapping x to y, to minimize the loss function $L(y, F(x))$. The squared error function is used as the loss function to compute the approximation function as $L\big(y, F(x)\big) = \big(y, F(x)\big)^2$.

The gradient boosting algorithm begins with creating an initial base learner $F_0(x)$ which is the constant function, and then it uses the steepest descent step to minimize the loss function. The gradient of loss function $L(y, F(x))$ can be computed by:

$$\tilde{F}(x) = \underset{F(x)}{\operatorname{argmin}} L_{y,x}(y, F(x))\tilde{y}_i = -\left[\frac{\partial L(y_i, F(x_i))}{\partial F(x_i)}\right]_{F(x)=F_{m-1}(x)}, i = 1, \dots, N \quad (13)$$

where $\tilde{y}_i$ denotes approximation value of the gradient of loss function $L(y, F(x))$. x_i is input variables. m is the number of iterations.

After that, GBM takes further steps to define the regression trees $h(x_i; \alpha)$ with α as a parameter of weaker learners. The parameter of the tree takes the following equations:

$$\alpha_m = \underset{\alpha,\beta}{\operatorname{argmin}} \sum_{i=1}^{N} [\tilde{y}_i - \beta h(x_i; \alpha)]^2 \quad (14)$$

where α_m is the parameters received at iteration m. β is the weights of each weak learner. Each regression tree is added to the current negative gradient. Next, the optimal length (ρ_m) is determined by the following equation:

$$\rho_m = \underset{\rho}{\operatorname{argmin}} \sum_{i=1}^{N} L(y_i - F_{m-1}(x_i) + \rho h(x_i; \alpha_m)) \quad (15)$$

Finally, the GBM model can be achieved by updating model $F_m(x)$ at each iteration m by the following equation:

$$F_m(x) = F_{m-1}(x) + \rho_m h(x; a_m) \quad (16)$$

In summary, the weak predictive trees are repetitively generated and added to predicting machine model until the prediction is closest to the actual event.

3.5 Forecasting Accuracy Evaluation

The accuracy of forecasting performance varies with the different forecasting models used, so an accuracy test is vital in choosing the best model (Mahmoud, 1984; Makridakis et al., 1983). Evaluating the forecasting performance is an important stage when deciding to select the best model that can predict the unknown values. Model performance can be generally tested by comparing the predicted outputs produced by each model with the actual output. Hua (2012) suggests that MAPE is a more appropriate measure to evaluate the performance accuracy of the models because it focuses on the absolute error values. Two tests are widely adopted in the literature to check the forecasting reliability: MAPE and RMSE. Therefore, the standard statistical approaches for forecasting performance evaluation used in this study are Mean Absolute Percentage Error (MAPE) and Root Mean Square Error (RMSE).

For single model comparison, a MAPE score of 10-20% is generally acceptable (Abidin & Jaffar, 2014; Fan et al., 2010; Hua & Pin, 2000; Wong et al., 2007). The widely accepted rule in literature is that the MAPE score of less than 10% is 'highly accurate' forecasting, and it is good forecasting for a MAPE score between 10-20% (Lewis, 1982). For multi-model comparison, the basic rule is the one with lower MAPE would be considered to have a better forecasting performance.

MAPE is one of the most common approaches used to evaluate forecasting performance. It is computed by:

$$MAPE = \frac{1}{h}\sum_{i=1}^{h}\frac{|e_t|}{Y_t} * 100 \tag{17}$$

where $|e_t|$ is the absolute forecast error term $(Y_i - \hat{Y}_i)$. Y_i is the observed and $\hat{Y}_i$ is the predicted values for h horizon. The generally accepted MAPE scores vary across the studies, but smaller than 20% would be considered to be acceptable.

RMSE used to evaluate the forecasting performance at each forecast horizon takes the following form:

$$RMSE = \sqrt{\frac{1}{h}\sum_{i=1}^{h}(Y_i - \hat{Y}_i)^2} \tag{18}$$

where Y_i is the observed and $\hat{Y}_i$ is the predicted values for h horizon. The lowest RMSE value indicates the highest forecasting accuracy where the error is the smallest. There is no clear-cut rule on what values of RMSE can be accepted as it is computed in error unit value. The rule of thumb in this study is that RMSE less than 2 indicates a good-fit model, which means the difference in construction time length between the predicted data and actual data should be less than 2 months.

3.6 Summary of Methodology Framework

A diagram summarising of methodology framework used in this study is provided in Figure 5.

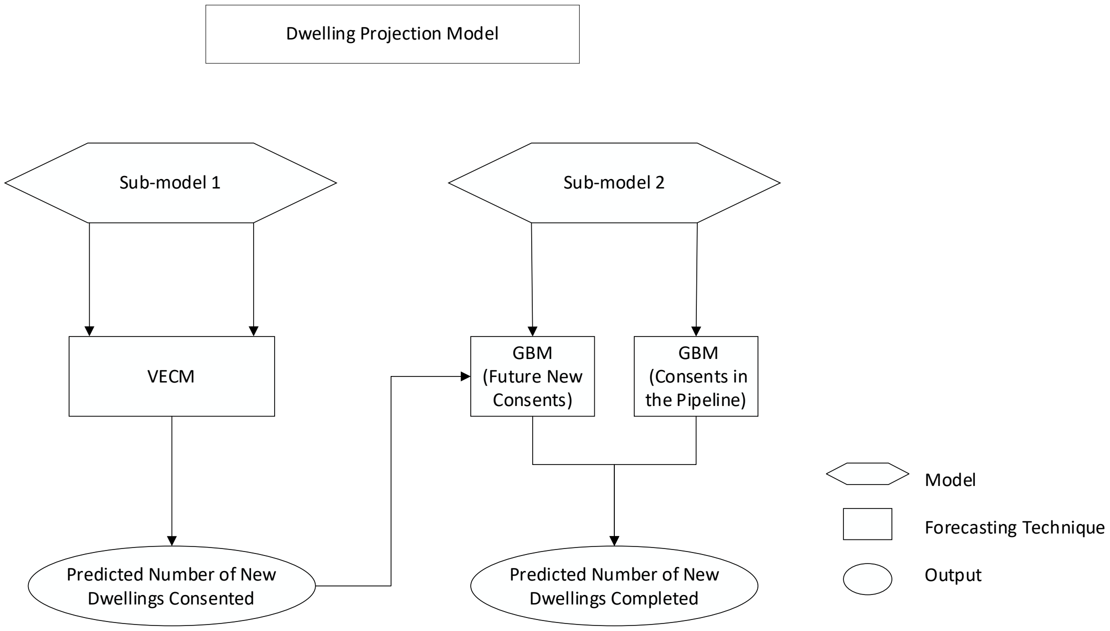

Figure 5 Summary of Forecasting Methodology

The dwelling projection has two sub-model 1. Sub-model 1 uses the VECM estimation method with a set of economic variables to predict the number of DCS in the short and medium term. Sub-model 2 employs GBM to predict the construction period from two different inputs: consents in the pipeline (under construction) and new dwellings consented (DCS output from Sub-model 1). The outputs are aggregated to compute the total number of CCC in the short and medium term.

SECTION 4 - EMPIRICAL RESULTS

This section presents the empirical results for both sub-models. Section 4.1 presents the results of Sub-model 1, and Section 4.2 illustrates the results of Sub-model 2.

4.1 Empirical Results of Sub-Model 1

In this section, the first part presents all empirical results of the candidate VECMs, while the second part presents the empirical results of the hybrid model. This section presents the empirical results, including unit root tests, cointegration tests, VECM equations, forecasting evaluation and ex-ante forecasts.

4.1.1 Unit Root Test:

To check whether the time series are stationary and identify the order of integration, the ADF test was applied to all selected variables. Table 1 shows the results of implementing the ADF test with lag order selection using the AIC method. The results show that all the series have a unit root where the null hypothesis is not rejected, meaning non-stationary series. Meanwhile, the test results indicate that all the series become stationary at least at a 10% significance level after taking the first difference, implying the order of integration is one or *I(1)*. This result is aligned with the assumption of VECM that the variables included should be of the same order of integration so that the cointegrating relationships can be investigated.

Table 1 Results of Unit Root Test

Variable	Level		Frist Difference	
	t-stat	p-value	t-stat	p-value
DCS	-1.93	0.602	-4.51***	0.010
AHP	-0.65	0.969	-3.23*	0.092
GDP	-0.81	0.817	-12.01***	0.000
LVL	-2.89	0.214	-3.89**	0.021
UNE	-0.80	0.956	-5.05***	0.010
INT	-0.87	0.798	-3.99**	0.020
POP	-2.46	0.388	-5.01***	0.001

*Note: ***, ** and * indicate statistically significant at 1%, 5% and 10%, respectively.*

4.1.2 Cointegration Test

After all the variables were confirmed to be *I(1)* series, the cointegration test was implemented for VECM-1 and VECM-2, whereas VECM-3 is just a replication of VECM-2 in which the interest rate is set as an exogenous variable. Table 2 shows the results of the

Johansen cointegration tests for each candidate VECM at the lag orders selected based on AIC.

Table 2 indicates the results of the Johansen cointegration test which is conducted to determine the number of cointegrating ranks among the variables in each model. The test results show that some cointegration equations exist among the variables. For VECM-1 which includes all variables, the Trace statistic is statistically significant at a 1% level at five cointegrating ranks, indicating that there are at most 5 cointegrating equations among the variables. The Max-eigen statistic shows that there are four cointegrating equations, but the statistic at five cointegrating ranks is very close to the 10% critical value. From these two types of statistics, it is concluded that there are at most five cointegrating ranks in VECM-1.

Table 2 Results of the Cointegration Test

VECM-1: ***DCS, AHP, GDP, LVL, UNE, INT*, *POP***

Cointegrating Rank	Trace Statistic	0.1 Critical Value	0.05 Critical Value	0.01 Critical Value	Max-eigen Statistic	0.1 Critical Value	0.05 Critical Value	0.01 Critical Value
r<=6	5.60	7.52	9.24	12.97	5.60	7.52	9.24	12.97
r<=5	18.68*	17.85	19.96	24.60	13.08	13.75	15.67	20.20
r<=4	43.25***	32.00	34.91	41.07	24.58**	19.77	22.00	26.81
r<=3	86.55***	49.65	53.12	60.16	43.3***	25.56	28.14	33.24
r<=2	145.74***	71.86	76.07	84.45	59.19***	31.66	34.40	39.79
r<=1	221.09***	97.18	102.14	111.01	75.35***	37.45	40.30	46.82
r=0	344.87***	126.58	131.70	143.09	123.78***	43.25	46.45	51.91

VECM-2: ***DCS, AHP, GDP, LVL, UNE, INT***

Cointegrating Rank	Trace Statistic	0.1 Critical Value	0.05 Critical Value	0.01 Critical Value	Max-eigen Statistic	0.1 Critical Value	0.05 Critical Value	0.01 Critical Value
r<=5	4.33	7.52	9.24	12.97	4.33	7.52	9.24	12.97
r<=4	17.01	17.85	19.96	24.60	12.68	13.75	15.67	20.20
r<=3	41.38***	32.00	34.91	41.07	24.37**	19.77	22.00	26.81
r<=2	76.48***	49.65	53.12	60.16	35.1***	25.56	28.14	33.24
r<=1	133.26***	71.86	76.07	84.45	56.78***	31.66	34.40	39.79
r=0	199.94***	97.18	102.14	111.01	66.68***	37.45	40.30	46.82

*Note: ***, ** and * indicate statistically significant at 1%, 5% and 10%, respectively.*

VECM-2 which excludes the population variable has similar results between both types of statistics. There are at most three cointegrating ranks among the variables where the Trace statistic is statistically significant at the 1% level and the Max-eigen statistic is statistically significant at the 5% level. Therefore, VECM-2 is concluded to have three cointegrating ranks.

4.1.3 Level-VAR Representation of VECMs

VECM-1 and VECM-2 are the candidate models which have a different subset of explanatory variables and cointegration relationships. VECM-1 was constructed with five lags, the deterministic trend and five cointegration ranks. Likewise, VECM-2 was performed with five lags, the deterministic trend and four cointegration ranks. In practice, both models were estimated in VECM whose estimates are in a differenced form. To generate forecasts of VECM, VECM is needed to transform into their equivalent level-VAR representation whereas the cointegration rank is important information for the transformation. The details of the transformation method can be seen in Pfaff (2008). The level-VAR representations of both models are obtained as follows:

VECM-1:

$$\begin{aligned} DCS_t = &-419.7861 - 0.1381DCS_{t-1} - 0.1668GDP_{t-1} \\ &+ 0.0012AHP_{t-1} - 0.0004LVL_{t-1} - 13.0680UNE_{t-1} \\ &+ 53.0204INT_{t-1} + 0.0535POP_{t-1} + \cdots \\ &- 0.0544DCS_{t-5} - 0.2192GDP_{t-5} - 0.0010AHP_{t-5} \\ &+ 0.0003LVL_{t-5} - 16.6529UNE_{t-5} + 2.8586INT_{t-5} \\ &- 0.0070POP_{t-5} + \epsilon_t \end{aligned} \tag{19}$$

VECM-2:

$$\begin{aligned} DCS_t = &-328.9262 - 0.1995DCS_{t-1} - 0.9110GDP_{t-1} \\ &+ 0.0012AHP_{t-1} + 0.0006LVL_{t-1} + 5.3185UNE_{t-1} \\ &+ 44.7112INT_{t-1} + \cdots \\ &- 0.2898DCS_{t-5} + 0.3417GDP_{t-5} - 0.0003AHP_{t-5} \\ &- 0.0010LVL_{t-5} - 9.5794UNE_{t-5} - 44.2599INT_{t-5} + \epsilon_t \end{aligned} \tag{20}$$

Overall, both equations were developed and used to generate forecasts. However, to find the best-fit model, these two models should be checked with several diagnostic tests and prediction performance evaluations.

4.1.4 Model Diagnostics

Several diagnostic tests were applied to candidate models to detect any violation of standard assumptions. Table 3 shows the results of the serial correlation Lagrange multiplier test, ARCH-LM heteroscedasticity test and Jarque-Bera normality test.

Table 3 Results of the Cointegration Test

Diagnostic test	VECM-1	VECM-2
LM p-value	0.301	0.937
ARCH-LM p-value	1.000	1.000
Jarque-Bera p-value	0.730	0.285

First, the serial correlation test of Lagrange multiplier (LM) was performed with up to eighth order and has the p-values are more than 0.01, suggesting that the null hypothesis of no serial correlation is not rejected. Hence, both models satisfy the standard assumption of no serial correlation. Second, the heteroscedasticity (ARCH-LM) test indicates the p-values of (1.0) for both models are higher than the rejection range (0.01-0.10), implying that the residuals are homoscedastic in the models. Finally, the Jarque-Bera normality tests also show that the null hypothesis of normality is not rejected where the p-values for both models are larger than the rejection range. In conclusion, there is no evidence of problems related to serial correlation, heteroscedastic and non-normal errors for both candidate models.

4.1.5 VECM-3 Development

As discussed in Section 3.3.1, VECM-3 is developed based on using interest rate as an exogenous variable in the VECM-2 system. To predict the future values of an exogenous variable, the ARIMA forecasting method was applied. The train data (2005Q1-2021Q3) was used to develop the ARIMA model. The ARIMA's order is selected automatically by AIC. The ARIMA (3,1,5) was obtained and then used to generate the forecast for 2021Q4-2026Q2. The forecasted values were compared with the test data, giving a MAPE score of 1.22 which is considerably low and acceptable. The ARIMA (3,1,5) model was used to compute the forecast for a period of 2026Q3-2050Q4. Finally, all the data are merged into a full dataset of 2005Q1-2050Q4 and then are fed into VECM-3 to generate the forecasts of other variables by using the coefficients of VECM-2.

4.1.6 Forecasting Output and Performance Evaluation

There are three candidate models where their level-VAR representations will be used to compute the out-of-sample forecasts, and then the forecasting performance was compared among the models using the MAPE score. Figure 6 presents the predicted values of dwellings consented of each VECM.

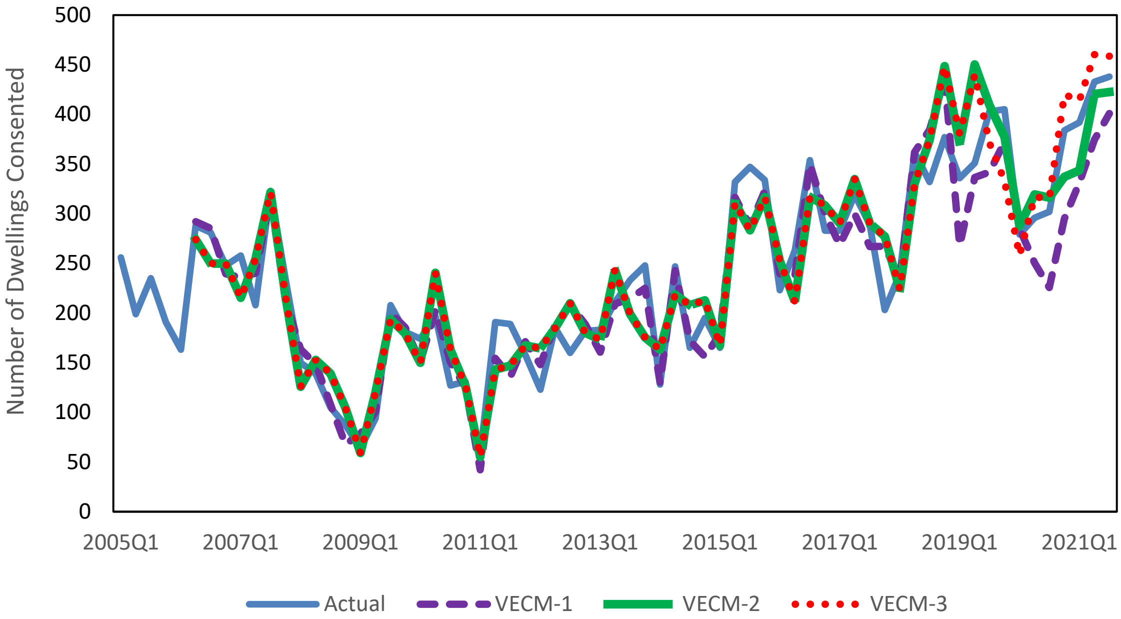

Figure 6 Results of Predicted Values of Dwellings Consented

The figure above reflects that the deviation between the actual value and predicted value of DCS varies across the candidate models. For in-sample comparison, all the candidate models tend to have a good fit with the train data, in particular the long-term trend, albeit with unparallel turning points. Figure 7 presents a closer look at out-of-sample forecasting disparities across the models over 2018Q4-2021Q3.

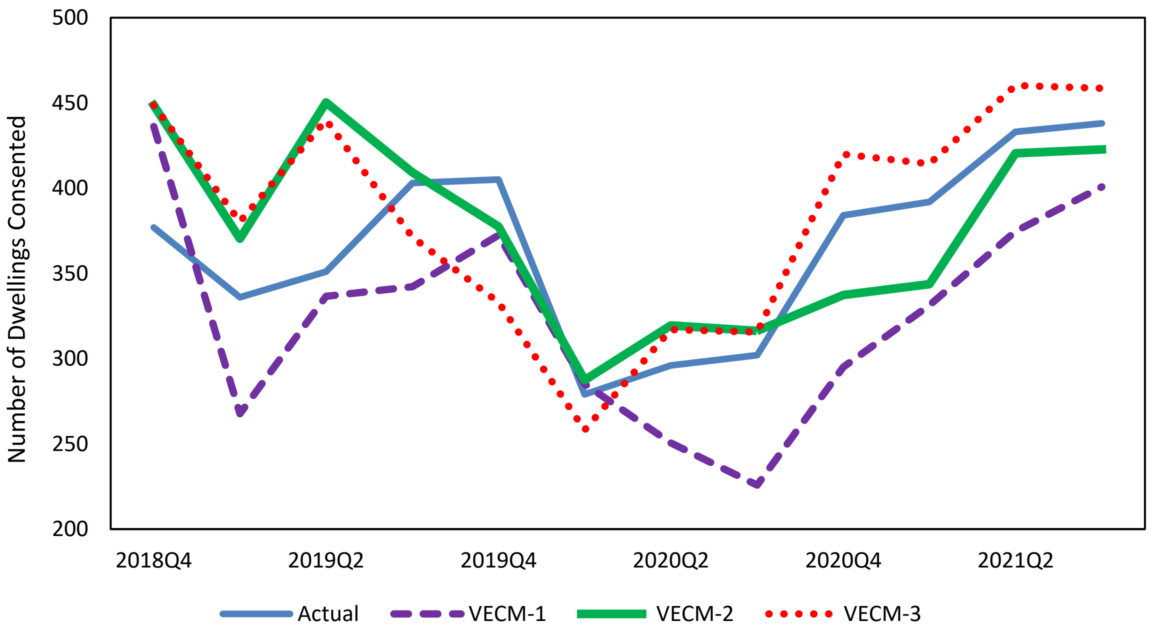

Figure 7 Results of Out-of-Sample Forecasts

The figure displays that VECM-1 has lower forecasted values than the actual values. VECM-2 and VECM-3 tend to have closer forecasted values to the actual values, but the former is relatively lower than the latter. To be more precise about their forecasting performance, the predictive adequacy of the candidate models is further evaluated by comparing them with the test data over the out-of-sample period (2018Q4-2021Q3) as shown in Table 4.

Table 4 Evaluation of Out-of-Sample Forecasting Performance

		VECM-1		**VECM-2**		**VECM-3**	
Quarter	Actual values	Predicted Values	Percentage Error	Predicted Values	Percentage Error	Predicted Values	Percentage Error
2018Q4	377	436	15.70%	449	19.00%	449	19.0%
2019Q1	336	268	-20.32%	370	10.23%	380	13.2%
2019Q2	351	337	-4.10%	450	28.29%	440	25.4%
2019Q3	403	342	-15.11%	409	1.53%	371	-7.9%
2019Q4	405	372	-8.06%	378	-6.78%	333	-17.8%
2020Q1	279	285	2.17%	287	3.01%	258	-7.5%
2020Q2	296	251	-15.33%	319	7.91%	317	7.1%
2020Q3	302	226	-25.24%	316	4.73%	316	4.5%
2020Q4	384	295	-23.16%	337	-12.17%	420	9.4%
2021Q1	392	331	-15.46%	344	-12.31%	414	5.7%
2021Q2	433	375	-13.47%	420	-2.89%	460	6.3%
2021Q3	438	401	-8.49%	423	-3.50%	459	4.7%
		MAPE =	**13.88%**	**MAPE =**	**9.36%**	**MAPE =**	**10.69%**

Table 4 shows that the MAPE score of each candidate model is less than 20% absolute percentage error, which is generally acceptable (Abidin & Jaffar, 2014, Lewis, 1980). In specific, VECM-1 has the highest MAPE score, 13.88% absolute percentage error, compared to the rest. VECM-2 which excludes the population variable achieves the lowest MAPE score (9.36%). On the other hand, VECM-3 setting interest rate as an exogenous variable does not improve the forecasting accuracy as its MAPE score rises to 10.69%. Hence, VECM-2 has the highest forecasting performance based on MAPE score evaluation.

Notwithstanding this, a good forecasting model may not guarantee to produce the most realistic output over the ex-ante forecasting period. This happens when the forecasted values could increase infinitely across the time, i.e., the number of dwelling consents can reach any number in the long run, while the supply capacity tends to be more constrained over time. To this end, the ex-ante out-of-sample forecast (2022-2031) was implemented on each candidate model to confirm whether the outputs are realistic at a rate that population growth can satisfy. The output of the forecast was aggregated on an annual basis as presented in Figure 8.

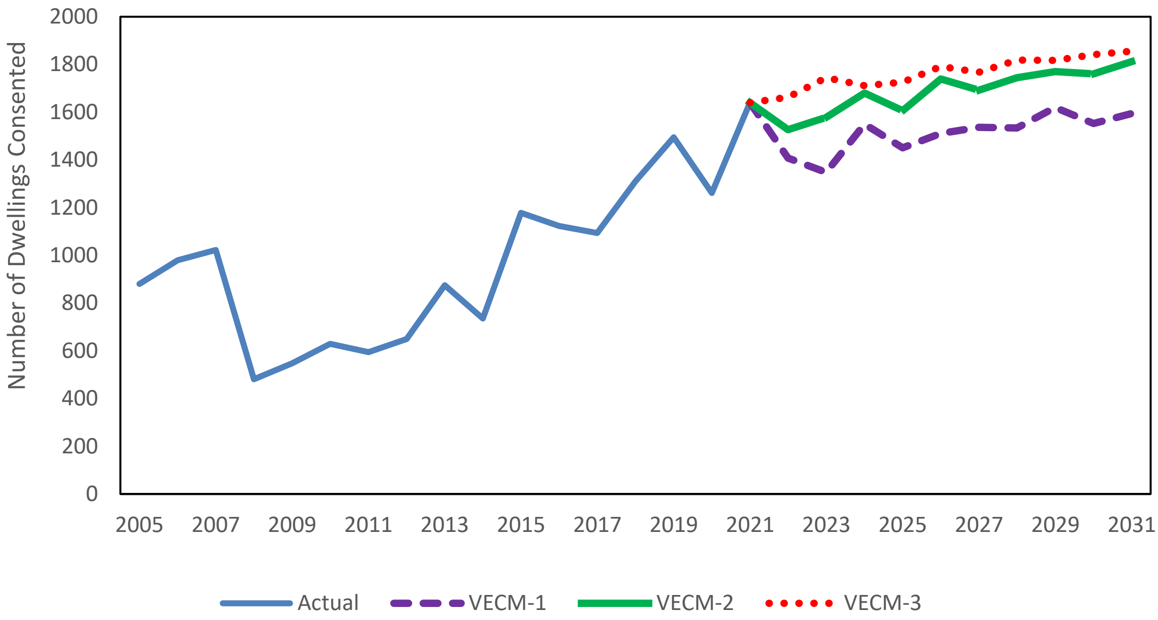

Figure 8 Results of Ex-ante Forecasts

VECM-1 and 2 tend to have a similar volatility pattern and trend despite a significant deviation. The former has the lowest forecasted values while the latter can reach up to 1,800 per year in 2031. VECM-3 has the highest forecasted values and is less volatile compared to the rest. Noticeably, both VECM-1 and VECM-2 show a considerable drop in the next two years whereas VECM-3 has an upward trend in the same period. Overall, the projected number of dwellings consented would start from 1,500-1,900 per year in the next decade, so all candidate models forecast relatively acceptable ranges of dwelling consents that could potentially happen in reality over the forecasting period.

4.2 Empirical Results of Sub-model 2

This section presents the results of GBM and total CCC predictions defined from two components: the first one is to produce the prediction of CCC number from existing dwellings consented in the pipeline (under construction) and the second one is to compute the prediction of CCC number from new dwellings consented (output from Sub-model 1). The details are shown as follows.

4.2.1 GBM Modelling and Forecasting Performance Evaluation

The GBM was trained to predict the construction period by using a historical dataset containing several variables, including typology, location, number of dwellings per consent, and consent value. The GBM was firstly trained multi-times by using different combinations of parameters, which are the number of trees, Gaussian distribution, learning rate, the depth of the trees, cross-validation and the minimum number of observations allowed in the tree terminal nodes. There were 81 candidate GBMs, and the most suitable candidate GBM was selected based on the RMSE and model fitness.

The selected candidate GBM was used to compute the predicted values of the construction period, and these values were used to compare with the test data to evaluate the GBM's forecasting performance. The result shows that the RMSE score is 0.58, meaning approximately an error of 0.58 months on average between the actual data and predicted data. This score is considered acceptable in practice. Therefore, the obtained GBM is then used to compute CCC on the two components.

4.2.2 Component 1: Existing Dwellings Consented in the Pipeline

This component focuses on predicting the number of CCC to be issued for the existing dwellings consented by using the obtained GBM. The existing dwellings consented data were fed into GBM to compute their construction period. The data of the construction period were summed with the latest date of the status at which existing dwellings consented were progressing to compute the date of those existing dwellings consented to be issued for CCC. Finally, the number of CCC in each year was computed as shown in Figure 9.

Figure 9 displays the prediction of the number of CCC to be issued for the existing dwellings consented. Overall, the period of all existing dwellings consented completing the construction is approximately two years, accounting for about 1,572 CCC. The majority of CCC will be issued within 2022, and it falls sharply in 2023. This implies that most of the existing dwellings consented will have their construction completed in 2022.

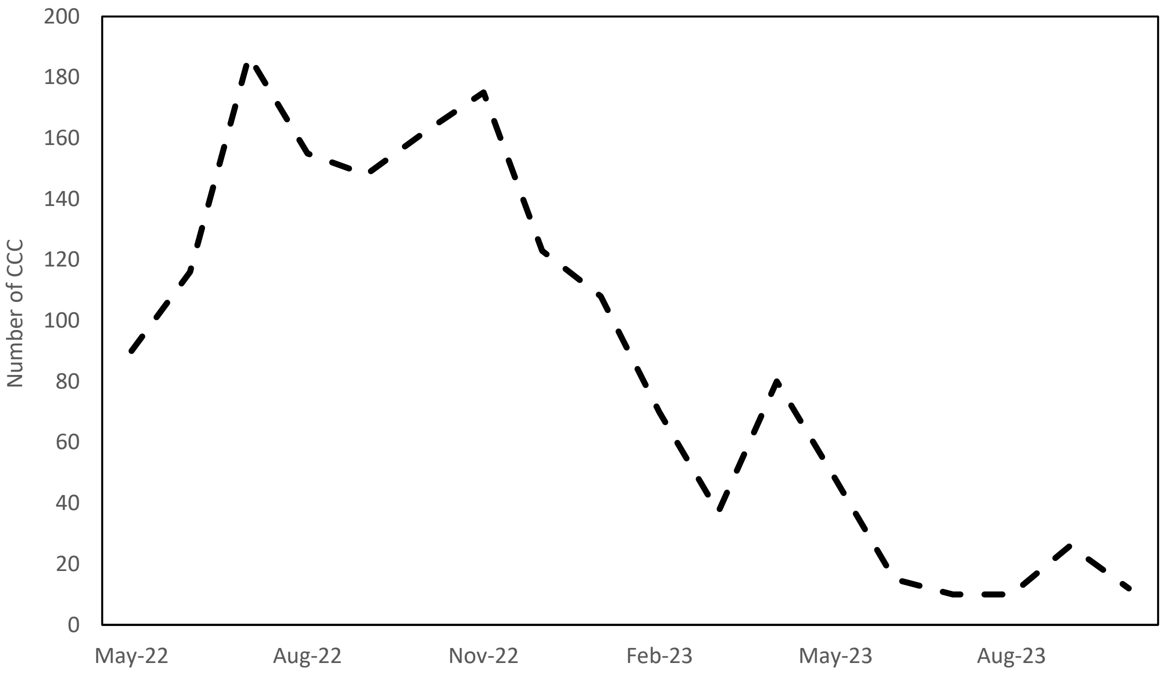

Figure 9 Results of Predicted Number of CCC from Existing Dwellings Consented

4.2.3 Component 2: New dwellings Consented from Sub-model 1

The second component is the prediction of the CCC number to be issued for the new dwellings consented forecasted by Sub-model 1. To consider both with and without population variable scenarios, both the forecasts of VECM-1 and VECM-2 are used to compute the forecasts of CCC, namely CCC-VECM-1 and CCC-VECM-2, respectively. The data of new dwellings consented were collected from VECM-1 and VECM-2 of Sub-model 1.

As data of new dwellings consented are on a quarterly and annual basis, the data were converted to monthly frequency by using a phasing technique to reflect the seasonal difference. The phasing technique uses the monthly proportion of quarterly values in a specific year as a proportion to distribute the predicted dwellings consented from Sub-model 1. In this case, selecting the year to be used is very crucial, so this study selected 2019 as it is the most recent and normal year without any economic disruption, e.g., Covid-19. To implement this, this study calculated the monthly proportion as the former is on a quarterly basis and the latter is on annual basis. The monthly proportions are shown in Column MV/QV of Table 5.

Table 5 Results of Monthly Proportion

Month	Monthly Value (MV)	Quarter Value (QV)	Annual Value (AV)	MV/QV
Jan	67	336	1495	0.20
Feb	99	336	1495	0.29
Mar	170	336	1495	0.51
Apr	81	351	1495	0.23
May	173	351	1495	0.49
Jun	97	351	1495	0.28
Jul	107	403	1495	0.27
Aug	153	403	1495	0.38
Sep	143	403	1495	0.35
Oct	146	405	1495	0.36
Nov	152	405	1495	0.38
Dec	107	405	1495	0.26

After this, the dwelling proportions by several categories were computed by using the historical CCC data. Table 6 presents the dwelling proportion by several categories, such as the typology, location, median dwelling number and value per consent.

Table 6 Results of Dwelling Proportion by Categories

Typology	Location	Median Dwelling/ Consent	Median Dwelling Value	Total Dwellings	Dwelling Proportion
Detached Single	Infill East	1.00	265,500	371	0.051
Detached Single	Infill West	1.00	220,184	516	0.071
Detached Single	Peacocke Stg 1	1.00	297,000	301	0.041
Detached Single	Rotokauri	1.00	335,000	177	0.024
Detached Single	Rototuna	1.00	379,900	1,492	0.204
Detached Single	Ruakura	1.00	371,975	518	0.071
Attached 2-4	Infill East	2.50	590,000	609	0.083
Attached 2-4	Infill West	3.00	500,000	920	0.126
Attached 2-4	Peacocke Stg 1	2.00	350,000	78	0.011
Attached 2-4	Rototuna	2.00	459,783	544	0.074
Attached 5-10	Infill East	6.00	1,096,000	637	0.087
Attached 5-10	Infill West	6.00	1,200,000	527	0.072
Attached 5-10	Rototuna	6.00	920,000	108	0.015
Attached ≥ 10	Infill East	13.00	1,965,000	142	0.019
Attached ≥ 10	Infill West	15.50	2,545,000	370	0.051

The result shows that the detached-single dwelling in Rototuna has the highest proportion of 0.204 whereas the attached 2-4 in Peacocke Stg 1 has the lowest proportion of 0.11.

The obtained proportions were used to distribute the new DCS data across the categories. For instance, if the monthly number of new DCS is 100, so the 0.204 of 100 dwellings consented are distributed to the Rototuna location as detached-single dwellings with the medium number and value per consent of 1.00 and $379,900, respectively. This method allows authors to prorate the data of new dwellings consented from the Sub-model and fed them into the candidate GBM to compute their construction period. Figure 10 shows the output of the CCC number to be issued of the new DCS data for both VECM-1 and VECM-2 from Sub-model 1 over the period of 2022-2031, denoted by CCC-VECM1 and CCC-VECM2, respectively.

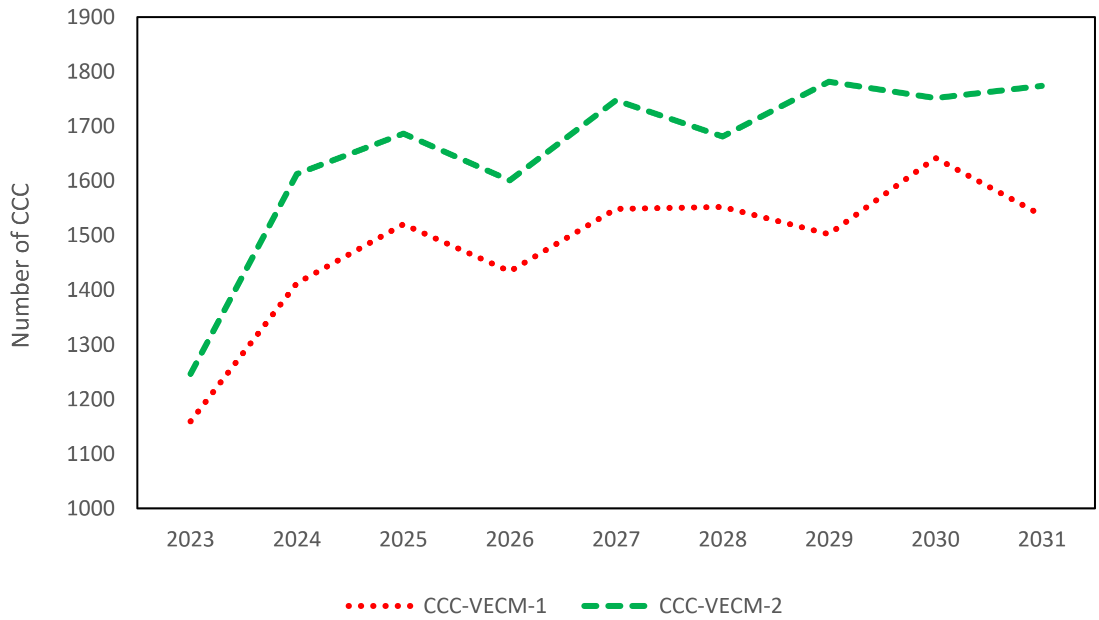

Figure 10 Results of Predicted Number of CCC from New Dwellings Consented

Figure 10 indicates that the number of new dwellings completed for CCC-VECM-1 and CCC-VECM-2 will be 1,159 and 1,246 in 2023 and gradually rise to 1,537 and 1,774 in 2031, respectively. The figure also shows that all dwellings consented in 2022 will have CCC issued from 2023.

4.2.4 Total Predicted CCC

The historical and predicted CCC number is shown in Figure 11. The difference between Figure 10 and 11 is the number of CCC for existing dwellings consented are added to that for new dwellings consented. The number of CCC for both CCC-VECM-1 and CCC-VECM-2 including both actual and predicted numbers will be 1,511 in 2022 and surges to 1,400-1800 between 2023-2031.

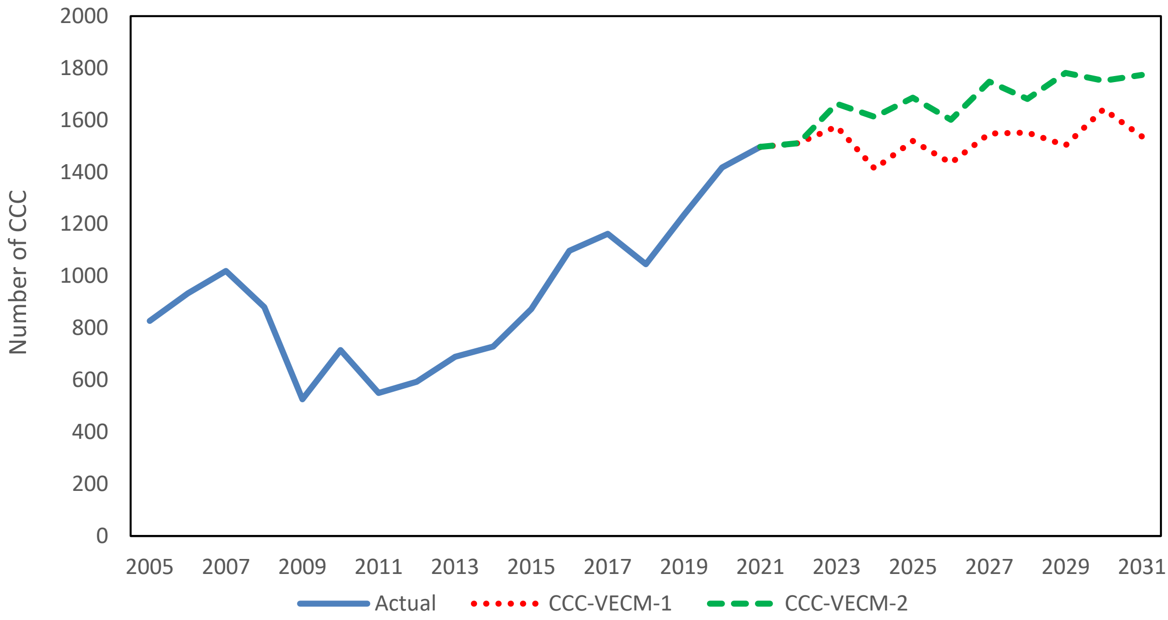

Figure 11 Results of Total Predicted Number of CCC

SECTION 5 - DISCUSSION AND POLICY IMPLICATION

Two sub-models are constructed by using a set of forecasting techniques. All the forecasting models generate reliable forecasts which are both close to the real world and achieve lower forecasting errors.

In Sub-model 1, the multivariate forecasting framework is adopted to predict the demand for dwellings consented. The VECM approach is employed to formulate the forecasting model, and it is considered accurate at capturing the movement of dwellings consented during the test-data period. Three VECMs are proposed and have MAPE scores between 9%-14%. VECM-1 which is the full model with the inclusion of all six variables (GDP, house price, land value, unemployment, interest rate, and population size) has the highest MAPE score (13.88%), whereas VECM-2 excluding the population variables achieves the lowest MAPE score (9.36%). This implies that the removal of the population variable improves the forecasting performance, but it does not necessarily mean that the population does not have a significant impact on the demand for new dwellings.

Ultimately, people populate houses, and population can never be entirely decoupled from dwelling demand. However, if the reliability of population variable is detracting from forecasting accuracy, it may be necessary to remove that variable. This is a question of data quality where, indeed, data of the population variable have been interpolated to a quarterly frequency. The interpolation might cause a loss of important information associated with the potential relationship between dwellings consented and population size, thus resulting in a higher MAPE. In addition, VECM-3 setting interest rate as an exogenous variable in the level-VAR representation of VECM-2 has a MAPE score of 10.69%, which is not very much different from that of VCEM-2.

The ex-ante forecasts of each model are constructed to compare the reliability of the model and to ensure that model will not predict the number of dwellings consented at a rate that population growth cannot satisfy. The ex-ante forecast is vitally important for the policymaker, in particular for HCC, to implement their infrastructure investment plans and urban growth strategies. This also means that a good model should be able to produce a realistic output that is practically achievable in the local housing industry. In Sub-model 1, VECM-1 tends to underestimate the dwelling demand compared with the recent movement in the housing market in Hamilton City where the GDP and house price has risen rapidly; and, with the new government policy promoting city intensification the housing market tends to be upsizing and expanding in the value and number of new consents. VCEM-2 is more likely able to generate the forecasts (1,500-1,800) closing to the city's expectations and matching with historical behaviour. It is also aligned with the available development capacity in the next decade.

On the other hand, VECM-3 seems to overestimate and has a distinctive pattern compared to the other two models. It has predicted the dwellings consented number to rise gradually in the next two years and then drop down slightly, albeit with an upward trend. This result seems to be overpredicted because the recent construction market is likely to be hit hard by New Zealand's tightened macroeconomic environment, including the rising interest rate and falling house prices, so an expected fall in the dwelling demand in the next few years is suggested.

For two potential candidate models, there are both advantages and disadvantages. VECM-1 is more theoretically sound as it takes account of the population variable, but it also becomes a limitation as the population variable has been interpolated and this could be a reason for the higher MAPE score. On the other hand, VECM-2 is more powerful as the MAPE score is lower but does not consider the population factor in the model. From an econometric aspect, it suggests that VECM-2 which uses original data and assumes the endogeneity for all variables (GDP, house price, land value, unemployment and interest rate) serves as a good candidate VECM for predicting the number of new dwellings consented in Hamilton City in the short and medium term. From theoretical and real-world perspectives, the population is an important factor in determining the demand for new dwellings, so the VECM-1 should be treated as another potential candidate model albeit with a higher MAPE score.

There is no such guarantee of which model will have the most accurate ex-ante forecast when the actual data for the future period are unavailable for comparison. It is noteworthy to keep it open for the choice of the models and re-evaluate once the actual data become available. Therefore, both VECM-1 and VECM-2 are considered to be potential forecasting models. The forecasted values from both models serve as two scenarios for future validation and improvement when future data become available. Also, as the VECM method is unable to predict the future shock and foresee the sudden change in macroeconomics, it is thus recommended to update the model regularly when needed. However, for VECM-1 the data of population estimates are updated less frequently and only available every five years via the Census. This suggests that VECM-2 is a more agile and reliable forecasting model as it can be updated more frequently.

In sub-model 2, the number of CCC is predicted by using the Gradient Boosting algorithm to train a model that can produce an accurate construction period. The GBM was obtained with an RSME of 0.58, which is reasonably acceptable. The GBM is then used to predict the CCC number from two components. First, the data of existing dwellings consented are fed into

GBM to produce the number of CCC. The result suggests that all the existing dwellings consented tend to get issued for CCC within two years; however, it does not necessarily mean that the construction period of each dwelling consented takes two years as some already have started construction years before, but all are predicted to get issued for CCC in the next two years.

For instance, some recent dwellings consented may take longer than two years to complete the construction. This could be due to the delay in material supply or COVID-19-related issues. In this case, the construction period trained by GBM will reflect the construction activity, e.g., if there are supply chain constraints or material shortages, this induces a longer construction period from consent granted to CCC issued, which can be reflected in the train data. Therefore, GBM utilises the train data to predict a construction period that is close to the real-world situation.

Second, the data of new dwellings consented from Sub-model 1 are fed into GBM to compute the number of CCC to be issued for those dwellings consented. To compute this, the phasing technique is introduced to convert quarterly or annually data into monthly data. Although this technique is naïve, it provides the best possible outcome that reflects the seasonal difference, while the local knowledge and judgement have been critically used in selecting the year used for phasing; this technique can be improved in the future.

The dwelling proportion by categories was calculated using historical CCC data and then used to prorate the new dwellings consented, but the future land supply in different growth locations is not considered in the computation of dwelling proportion, e.g., Rototuna capacity might not be available at a certain point in the future. The main purpose of this exercise is to capture the diversity of the construction length based on the different categories assuming the building construction under the most common urban growth patterns will carry this diversity in the long run. Since the algorithm is constructed to be re-trained on a monthly basis to include the latest month data, any major change in the construction industry will be captured gradually, e.g., new construction technology or pre-fabricated dwellings will have a much shorter construction time reflected during the algorithm training.

The final output from Sub-model 2 indicates that the number of CCC will have an upward trend over the forecasting period, implying increasing housing demand, especially in the short and medium term. Therefore, the housing supply and infrastructure capacity planning need to take account of this higher dwelling forecast.

The proposed forecasting methodology can serve as a reasonably reliable approach for forecasting housing demand in Hamilton City as well as other regions. The preliminary ex-ante forecast results provide a valuable contribution to local government in examining the future local housing market growth and especially in designing effective infrastructure investment plans and urban growth strategies. Practically, it will be used internally as an input to HCC's growth modelling environment, including the Growth Model, the Three-Waters Models, the Transport Model, the Development Contribution Model and the Asset Management Models. Furthermore, these models provide fundamental inputs to Council's long-term planning, infrastructure investment timing, land use and zoning decisions, and to its funding and financing activities.

The proposed models can also serve as a foundation for forecasting dwelling demand in other regions, in particular for other local councils, while a number of modifications are suggested based on characteristics of the market, purposes of research, location, data availability and local policies and regulations.

Nevertheless, amid unforeseeable changes in the economic environment and with more data available in the future, the model should be updated at various times such that those unforeseeable economic impacts can be fully accounted for, and more reliable forecasts can be achieved. HCC will review and update the model at different timeframes: Sub-model 1 will be reviewed and updated annually to capture recent economic conditions, and Sub-model 2 will be updated monthly when monthly data are available as this enables models to capture the latest situation in the construction market. It is important to note that the forecast results provided in this study are solely estimated based on the proposed methodology using the data input available and collected at the time of the study conducted, so the forecast results may vary when using a different approach and dataset in the future review.

SECTION 6 - CONCLUSION

This study was motivated by the need for forecasting models that can predict the number of dwellings consented and completed in the short and medium term for Hamilton City. This study delivers a methodology framework using a combination approach. It consists of two sub-models using different forecasting techniques and combining the outputs to generate the forecasts of dwellings completed.

In Sub-model 1, an econometric framework was developed by using VECMs with a set of economic variables that can explain the demand for dwellings consented. The candidate models were evaluated using MAPE scores. VECM-1 which is the full model including the population variable has the highest MAPE score of 13.88%, while VECM-2 removing the population variable from the model has the lowest MAPE score of 9.36%. VECM-3 assuming the exogeneity function of the interest rate variable has a MAPE score of 10.69%. VECM-1 and VECM-2 are considered potential candidate models based on econometric and practical perspectives. The capabilities of these models are showcased by applying their estimation to generating the ex-ante forecasts of new dwellings consented. The forecasting models in Sub-model 1 have been validated to be capable of generating a reliable dwelling consented projection.

Sub-model 2 was developed to forecast the number of dwellings completed using the consent data and output from Sub-model 1. A gradient boosting algorithm was implemented to predict the timeframe of each consent to be completed based on several factors in the consent information, including location, typology, dwelling number and value. The model performance was evaluated by using RMSE, showing that the error was acceptably low. The forecasts were combined from two components: the first component was constructed on the existing dwellings consented in the pipeline, and the second was built on the future dwellings consented which are the outputs from two VECM models, with and without a population variable. The final forecast output of Sub-model 2 has been shown reasonably acceptable and practically useful.

Overall, the proposed combination approach in this study is robust and applicable as it encompasses a wide array of modelling techniques and utilises all available data in both local and national statistics. The methodology developed in this study provides valuable practical use for HCC to forecast future housing demand which serves as an input in designing effective development strategies and planning. Furthermore, it can be used as a research direction for other local governments in developing their own dwelling projection. Notwithstanding this, this study remains subject to further improvement in modelling that can incorporate more economic variables and use a longer span of data period. A comparison of using different machine learning techniques is recommended in future studies.

SECTION 7 - REFERENCE

Abidin, S.N., & Jaffar, M.M. (2014). Forecasting share prices of small size companies in Bursa Malaysia using Geometric Brownian Motion. *Applied Mathematics & Information Sciences*, 8, 107 112.

Akintoye, A., & Skitmore, M. (1994). Models of UK private sector quarterly construction demand. *Construction Management and Economics*, 12, 3-13.

Anderson, R. G., Hoffman, D. L., & Rasche, R. H. (2002). A vector error-correction forecasting model of the US economy. *Journal of Macroeconomics*, 24(4), 569-598.

Asteriou, D., & Hall, S. (2011). *Applied Econometrics*. Palgrave Macmillan, 265-286.

Ayub, S., & Jafri, Y.Z. (2020). Comparative study of an ANN-ARIMA hybrid model for predicting Karachi stock price. *American Journal of Mathematics and Statistics*, 10, 1-9.

Babu, C.N., & Reddy, B.E. (2014). A moving-average filter-based hybrid ARIMA-ANN model for forecasting time series data. *Applied Soft Computing*, 23, 27-38.

Baltagi, B. (2011). *Econometrics*. Springer Berlin, Heidelberg. 10.1007/978-3-642-20059-5.

Bon, R. (1989). *Building as an Economic Process*. Prentice Hall, Englewood Cliffs, New Jersey.

Bork, L., & Moller, S. (2016). Housing price forecastability: A factor analysis. *Real Estate Economics*, 46(3).

Box, G. E. P., & Jenkins, G. M. (1976). *Time series analysis: Forecasting and contr*ol. San Francisco: Holden-Day.

Breusch, T.S. (1978). Testing for autocorrelation in dynamic linear models. *Australian Economic Papers*, 17, 334-355.

Briscoe, G. (1988). *The Economics of the Construction Industry*. Mitchell, London.

Büyüksahin, Ü.C., & Ertekin, S. (2019). Improving forecasting accuracy of time series data using a new ARIMA-ANN hybrid method and empirical mode decomposition. *Neurocomputing*, 361.

Chan, S. L. (2001). Empirical tests to discern linkages between construction and other economic sectors in Singapore. *Construction Management and Economics*, 19(4), 355-363.

Chatfield, C. (2000). *Time-Series Forecasting*. Chapman and Hall/CRC.

Dickey D. A., & Fuller W. A. (1979). Distribution of the estimators for autoregressive time series with a unit root. *Journal of the American Statistical Association*, 74(366a), 427-431.

Drought, S., & McDonald, C. (2011). Forecasting house price inflation: a model combination approach. *Reserve Bank of New Zealand. Discussion Paper Series, ISSN 1177-7567*.

Engle, R. F., & Granger C. W. (1987). Co-integration and error correction: Representation, estimation, and testing. *Journal of the Econometric Society,* 55(2), 251-276.

Engle, R. F. (1982). Autoregressive conditional heteroscedasticity with estimates of the variance of United Kingdom inflation. *Econometrica*, 50(4), 987–1007.

Eswaran, C., & Logeswaran, R. (2012). An enhanced hybrid method for time series prediction using linear and neural network models. *Applied Intelligence*, 37(4), 511–519.

Faghih, S. A.M., Gholipour, Y. & Kashani, H. (2021). Time series analysis framework forforecasting the construction labor costs. *KSCE Journal of Civil Engineering*, 25, 2809–2823.

Faghih, S. A. M., & Kashani, H. (2018). Forecasting construction material prices using vector error correction model. *Journal of Construction Engineering and Management*, 144.

Fan, R. Y. C., Ng, S. T., & Wong, J. M. W. (2010). Reliability of Box-Jenkins model for forecasting Hong Kong construction demand covering times of economic austerity. *Construction Management and Economics*, 28(3), 241-254.

Fan, R., Ng, S., & Wong, J. (2011). Predicting construction market growth for urban metropolis: An econometric analysis. *Habitat International*, 35, 167-174.

Figueiredo, E., Figueiras, J., Park, G., Farrar, C.R., & Worden, K. (2011). Influence of the autoregressive model order on damage detection. *Computer-Aided Civil and Infrastructure Engineering*, 26(3), 225–38.

Friedman, J. H. (2001). Greedy function approximation: A Gradient Boosting machine. *The Annals of Statistics*, 29(5), 1189–1232.

Gholamhosseini, H., Luo, D., & Reynolds, K. (2006). The comparison of different feed forward neural network architectures for ECG signal diagnosis. *Medical Engineering & Physics*, 28, 372-378.

Godfrey, L. G. (1978). Testing for higher order serial correlation in regression equations when the regressors include lagged dependent variables. *Econometrica*, 46(6), 1303–1310.

Gujarati, D.N. (1995). *Basic Econometrics (3rd Ed.)*. McGraw-Hill International Editions. New York.

Gupta, R., Kabundi, A., &Miller, S.M(2011). Forecasting the US real house price index: Structural and non-structural models with and without fundamentals. *Economic Modelling*, 28(4), 2013–2021.

Gyourko, J. & Saiz, A. (2006) Construction costs and the supply of housing structure. *Journal of Regional Science*, 46(4), 661–80.

Hallac, D., Leskovec, J., & Boyd, S. (2015). Network Lasso: Clustering and Optimization in Large Graphs. KDD: Proceedings/International Conference on Knowledge Discovery & Data Mining. *International Conference on Knowledge Discovery & Data Mining*, 2015, 387–396.

Hamilton, J. (1994). Time Series Analysis. Princeton University Press, New Jersey.

Hampson, K., & Brandon, P. (2004). Construction 2020: A vision for Australia's property and construction industry.

Hillebrandt, P. (1984). *Analysis of the British Construction Industry*. London: Macmillan.

Hindle, R. D. (1993). The effects of the short-term business cycle on the construction industry. Paper presented at the CIB W55 Conference, Lisbon, September 1993.

Hua, G. B. (2000). Evaluating the performance of combining neural networks and genetic algorithms to forecast construction demand: the case of the Singapore residential sector. *Construction Management and Economics*, 18(2), 209-217.

Hua, G. B. (2012). Modelling sectoral construction demand and its relationship with economic indicators. *International Journal of Construction Education and Research*, 8, 223-240.

Hua, G. B., & Pin, T. H. (2000). Forecasting construction industry demand, price and productivity in Singapore: The Box-Jenkins approach. *Construction Management and Economics*, 18(5), 607–618.

Hua, G. B. (1996). Residential construction demand forecasting using economic indicators: a comparative study of artificial neural networks and multiple regression. *Construction Management and Economics*, 14(1), 25–34.

Jarque, C. M., & Bera, A. K. (1987). A test for normality of observations and regression residuals. In*ternational Statistical Review,* 55(2), 163–172. https://doi.org/10.2307/1403192

Jiang, B., Athanasopoulos, G., Hyndman, R. J., Panagiotelis, A., & Vahid, F. (2017). Macroeconomic forecasting for Australia using a large number of predictors. *Monash Econometrics and Business Statistics Working Paper No. 2/17*. Monash University, Department of Econometrics and Business Statistics.

Jiang, H., & Liu, C. (2014). A panel vector error correction approach to forecasting demand in regional construction markets. *Construction Management and Econo*mics, 32(12).

Jiang, H. & Liu, C. (2011). Forecasting construction demand: a vector error correction model with dummy variables. *Construction Management and Economics*, 29(9), 969–79.

Jiang, H., Xu, Y., & Liu, C. (2013). Construction price prediction using vector error correction models. *Journal of Construction Engineering and Management*, 139(11).

Johansen, S. (1988). Statistical analysis of cointegration vectors. *Journal of Economic Dynamics and Control*, 12, 231-254.

Johansen, S. (1991). Estimation and hypothesis testing of cointegration vectors in Gaussian vector autoregressive models. *Econometrica*, *59(6),* 1551–1580.

Johansen, S., & Juselius, K. (1990). Maximum likelihood estimation and inference on cointegration with applications to the demand for money. *Oxford Bulletin of Economics and Statistics, 52*, 169-210.

Johansen, S., Mosconi, R., & Nielsen, B. (2000). Cointegration analysis in the presence of structural breaks in the deterministic trend. *The Econometrics Journal, 3*, 216-249.

Khashei, M., & Bijari, M. (2010). An artificial neural network (p, d, q) model for timeseries forecasting. *Expert System with Applications, 37*, 479-489.

Khashei, M., & Bijari, M. (2011). A new hybrid methodology for nonlinear time series forecasting. *Modelling and Simulation in Engineering.*

Kim, J. M., Cho, C., & Jun, C. (2022). Forecasting the price of the cryptocurrency using linear and nonlinear error correction model. *Journal of Risk Financial Management*, *15(2)*, 74.

Kuzin, V., Marcellino, M., & Schumacher, C. (2013). Pooling versus model selection for nowcasting GDP with many predictors: Empirical evidence for six industrialized countries. *Journal of Applied Econometrics, 28(3)*, 392–411.

Lewis, C.D. (1982). *Industrial and business forecasting methods*. London: Butterworths.

Lewis, T. M. (2009). *Quantifying the GDP-construction relationship*. In L. Ruddock (Ed.), Economics for the modern built environment (pp. 34-59). London: Taylor & Francis.

Makridakis, S., Wheelwright, S. C., & McGee, V. E. (1983). *Forecasting: Methods and Applications.* Wiley, New York.

Meen, G. (1996). Spatial aggregation, spatial dependence and predictability in the UK housing market. *Housing Studies*, *11(3),* 345–72. Knight, K., & Fayek, A. R. (2002). Use of fuzzy logic for predicting design cost overruns on building projects. *Journal of Construction Engineering and Management*, *128(6),* 503–12.

Merkies, A. & Poot, H. J. (1990). A case-study of forecasting building activity in the Netherlands and New Zealand, *The CIB90 Conference on Building Economics and Construction Management*. Sydney, Australia, 201-216.

Cameron P. M., & Cochrane W. (2021). 2018-base Population, Family and Household, and Labour Force Projections for the Waikato Region, 2018-2068. Waikato Regional Council Technical Report 2021/22.

Milunovich, G. (2020). Forecasting Australia's real house price index: A comparison of time series and machine learning methods. *Journal of Forecasting, 39*, 1098-1118.

Moon T., & Shin D. (2018). Forecasting model of construction cost index based on VECM with search query. *KSCE Journal of Civil Engineering, 22(8),* 2726-2734.

Morikawa, M. (2020). Uncertainty in long-term macroeconomic forecasts: Ex post evaluation of forecasts by economics researchers. *The Quarterly Review of Economics and Finance.*

Ng, S. T., Fan, R. Y. C. & Wong, J. M. W. (2011). An econometric model for forecasting private construction investment in Hong Kong. *Construction Management and Economics, 29(5),* 519 534.

Nie, P., & Roccotelli, M., Fanti, M., & Ming, Z. F., & Li, Z. (2021). Prediction of home energy consumption based on gradient boosting regression tree. *Energy Reports, 7*, 1246-1255.

Ofori, G. (1990). *The Construction Industry: Aspects of its Economics and Management*, Singapore University Press, National University of Singapore.

Ofori, G., & Han, S. S. (2003). Testing hypotheses on construction and development using data on China's provinces, 1990-2000. *Habitat International, 27(1),* 37-62.

Parot, A., & Michell, K., & Kristjanpoller, W. (2019). Using artificial neural networks to forecast exchange rate, including VAR-VECM residual analysis and prediction linear combination. *Intelligent Systems in Accounting, Finance and Management, 26*.

Patterson, D. W. (1996). *Artificial neural networks - theory and applications*.

Pfaff, B. (2008). VAR, SVAR and SVEC Models: Implementation Within R Package VARS. *Journal of Statistical Software*, *27*(4), 1–32.

Rapach, D. E., & Strauss, J. K. (2007). Forecasting real housing price growth in the Eighth District states. *Regional Economic Development*, *3*(2), 33–42.

Said E., & Dickey, A. (1984). Testing for unit roots in autoregressive-moving average models of unknown order. *Biometrika*, *71*(3), 599–607.

Shahandashti S. M., & Ashuri, B. (2013). Forecasting engineering news-record construction cost index using multivariate time series models. *Journal of Construction Engineering and Management, 139(9)*, 1237-1243.

Stock, J. H., & Watson, M. W. (2012). Generalized shrinkage methods for forecasting using many predictors. *Journal of Business & Economic Statistics*, *30(4),* 481–493.

Sumathi, S., & Paneerselvam, S. (2010). Computational Intelligence Paradigms: Theory & Applications using MATLAB.

Tan, H. B., & Baharumshah, A. Z. (1999). Dynamic causal chain of money, output, interest rate and prices in malaysia: evidence based on vector error- correction modelling analysis. *International Economic Journal, 13,* 103-120.

Tan, W. (1989). Subsector fluctuations in construction. *Construction Management and Economics, 7*, 41-51.

Tang, J. C. S., Karasudhi, P., & Tachopiyagoon, P. (1990). Thai construction industry: demand and projection. *Construction Management and Economics, 8(3),* 249–257.

Timmermann, A. (2006). *Chapter 4 Forecast Combinations.* Handbook of Economic Forecasting, 1, 135-196.

Truong, Q., Nguyen, M., Dang, H., & Mei, B. (2020). Housing price prediction via improved machine learning techniques. *Procedia Computer Science, 174*, 433-442.

Tse, R. Y. C., Ho, C. W. & Ganesan, S. (1999). Matching housing supply and demand: an empirical study of Hong Kong's market. *Construction Management and Economics, 17(5),* 625–33.

Ulke, V., & Sahin, A., & Subasi, A. (2018). A comparison of time series and machine learning models for inflation forecasting: empirical evidence from the US. *Neural Computing and Applications*.

van Aarle, B., Boss, M., & Hlouskova, J. (2000). Forecasting the euro exchange rate using vector error correction models. *Review of World Economics, 136(2),* 232–258.

Wang, L., Zou, H., Su, J., Li, L. & Chaudhry, S. (2013). An ARIMA-ANN hybrid model for time series forecasting. *System Research and Behavioural Science,* 30(3), 244-259.

Wong, J. M., Chan, A., & Chiang, Y. H. (2007). Forecasting construction manpower demand: A vector error correction model. *Building and Environment, 42*, 3030-3041.

Wong, J. & Ng, S. (2010). Forecasting construction tender price index in Hong Kong using vector error correction model. *Construction Management & Economics, 28,* 1255-1268.

Xu J. W., & Moon S. (2013). Stochastic forecast of construction cost index using a cointegrated vector autoregression model. *Journal of Management in Engineering 29(1),* 10-18.

Yoon, J. (2021). Forecasting of real GDP growth using machine learning models: Gradient boosting and random forecast approach. *Computational Economics, 57*, 247–265.

Yu, L., Wang, S., & Lai, K. K. (2009). A neural-network-based nonlinear metamodeling approach to financial time series forecasting. *Applied Soft Computing, 9*, 563-574.

Zhang, P. (2003). Time series forecasting using a hybrid ARIMA and neural network model. *Neurocomputing, 50*, 159-175.

FURTHER INFORMATION

Hamilton City Council
Garden Place, Private Bag 3010, Hamilton

/HamiltonCityCouncil

@CouncilHamilton

07 838 6699

hamilton.govt.nz