Int. J. Advance Soft Compu. Appl, Vol. 13, No. 2, July 2021
Print ISSN: 2710-1274, Online ISSN: 2074-8523

Exploring Uncertainty in Land Price Through Learning Models

Changro Lee

Kangwon National University, Chuncheon, Republic of Korea
e-mail: spatialstat@naver.com

Abstract

This study proposes a practical solution for estimating the uncertainty in predicted values. One of the most crucial aspects of prediction is reserving one's judgment in the case of uncertainty and signaling an appropriate warning message to stakeholders. However, despite the increased popularity of machine learning models, measuring uncertainty in predicted values has not been the main concern, with land valuation being no exception to this practice. In this study, two very different machine learning models, a random forest and a Monte Carlo dropout neural network were used to estimate land prices. Gyeonggi province, South Korea, was chosen for the analysis, and variations in estimates from these models were exploited to measure predictive uncertainty. The uncertainty in land price was found to differ systematically depending on the land use category, with the largest magnitude of uncertainty for forest, followed by farming land, residential sites, and commercial sites. In general, the uncertainty in agricultural land price is higher than that in urban land price. Although property prices have been modeled extensively in the literature by a variety of methods, measurements of the uncertainty in predicted prices are limited. This study demonstrates that the uncertainty in land price can be described efficiently through machine learning models, and it is expected that the empirical findings identified in this study can be applied directly in valuation practices, promoting the rapid adoption of machine learning methods in the real estate industry.

Keywords: *uncertainty, land price, random forest, neural network, Monte Carlo dropout*

1 Introduction

Land valuation is used extensively by various stakeholders, including governments for property tax assessment, banks for approval of collateral loans, and financial institutions for management of real estate portfolios. Notable methodological advances have been made in recent years to enhance the predictive accuracy of valuation models, with the application of machine learning methods as a representative example. Land valuation is an emerging area in the real estate industry, in which machine learning models such as neural networks have been applied rapidly.

However, machine learning models have not paid due attention to prediction uncertainty. Thus, the results do not appear to be sufficient to convince governments, banks, and financial institutions to trust the accuracy of the predicted values (Conway, 2018). Representing prediction uncertainty is indispensable for real-world tasks (Ghahramani, 2015). In real-world applications, rather than simply requiring predictions, the assurance of the nature of these predictions is also desired. A crucial part of learning involves reserving one's judgment in the case of uncertainty. Surprisingly, describing uncertainty has not been the aim of many machine learning applications, including those used for land valuation.

This study attempts to measure the uncertainty in predicted land prices. First, machine learning models capable of capturing uncertainty in predictions are chosen: a random forest and a Monte Carlo dropout neural network (hereafter, RF and MC dropout neural network, respectively). These two models are algorithms that are trained in very different manners. Then, variations in predicted prices from these models are used to estimate the uncertainty in land price. Finally, the estimated uncertainty in land price is compared across land use categories, and its implications are discussed.

The approach taken in this study to quantify the uncertainty associated with the predicted price would significantly promote the adoption of machine learning tools in the real estate industry, which is one of the main contributions of this work. In addition, the empirical results found in this study, such as different degrees of uncertainty per land use category, could be used directly in valuation practices, which could improve the overall asset adequacy held by financial institutions.

The remainder of this paper is organized as follows. Section 2 presents background information on uncertainty in property valuation and the approach for estimating predictive uncertainty. Section 3 describes land price dataset, RF, and MC dropout neural network. Then, the uncertainty estimation results and implications are discussed in Section 4, and finally, a summary of the study and conclusions are presented in Section 5.

2 Literature Review

2.1 Land Price Uncertainty in Property Valuation

Rosen (1974) demonstrated that goods possess various attributes that combine to form bundles of characteristics and the overall price. This theory has served as a cornerstone for the valuation of all heterogeneous goods, such as vehicles and houses (Lee & Park, 2021). In the 1990s and 2000s, valuations from a spatial perspective have been attempted (Wyatt, 1996; McCluskey et al., 2000; Chica-Olmo, 2007) because location is the first and foremost price determinant. Since 2010, with the boom of big data and artificial intelligence, machine learning-based valuations (Yu et al., 2018; Wang et al., 2019) have been performed. This trend is still an active research area in the property valuation literature.

In valuations prepared for the purposes of loans or insurance, predicted prices are usually provided as single figures, and these figures are often accepted as confident and certain prices by stakeholders. However, from the both valuation theory and practice perspectives, a single figure is unrealistic (Kucharska-Stasiak, 2013; Mooya, 2016). Kucharska-Stasiak (2013) stated that the uncertainty in property valuation is exceptionally high, owing to both the characteristics of real estate (e.g., fixed location, long useful life, variations in physical features) and the characteristics of the real estate market (e.g., low efficiency, low elasticities of supply and demand). Although the valuer must provide a single figure, a description should be developed to explain the uncertainty in the final figure (Mallinson & French, 2000).

In the past, property valuation tended to focus on estimating house prices because housing problems are of great concern to policymakers and householders. However, studies on land valuation are less prevalent in the literature, even though land prices are likely to show more variability in their actual measured market values, compared to house prices.[1] A house is a property whose main function is to provide residential services to the occupants, and thus, turning its current use (residence) to another is difficult. By contrast, land is a property whose use has not yet been determined; thus, it has the potential to be developed for diverse purposes, leading to more variability in its estimated price.

Therefore, the consideration of uncertainty becomes especially critical for land valuation. Tax assessment or collateral valuation for land would be typical applications in which confidence is a key factor in decision-making, and thus, an indication of the degree of uncertainty would greatly assist many decision-makers

[1] Considering this variability in land price, IAAO (2013) suggested a lower (more generous) performance standard for land valuation as compared with that for house valuation.

and improve the credibility of the valuation work (Mallinson & French, 2000; Ghahramani, 2015).

2.2 Approach for Estimating Predictive Uncertainty

The standard approach for measuring uncertainty in predictions is to use a Bayesian estimation method, which essentially learns a distribution over parameters, such as weights in neural networks. Such an approach, however, requires excessive computation, and thus cannot be adopted easily using the today's ordinary machine-learning framework, possibly limiting its practical use (Jin et al., 2019).

An ensemble approach is a more practical alternative for estimating predictive uncertainty. The ensemble approach typically involves some randomization, either in the base learners themselves or in the data used to train each base learner. The idea behind this approach is simple and intuitive: use an ensemble of individual models to obtain multiple predictions, and use their empirical variance as an approximate measure of uncertainty (Lakshminarayanan et al., 2017).

RF is an approach that applies a *bagging* procedure to decision tree building algorithms. It randomly chooses samples for each tree's training and randomly selects a subset of features when selecting each split in each decision tree (Breiman, 2001). RF is the most widely used and practically effective machine learning method. RF averages the individual trees' predictions to obtain an overall prediction, and Fig. 1 shows an example of RF with 500 decision trees. The confidence of the overall prediction can be measured by using the standard deviation of predictions across the trees. The standard deviation indicates the relative confidence of predictions, and observations with higher standard deviations require caution when using the predicted values, because individual decision trees produce substantially different estimates for them, rather than consistent ones.

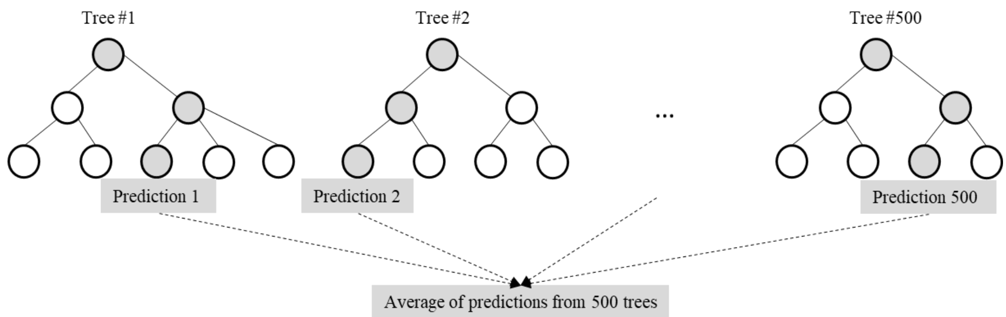


Fig. 1 Random forest composing of 500 decision trees

An ensemble approach such as RF requires a combination of many base learners to build an ensemble learner. Hence, the ensemble approach is typically implemented in learning schemes in which the baser learners can be created and fitted quickly. For this reason, RF is frequently used in the ensemble approach because base learners in RF, decision trees, can be easily specified and quickly fitted to the training data. Owing to this advantage of RF, this study adopts the approach for estimating the uncertainty in predictions.

A neural network can be viewed as a generalization of linear models that perform multiple processing stages to predict a target. It is a multilayered network of neurons (also known as nodes) and can be represented as follows:

$$f(x) = \sigma(b + w'x) \tag{1}$$

where b and w indicate the bias and slope weights, respectively. The activation function σ is employed here to capture the nonlinearity inherent in the data. An example of a neural network is shown in Fig. 2 (a), which presents a diagram of a simple neural network with one input layer, one hidden layer with six neurons, and one output layer.

When a dropout technique is applied to a neural network, it can also be considered as an ensemble learner. Dropout simply refers to turning off certain neurons at each training step, as shown in Fig. 2 (b). Each neuron has a probability p of being ignored, called the dropout rate. In practice, the dropout rate is typically set between 0.0 and 0.5 (approximately 50% of all neurons will be turned off). Dropout is essentially a regularization technique, that is, it helps to prevent overfitting. In every training iteration, the neurons are randomly dropped out in each layer, such that a different set of neurons is dropped out each time. Hence, each time the network's architecture is slightly different and the outcome can be interpreted as an average of many different neural networks, which results in better network generalization.

Gal and Ghahramani (2016) proposed using MC dropout to estimate a network's predictive uncertainty by using dropout at test time. Different networks with different neurons dropped out can be treated as Monte Carlo samples from the space of all available models. The neurons are simply dropped out at inference time (test time), and many predictions are obtained from each model. Then, the distributions of these predictions may be analyzed to gauge the uncertainty in the predictions. In effect, the dropout technique provides an inexpensive approximation to training an ensemble of exponentially many neural networks (Jin et al., 2019).

Although neural networks are generally known to be time-consuming to train, thus hindering their adoption in ensemble schemes, the MC dropout neural network is extremely simple and efficient for creating and fitting the data. Hence, this study employs the MC dropout neural network to estimate predictive uncertainty.

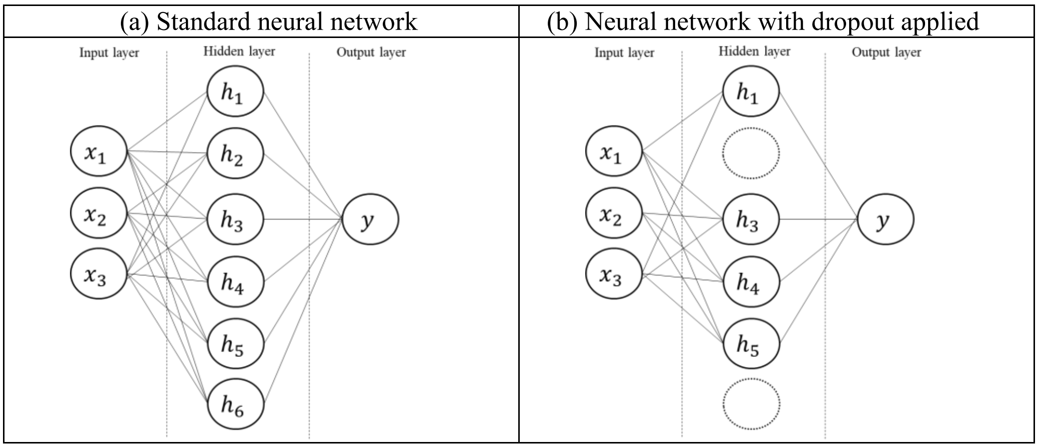


Fig. 2 Example of a neural network architecture

Although RF is the most popular algorithm in machine learning, few studies have been reported that used it explicitly to estimate uncertainty. Attempts to predict property prices by RF have been made in several studies (Antipov & Pokryshevskaya, 2012; Hong et al., 2020; Yilmazer & Kocaman, 2020), but few have utilized it to derive the inherent uncertainty in predicted prices.

Attempts to estimate property prices via neural networks can also be found in numerous studies (Peterson & Flanagan, 2009; Amri & Tularam, 2012; McCluskey et al., 2013; Morano & Tajani, 2013; Sampathkumar et al., 2015). However, these studies have not paid sufficient attention to the inherent uncertainty in predicted prices. In contrast to property valuation, uncertainty is of great interest for safety-critical domains such as healthcare, and studies using neural networks for estimating uncertainty are abundant in this area: medical images are frequently analyzed using MC dropout neural networks for measuring uncertainty in medical detections such as brain tumors or pulmonary nodules (Jungo et. al, 2017; Ozdemir et al., 2017; Zhao et al., 2018; Camarasa et al., 2020).

This study is different from previous studies concerning for the following three aspects. First, it attempts to describe uncertainty in land prices by applying confidence-measuring models to the real estate industry. Second, most machine-learning applications have been observed in unstructured data areas, such as image processing. Few attempts have been made to apply this popularity-gaining technique to structured data-intensive areas, such as the real estate industry, where tabular format data with rows and columns are dominant. This study is an early attempt to exploit structured data to estimate the uncertainty of land prices. Finally, the degree of uncertainty per land use category is interpreted in-depth to provide practical insights to land valuation stakeholders.

3 RF and the MC Dropout Neural Network

3.1 Land Price Dataset in Gyeonggi province

Gyeonggi province, with a population of over 13 million, is the most populous of the nine provinces in South Korea. It is well known for dynamic market activities, such as increasing property prices. As shown in Fig. 3, Seoul, the capital of the nation, is in the heart of the province, and *Gyeonggi* itself can be translated as "the area surrounding the capital" in Korean. The province is also the fastest-growing region of the nine provinces: its population growth rate in 2019 was 1.25%, whereas the rate of Seoul for the same year was −0.38% (KOSIS, 2020).

Land prices vary dramatically across Gyeonggi province, because the province has mixed characteristics of urban and rural landscapes. In particular, land prices in Gyeonggi province are well-known for their high growth premium due to the expected development potential of the region, as it has excellent accessibility to Seoul. Thus, for farming land with identical agricultural productivity, lots in Gyeonggi province command a considerably higher price than those in other provinces. It is a fast-growing region characterized by diverse landscapes, both urban and rural. Thus, it serves as a good study area for investigating price uncertainty of a variety of land uses.

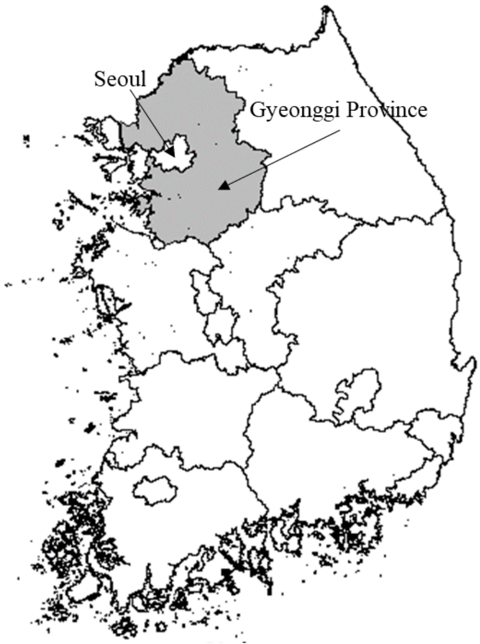


Fig. 3. Location of Gyeonggi province

The dataset adopted in this work was collected from a public website managed by the local government. It consists of land attributes of 47,460 sample sites. The attributes include site area, land use category, zoning, and so forth.[2] It also includes land sales prices, and these prices are used as a reference for various administrative purposes such as property taxation and compensation in the eminent domain. The dataset was updated on 1st January 2020. Table 1 lists the descriptive statistics of the dataset. The median price for sites in the dataset is 110,000 KRW per square meter, and the median area is 1,326.0 m^2. In terms of land use, farming land and residential sites are frequent categories, whereas forest and commercial sites are less frequent. The 47,460 locations in Table 1 were randomly divided into training data (80%, 37,968 locations) and test data (20%, 9,492 locations) for subsequent analysis.

Table 1: Descriptive statistics of the 47,460 locations

	Min.	Mean	Median	Max.
Price (KRW/m^2)	500	490,022	110,000	13,200,000
Site area (m^2)	18.2	3,629.9	1,326.0	118,348.0
Land use	forest (6,267), farming land (22,357), residential site (12,585), commercial site (6,251)			

Factors affecting land prices are practically countless, and determining input variables is always a compromise between theoretical aspects and data availability. Six input variables were used to estimate the land price in this study: site area, site zoning, site use, slope of the site, site shape, and road width (m) onto which the lot abuts. The target variable is the price per square meter. The input variables were standardized to have a mean of zero and a standard deviation of one. The target variable was min-max scaled to have a minimum value of zero and a maximum value of one.

3.2 Training RF

As explained earlier, RF is an ensemble learning method that constructs multiple decision trees at training time, and outputs the mean prediction of the individual decision trees for regression tasks.[3] The key aspect of RF performance is to train on different random subsets of the training dataset and select features among a random subset of features for each decision tree.

[2] The dataset includes 12 variables in total: lot identifier, county code, street address, assessed price, neighborhood description, land sales price, site area, site zoning, site use, slope of the site, site shape, and road width (m). One location was deleted from the dataset before the analysis, owing to missing values.

[3] For classification tasks, RF outputs the class given by the mode of the individual decision trees.

More formally, through the predictions of T individual decision trees, T output vectors, $\hat{\boldsymbol{y}}^{(1)}, \hat{\boldsymbol{y}}^{(2)}, \ldots, \hat{\boldsymbol{y}}^{(T)}$, are obtained. Their predictive mean and standard deviation can then be represented as follows:

$$mean\ \ E(\hat{\boldsymbol{y}}) = \frac{\sum_{k=1}^{T} \hat{\boldsymbol{y}}^{(k)}}{T} \tag{2}$$

$$standard\ \ deviation = \sqrt{\frac{\sum_{k=1}^{T} [\hat{\boldsymbol{y}}^{(k)} - E(\hat{\boldsymbol{y}})]^2}{T}} \tag{3}$$

Equation (2) can be understood as the expected output given input $\boldsymbol{x}$, and RF generally yields a better model overall than individual decision trees. Equation (3) can be used to measure the confidence of RF in its prediction. In this study, RF was implemented based on 500 trees (T = 500)[4] using aforementioned six input variables. The larger the standard deviation is, the greater the uncertainty of RF in its prediction.

3.3 Training the MC Dropout Neural Network

The neural network type used in this study is a fully connected layer neural network or a dense neural network. The MC dropout technique applies to the neural network, and the implementation details are as follows. The network has three fully connected hidden layers, each containing 64 neurons, with dropout layers interleaved. The output layer is a linear layer with one output neuron representing land price. A suitable dropout rate p needs to be identified, and the prediction performance under various dropout rates (p = 0.0, 0.1, 0.2, 0.3, 0.4, 0.5) was reviewed using the usual cross-validation process. A dropout rate of p = 0.1 was chosen empirically for the network used in this study. Each network was trained for 15 epochs using the six input variables explained earlier, with the mean squared error (MSE) and the mean absolute error (MAE) adopted as the loss functions, as shown in the following equations.

$$MSE = \frac{1}{n} \sum_{i=1}^{n} (\hat{y} - y)^2 \tag{4}$$

$$MAE = \frac{1}{n} \sum_{i=1}^{n} |\hat{y} - y| \tag{5}$$

where $\hat{y}$ is the estimated price from the network and y is the observed price. For each MC dropout network, 500 Monte Carlo samples were drawn to estimate the predictive mean and standard deviation.

---

[4] In general, the more trees you employ, the better the results. However, at a certain threshold, the improvement in prediction performance from using more trees is lower than the computational cost. Typical values for the number of trees is between 50 to 500. In this study, 500 trees were set for learning.

4 Results and Discussion

4.1 Results

The dataset of 47,460 samples was randomly split into a training dataset (80%) and a test dataset (20%).[5] This split was repeated 100 times, and the RF and the MC dropout neural network were fitted to the corresponding split dataset 100 times.[6] Table 2 shows the model performance resulting from these multiple fittings. As shown in the table, the MSE and MAE of RF are lower than those of the MC dropout neural network, indicating that the former outperforms the latter. However, the difference between the two models seems to be negligible.

Table 2: Comparison of model performance

	Random forest	MC dropout neural network
MSE	0.0013–0.0015	0.0014–0.0016
MAE	0.0152–0.0209	0.0176–0.0215

The goodness-of-fit of the two models for the test dataset is visually shown in Fig. 4 (one example result of the 100 fittings). Both the predicted and observed prices were normalized to have values between zero and one. The predicted prices appear to follow the observed prices approximately well. However, the range of predicted prices becomes narrower than that of observed prices, meaning that both models have difficulty capturing the entire distribution of prices.

[5] A cross-validation approach was used during hyperparameter tuning, and thus, the final prediction performance was evaluated based on the train-test split approach.
[6] More than 100 repetitions did not produce significantly different result from Table 2.

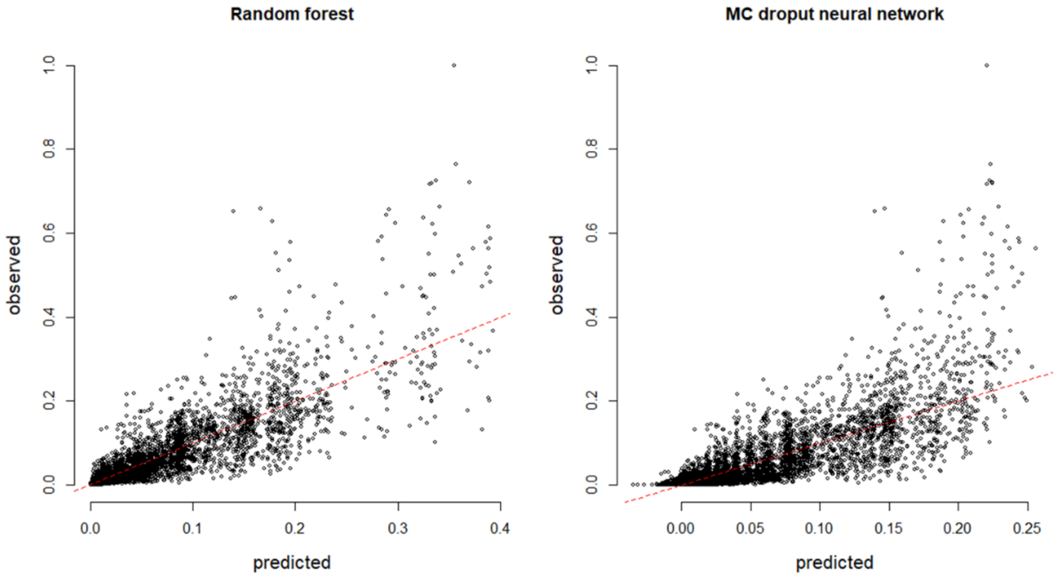
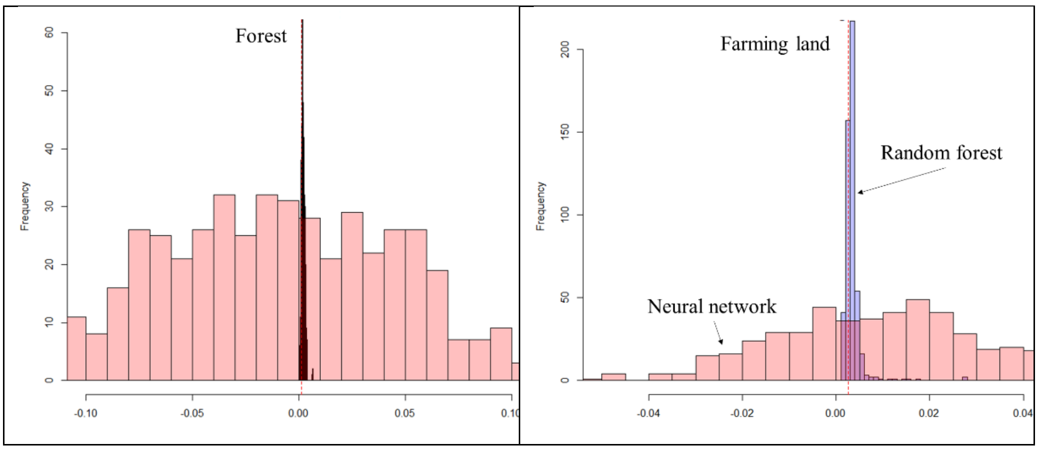


Fig. 4. Goodness-of-fit in RF and the MC dropout neural network for the test dataset (one example result of the 100 fittings)

Fig. 5 shows the distribution of predicted prices from the RF and the MC dropout neural network. Each sub-figure in the figure represents a site selected randomly from a test dataset for each land use category: forest, farming land, residential site, and commercial site. The vertical dotted line in each sub-figure indicates an observed price for the selected site (price values were standardized to have a minimum value of zero and a maximum value of one, as explained earlier). As shown in the figure, the price distribution from the MC dropout neural network is much wider than that from RF, indicating that the MC dropout neural network tends to exhibit more variability in its outputs. In contrast, RF tends to produce prices showing less variability, and they are centered around the true observed price. The wider distribution of predicted prices from the MC dropout neural network indicates that the network has difficulty in producing consistent point estimates.



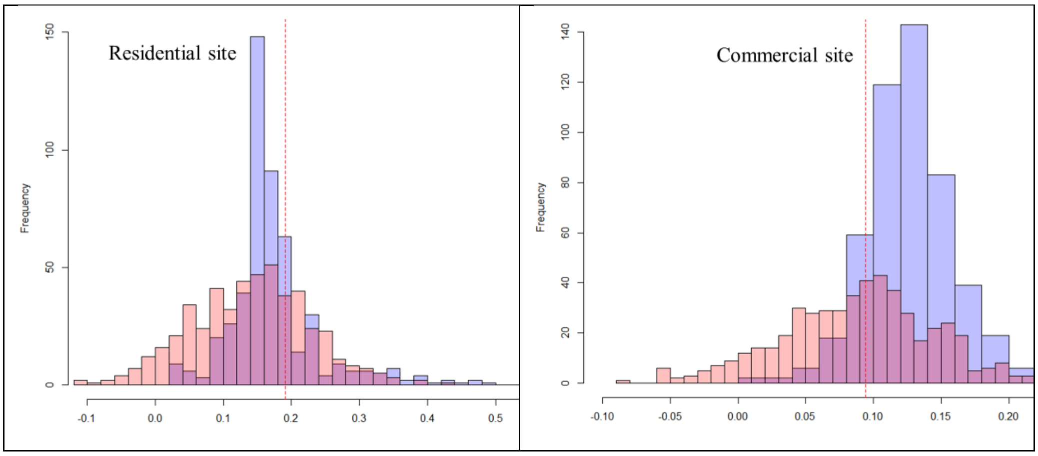


Fig. 5. Distribution of predicted prices from RF and the MC dropout neural network (one example per land use category)

Overall, RF shows more accurate goodness-of-fit and less variability in its predicted price than the MC dropout neural network, and hence it can be concluded for the region chosen for this study that RF outperforms the neural network in terms of both bias and variance.

4.2 Uncertainty Varied by Land Use Category

Table 3 shows standard deviations of predictions from RF and the MC dropout neural network. The standard deviations were calculated for the test dataset 100 times. They were computed considering the ratios of predicted prices over observed prices rather than the raw predicted prices. Standard deviations calculated for the original predicted prices tend to increase as the price level increases. Thus, for example, the standard deviation of prices for commercial sites is almost always larger than that of prices for forest land. Standard deviations computed for the ratios are frequently used to evaluate property valuation performance[7], and this study adopts this index to analyze uncertainty in predicted prices.

Table 3: Standard deviations of the ratios of predicted prices over observed prices

Land use	Random Forest		MC dropout neural network	
	Mean	Median	Mean	Median
Forest	1.16–1.35	0.60–0.64	89–157	40–44
Farming land	0.45–0.46	0.30–0.32	7.2–8.1	4.7–5.8

[7] Ratios calculated in this manner are used extensively in property assessment to provide measures of assessment performance, and analyses of such ratios are often called ratio studies (IAAO, 2013).

Residential site	0.38–0.44	0.27–0.28	1.4–1.7	0.9–1.1
Commercial site	0.37–0.42	0.25–0.27	0.8–0.9	0.4–0.6

The overall standard deviations of the MC dropout neural network are much larger than those of RF, which was already confirmed from Fig. 5. Although both the goodness-of-fit and variability of predicted prices were found to be different between the two models in the previous section, the rank of variability was consistent. For both RF and the MC dropout neural network, the uncertainty of predicted prices was largest for the forest, followed by farming lands such as dry fields and paddy fields. A residential site such as land for a single-family house or an apartment complex was ranked 3rd in terms of uncertainty, and finally a commercial site was identified to be the least uncertain in the price estimate.

In short, the uncertainty in price was found to vary systematically depending on the land use category: price uncertainty in agricultural land such as forest and farming land is higher than that of urban lands such as residential and commercial sites.

If landowners have perfect information and the land market is transparent, the price of land theoretically equals the present value of anticipated future income generated from the land: the present value of sales income from agricultural produce for agricultural land, and the present value of rental income from leases for urban land. However, all the land has one more component to consider when estimating its market value, and it is the possibility of conversion to different uses. In general, land is likely to be converted to more intensive use, typically developed from agricultural use to urban use. Thus, the land price in a region characterized by both urban and agricultural landscapes, such as Gyeonggi province, consists of two components: the present value of current use and future use after conversion. Such value of expected future use represents a growth premium for the site. The growth premium is especially high for forest and farming land because they have much higher potential for conversion to urban land use than residential and commercial sites that are already utilized intensively.

It is well known that agricultural land in Gyeonggi province always sells for a high growth premium, which originates from the expectation of land conversion to urban use because the province has excellent accessibility to Seoul compared to other provinces. In a rapidly growing region with great accessibility to an urban center, the growth premium may easily account for half of the average price of land (Capozza & Helsley, 1989).

5 Conclusion

This study attempted to describe uncertainty in land valuation by applying two machine learning methods, RF and the MC dropout neural network. Variabilities from the base learners in RF and the MC dropout neural network were exploited to estimate the uncertainty in predicted land prices. As for RF, standard deviations from the 500 individual decision trees were used as a proxy for uncertainty. Similarly, standard deviations from 500 different networks generated with different neurons dropped out were used to estimate the uncertainty in the predicted price. Gyeonggi province, a region with diverse land uses and a growing population, was chosen for the analysis.

Based on the analysis results, the following findings are noteworthy: RF shows more accurate goodness-of-fit and less variability in its predicted price than the MC dropout neural network, indicating that RF is superior to the neural network in terms of both bias and variance. The uncertainty in price was systematically different, depending on the land use category. The uncertainty was largest for forest, followed by farming land, residential sites, and commercial sites. In general, price uncertainty in agricultural lands such as forest and farming land is higher than that of urban land. A land price can be decomposed into current use value and future use value after conversion, and the latter is often termed a growth premium. The growth premium is especially high for forest and farming land compared to residential and commercial sites. Gyeonggi province has excellent accessibility to Seoul compared to other provinces, and thus, the growth premium in this region may account for a considerable portion of the land price.

The most significant concern with the property valuation model is that an appropriate tool to gauge the confidence for the predicted prices is not easily available. In real-world land valuations such as the property tax assessment, the approach proposed in this study could be used as follows: first, a valuation model such as RF would be trained on the basis of the sales dataset. Second, this trained model would be used to predict land prices in the jurisdiction, including those that are untraded. Third, an index such as a standard deviation obtained from individual decision trees would be used to represent the uncertainty in the predicted price of each site. Predicted sites in which the values are characterized by high uncertainty may be sent to an appraiser for reappraisal or field inspection. By adopting this process, the value assessed by the government would inspire the taxpayer's confidence, reducing unnecessary administrative costs such as tax appeals.

The findings of this study can also be directly utilized in private financial institutions: banks should exercise more caution when approving collateral loans for agricultural land because it is more likely to show volatility in collateral value than urban land during the loan period.

The degree of price uncertainty differed depending on the land use category, as demonstrated in this study. However, price uncertainty is caused by multiple factors, such as the regional characteristics and economic cycles of the real estate market.

Mature regions in the stable phase of a market cycle will command less uncertainty in the price, while rapidly growing regions in the booming stage of a market cycle will show more uncertainty in the price. The cause of price uncertainty needs to be investigated from various perspectives in future studies.

The approach taken in this study is expected to be widely used in land valuation practice, promoting the rapid adoption of machine learning methods in the real estate industry.

References

1. Amri, S., & Tularam, G. A. (2012). Performance of multiple linear regression and nonlinear neural networks and fuzzy logic techniques in modelling house prices. Journal of Mathematics and Statistics, 8(4), 419-434.

2. Antipov, E. A., & Pokryshevskaya, E. B. (2012). Mass appraisal of residential apartments: An application of Random forest for valuation and a CART-based approach for model diagnostics. Expert Systems with Applications, 39(2), 1772-1778.

3. Breiman, L. (2001). Random forests. Machine learning, 45(1), 5-32.

4. Camarasa, R., Bos, D., Hendrikse, J., Nederkoorn, P., Kooi, E., van der Lugt, A., & de Bruijne, M. (2020). Quantitative Comparison of Monte-Carlo Dropout Uncertainty Measures for Multi-class Segmentation. In Uncertainty for Safe Utilization of Machine Learning in Medical Imaging, and Graphs in Biomedical Image Analysis (pp. 32-41). Springer, Cham.

5. Capozza, D. R., & Helsley, R. W. (1989). The fundamentals of land prices and urban growth. Journal of urban economics, 26(3), 295-306.

6. Chica-Olmo, J. (2007). Prediction of housing location price by a multivariate spatial method: Cokriging. Journal of Real Estate Research, 29(1), 91-114.

7. Conway, J. J. E. (2018). Artificial intelligence and machine learning: Current applications in real estate (Doctoral dissertation, Massachusetts Institute of Technology).

8. Gal, Y., & Ghahramani, Z. (2016, June). Dropout as a bayesian approximation: Representing model uncertainty in deep learning. In international conference on machine learning (pp. 1050-1059).

9. Ghahramani, Z. (2015). Probabilistic machine learning and artificial intelligence. Nature, 521(7553), 452.

10. Hong, J., Choi, H., & Kim, W. S. (2020). A house price valuation based on the random forest approach: the mass appraisal of residential property in South Korea. International Journal of Strategic Property Management, 24(3), 140-152.

11. IAAO. (2013). Standard on Ratio Studies. International Association of Assessing Officers.

12. Jin, B., Li, D., Srinivasan, S., Ng, S. K., Poolla, K., & Sangiovanni-Vincentelli, A. (2019, June). Detecting and diagnosing incipient building faults using uncertainty information from deep neural networks. In 2019 IEEE International Conference on Prognostics and Health Management (ICPHM) (pp. 1-8). IEEE.

13. Jungo, A., McKinley, R., Meier, R., Knecht, U., Vera, L., Pérez-Beteta, J., ... & Reyes, M. (2017, September). Towards uncertainty-assisted brain tumor segmentation and survival prediction. In International MICCAI Brainlesion Workshop (pp. 474-485). Springer, Cham.

14. KOSIS. (2020). Statistics on population and household, Korean Statistical Information Service.

15. Kucharska-Stasiak, E. (2013). Uncertainty of property valuation as a subject of academic research. Real Estate Management and Valuation, 21(4), 17-25.

16. Lakshminarayanan, B., Pritzel, A., & Blundell, C. (2017). Simple and scalable predictive uncertainty estimation using deep ensembles. In Advances in neural information processing systems (pp. 6402-6413).

17. Lee, C., & Park, K. K. H. (2021). Forecasting trading volume in local housing markets through a time-series model and a deep learning algorithm. Engineering, Construction and Architectural Management.

18. Mallinson, M., & French, N. (2000). Uncertainty in property valuation–The nature and relevance of uncertainty and how it might be measured and reported. Journal of Property Investment & Finance, 18(1), 13-32.

19. McCluskey, W. J., Deddis, W. G., Lamont, I. G., & Borst, R. A. (2000). The application of surface generated interpolation models for the prediction of residential property values. Journal of Property Investment & Finance.

20. McCluskey, W. J., McCord, M., Davis, P. T., Haran, M., & McIlhatton, D. (2013). Prediction accuracy in mass appraisal: a comparison of modern approaches. Journal of Property Research, 30(4), 239-265.

21. Mooya, M. M. (2016). Real Estate Valuation Theory. Springer Books.

22. Morano, P., & Tajani, F. (2013). Bare ownership evaluation. Hedonic price model vs. artificial neural network. International Journal of Business Intelligence and Data Mining, 8(4), 340-362.

23. Ozdemir, O., Woodward, B., & Berlin, A. A. (2017). Propagating uncertainty in multi-stage bayesian convolutional neural networks with application to pulmonary nodule detection. arXiv preprint arXiv:1712.00497.

24. Peterson, S., & Flanagan, A. (2009). Neural network hedonic pricing models in mass real estate appraisal. Journal of real estate research, 31(2), 147-164.

25. Rosen, S. (1974). Hedonic prices and implicit markets: product differentiation in pure competition. Journal of political economy, 82(1), 34-55.

26. Sampathkumar, V., Santhi, M. H., & Vanjinathan, J. (2015). Evaluation of the trend of land price using regression and neural network models. Asian Journal of Scientific Research, 8(2), 182.

27. Wang, F., Zou, Y., Zhang, H., & Shi, H. (2019, October). House price prediction approach based on deep learning and arima model. In 2019 IEEE 7th International Conference on Computer Science and Network Technology (ICCSNT) (pp. 303-307). IEEE.

28. Wyatt, P. (1996). Using a geographical information system for property valuation. Journal of Property Valuation and Investment.

29. Yilmazer, S., & Kocaman, S. (2020). A mass appraisal assessment study using machine learning based on multiple regression and random forest. Land Use Policy, 99, 104889.

30. Yu, L., Jiao, C., Xin, H., Wang, Y., & Wang, K. (2018). Prediction on housing price based on deep learning. International Journal of Computer and Information Engineering, 12(2), 90-99.

31. Zhao, G., Liu, F., Oler, J. A., Meyerand, M. E., Kalin, N. H., & Birn, R. M. (2018). Bayesian convolutional neural network based MRI brain extraction on nonhuman primates. Neuroimage, 175, 32-44.