

## FORECASTING SPATIALLY CORRELATED TARGETS: SIMULTANEOUS PREDICTION OF HOUSING MARKET ACTIVITY ACROSS MULTIPLE AREAS

Changro LEE\*

*Department of Real Estate, Kangwon National University, 1 Kangwondaehak-gil,  
Chuncheon, 24341, Gangwon-do, Republic of Korea*

Received 08 November 2021; accepted 27 March 2022

**Abstract.** This study involved the development of an approach to forecast house prices and trading volumes across multiple areas simultaneously. Spatially correlated targets, such as house prices, can be predicted more accurately by leveraging the correlations across adjacent areas. A multi-output recurrent neural network, a deep learning algorithm specifically developed to analyze sequence data, was utilized to forecast the house prices and trading volumes in the four chosen study areas. The forecasting accuracy of future house prices in one of the four geographical areas clearly improved; this area was found to be a price-lagging area, and the forecasting accuracy of this area significantly increased by exploiting the information of a price-leading area. As for the prediction of trading volumes, the difference in performance between the multi-output recurrent neural network and conventional models was very small. The results of this study are expected to promote the use of deep learning to predict the housing market activity through a simultaneous forecasting framework.

**Keywords:** multi-output neural network, simultaneous prediction, correlation, house price, trading volume.

### Introduction

Prices and trading volumes in the real estate market are two fundamental issues that have been actively studied and reported in the literature. Rising house prices imply a decrease in the affordability of housing for tenants, but encourage homeowners to spend more, thereby stimulating the economy. In contrast, falling house prices increase the risk of houses being valued less than the homeowners' outstanding mortgages, increasing the probability of default. Hence, monitoring the real estate market from various perspectives has always been a crucial matter for both policymakers and households. Trading volumes, which indicate the intensity of trading activity in the real estate market, are another important aspect of the market. An increase in trading volumes means a rising market, and a decrease in trading volumes is expected to result in a falling market. Accurate forecasting of trading volumes can help the government gauge housing demand in the future, enabling it to design housing supply policies in a relevant manner. The ability to accurately predict trading volumes could also bring valuable business opportunities to the attention of real estate developers and constructors.

However, empirical predictions of prices and trading volumes have been mainly considered in economics under the time-series model framework. Property prices and trading volumes are spatial phenomena, and thus need to be analyzed from the perspective of geographical association, that is, the spatial correlation of the real estate market activity should be fully utilized to enhance predictive accuracy. In this study, spatially correlated targets are forecast using deep learning. Property prices and trading volumes in the housing market of South Korea are predicted using a recurrent neural network (RNN). In particular, a multi-output RNN that can predict prices or trading volumes in multiple areas simultaneously is exploited. The results are compared with those obtained with a traditional time series model and a single-output RNN.

A neural network is a de facto standard model for implementing deep learning approaches and is widely used in various application areas, from computer vision to natural language processing. Since neural networks specifically developed for analyzing sequence data, that is, RNNs, were proposed by the academic community, correlated targets in attribute space have been extensively

\*Corresponding author. E-mail: [spatialstat@naver.com](mailto:spatialstat@naver.com)

predicted by this class of networks. The simultaneous prediction of the cost of healthcare and the length of the hospital stay is a typical example, as is evident from the healthcare literature. However, to the best of the authors' knowledge, correlated targets in physical space have not been predicted with the aid of an RNN framework. This study attempts to make predictions for geographically correlated targets by exploiting the multi-output architecture of a neural network. The results of this study are expected to promote the application of deep learning algorithms for forecasting spatially correlated phenomena.

The remainder of this paper is organized as follows. Section 1 explains the spatial autocorrelation and spillover effects, and a deep learning approach for predicting targets with these properties. Section 2 describes the chosen study areas, dataset, and neural network architecture specifically developed for forecasting house prices and trading volumes. Section 3 presents the results and implications of this study. Finally, a summary and the limitations of the study are provided in the last section.

## 1. Literature review

### 1.1. Spatial autocorrelation and spillover effects

Spatial autocorrelation and spillovers are well-known characteristics of spatial phenomena. Spatial autocorrelation can be succinctly described by the first law of geography (Tobler, 1970): everything is related to everything else, but near things are more related than distant things. That is, spatial autocorrelation indicates that the values of a variable located in geographic areas that are close to each other show a similar pattern (Ismail, 2006), thus, the information in nearby areas can help to predict the values in a target area. Spatial spillover indicates that a local trend in one area can spread to nearby areas promptly, and this effect has frequently been investigated in studies on the effects of infrastructure development (Yilmaz et al., 2002; Tong et al., 2013; Xie et al., 2016).

Spatial regressions are well-known models in which spatial autocorrelation and spillovers among areal units are specified using a spatial weight matrix (Moscone & Tosetti, 2014). These regressions have been extensively utilized in the literature for investigating spatial issues such as land use, air pollution, crime geography, population movement, and regional economic growth (Griffith & Wong, 2007; Fang et al., 2015; Ahmar & Aidid, 2018; Dai et al., 2018).

Because prices and trading volumes in the real estate market are also spatial phenomena, these two effects have been subjected to extensive analysis, as reported in the literature. In other words, the values in adjacent areas have been vigorously exploited to accurately predict house/land prices and trading volumes in the target area. Conway et al. (2010) analyzed the greenspace contribution to housing prices by incorporating spatial autocorrelation into a pricing model. DeFusco et al. (2018) estimated the extent to which the American housing boom

spread from one market to another via spatial spillovers. Yang et al. (2018) also investigated 69 Chinese cities by applying a spillover index and concluded that city-level monthly housing prices are highly interdependent. As for forecasting trading volumes, Lee et al. (2019) predicted the volumes of apartments in South Korea, and Temur et al. (2019) forecasted the house trading volume in Turkey by considering spatial autocorrelation or spillover effects. Other similar studies utilizing spatial autocorrelation/spillover effects have been reported in the literature (Rambaldi & Rao, 2011; Cellmer, 2013; Taltavull de La Paz et al., 2017; Wang et al., 2019).

This study follows the prediction approach adopted in previous studies and attempts to predict prices and trading volumes in a target area by employing the values of nearby areas.

### 1.2. Deep learning approach and a multi-output RNN

Deep learning approaches have been applied in the real estate industry since the AI boom that occurred roughly in 2010, to detect building defects (Perez et al., 2019), quantify urban qualities such as the visual impression of a neighborhood using street images and satellite images (Law et al., 2019), and identify fraud in property acquisition tax returns (Lee, 2021).

In the above-mentioned studies, in which a deep learning approach was utilized, the de facto standard implementation model was a neural network. This model consists of several layers capable of learning the nonlinear relationship inherent in the input data and demonstrated excellent performance in processing unstructured data such as images and free-form texts. Predicting prices and trading volumes is no exception to this line of research, and many studies have shown that neural networks outperform conventional models in monitoring and forecasting the real estate market. For example, Poursaeed et al. (2018) and Kang et al. (2021) fed imagery data to neural networks for the analysis of the real estate market. Poursaeed et al. (2018) used a large dataset of photos of house interiors and exteriors to predict house prices. Kang et al. (2021) used street-view images and house photos to model house price appreciation. In addition to imagery data, free-form texts have also been actively utilized recently, whereby studies on free-form texts are often referred to as sentiment analysis. Beracha et al. (2019) and Shi et al. (2021) fed news articles and social media messages into neural networks to estimate price trends for future real estate markets, thereby demonstrating that media-expressed sentiment can help predict real estate returns. Overall, there is a noticeable trend in real estate market analysis of the exploitation of multi-source data, as evidenced by image analysis and sentiment analysis within a neural network framework (Zou & Wang, 2021).

However, most studies in which neural networks were applied to the real estate market have focused on single-output prediction. That is, they attempted to forecast prices or trading volumes in a single study area. Few studies

have predicted the activities of the real estate market in multiple areas simultaneously. One attractive benefit of a neural network is specification flexibility (Lakshmanan et al., 2020); thus, it can be designed to produce multiple outputs simultaneously by using a shared learning process. With correlated targets, the multi-output model enables more accurate prediction by sharing parameters across different tasks (Bakker & Heskes, 2003; Ben-David & Schuller, 2003; Li et al., 2019; Xu et al., 2019).

In contrast to the paucity of real estate literature on multi-output prediction, the utilization of multi-output neural networks is not rare in the healthcare industry, and Futoma et al. (2017) modeled clinical time series with a multi-output neural network for early sepsis detection. Cui et al. (2018) simultaneously predicted the cost of healthcare and the length of the hospital stay by leveraging the correlations between the two measures. He et al. (2021) designed a neural network with multi-output architecture to simultaneously predict inpatient flow and the length of stay because they share common characteristics such as the recovery status and surgery types. Briefly, healthcare researchers have long attempted to predict the consumption of healthcare resources by linking related variables together through a multi-task forecasting framework.

In the real estate literature, few studies on multi-output prediction have been reported, which motivated us to employ a multi-output neural network for predicting prices and trading volumes in several areas simultaneously. Four study areas were chosen for the analysis in this study, and the Pearson correlation coefficients of prices between the areas are in the range 0.73–0.95, and those of the trading volumes are 0.21–0.75. The magnitude of these coefficients indicates moderate to strong relations between the study areas, and seems to suggest that both multi-inputs and multi-outputs should be utilized for forecasting. That is, the values in the target and adjacent areas are used as the input, and the values in these areas are predicted simultaneously.

As demonstrated in the panel on the left in Figure 1, prior studies attempted to predict a target area using information from adjacent areas, including the target area. In contrast, this study forecasts multiple target areas using information from adjacent areas, including the target area, as depicted in the panel on the right in the figure.

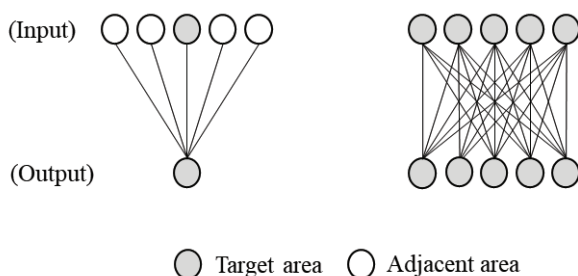


Figure 1. Prediction of the target and adjacent areas

## 2. Methodology

### 2.1. Study area and dataset

South Korea comprises eight provinces and nine metropolitan cities. Two provinces and two metropolitan cities were chosen for the analysis: Gyeonggi Province, Gangwon Province, Seoul City, and Incheon City. As shown in Figure 2, the four study areas are located around the capital city, Seoul, and this region is often referred to as the *capital region*. This region is frequently highlighted in the mass media because of the recent rapid rise in housing prices and trading volumes. Hence, the region has always been an area of intensive interest to both the government and private real estate investors, and thus was selected for the study.<sup>1</sup>

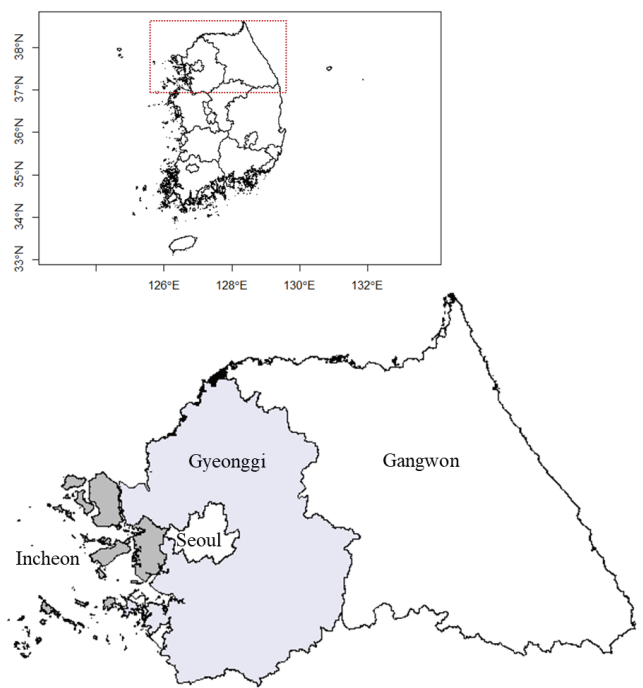


Figure 2. Study area

The two datasets that were used in the study were collected from a public website operated by Statistics Korea: house price data and house trading volume data. These data have been publicly disclosed on a monthly basis since January 2006, when the *Act on Reporting Real Estate Transactions* was enforced. Both the price and trading volume data cover the 186 months from January 2006 to June 2021.

Figures 3 and 4 show the monthly trend in house prices and trading volumes for these 186 months starting in January 2006. As indicated in Figure 3, the prices tended

<sup>1</sup> Local governments in South Korea can reinforce their own real estate policies such as property tax relief and adjustment of tax rates, thereby shaping administrative division-based submarkets. In addition, the datasets used in the study were published at the scale of provinces and metropolitan cities. Considering these, the current administrative division was adopted for the market analysis.

to increase after the 150<sup>th</sup> month (July 2018). The soaring house prices during this period were commonly observed in other countries as well, and they have been increasing at a faster rate since the start of the COVID-19 pandemic. In contrast to the behavior of house prices, the trading volumes in Figure 4 do not show any noticeable upward or downward tendency during the study period. The data pertaining to the latter part of the study period (the 18 months from January 2020 to June 2021) was reserved to test the performance of the model.

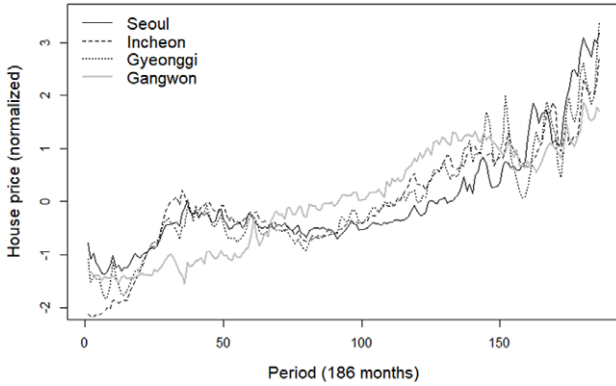


Figure 3. Fluctuations in house prices between Jan. 2006 and Jun. 2021

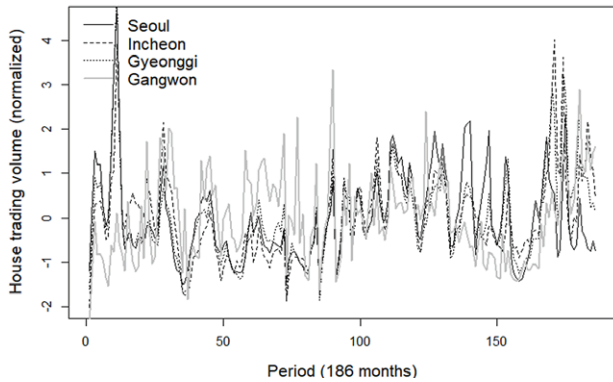


Figure 4. Fluctuations in house trading volumes between Jan. 2006 and Jun. 2021

### 2.2. Specification of RNN

RNNs are a class of neural networks that is specifically designed for learning from time-series data. RNNs are characterized by a recurrent neuron that produces outputs and sends them back to the neuron itself, thereby helping the RNN memorize past patterns of the time series data (Géron, 2019). An RNN was employed for the analysis because the dataset used in this study contains typical time-series data.

More specifically, a long short-term memory network (LSTM) is utilized to predict house prices and trading volumes. LSTM is a special class of RNN capable of back-propagating the error more efficiently than other types of RNNs. The theoretical concept underpinning

RNNs was proposed by Hochreiter and Schmidhuber (1997), and is extensively used for learning patterns from sequential data.

The length of the time series used in the study was 186 months, which is relatively short and may be insufficient to enable the relevant pattern to be learnt from the data. Hence, a multi-output LSTM was employed to mitigate this weakness. By leveraging the correlations across multiple target values, the multi-output LSTM can complement the problem resulting from the use of a small dataset. The underlying expectation here is that there exist common drivers related to the fluctuations in prices and trading volumes that can be shared while the neural network is in the process of being trained.

Figure 5 shows the architecture of the neural network used in this study. The input values of the four study areas were fed into an initial dense layer and then forwarded to an LSTM layer followed by a dropout layer. The LSTM layer is designed to learn the entire sequence of data. The dropout layer randomly sets previous neurons to zero at a certain ratio (20% in this study) to prevent overfitting. Then, the combination of the LSTM and dropout layers is repeated twice more, and finally, a multi-output dense layer is added to the architecture. The implementation details are as follows: The lookback period was specified to be 12 months. An adaptive moment estimation (Adam) optimizer and Glorot initialization with a uniform distribution were used. The multi-output LSTM was trained for 200 epochs with the mean squared error (MSE) adopted as the loss function, and the batch size was 32.

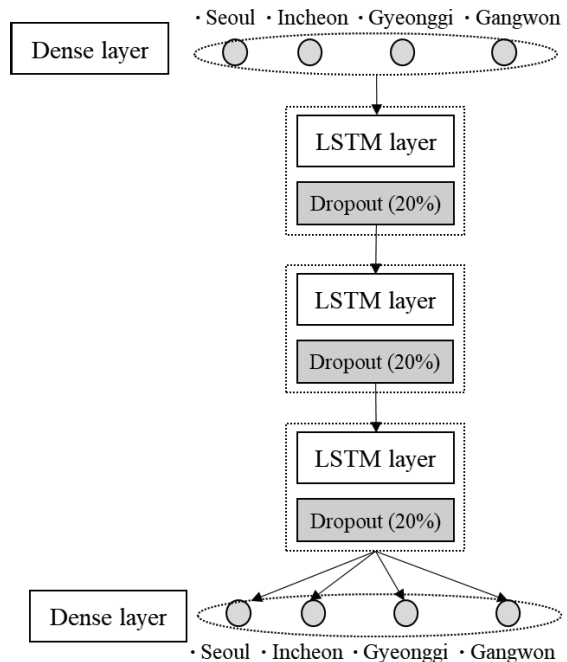


Figure 5. Neural network architecture

### 3. Discussion

#### 3.1. Results

Table 1 presents the predictive accuracy of house prices obtained from the test data covering the period from January 2020 to June 2021 (18 months). For comparison purposes, the results from an autoregressive integrated moving average (ARIMA) model and a single-output LSTM are provided in the table. The MSE, which was used as a metric to compare the performance of the models, was calculated using the following equation:

$$MSE = \frac{1}{n} \sum_{i=1}^n (y - \hat{y})^2, \tag{1}$$

where:  $y$  denotes the observed value and  $\hat{y}$  denotes the estimated value from the models. The results in Table 1 indicate that the three models perform consistently with little difference for each area that was studied except for Gangwon Province. The MSEs for this province decreased dramatically, from 0.02–0.03 to less than 0.01.

Table 1. MSEs for the house price predictions

| Prediction type         | Model | Seoul | Incheon | Gyeonggi | Gangwon                |
|-------------------------|-------|-------|---------|----------|------------------------|
| Single prediction       | ARIMA | 0.15  | 0.04    | 0.03     | 0.02                   |
|                         | LSTM  | 0.17  | 0.04    | 0.05     | 0.03                   |
| Simultaneous prediction | LSTM  | 0.18  | 0.03    | 0.04     | <0.01<br>(0.004-0.006) |

The improvement in the predictive accuracy for Gangwon Province is clearly confirmed in Figure 6, where the predicted values from the multi-output LSTM closely approximate the observed values in the 18-month period between January 2020 and June 2021. In contrast, there exists an apparent divergence between the observed values and the predicted values from the ARIMA or the single-output LSTM.<sup>2</sup>

Table 2 presents the predictive accuracy for the trading volume of houses. In contrast to the prediction of house prices, the performance of the three models differed little for predicting the trading volumes in all the study areas, which is consistent with the magnitude of the correlation

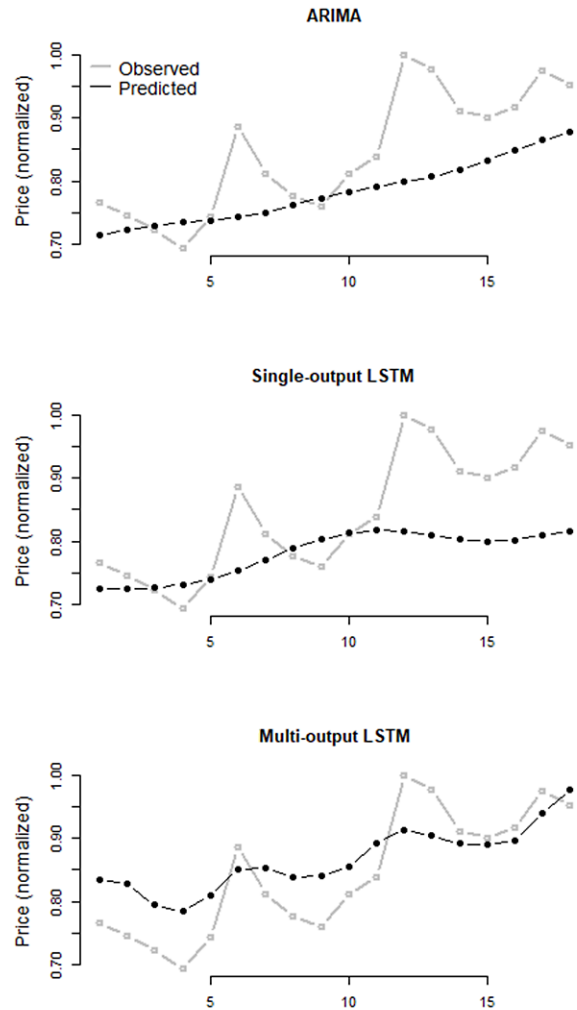


Figure 6. Goodness-of-fit for house prices in Gangwon Province

coefficients explained earlier.<sup>3</sup> The prediction results for house prices are discussed in detail hereafter.

Table 2. MSEs for the house trading volume predictions

| Prediction type         | Model | Seoul | Incheon | Gyeonggi | Gangwon |
|-------------------------|-------|-------|---------|----------|---------|
| Single prediction       | ARIMA | 0.03  | 0.07    | 0.04     | 0.04    |
|                         | LSTM  | 0.03  | 0.07    | 0.05     | 0.05    |
| Simultaneous prediction | LSTM  | 0.02  | 0.08    | 0.06     | 0.03    |

#### 3.2. Implications

The multi-output LSTM demonstrated outstanding performance in the prediction of house prices in Gangwon Province, as is clear from Table 1. In the initial stage of

<sup>2</sup> As indicated in Figure 6, models including LSTMs suffer from capturing the cyclic pattern in house prices. This deficiency can be attributed to the small size of the dataset (186 months in this study). Thus, this problem can be mitigated by a sufficient increase in the dataset size. If this option is not available, incorporating a temporal correlation manually into the model component can be an alternative. For example, a temporal correlation of data can be estimated using ARIMA and employed as an additional input variable into LSTM. Although this combinational approach is beyond the scope of this study, it is worth exploring as a hybrid method in a future study.

<sup>3</sup> The correlation coefficient was 0.73–0.95 for house prices and 0.21–0.75 for trading volumes, as explained in subsection 2.2.

data exploration, the potential for improving the performance for estimating house prices by using a simultaneous prediction approach was indicated by the high correlation coefficients, which were in the range 0.73–0.95. However, an increase in the predictive accuracy of house prices was realized only in the province of Gangwon.

The improved performance in Gangwon Province may be attributed to the fact that this province is an area known to lag in terms of price fluctuations that tend to follow the trends in price-leading areas such as Seoul, Incheon, and Gyeonggi. However, the tendency of price movements in Gangwon Province to lag behind is not clear from the fluctuations in house prices in Figure 3.

A Granger causality test was implemented to validate the statement that Gangwon Province is a price-lagging area in the capital region. A time-series  $X$  is said to Granger-cause  $Y$  if regressing later values of  $Y$  on earlier values of itself and  $X$  is more predictive of  $Y$  than regressing on earlier values of  $Y$  alone (Granger, 1969).

The order of causality, that is, the lookback period, was specified to be 12, which is the same as that used for training the LSTM. Table 3 provides the results of the Granger causality test. The price fluctuations in Gyeonggi Province and Seoul appeared to help predict future prices in Gangwon Province, as indicated by the  $F$ -statistic and corresponding  $p$ -value (0.008 for Gyeonggi, and 0.05 for Seoul). The  $F$ -statistic in the case of Gyeonggi Province, being a cause, is 2.36 and significant with a  $p$ -value of 0.008, implying that the prices in Gyeonggi Province noticeably contributed to improving the prediction of prices in Gangwon Province. The next highest  $F$ -statistic in terms of magnitude is Seoul and then Incheon, in this order. The order of the magnitude of the  $F$ -statistic in Table 3 corresponds to the geographical distance of the area concerned from Gangwon Province, as shown in Figure 2. This consistency between the Granger causality test and the physical distance reasonably supports the idea that Gangwon Province is a price-lagging area, and utilization of a multi-output LSTM helped to enhance the predictive accuracy by leveraging the price information across the price-leading areas in the immediate vicinity.

Table 3. Result of Granger causality test

| Cause ( $X$ ) → Effect ( $Y$ ) | $F$ -statistic | $p$ -value |
|--------------------------------|----------------|------------|
| Gyeonggi → Gangwon             | 2.36           | 0.008      |
| Seoul → Gangwon                | 1.79           | 0.050      |
| Incheon → Gangwon              | 1.63           | 0.084      |

## Conclusions

Four study areas in the capital region were chosen to predict house prices and trading volumes. An RNN, a neural network specially designed to analyze time series data, was utilized, and a multi-output LSTM was implemented to exploit the correlations in house prices and trading volumes across adjacent areas, thereby alleviating the problems that

arise from using a small dataset. The multi-output LSTM outperformed baseline models such as the ARIMA model and single-output LSTM for predicting house prices in Gangwon Province. Our approach showed little improvement for predicting house prices in the other three areas: Seoul, Incheon, and Gyeonggi Province. As for predicting the trading volumes, the performance of the three models was similar with little difference among them.

The improved performance in the prediction of house prices in Gangwon Province can be attributed to the leading-lagging relationship between adjacent areas. The Granger causality test was employed to show that Gangwon Province was a price-lagging area in terms of house prices in the capital region, and this reality was exploited by the multi-output LSTM, thereby enhancing the accuracy of the prediction of house prices in Gangwon Province. The Granger causality test is a statistical test that has been widely used for forecasting time series in the following manner, for example, in testing whether one time series, such as house prices, can be useful in predicting another time series, such as house trading volumes, or vice versa. This study extended the use of the Granger causality test to geographical settings, demonstrating that it can be effectively used to identify spatially leading and lagging areas.

Monitoring and forecasting tendencies in the real estate market has always been a primary issue for both policymakers and private investors. Governments check and track movements in the prices and trading volumes in the market, and often require predictions because these trends indicate the health and efficiency of the market. These trends could help the government to identify an emerging bubble in house prices or a decrease in buyers' ability to afford houses in advance. Real estate developers keep an eye on prices and trading volumes to estimate the level of construction activity in the following year. Investors employ different portfolio management strategies depending on the status of the market, which is usually determined by the degree of market activity, for example in the form of price and volume movements. Hence, an improvement in the forecasting accuracy of the future market can provide useful guidelines for both the government and private sector companies.

The multi-output LSTM approach adopted in this study can be flexibly utilized in tasks beyond the housing market. For example, predicting prices and trading volumes of commercial properties is a major concern for office and retail managers. Forecasting land market activity is critical for new town developers. Hence, the study results are expected to promote the application of deep learning-based approaches in the overall real estate industry.

The proposed multi-output LSTM approach can be applied to different study areas and tasks on a larger spatial scale. In the case of application to other study areas, a clear leading-lagging relationship between neighboring areas may not be present. This could lead to little improvement in the prediction by the multi-output LSTM. In addition, the LSTM layers used in the network are computation-

ally intensive, and thus the computation could be considerably time-consuming to perform if the multi-output LSTM were applied to tasks on a nationwide or global scale. Hence, generalizing the findings of this study to other provinces and spatially scaling up prediction tasks would be necessary in the future.

### Funding and disclosure statement

The author declares no funding and no conflict of interest.

### References

- Ahmar, A. S., & Aidid, M. K. (2018). Crime modeling using spatial regression approach. *Journal of Physics: Conference Series*, 954(1), 012013. <https://doi.org/10.1088/1742-6596/954/1/012013>
- Bakker, B. J., & Heskes, T. M. (2003). Task clustering and gating for Bayesian multitask learning. *Journal of Machine Learning Research*, 4, 83–99.
- Ben-David, S., & Schuller, R. (2003). Exploiting task relatedness for multiple task learning. In *Learning theory and kernel machines* (pp. 567–580). Springer. [https://doi.org/10.1007/978-3-540-45167-9\\_41](https://doi.org/10.1007/978-3-540-45167-9_41)
- Beracha, E., Lang, M., & Hausler, J. (2019). On the relationship between market sentiment and commercial real estate performance – a textual analysis examination. *Journal of Real Estate Research*, 41(4), 605–638. <https://doi.org/10.22300/0896-5803.41.4.605>
- Cellmer, R. (2013). Use of spatial autocorrelation to build regression models of transaction prices. *Real Estate Management and Valuation*, 21(4), 65–74. <https://doi.org/10.2478/remav-2013-0038>
- Conway, D., Li, C. Q., Wolch, J., Kahle, C., & Jerrett, M. (2010). A spatial autocorrelation approach for examining the effects of urban greenspace on residential property values. *The Journal of Real Estate Finance and Economics*, 41(2), 150–169. <https://doi.org/10.1007/s11146-008-9159-6>
- Cui, L., Xie, X., Shen, Z., Lu, R., & Wang, H. (2018). Prediction of the healthcare resource utilization using multi-output regression models. *IJSE Transactions on Healthcare Systems Engineering*, 8(4), 291–302. <https://doi.org/10.1080/24725579.2018.1512537>
- Dai, Z., Guldmann, J. M., & Hu, Y. (2018). Spatial regression models of park and land-use impacts on the urban heat island in central Beijing. *Science of the Total Environment*, 626, 1136–1147. <https://doi.org/10.1016/j.scitotenv.2018.01.165>
- DeFusco, A., Ding, W., Ferreira, F., & Gyourko, J. (2018). The role of price spillovers in the American housing boom. *Journal of Urban Economics*, 108, 72–84. <https://doi.org/10.1016/j.jue.2018.10.001>
- Fang, C., Liu, H., Li, G., Sun, D., & Miao, Z. (2015). Estimating the impact of urbanization on air quality in China using spatial regression models. *Sustainability*, 7(11), 15570–15592. <https://doi.org/10.3390/su71115570>
- Futoma, J., Hariharan, S., Heller, K., Sendak, M., Brajer, N., Clement, M., & O'Brien, C. (2017, November). An improved multi-output Gaussian process rnn with real-time validation for early sepsis detection. In *Machine Learning for Healthcare Conference* (pp. 243–254). PMLR.
- Géron, A. (2019). *Hands-on machine learning with Scikit-Learn, Keras, and TensorFlow: concepts, tools, and techniques to build intelligent systems*. O'Reilly Media.
- Granger, C. W. (1969). Investigating causal relations by econometric models and cross-spectral methods. *Econometrica: Journal of the Econometric Society*, 37(3), 424–438. <https://doi.org/10.2307/1912791>
- Griffith, D. A., & Wong, D. W. (2007). Modeling population density across major US cities: a polycentric spatial regression approach. *Journal of Geographical Systems*, 9(1), 53–75. <https://doi.org/10.1007/s10109-006-0032-y>
- He, L., Madathil, S. C., Servis, G., & Khasawneh, M. T. (2021). Neural network-based multi-task learning for inpatient flow classification and length of stay prediction. *Applied Soft Computing*, 108, 107483. <https://doi.org/10.1016/j.asoc.2021.107483>
- Hochreiter, S., & Schmidhuber, J. (1997). Long short-term memory. *Neural Computation*, 9(8), 1735–1780. <https://doi.org/10.1162/neco.1997.9.8.1735>
- Ismail, S. (2006). Spatial autocorrelation and real estate studies: a literature review. *Malaysian Journal of Real Estate*, 1(1), 1–13.
- Kang, Y., Zhang, F., Peng, W., Gao, S., Rao, J., Duarte, F., & Ratti, C. (2021). Understanding house price appreciation using multi-source big geo-data and machine learning. *Land Use Policy*, 111, 104919. <https://doi.org/10.1016/j.landusepol.2020.104919>
- Lakshmanan, V., Robinson, S., & Munn, M. (2020). *Machine learning design patterns*. O'Reilly Media.
- Law, S., Paige, B., & Russell, C. (2019). Take a look around: using street view and satellite images to estimate house prices. *ACM Transactions on Intelligent Systems and Technology (TIST)*, 10(5), 1–19. <https://doi.org/10.1145/3342240>
- Lee, C. (2021). Deep learning-based detection of tax frauds: an application to property acquisition tax. *Data Technologies and Applications*. <https://doi.org/10.1108/DTA-06-2021-0134>
- Lee, K., Kim, H., & Shin, D. H. (2019). Forecasting short-term housing transaction volumes using time-series and internet search queries. *KSCE Journal of Civil Engineering*, 23(6), 2409–2416. <https://doi.org/10.1007/s12205-019-1926-9>
- Li, Y., Li, K., & Tong, S. (2019). Adaptive neural network finite-time control for multi-input and multi-output nonlinear systems with positive powers of odd rational numbers. *IEEE Transactions on Neural Networks and Learning Systems*, 31(7), 2532–2543. <https://doi.org/10.1109/TNNLS.2019.2933409>
- Moscone, F., & Tosetti, E. (2014). Spatial econometrics: theory and applications in health economics. *Encyclopedia of Health Economics*, 3, 329–334. <https://doi.org/10.1016/B978-0-12-375678-7.00719-7>
- Perez, H., Tah, J. H., & Mosavi, A. (2019). Deep learning for detecting building defects using convolutional neural networks. *Sensors*, 19(16), 3556. <https://doi.org/10.3390/s19163556>
- Poursaeed, O., Matera, T., & Belongie, S. (2018). Vision-based real estate price estimation. *Machine Vision and Applications*, 29(4), 667–676. <https://doi.org/10.1007/s00138-018-0922-2>
- Rambaldi, A. N., & Rao, D. P. (2011). *Hedonic predicted house price indices using time-varying hedonic models with spatial autocorrelation*. School of Economics, University of Queensland.
- Shi, H., Ma, Z., Chong, D., & He, W. (2021). The impact of Facebook on real estate sales. *Journal of Management Analytics*, 8(1), 101–112. <https://doi.org/10.1080/23270012.2020.1858985>
- Taltavull de La Paz, P., López, E., & Juárez, F. (2017). Ripple effect on housing prices. Evidence from tourist markets in Alicante, Spain. *International Journal of Strategic Property Management*, 21(1), 1–14. <https://doi.org/10.3846/1648715X.2016.1241192>

- Temur, A. S., Akgün, M., & Temur, G. (2019). Predicting housing sales in Turkey using ARIMA, LSTM and hybrid models. *Journal of Business Economics and Management*, 20(5), 920–938. <https://doi.org/10.3846/jbem.2019.10190>
- Tobler, W. R. (1970). A computer movie simulating urban growth in the Detroit region. *Economic Geography*, 46(2), 234–240. <https://doi.org/10.2307/143141>
- Tong, T., Yu, T. H. E., Cho, S. H., Jensen, K., & Ugarte, D. D. L. T. (2013). Evaluating the spatial spillover effects of transportation infrastructure on agricultural output across the United States. *Journal of Transport Geography*, 30, 47–55. <https://doi.org/10.1016/j.jtrangeo.2013.03.001>
- Wang, W. C., Chang, Y. J., & Wang, H. C. (2019). An application of the spatial autocorrelation method on the change of real estate prices in Taitung City. *ISPRS International Journal of Geo-Information*, 8(6), 249. <https://doi.org/10.3390/ijgi8060249>
- Xie, R., Fang, J., & Liu, C. (2016). Impact and spatial spillover effect of transport infrastructure on urban environment. *Energy Procedia*, 104, 227–232. <https://doi.org/10.1016/j.egypro.2016.12.039>
- Xu, D., Shi, Y., Tsang, I. W., Ong, Y. S., Gong, C., & Shen, X. (2019). Survey on multi-output learning. *IEEE Transactions on Neural Networks and Learning Systems*, 31(7), 2409–2429. <https://doi.org/10.1109/TNNLS.2019.2945133>
- Yang, J., Yu, Z., & Deng, Y. (2018). Housing price spillovers in China: a high-dimensional generalized VAR approach. *Regional Science and Urban Economics*, 68, 98–114. <https://doi.org/10.1016/j.regsciurbeco.2017.10.016>
- Yilmaz, S., Haynes, K. E., & Dinc, M. (2002). Geographic and network neighbors: spillover effects of telecommunications infrastructure. *Journal of Regional Science*, 42(2), 339–360. <https://doi.org/10.1111/1467-9787.00262>
- Zou, S., & Wang, L. (2021). Detecting individual abandoned houses from google street view: a hierarchical deep learning approach. *ISPRS Journal of Photogrammetry and Remote Sensing*, 175, 298–310. <https://doi.org/10.1016/j.isprsjprs.2021.03.020>