

Identification of Commercial Real Estate Investment Targets Using Spatial Interpolation of Rent Prices

Adam Green

2023-08-15

Executive Summary

Motivation

The COVID-19 pandemic has had major global economic impacts, one of the most significant being the rise of work-from-home and the overall residential-based labour market. While much of the world has returned to normal, commercial real-estate markets continue to experience the effects of poor commercial office-space demand (Gupta, 2022). Strain has also grown in the residential real-estate market as rent prices explode in urban areas. The culmination of these issues has been particularly intense in Toronto, Ontario where downtown office vacancies have reached 15.3% in the first quarter of 2023 (The Canadian Press, 2023). Additionally, one bedroom condominium rent has increased 15.1% over the past year (Toronto Residential Real Estate Board, 2023). As office demand drops, residential demand rises, and interest rates rise to combat inflation, commercial real estate owners will likely look to offload properties to avoid defaulting on property loans as their properties stop appreciating in value (Glaze, 2023). The outcome could be discounted commercial buildings available for purchase in areas of high population density. It is likely some of these buildings will be in areas of mixed zoning allowing conversion of commercial office space to residential dwellings. This will provide real estate investors a unique opportunity to generate significant return on investment if they can identify office buildings with the highest potential residential rent price at the lowest sale price. Given the current real estate environment, this research aims to capitalize on residential conversion through identifying commercial offices, currently for sale, that could return the highest possible rent if converted to residential dwellings. To meet this objective, the current work presents a series of models for estimating potential rent prices at commercial locations across downtown Toronto based on several spatial features of interest.

Data

Data pertaining to rental apartment listings, commercial office listings and a range of spatial features was collected. Rental apartment listings and associated information was gathered from Rentals.ca including rent price, location, square footage, bedroom count and bathroom count while commercial office listings were gathered from Zolo.com with additional information related to office size, location and sale price. These listings datasets represent the basis of this work as predictive models were trained using apartment listings to make predictions for commercial listings. Beyond listings of interest, spatial features of interest included location of grocery stores, green space, planned affordable housing developments and subways, in addition to crime rates across downtown Toronto. Distance from each listing to the nearest grocery store, nearest green space, etc. was calculated and included in the predictive models. These features were hypothesized to have impacts on residential rent prices with inclusion in modelling achieving greater accuracy in rent price prediction.

Methods

Using the datasets described above, several predictive models were trained. The first group of models was created to interpolate rent prices across downtown Toronto at singular spatial locations. These models were used to make point estimates of residential rent for each commercial office located in a mixed zoning area. Following this, the locations with the top five highest predicted rent/sale price ratio were compiled to highlight premier investment targets. Comparatively, the second model of interest was trained to predict rent prices across Toronto at the neighborhood level. Commercial office spaces located in the neighborhoods with the highest predicted rent prices were extracted as the most viable investment targets for comparison against initial point-wise prediction results. The outlined models also included bedroom count and apartment size as predictors, in addition to spatial features.

Results

Findings of spatial interpolation of rent among commercial investment targets indicated that office space located at 37 Madison Avenue, 31 Elm Street, and 407-80 Queen Street East would retrieve the highest potential rent price if converted to residential dwellings. Comparatively, when office space sale price was taken into account, locations at 507-800 Bathurst Street and 308-120 Carlton Street were predicted to retrieve the highest return on investment. Following these findings, identified locations were explored for suitability for conversion from office space to residential space. Suggested investment targets appeared adequate for residential conversion with the space on Bathurst Street already located within a partially residential building. Alternative modelling methods had results congruent with spatial interpolation. Despite some positive results, there was a high degree of prediction variance during spatial interpolation that could heavily impact outcomes. Overall the proposed models provide a general indication top commercial investment targets that could be helpful in understanding the rent value at locations in downtown Toronto.

Limitations

Despite model successes, several limitations were also present that could impact findings. First, access to commercial and residential datasets was a notable limitation. Compilation of residential listings was limited by slow computation times resulting in a reduction in the area of focus to just downtown Toronto. Additionally, commercial office listings were plagued by missing data and artefacts leading to the removal of ~75% of initially gathered listings. These limitations reduce generalizability of proposed models and overall findings. Investment targets of interest may not represent the overall best investment options, just those from where data was available.

Additionally, the outcomes of this research are limited by having no method to evaluate the actual potential for a given commercial office space to be converted to a residential dwelling. While the work identified ideal investment targets, findings are based only on predicted rent. It is possible that some of the top investment targets are unfeasible for residential conversion given their physical structure/appearance. Certain office spaces can not realistically be converted to residence while maintaining an aesthetic appeal that would attract renters. For this reason, findings must be used strategically to initially identify investment targets before further research is completed to understand the true potential of residential conversion. Finally, the findings of this work are limited by the inability to assess accuracy of rent predictions on commercial buildings. As true residential rent prices for these locations do not exist, the accuracy of predicted rent is unknown. Efforts were made to evaluate model accuracy using the apartment listing dataset, however, the outcomes of the analysis cannot be generalized fully to predictions made on commercial office spaces. Overall, limitations encountered in this work impact the ability to generalize findings, however, in the context outlined in the full report, findings are meaningful in providing initial indication of profitable commercial investment targets for the purpose of residential conversion in downtown Toronto.

Report

Introduction

The COVID-19 pandemic has had major global economic impacts, one of the most significant being the rise of work-from-home and the overall residential-based labour market. While much of the world has returned to normal, commercial real-estate markets continue to experience the effects of poor commercial office-space demand⁹. Strain has also grown on the residential real-estate market as rent prices explode in urban areas. The culmination of these issues has been particularly intense in Toronto, Ontario where downtown office vacancies have reached 15.3% in the first quarter of 2023¹⁰. Additionally, one bedroom condominium rent has increased 15.1% over the past year ¹¹. *Figure 1* provides indication of the trends in Toronto residential rent over the last 30 years. As office demand drops, residential demand rises and interest rates rise to combat inflation, commercial real estate owners will look to offload properties to avoid defaulting on their property loans as their properties stop appreciating in value¹². The outcome could be discounted commercial buildings available for purchase in areas of high population density. It is likely some of these buildings will be in areas of mixed zoning allowing conversion of commercial office space to residential dwellings. This will provide real estate investors a unique opportunity to generate significant return on investment if they can identify office buildings with the highest potential residential rent price at the lowest sale price. The current work aims to analyze the Toronto rent market to identify the influence of geo-spatial features on apartment rent prices. The project consists of two primary objectives. The first is to train a spatial interpolation model to predict Toronto rent prices given a range of characteristics including location, proximity to amenities, square footage, proximity to green space and proximity to transportation. The second objective is to implement the model for inference in the prediction of rent prices for commercial buildings if converted to a residence. The second objective will specifically target areas where zoning laws permit the conversion of office space to residence. Overall, the outcomes of this research will be a trained model capable of predicting rent price given characteristics of a real estate location, as well as a comprehensive understanding of the most valuable office space from a residential perspective.

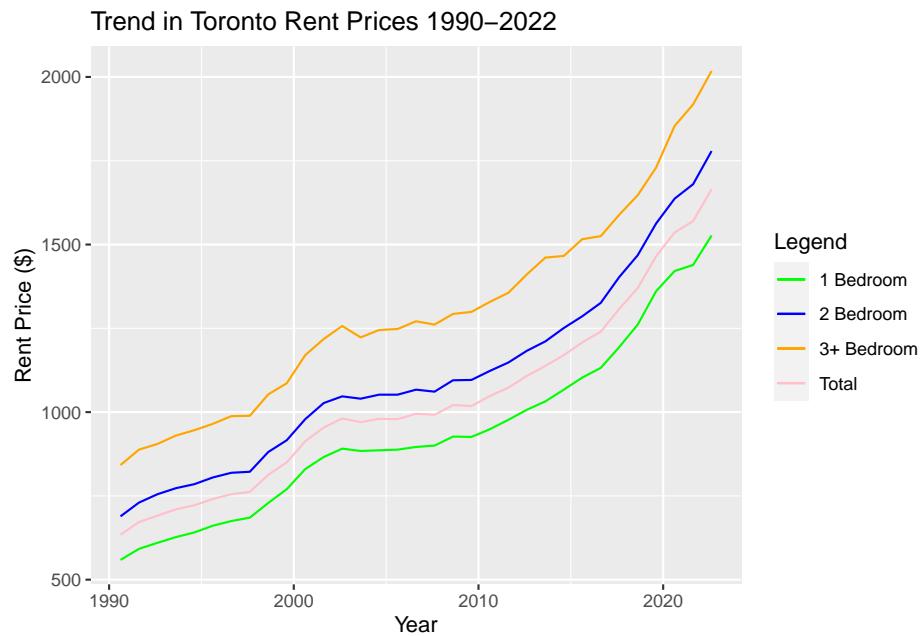


Figure 1: Trend in rent prices across Toronto over the last 30 years, separated by number of bedrooms

Data Sources

The data explored in this report spans several separate datasets as the final listing price analysis model must incorporate rental listings and their associated features with datasets related to various amenities and spatial features including green space, crime rates, grocery stores and affordable housing developments. Additionally, inference regarding rental prices will be made on a separate dataset pertaining to commercial office spaces and their features. Datasets are described below. It should also be noted that the CRS used for this analysis was UTM zone 17N (EPSG:2958) while modelling was completed with data transformed to a projected CRS, MTM zone 8 (EPSG:2017).

Rental Listings

Rental listings used in this analysis were gathered from rentals.ca¹ for a region within Toronto focused on downtown and the close neighbouring area. The bounding box used to select this area is noted in the associated analysis code. A subsection of Toronto was selected for this analysis due to the high density of rental listings across the city and computation rate issues related to the scraping process. Positive results in this research could warrant further exploration of a larger region of Toronto. The data gathered from rentals.ca included 1748 viable listings with associated location, price, bedroom count, bathroom count and square footage. Price was the response variable of interest in this analysis with values ranging from \$800 - \$41000 and an average of \$3019. Visualization of gathered listing locations is provided in *Figure 2*. Additional cleaning steps undertaken to prepare this data for EDA and modelling are detailed below.

**Locations of Available Rental Listings within
Downtown Toronto and the Surrounding Area**

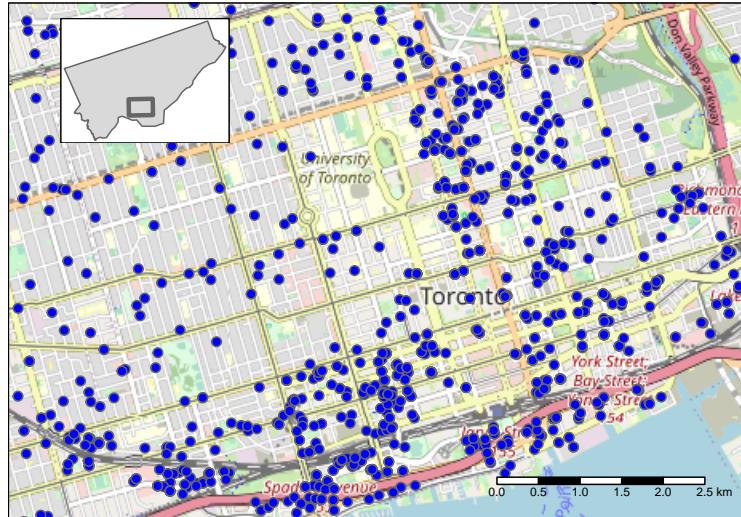


Figure 2: The distribution of rental apartment listings retrieved across downtown Toronto to be used in interpolation of rent

Commercial Listings and Zoning Data

The final dataset of interest gathered for this analysis included the locations of 324 commercial office spaces available for purchase across all of Toronto. This data was scraped from Zolo.ca⁷ and includes the listing location as well as the sale price and the size of the office space. The rent prediction model fit on residential

listings will be used for inference on the commercial spaces dataset as the final outcome of this research. The distribution of commercial office spaces across the area of interest is provided in *Figure 3*, however the listings displayed in this image represent only those residing in mixed zoning areas. 49 final listings are included in this dataset.

Beyond Specific commercial listings, data pertaining to by-law zoning across Toronto was also gathered for use in the model inference stage. A key component of commercial office space rent analysis is that the office space must be eligible to be converted into residential space. To ensure this, zoning areas were retrieved from Toronto's open data portal⁶ and will be used to extract commercial office spaces for sale that reside in mixed-zoning areas.

Zoning Regions of Downtown Toronto, with Locations of Office Space For Sale

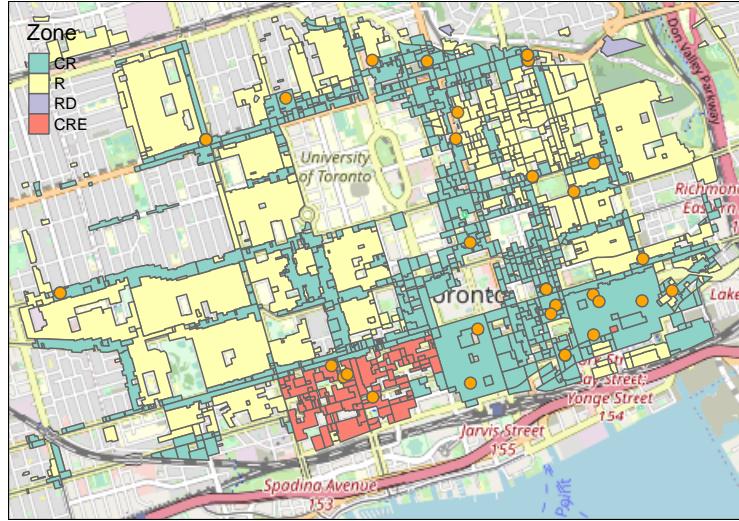


Figure 3: Distribution of commercial office locations within mixed zoning areas in downtown Toronto, to be used for inference during modeling

Spatial Features

As previously noted, several spatial features of interest were retrieved based on their perceived relationship with rental listing prices. Shape files pertaining to crime rates², planned affordable housing developments³, subway lines⁴ and green space⁵ were retrieved from the City of Toronto's open data portal. Additionally, locations of grocery stores in Toronto were gathered as the final spatial feature of interest in modelling. Grocery store data was retrieved from Open Street Map. These features are visualized in *Figure 4*. While the locations of these features were retrieved for the entirety of Toronto, this analysis restricts the features to only downtown Toronto and the surrounding area.

Data Cleaning and Preprocessing

Several data cleaning steps were undertaken prior to data exploration and initial modelling. First, regarding rental listings, initial scraped data required text processing to extract only key values, as well as removing duplicate rows and rows that represented listings for multiple apartments across a range of prices. Following this process, apartment price values were capped at \$9000 and apartment size was restricted between 100sqft and 4500sqft. These adjustments were made as it was observed that artefacts in the data tended to appear

among listings with values outside of these cut-offs. This alteration ultimately aims to eliminate inaccuracies in the data. Further, the apartment size column included approximately 25% missing values. To manage this issue, square footage values were imputed based on the mean size of apartments that had the same number of bedrooms. The pre-processing methods removed approximately 125 observations from the final dataset. The final data alteration of note was removal of listings located at the same listed location. Including multiple listings from the same location was found to disrupt kriging attempts so duplicate locations were removed. This process resulted in removal of approximately 600 observations.

Beyond cleaning related to the listings dataset, data pertaining to spatial features underwent simple pre-processing to convert features to a consistent CRS matching the listings dataset. Data pertaining to Toronto zoning was filtered to include only zones of interest, such as those falling under the residential and commercial zoning umbrellas. Lastly, data related to commercial listings underwent pre-processing similar to the residential listings data after scraping was complete. This process extracted only necessary values into a concise dataset while missing values were dropped from the dataset as Zolo.ca masked the features of certain office spaces. The commercial listings were also filtered base don the observation window of interest surrounding the Toronto downtown area. As the eventual interpolation model for rent price will be fit based only on the downtown and surrounding area, inference will only be made on commercial buildings in this region as well. The final commercial listings dataset contained 96 observations of listing location, price and size.

Methods

This work presents a series of models to be used in the exploration of potential residential rent price among commercial office spaces. Following exploratory analysis, the primary model investigated was a universal kriging model that included previously defined predictors in an effort to interpolate rent prices with greater accuracy than ordinary kriging. Four separate kriging models were created based on findings during exploratory analysis and model evaluation. Based on outcomes of exploratory analysis, the two feature sets of interest were the full set with all previously outlined predictors, and the reduced set which included only bedroom count, size, proximity to affordable housing developments and proximity to a grocery store. Two models were created to predict rent price outright using these feature sets while two models were created to predict the logarithm of rent price. A summary of models under investigation is provided in *Table 1*. The decision to attempt kriging using the the logarithm of price was made due to negative price estimates being made by the original models. Using the logarithm of price ensured that all estimates remained above 0. The variogram used in these models also included an error term to promote smoothing of estimates during kriging. After fitting the kriging model, the model was subsequently used to interpolate rent prices across all commercial office space locations of interest. This process required configuration of the commercial listings dataset to include the same feature set as that of the apartment listings. Following final predictions of potential rent price, commercial locations were ranked based on absolute predicted rent price as well the ratio of rent price to building sale price to determine final investment targets. Kriging attempts were verified using 5 fold cross-validation based on the original listings dataset to provide indication of overall model efficacy. Beyond kriging, modelling efforts also involved use of an INLA model to predict rent prices across Toronto at the neighborhood level, with neighborhood boundaries defined by the boundaries used in the crime rates dataset. This model was meant to provide alternative results against which to compare the outcomes of kriging. Initially, areal estimation of rent price was intended for the entirety of Toronto instead of restricting prediction to the downtown area. Issues and inaccuracies in modelling ultimately led to use of just the downtown area for this model as well. A single INLA model iteration was observed based on the same feature set used in the reduced kriging model and a besag model applied for approximation of structured spatial random effects. After predicting rent across neighborhoods in the downtown Toronto area, commercial office locations situated in the regions of highest predicted rent were extracted for final analysis. Motivated by findings of the INLA model, a final generalized linear model based on a gamma family distribution was also developed to observe predictive ability of a model using the features of interest without the inclusion of a spatial structured random effects term. This model was again fit using the full feature set and used to predict commercial locations with the top rent to sale price ratio. Final results produced using

Table 1: Models explored in this work included the predictor set, response variable and purpose in this work

Model	Features	Response.Variable	Purpose	Title
Kriging Full, Standard Model	All Predictors	Rent Price	Primary Observation	Full Standard
Kriging Reduced, Standard Model	Reduced Set	Rent Price	Primary Observation	Reduced Standard
Kriging Full, Log Model	All Predictors	Log of Rent Price	Primary Observation	Full Log
Kriging Reduced, Log Model	Reduced Set	Log of Rent Price	Primary Observation	Reduced Log
INLA	Reduced Set	Rent Price	Comparison Model	INLA
GLM	All Predictors	Rent Price	Comparison Model	GLM

the three model types were compared to obtain final conclusions.

Results

Results of the outlined modelling process are separated into sections below including exploratory analysis outcomes, ordinary kriging, universal kriging and areal modelling with INLA.

Exploratory Analysis and Spatial Association of Predictors

The compound plot provided in *Figure 4* explores the distribution and observation density of spatial features of interest. Plots 4A and B detail the locations of grocery stores and planned affordable housing developments in addition to the intensity of observations across the region of interest. Regarding grocery stores, observations appear clustered with areas of high intensity located in the corners of the downtown area with relatively low density through the interior of the region of interest. Based on this plot, it is clear that proximity to grocery stores as well as the number of nearby grocery stores will vary across rental listings, motivating this feature for inclusion in the proposed predictive model. Additionally, planned affordable housing developments vary across the region of interest with highest intensity across the north east of the region and lowest intensity across the south west. It is expected that rent prices will be impacted by the density of these affordable housing developments as low-cost housing may drive down rental market value. Despite this intuitive outcome, some studies have actually found the opposite with affordable housing developments increasing surrounding rental prices as much as 6.5%⁸. Regardless of the net positive or negative affect on surrounding rental prices, affordable housing developments will be included as a spatial feature in final modelling efforts to capture possible variation in rental prices stemming from these projects.

Finally, plots 4C and D indicate the distribution of green space and crime rates, respectively, across downtown Toronto. Dispersion of green space is somewhat uniform across the region, however, it appears the waterfront area has a slightly reduced quantity and average size of green space. This variation promotes inclusion of proximity to green space as a feature during predictive modelling. Variation in crime rates per 1000 people across the region varies greatly with notable increases in crime rates within the University, Downtown Yonge East and Moss Park regions. Plot 4D was created using only reported robberies due to the unavailability of aggregate crime statistics. It should be noted that crime rates appear to correlate positively with affordable housing developments. As well, planned affordable housing developments and grocery stores were found to correlate closest with rent prices compared to other predictors. Overall, the spatial features explored in *Figure 4* provide initial indication of potential sources of spatial variation in rental listing prices. This variation motivates inclusion of these variables in predictive modelling efforts to improve accuracy by accounting for a higher proportion of spatial variation in listing prices.

Variogram and Ordinary Kriging

Following observation of spatial features, a baseline ordinary kriging model was created to provide initial indication of interpolated rent across downtown Toronto. To achieve this, a variogram model was fit to



Figure 4: Distribution of spatial features of interest across downtown Toronto. Plots detailing grocery stores and planned affordable housing developments provide the density of these features in relation to observed listing locations

evaluate variance in rental price between listings over a range of separation distances. A nugget was apparent in the variogram model indicating that subsequent modelling efforts should include an error term to smooth rental price estimates at the point of observation. Rent price interpolation across downtown Toronto using the ordinary kriging model is visualized in *Figure 5*. This figure provides initial indication of rent variation across Toronto, however, noise is apparent as there is high variance around observed listings. This noise further motivates inclusion of additional predictors and an error term in a final universal kriging model. Additionally, this variation promotes the use of a secondary areal model to estimate rent across Toronto neighbourhoods as opposed to interpolation of prices using point estimates. Modelling in this way could improve the accuracy of rent price prediction if the universal kriging model was found ineffective. Overall, indicators suggested universal kriging would be able to improve upon ordinary kriging efforts to provide more accurate estimates of commercial office residential rent.

Predicted Listing Prices Across Downtown Toronto Using Ordinary Kriging

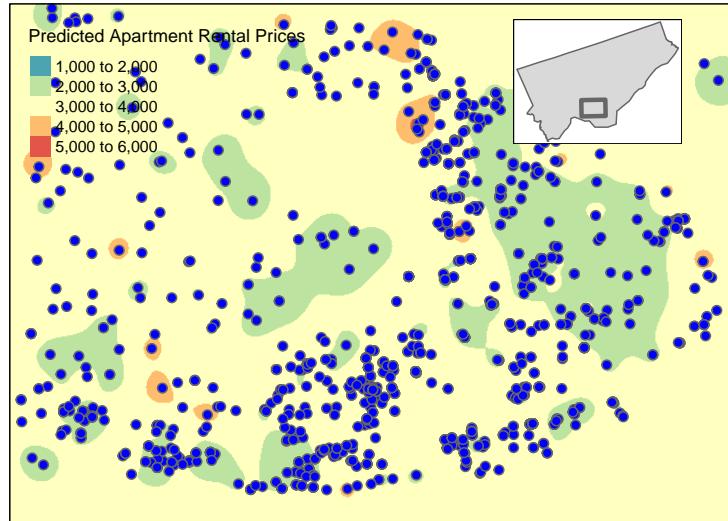


Figure 5: Predicted rent prices across downtown Toronto using ordinary kriging

Universal Kriging

Following ordinary kriging, the main models under consideration in this work were created. Two universal kriging model versions were explored, each with a variogram that included an error term to improve smoothing of interpolated rent price values. The first model version included all previously mentioned predictors to interpolate rent prices while the second model included only a subset of variables, motivated by findings from the exploratory analysis. After fitting universal kriging models, each was used to interpolate rent price at commercial office locations. Results of this process showed generally reasonable rent price predictions for most locations with predicted values for the full standard model ranging from \$1944 to \$15733 a mean of \$8041 and a standard deviation of 247.9.

While predicted values were reasonable, further analysis of model accuracy using cross validation indicated that models occasionally predicted negative rent price values. This result motivated the use of a log transformation in the response variable after alternative strategies failed to limit predictions above 0. Secondary models using the same predictor sets were trained to predict the logarithm of the rent price before being used in interpolation of rent at commercial locations. Use of a log transformation appeared successful in maintaining reasonable predictions for commercial office rent, and predicting positive values when tested in a cross validation process, when predictions were transformed back to their original scale. It should also be

noted that all models of interest tended to estimate high rent prices among commercial locations. *Figure 6* provides the locations of the office spaces with the top five highest predicted rent and the highest predicted ratio of rent to sale price based on standard models while *Figure 7* provides the top locations based on cost to rent ratio using the log models. Top locations based on absolute rent price were the same between model types. As well, rankings remained the same for both the full and reduced feature set models. Top investment locations based on rent to cost ratio differ meaningfully from those with the highest predicted rent and differ between models. Top investment locations based on absolute predicted rent as well as rent to cost ratio are of value as investors looking for a property to hold long term may benefit more from a higher absolute rent price. Comparatively, those who want to break even as fast as possible would find more value in the rent to cost ratio.

Locations of Office Space with Top Five Highest Predicted Rent Price Using Standard Model



Locations of Office Space with Top Five Highest Predicted Rent to Cost Ratio Using Standard Model



Figure 6: Top investment locations based on absolute predicted rent price (top) and ratio of predicted rent to sale price (bottom) for kriging models predicting outright price.

Model Validation

Following prediction of optimal commercial office investment locations, cross validation of kriging models was completed. This cross validation process provided the only means to assess model accuracy as truth values do not exist for commercial targets to allow validation in that form. 5 fold cross validation was used

Locations of Office Space with Top Five Highest Predicted Rent to Cost Ratio Using Log Model



Figure 7: Top investment locations predicted based on ratio of predicted rent to sale price for the log models.

with outcomes displayed in *Figure 8* below. The figure indicates the magnitude of the residual at each apartment location for the standard and log models containing all predictors. Residuals appear inconsistent in magnitude across downtown Toronto with some spatial auto-correlation appearing across regions of the observed area. Notably, the areas near the predicted optimal investment locations appear to have generally small residuals compared to regions such as the north edge of the downtown. As well, residuals appear small in the south and eastern areas of downtown with larger residuals appearing moving away from downtown to the north and west. While there appears to be definite spatially correlated variance in the magnitude of residuals, predictions appear generally accurate with the mean of the absolute value of residuals for the standard model observed at 828.88. This means predictions of apartment rent price were normally within \$800 of the true value. While this error rate may not be acceptable in some instances, for the current work, this mean absolute error is reasonable.

INLA Model

Beyond efforts in predicting potential rent at specific commercial point locations, alternate results were found when INLA was used to estimate rent across Toronto neighborhoods. Initial plans centered around the use of INLA for prediction of rent prices across all neighborhoods of Toronto, however, model inconsistencies and evidence of significant differences in the geographic structure of Toronto amenities outside of the downtown area were identified. As a result, it was decided that only the downtown area would be assessed, primarily for the purpose of comparing results to those of universal kriging models. The INLA model designed in this work included the reduced feature set, averaged over neighborhoods of Toronto, to predict outright rent price across the neighborhoods. The model was found to have an almost negligible spatial structured random effects term indicating the high degree of variation captured by the included predictors.

Based on neighborhood rent estimations visualized in *Figure 9*, rent prices for commercial locations were extracted and ranked. Findings from this process indicated locations at 101-800 Bathurst Street and 415-120 Carlton Street to have the highest absolute predicted residential rent. The results of the INLA model showed a very narrow prediction range, varying no more than \$500 among neighborhoods. Beyond INLA, rent price modelling using a GLM found results similar to those obtained using kriging models. Commercial locations with the highest predicted rent were found to be 37 Madison Avenue, 31 Elm Street and 404-80 Queen

Magnitude of Residual at Listings Across Toronto (Standard Model)



Magnitude of Residual at Listings Across Toronto (Logarithm Model)

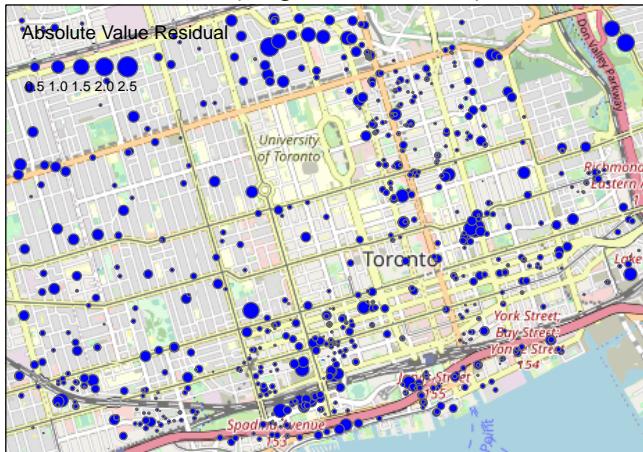


Figure 8: Spatial variation in residuals of kriging models fit using the full feature set. The image presented on the top provides residuals from the model using price as the response while the image on the bottom indicates residuals based on $\log(\text{price})$.

Street East. As this model was intended purely for comparison of results against kriging, limited emphasis was placed on feature selection or achieving maximum accuracy. Instead, the model was used to support or refute results of kriging.

Fitted Values Predicted With INLA for Rent Price Across Neighborhoods

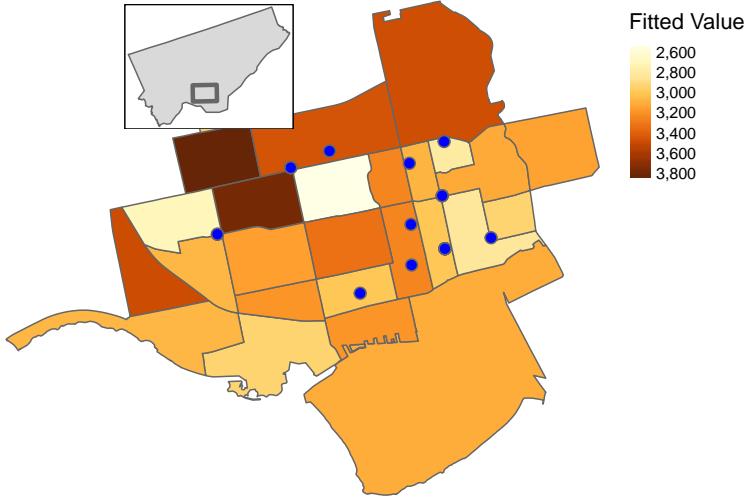


Figure 9: Predicted rent price across neighborhoods in downtown Toronto. This image provides indication of neighborhood size as well as variation in predictions.

Discussion

Based on the universal kriging results presented above, it can be concluded that the proposed models have potential in the identification of commercial office investment targets. While all four kriging models provided similar overall results, the models that included all predictors performed best. This conclusion is based primarily on the outcomes of cross validation. Cross validation supported findings that all four models are viable in predicting rental prices with the lowest mean absolute error achieved by the standard and log full models. This supports the selection of spatial predictors in this work as the inclusion of subway proximity, green space proximity and crime rates had a positive impact on model outcomes. As well, the positive outcomes recognized in the cross validation process promote future work using universal kriging in estimation of apartment rent across across Toronto. This use case could be valuable for renters and landlords, especially those with property in areas where models predicted with low error. Although most areas of downtown Toronto retained somewhat low error during prediction, areas to the north and west had generally higher error, possibly due to reduced listing density with high price variance in these areas.

Despite the positive initial results, outcomes of the kriging process are not without uncertainty. The standard full model had certain instances of predicting negative price values during cross validation. While this problem did not impact prediction of rent at commercial locations, it is an issue that indicates underlying model inaccuracies. During model testing, several strategies were tested to rectify this issue including capping the upper extremes of predictors and altering the predictor set used in model fitting. These strategies were unsuccessful and motivate a problem to be explored in future work. As well, some of the predicted rent values obtained from kriging models were extremely unrealistic. While investment target rankings were consistent the actual rent predictions need further assessment. Additionally, while the cross validation process was useful in obtaining a general sense of model accuracy, the findings of the process cannot be guaranteed to

translate into accuracy in prediction of rent at commercial locations. As a result, model accuracy cannot be ensured for those using the current model to identify investment targets.

Beyond understanding the viability of the universal kriging models, results of the areal INLA model and the GLM allowed further analysis of proposed models through comparison of outcomes. The INLA model performed poorly overall with errors arising when attempting to include the full feature set as predictors. As previously noted, the spatial structured random effect term of the model added almost nothing to fitted values and bedroom count was indicated as the only significant predictor in the model. It is likely that further development of the INLA model would be required to achieve meaningful results. Issues in the aggregation process for spatial predictors when aggregating listings across neighborhoods may be partially responsible for the poor model outcomes. The large size of neighborhoods may have negatively impacted results by capturing areas with great range in socioeconomic status.

As expected, commercial office rent predictions of the INLA model were not consistent with those observed during the kriging process. This outcome may be a sign of strength in the kriging models as they appear to be capturing variation between listing prices at a lower level than the INLA model. For this reason universal kriging is a more effective strategy in handling the problem of rent interpolation in downtown Toronto, compared to an areal approach such as the INLA model observed here. In addition to INLA results, rent prediction using a GLM produced results similar to kriging when predicting rent price among commercial locations. This finding reinforces the potential of kriging in rent prediction as it indicates that the listings suggested as optimal investment targets by the model were less likely to have occurred simply due to chance. This result also promotes further development of the GLM model through feature elimination and cross validation as an alternative to the kriging models.

Following comparison of kriging to outcomes of extended modelling, the final analysis completed explored conversion potential of commercial office investment targets. *Figure 10* shows two images of top identified locations on Bathurst street where the office is located on an upper level of the building and on Elm Street. Observation of the interior of these buildings and other identified locations provided indication that these offices could very likely be converted to residential dwellings. The identified offices were spacious with natural light and layouts that would not restrict a residence. This additional exploration further motivates the analysis and investment strategy presented in this work. Overall, the universal kriging models that included the full feature set were most successful in interpolating potential residential rent prices at commercial office locations. While there remains some inconsistencies in model predictions and accuracy, further exploration into the causes of these issues will allow for guarantees regarding model accuracy.

Limitations and Future Work

While the objectives of this work were met, successfully identifying several investment targets, several limitations were identified that reduce generalizability of the current work and promote future research. First, generalizability of this research was negatively impacted by limitations in the gathered real estate listing data. While the scraping process was generally successful in compiling both apartment rental listings and commercial office sale listings, retrieval was largely reliant on accuracy of data on listing websites. As a result, apartment listings were found to contain artefacts where original data was input incorrectly. Although certain pre-processing steps were undertaken to correct these errors such as filtration and capping of apartment size and price variables, it is likely the some inconsistencies remained. Further, related to commercial listings, restrictions from the Zolo.com site resulted in large quantities of missing data as well as the inherent issue of missing bedroom counts among commercial listings. It is likely that solutions to remedy these issues were imperfect, leading to some inconsistencies in the interpolation/prediction process for estimating rent. Overall, the data limitations inherent to the sites from which listing data was retrieved was a factor that could negatively impact accuracy of modelling outcomes.

Beyond general data inaccuracies, this work is dependent on the locations of current rental apartment listings. The locations of these listings at any given time are arbitrary, resulting in final models that may ultimately be accurate in some locations of downtown Toronto and inaccurate in others due to availability of listings. This outcome depends solely on the rental locations available during model fitting. Future work on this topic



Figure 10: Images of top predicted office spaces for investment on Bathurst Street (left) and Elm Street (right)

could benefit from expanding the training dataset through incorporation of listings from a variety of rental sites instead of single sourced data.

Further, the outcomes of the current work are limited by the inability to truly assess model accuracy in predicting potential rent prices of converted office space. Despite positive findings related to potential investment targets, cross validation during the kriging process was the only form in which model accuracy could be assessed. As model inference was performed on commercial properties that do not have a true rent price associated with them, validation of inferences made at these locations was impossible. The accuracy assessment completed using apartment listings and cross validation may not generalize to commercial properties which reduces the certainty of results.

Additionally, a major limitation of this work is the inability to observe information related to the difficulty or outright ability to convert office space to residential space. The current model can indicate the potential rent return for office space, but cannot ensure the office space can reasonably be converted to residential space. As a result, the models proposed in this work exist as initial investment indicators rather than a more comprehensive analysis solution. To achieve the most benefit out of this work, additional research to explore an identified location is required. The inability to assess conversion potential in an office location greatly limits the practicality of the model outcomes. Future work could explore a supporting model using a convolutional neural network to estimate the renovation cost of an office based on previous instances of residential conversion if such data is available. This model could support the current work in providing an additional dimension of cost to help estimate profitability of a property.

Finally, further limiting application of findings is the restriction of analysis to the downtown area. While the initial intention of this work was to extend modelling to include all of Toronto, limitations in computation speed as well as significant differences in the distribution of apartment listings throughout Toronto motivated use of just the downtown area. Although this decision likely allowed for more accurate predictions near downtown, it also reduced generalizability of the model. Focusing on just the downtown area was successful, however, future work should look to explore extending the model to other areas of Toronto gradually to understand the trade-off between increased training data and the dissimilarity of various regions of Toronto.

Conclusion

Overall, the universal kriging models presented in this work show promise in applications of rent price prediction in urban areas of high population density. Of the four models presented in this work, those including all of bedroom count, size, grocery store proximity, proximity to planned affordable housing developments, proximity to green space, proximity to the subway and crime rate as predictors, performed best. Kriging models predicting rent price outright as well as the logarithm of rent price were found to have general success in interpolating rent across commercial office spaces to identify targets for conversion to residential dwellings. Model inference identified commercial locations at 37 Madison avenue, 31 Elm Street and 407-80 Queen Street East as those with the highest predicted residential rent price while locations at 507-800 Bathurst Street and 308-120 Carlton Street had the highest predicted rent to sale price ratio. As well, comparison of kriging model results with those achieved using a GLM indicated a high degree of similarity, further supporting the strength of the universal kriging models. This work was successful in generating a spatial model to be used in the prediction of potential rent prices of commercial office space if converted to residential space. With further tuning and deeper evaluation of model accuracy, the models presented in this research could have meaningful applications in general rent price analysis for apartments as well as larger scale analysis of commercial office investment targets.

References

Data Sources:

1. Rentals.ca. (2023). Toronto. Retrieved from <https://rentals.ca/toronto?bbox=-79.43203,43.63689,-79.35401,43.67521>
2. Toronto Open Data. (2023). Neighborhood Crime Rates. Retrieved from <https://open.toronto.ca/dataset/neighbourhood-crime-rates/>
3. Toronto Open Data. (2023). Upcoming and Recently Completed Affordable Housing Units. Retrieved from <https://open.toronto.ca/dataset/upcoming-and-recently-completed-affordable-housing-units/>
4. Toronto Open Data. (2023). TTC Subway Shapefiles. Retrieved from <https://open.toronto.ca/dataset/ttc-subway-shapefiles/>
5. Toronto Open Data. (2023). Green Spaces. Retrieved from <https://open.toronto.ca/dataset/green-spaces/>
6. Toronto Open Data. (2023). Zoning By-law. Retrieved from <https://open.toronto.ca/dataset/zoning-by-law/>
7. ZOLO. (2023). Toronto Office Spaces for Sale. Retrieved from <https://www.zolo.ca/toronto-real-estate/commercial-office>

Additional References

8. Holder, s. (2022). What Does Affordable Housing Do to Nearby Property Values?. *Bloomberg*. Retrieved from <https://www.bloomberg.com/news/articles/2022-05-02/does-affordable-housing-lower-property-values>
9. Gupta, A., Mittal, V., & Van Nieuwerburgh, S. (2022). Work from home and the office real estate apocalypse (No. w30526). National Bureau of Economic Research.
10. The Canadian Press. (2023). Canada's national office vacancy rate hit an all-time high in Q1: CBRE. Retrieved June 9th, 2023 from <https://financialpost.com/real-estate/property-post/canada-national-office-vacancy-rate-all-time-high-q1-cbre>

11. Toronto Residential Real Estate Board. (2023). Rental Market Report. Retrieved June 6th, 2023 from <https://trreb.ca/index.php/market-news/rental-market-report>
12. Glaze, A. (2023). How Rising Interest Rates Can Hurt—Or Help—Businesses. Retrieved June 9th, 2023 from <https://www.forbes.com/sites/forbesfinancecouncil/2023/01/31/how-rising-interest-rates-can-hurt-or-help-businesses/?sh=e6adcb24f0eb>
13. Bank of Canada. (2023). Price Check Inflation in Canada. Retrieved June 9th, 2023 from <https://www.bankofcanada.ca/2019/02/price-check-inflation-in-canada/>

Images

<https://www.acotoronto.ca/building.php?ID=4169>

<https://www.realtor.ca/real-estate/25521710/507-800-bathurst-st-e-toronto-annex>