university of groningen

faculty of spatial sciences

Local spatio-temporal regression kriging for property price predictions

Abstract

This thesis proposes a hedonic regression model that accounts for spatio-temporal dependence and heterogeneity in property transaction prices. The considered methodology extends upon purely spatial geostatistical methods and applies local spatio-temporal regression kriging (LSTRK). We account for spatio-temporal dependence by specifying dual structure variograms. Since such variograms consider distances between properties in space and time simultaneously, they allow for the estimation of a single covariance structure that controls for spatio-, temporal- and joint dependence. We account for heterogeneity of both attribute prices and dependence structures by performing the analysis per individual submarket. To empirically examine the prediction accuracy of our model, out-of-sample forecasts are made on residential housing prices. We consider a housing transaction dataset of 57,154 properties in Mecklenburg County, North Carolina. Our data is split into 7 spatial zones and 5 temporal zones, providing us with a total of 35 submarkets. For each submarket, in-sample- and out-of-sample subsets are obtained. These allow for model training and prediction assessment respectively. In terms of mean absolute error, median absolute error, mean squared error and root mean squared error, the proposed methodology outperforms traditional hedonic models consistently. Hence, empirical results suggest that LSTRK indeed provides more accurate predictions than traditional OLS models.

Keywords Spatio-temporal dependence · Spatio-temporal heterogeneity · House price prediction

Author: P.L.Weenink
Supervisor: dr. X.Liu
Assessor: prof. dr. ir. A.J. van der Vlist
E-mail: p.l.weenink@student.rug.nl

Date: August 22, 2022

university of groningen

faculty of spatial sciences

University of Groningen

Local spatio-temporal regression kriging for property price predictions

Master's Thesis

To fulfill the requirements for the degree of
Master of Real Estate Studies
at University of Groningen under the supervision of
dr. X.Liu (Real Estate Studies, University of Groningen)
and
prof. dr. ir. A.J. van der Vlist (Real Estate Studies, University of Groningen)

P.L.Weenink(s3246620)

August 22, 2022

Disclaimer

Contents

List of Figures

List of Tables

Introduction

Real estate markets remain inefficient due to their heterogeneous nature (Gatzlaff and Tirtiroğlu 1995). Limited transactions, lack of centralized trading and information asymmetry lie at the core of this market inefficiency. Given this illiquid nature, property prices are not consistently reflected and actors have to rely on appraisals in order to get an understanding of where supply and demand meet (Mooya 2016). These actors comprise a number of individuals, institutions and companies in all layers of society. E.g. tax authorities rely on properties value estimates for levying property taxes; mortgage providers conduct collateral valuations to qualify borrowers for their mortgage applications; real estate portfolio managers and investors make investment and sales decisions based on recurrent evaluations of their portfolios; and individuals want to know their property value before setting up list prices (Liu 2012). Hence, efficient and accurate valuation models are necessary to let real estate markets function effectively (N. Crosby et al. 2018).

Valuations rely on advanced statistical methods that utilize actual market data to fit model parameters. Once trained, such models can be used to make out-of-sample price predictions (Leung et al. 2010). Given the uniqueness of individual properties and irregularity of transactions, statistical models frequently utilize the hedonic framework (Kuntz and Helbich 2014). This framework has been pioneered by Rosen (1974) and considers properties as heterogeneous goods, that can be decomposed into utility bearing bundles of attributes. The market equilibrium dictates that property prices can be determined by the total amount of utility-generating characteristics intrinsic to these houses (Kuntz and Helbich 2014; Liu 2012). Accordingly, hedonic regression techniques can be used to estimate these implicit attributes' prices and utilize these to predict property prices of any desired composition. Traditionally such regression techniques utilize ordinary least squares (OLS) in which structural property characteristics, location attributes and the time of sales are independent variables, and log transaction prices are the dependent variable (Liu 2012).

Although OLS is overall considered as the best linear unbiased estimator (BLUE), this only holds true if all OLS assumptions are met (Harville 1976). However, for the property market this is frequently not the case. According to Liu (2012) the assumption that there is no correlation in the error term is frequently violated. This causes parameter estimates to be inefficient and erroneous. Reason for this error correlation can be found in the spatial and temporal domain. Liu (2012) argues that there are four distinct phenomena that affect real estate markets simultaneously. These include: spatial dependence, temporal dependence, spatial heterogeneity and temporal heterogeneity. Spatial dependence occurs because houses within neighbourhoods tend to be homogeneous and share location amenities. Hence, such properties are highly substitutable and thus similar in prices. Temporal dependence occurs because earlier transactions contain information to later transactions. As such they can proxy for general market trends or changes in institutional settings. Spatial- and temporal heterogeneity happen because supply and demand of property characteristics shift over space and time. Therefore, prices intrinsic to characteristics may differ per submarket and period.

Since these four phenomena affect prediction accuracy, Liu (2012) stresses that a traditional hedonic pricing model is not sufficient and one needs to account for these effects. While augmentations of traditional hedonic OLS models have been considered to account for space-time trends, for example by inclusion of interactive space and time dummies or by clustering space and time errors, several publications indicate that dedicated space-time models provide more reliable results (e.g. Liu 2012; Chica-Olmo et al. 2019; Hayunga and Kolovos 2015). Considered approaches generally build upon spatial models and transform these into space-time models trough natural extensions. As such, these models do not merely consider separations between properties in space, but also in time (Kyriakidis and Journel 1999; Case et al. 2004). Hence, space-time models can be grouped within the same two modelling paradigms that are commonly acknowledged for purely spatial applications. The first uses weight matrices to reflect the behaviour of the error term and is commonly referred to as spatial econometrics. This approach aims to model the process that generates the error term explicitly, and requires an a priori specification of the dependence structure (Dubin 1998b; Chica-Olmo et al. 2019). The second approach utilizes a functional form to model the covariance structure of the error term directly and is known under the umbrella term geostatistics. Typically, functional forms are considered that expect dependence between properties to fall as their separations increase. Both methods estimate regression coefficients and model parameters by means of maximum likelihood methods (Dubin 1998b). Commonly considered spatial econometric models involve mixed regressive-autoregressive models, spatial autoregressive models, conditional autoregressive models, cluster analysis, spatial lag models, spatial error models and autoregressive disturbance models (Elhorst 2014). Spatial geostatistical methods involve a number of interpolation methods including triangular networks, inverse distance weighting and several kriging approaches (Hengl 2009).

When the extension of spatial models into space-time models is considered, especially models that fall into the econometric modelling paradigm have gained attention within the real estate literature. Considered approaches include spatio-temporal autoregressive models (STAR) plus similar variations (e.g. Pace et al. 1998; Clapp 2004; Sun et al. 2005; Pace et al. 2000; Liu 2012) and geographically and temporally weighted regression (e.g. Huang et al. 2010; B. Wu et al. 2014; Yao and Fotheringham 2015). Although these methods have proven to be successful in making accurate out-of-sample forecasts, their use has also been criticized. According to Hayunga and Kolovos (2015), spatial linear models have significant issues that are not mitigated by merely including an temporal component within the weight matrix. It is argued that inclusion of lagged dependent variables serve as a form of linear smoothing. Hence, spatio-temporal relations might be found unconditionally, even if these are not present (McMillen 2012). Furthermore, it is argued that spatial econometric models are not suitable for causal analysis (Gibbons and Overman 2012) and that there are several issues with autoregressive models, including an implausible normality assumption, non linearity and endogeneity (Pinkse and Slade 2010). Moreover, it is stated that modelling entire dependence structures is improbable using linear methods (Pinkse and Slade 2010).

Meanwhile, geostatistical extensions into space-time models have gained less attention. While such models have gained popularity in a number of fields, they remain relatively unknown in social sciences. This lack of attention is striking as geostatistical methods have proven to be valuable alternatives to econometric approaches in the purely spatial extent (Chica-Olmo et al. 2019). Although in this purely spatial extent no approach is endorsed over the other, researchers could favour certain modelling choices dependent upon the complexity of the implementation process, and the working data available (Liu 2012). Hence, a wide set of modelling options that encompass both econometrics and geostatistics is deemed favourable to fit specific situations (Fischer and Getis 2010). While numerous geostatistical methods have been discussed in earlier research (e.g. Bourassa et al. 2007; Bourassa et al. 2008; Yoo and Kyriakidis 2009; Kuntz and Helbich 2014; Montero and Larraz 2011; Hayunga and Kolovos 2015; Chica-Olmo 2007; Chica-Olmo et al. 2013; Case et al. 2004), it is remonstrated by Palma et al. (2019) that these studies mainly consider a purely spatial or temporal context, and fail to capture space time trends. This lack of joint space-time geostatistics for real estate is also stressed by Kuntz and Helbich (2014), who explicitly mention the lack of accounting for temporal dependence in their geostatistical model and suggest the use of space-time geostatistics for future research.

Despite general lack of attention, a handful of studies do consider the use of space-time geostatics in relation to real estate. Overall, two extensions are acknowledged. The first involves a multivariate geostatistical approach and mimics a temporal dimension by considering multiple spatial models at different points in time. These are then jointly analysed by means of co-kriging (e.g. Chica-Olmo et al. 2019). The second approach considers time as a true extra dimension. By not merely considering variance in space, but also separation in time, it specifies a single covariance structure that includes both effects simultaneously (e.g. Palma et al. 2019; Li and Revesz 2004). While both extensions improve upon purely spatial methods, the general literature endorses the second approach. Reason for this is that it allows for predictions in between and beyond observations in time. Therefore, it is considered to be more flexible (Heuvelink et al. 2017; Gräler et al. 2016).

While both methods have been considered in the field of real estate, to our knowledge only the first method has been applied in combination with the hedonic framework (e.g. Chica-Olmo et al. 2019). Therefore, it is of interest to see if the more advanced dual structure covariance method, used in the second method, can be integrated with the hedonic framework. The pros and cons of such dual structure models are discussed by Heuvelink et al. (2017). It is argued that this approach has all the advantages of conventional geostatistics. As such, it makes optimal predictions given the available data; quantifies prediction accuracy; takes covariate information into account; and offers modelling flexibility. However, since it considers both space and time in a joint way, it offers a more optimal solution for modelling dynamic spatial variables. It is this ability to create a joined covariance structure, that makes spatio-temporal kriging a powerful tool that distinguishes it from other methods. For example, it can be used for a dynamic selection of local neighbours (in both space and time) that should hypothetically be most correlated[1], and allow for more precise interpolation of the regression residuals. Nevertheless, there are important limitations. Space-time geostatistics are not reliable when extrapolating beyond the the space-time domain of the sampling data. Hence it should not be used to predict future prices, without inclusion of driving forces. Moreover, space-time geostatistics rely on a number of simplifying assumptions that can be difficult to meet. Especially the stationarity assumption, that requires the spatio-temporal (co)variance to be equal, should be considered. Furthermore, modelling of a joint space time covariance structure is more complex than its purely spatial form and additional effort is required to obtain reliable results. Also it can be computationally heavy when large sampling sizes are considered. Given these considerations, space-time geostatistics with a dual covariance structure would be applicable in situations where out-of-sample predictions are required for historical-till-present property prices, and where stationarity can be achieved.

In contribution to the acknowledged econometric methods, we examine the use of space-time geostatistics for property price valuations. More specifically, we consider the use of local spatio-temporal regression kriging (LSTRK). We account for spatio-temporal heterogeneity by solving our to be defined model based upon submarket sample data subsets at specific time frames and take into account spatio-temporal dependence by allowing errors to be correlated and interpolating the residuals. The proposed methodology builds upon existing specifications of kriging methods (Heuvelink et al. 2017). However, it extents these into a space-time model by considering a joint spatio-temporal covariance structure that considers spatial-, temporal- and spatio-temporal semi-variance. For our dual structure covariance specification we consider a space-time variogram. Similar to the purely spatial case, space-time variograms consider semi-variance as a function of separation. However, instead of merely considering separation in space, they also consider separation in time (Venkatachalam and Kumar 2017). In other words, we determine the dual covariance structure based on the locations and times of sales of individual properties, in relation to the location and time of sale of any other property within our data extent. As such, the proposed method distinguishes itself from other applied space-time methods, such as Chica-Olmo et al. (2019) who consider cross-variograms; Liu (2012) that applies sparse weight matrices and Case et al. (2004) that consider prior sales subsets. To our knowledge, we are the first to apply dual structure space-time variography with purpose of real estate valuations. In theory such a specification should outperform other geostatistical methods, as it allows spatial and temporal information to interact while considering all relevant information (Gräler et al. 2013). Nevertheless, we stress the difficulties in specifying a reliable covariance structure and the additional assumptions it comes with. Therefore, it remains to be seen if the proposed methodology is indeed able to compete with the modus operandi (e.g. Mayer et al. 2019).

[1] The accuracy is dependent upon how well the theoretical variogram fits the data.

To qualify the added value of LSTRK in real estate valuations, we ask the following research question:

> ***Does LSTRK provide adequate property valuations and is it theoretically and practically warranted?***

We consider two sub-questions to answer our main research question. In the first sub-question we focus on the predictive power of LSTRK. We use traditional hedonic OLS models as benchmark. Hence, our first sub-question is:

> *How does LSTRK perform compared to traditional hedonic models in terms of out-of-sample property price forecast accuracy?*

To answer this question we establish a system of equations, that is used to fit in-sample models and make out-of-sample forecasts. Out-of-sample forecasting should offer a valid indication of model performance, as it simulates a realistic scenario in which earlier transactions are used to estimate present prices. In addition to LSTRK we consider several OLS specifications that are used for model comparison. These include OLS with trend surface and OLS with interacting space and time dummies and OLS with inverse distance weighting (IDW) errors.

Next, we consider the practicality of LSTRK. While LSTRK relaxes some of the assumptions typically associated with OLS, it relies on alternative assumptions to do so. Furthermore, it requires the specification of various parameters and the inversion of large non-sparse covariance structures, which can be demanding in terms of computing power (Heuvelink et al. 2017). In this regard LSTRK might be considered unpractical, even if able to improve predictions. Hence, our second sub-question is:

> *How performable is LSTRK in terms of model complexity, parameter specification and computing power?*

To answer this question we rely upon the existing literature and consider the modelling process of our own empirical analysis.

To conduct this research, the gstat package in R is used. This is one of the few software solutions that supports spatio-temporal kriging and dual structure space-time variography (Heuvelink et al. 2017). Research steps and modelling decisions are discussed more thoroughly in the methodology section. To validate results and train our model, an extensive dataset with residential property transactions in North Carolina is used. Information contains sales dates, property locations and structural property characteristics. A comprehensive overview of the data to be used is provided in the data section.

The remainder of this paper is organized as follows. Section 2 describes our conceptual model and section 3 our empirical approach. Section 4 describes the data and the exploratory analysis. Section 5 presents the results, and section 6 concludes.

Theoretical Framework

Several methods have been proposed to account for spatio-temporal dependence and spatio-temporal heterogeneity effects. According to Yao and Fotheringham (2015), this is achieved by global- and local spatio-temporal models respectively. Probably the most popular method is the spatio-temporal autoregressive model (STAR) that was introduced to the field of real estate by Pace et al. (1998). This method uses a weighting matrix to incorporate spatio-temporal dependence in the error specification. The idea is that target house price are not only estimated by global transactions that happened earlier, but also by prior transaction of local neighbours. In the original specification this model accounts for spatial-temporal autocorrelation but does not consider the varying preferences of submarkets or time periods, known as heterogeneity. To account for this, the STAR model is adapted in a number of studies. such as Clapp (2004), who complements the STAR model with a semi-parametric approach to obtain a local regression model that uses neighbourhood price indices, and Liu (2012) who includes submarket dummies and estimates hedonic regression coefficients for each specified time slice. Moreover, several studies have considered how weight matrices for STAR models and variations should be specified. Sun et al. (2005) considers a second order spatial weight matrix to account for within building property differences. Smith and P. Wu (2009) propose natural neighbours to be identified via a Hadamart metrical product between the spatial and temporal weighting matrices, instead of considering separate matrices. Dubé and Legros (2013) and Dubé and Legros (2014) consider the use of exogenous threshold values for the weighting matrix specification. Xiao (2020) proposes a data driven approach for threshold parameters for the estimation of a spatio-temporal weighting matrix. Moreover, several applications of the STAR model and similar variations have been used with purpose of case studies.

While the use of neighbourhood indices or space/time dummies, within STAR model applications, offer a relatively easy-to-use solution to account for spatio-temporal heterogeneity, these methods do not allow coefficients to vary over space and time and thus might bias local coefficients. Hence, several studies consider a different approach and estimate local parameters by means of geographically and temporally weighted regression models. Helbich and Griffith (2016) consider spatial expansion, moving windows, geographically weighted regression and eigenvector spatial filtering, however do not consider temporal heterogeneity. Crespo (2009) and Huang et al. (2010) consider the use of spatio-temporal geographically weighted regression (GTWR), through inclusion of a spatio-temporal kernel function in local model calibration. B. Wu et al. (2014) extend the GTWR approach by accounting for autocorrelation effects. Yao and Fotheringham (2015), examines the use of a mixed model approach that accounts for spatio-temporal relationships at both local and global scales by means of semi-parametric GWR model.

Also within the geostatistical framework several approaches have been considered for property price predictions. Dubin (1998a) applies simple regression kriging in combination with variogram modelling to account for spatial autocorrelation. This model is adapted to a local model with nearest neighbours and spatio-temporal subsets, in a follow study presented in Case et al. (2004). Chica-Olmo (2007), Montero-Lorenzo and Larraz-Iribas (2012), and Chica-Olmo et al. (2013) apply cokriging with auxiliary variables for different datasets with non-overlapping sampling points. Yoo and Kyriakidis (2009) apply area-to-point kriging with external drift to account for spatial dependence and heteroscedasticity. Montero and Larraz (2011) consider various interpolation methods including inverse distance weighting, 2-D shape functions for triangles, ordinary kriging and ordinary cokriging. Kuntz and Helbich (2014) consider detreded- and universal kriging, and their multivariate extensions of detreded- and universal cokriging. Hayunga and Kolovos (2015) introduces the use of Bayesian Maximum Entropy to account for both spatio-temporal autocorrelation and heterogeneity. Kobylińska and Cellmer (2016) consider indicator kriging. N. Crosby et al. (2018) applies ordinary kriging, but utilizes road distance and travel times as separation measure instead of the usually applied euclidean distance. Chica-Olmo et al. (2019) considers regression-cokriging with a multi-equation model in which cross-correlation is used to account for temporal dependence. Overall kriging methods are favoured over other interpolation methods.

Table 1: Literature overview on methodologies that consider spatial or temporal dependence and heterogeneity within real estate markets

			Dependence		Heterogeneity		
	Model	Focus	Space	Time	Space	Time	Article
Spatial Econometrics	STAR	Intro STAR to RE	✓	✓	✗	✗	(Pace et al. 1998)
	LRM	Semi-Parametric STAR	✓	✓	✓	✓	(Clapp 2004)
	2STAR	Multi Unit Housing	✓	✓	✓	✗	(Sun et al. 2005)
	STAR	Joint ST Weight Matrix	✓	✓	✗	✗	(Smith and P. Wu 2009)
	STAR	Office Prices	✓	✓	✓	✓	(Nappi-Choulet and Maury 2009)
	STAR	Threshold Values WM	✓	✓	✗	✗	(Dubé and Legros 2013)
	STAR	Threshold Values WM	✓	✓	✗	✗	(Dubé and Legros 2014)
	STAR	STAR + Heterogeneity	✓	✓	✓	✓	(Liu 2012)
	NLSTARAR	Data Driven WM	✓	✓	✗	✓	(Xiao 2020)
	GTWR	ST kernel function	✗	✗	✓	✓	(Crespo 2009)
	GTWAR	Local Autoregression	✓	✓	✓	✓	(B. Wu et al. 2014)
	GWR	Semi-Parametric GWR	✓	✓	✓	✓	(Yao and Fotheringham 2015)

	SDPD SARMA SARAR	Mod. Comparison	✓	✓	✓	✓	(Otto and Schmid 2018)
	GWR	Moving Clusters	✓	✗	✓	✓	(Kopczewska and Ćwiakowski 2021)
	GTWR	PCA Clusters	✓	✓	✓	✓	(Soltani et al. 2021)
	HTM	Kalman Filter	✓	✓	✓	✗	(Francke and Vos 2004)
Geostatistics	SRK	Coeff Estimation	✓	✗	✗	✗	(Dubin 1998a)
	3D-SF	ST Interpolation	✓	✓	✗	✗	(Li and Revesz 2004)
	LSRK	RK With Local Subsets	✓	✗	✓	✓	(Case et al. 2004)
	RCK	Irregular Variables	✓	✗	✗	✗	(Chica-Olmo 2007)
	RCK	CRE valuation	✓	✗	✗	✓	(Montero and Larraz 2011)
	RCK	Coregionalization	✓	✗	✗	✗	(Chica-Olmo et al. 2013)
	A2PKED	Apartment Valuations	✓	✗	✓	✗	(Yoo and Kyriakidis 2009)
	RCK	Interpolation Comparison	✓	✗	✗	✗	(Montero and Larraz 2011)
	RCK/UCK	Kriging Comparison	✓	✗	✗	✗	(Kuntz and Helbich 2014)
	BME	RE Price Index	✓	✓	✓	✓	(Hayunga and Kolovos 2015)
	IK	Price Probability	✓	✗	✓	✗	(Kobylińska and Cellmer 2016)
	RCK/UCK	Temporal Cross Variogram	✓	✓	✗	✓	(Chica-Olmo et al. 2019)
	URK	Model Comparison	✓	✗	✗	✗	(Derdouri and Murayama 2020)
	STOK	Number of Transactions	✓	✓	✗	✗	(Palma et al. 2019)

Despite the large amount of literature on real estate valuation methodology, an overview of the most cited publications in the last two decades (table 1) indicates that little research accounts for spatio-temporal dependence and spatio-temporal heterogeneity simultaneously. Especially within the geostatistical modelling paradigms applications that consider both space and time are scarce. Hence, we examine these more thoroughly.

Li and Revesz (2004), consider spatio-temporal interpolation for property prices per square feet in two different ways. One adds a time dimension, using a multivariate approach in which a spatial variable is considered at various times and these are jointly analysed trough cokriging. The other treats time as an extra dimension so that variograms are calculated at distances in space and time, and kriging also predicts at times within observations. The firmer is classified as the deduction method, and the latter as the extension method. The extension method is favoured as it is less prone to errors and offers more flexibility. While providing an advanced interpolation method, the proposed model does not account for different property characteristics.

Case et al. (2004) compares four valuation models of which one could be considered geostatistical. The considered method applies local regression kriging based upon the 200-300 closest observations sold within the last three years, for each out-of-sample estimation point. To solve regression coefficients, maximum likelihood is applied based upon a spatial variogram. The model accounts for spatial autocorrelation by allowing the errors to be correlated and interpolating the estimated residuals, and for spatial- and temporal heterogeneity by using local subsets. Moreover, a sale date and geographic trend are included to account for temporal effects (within the three-year sales window) and spatial effects respectively. While this approach incorporates a temporal dimension, it does not include temporal correlation in the variogram specification and thus fails to account for temporal dependence in an optimal way.

H. Crosby et al. (2016) consider a four-stage model that converts several spatio-temporal sales points into a single space-time cube with temporal singularity; applies universal-kriging to identify spatial dependencies; includes parameters common to the hedonic framework; and implements these into Gaussian Process Regression.

Hayunga and Kolovos (2015) examine the use of Bayesian Maximum Entropy. They consider the residential housing market in Texas and aim to build price indices at inter- and intra-neighbourhood levels. Similar to Li and Revesz (2004) and H. Crosby et al. (2016) they consider separation in space and time in a joint way. The introduction of BME allows for rigorous integration of secondary soft data. The considered approach is favoured over alternative models and especially the use of a joint spatio-temporal geometry is routed for. It is argued that considering these components separately might lead to insufficient representation and specious interpretation of spatio-temporal phenomena.

Chica-Olmo et al. (2019) incorporates space-time geostatistics with the hedonic framework trough universal co-kriging and regression co-kriging. To account for spatial dependence and temporal heterogeneity they apply a deduction method so that direct spatial variograms are considered for the residuals of each considered time slice. To account for temporal dependence, cross-variograms are estimated in between the previous specified direct variograms. Both models are favoured over the benchmark OLS model, with slightly better performance of regression co-kriging.

Palma et al. (2019) analyse the number of normalized real estate transactions in Italy with a spatio-temporal random field model. The drift is modelled through an exponential smoothing model with a autoregressive parameter. Residuals are estimated through spatio-temporal ordinary kriging. To do so they use an extension method, similar to Li and Revesz (2004), H. Crosby et al. (2016), and Hayunga and Kolovos (2015). It is found that accounting for spatio-temporal effects in a joint way improves prediction results and provides better results than traditional methods.

The considered papers indicate the potential of joint space-time kriging. However, no application is found that focus on individual property price predictions. Hence, we consider some fundamental papers about space-time geostatistics briefly.

Miller (1997) has been the first to consider spatio-temporal interpolation in GIS, by utilizing an ordinary kriging model and extending its variogram function. Whereas traditional variograms only consist of two dimensional spatial data, the extended variogram includes a third (vertical) and fourth (temporal) dimension. Similar to traditional variograms the four dimensional model is approximated by the average squared difference between paired values. By treating time as a measurable extension to space, the spatial-temporal variogram indicates over what distances, and to which time extent, values at a certain point influence values at adjacent points.

The use of multi dimension variogram models, has since then be explored in a number of studies. A comparison of different variography methods to include both space and time in the computation of covariance matrices is made by Gräler et al. (2013). Performance of the different methods is compared through out-of-sample predictions of PM-10 concentrations in Europe, by means of multiple linear regression, direct kriging and regression kriging. The considered variography measures involve: separate daily variograms, daily evolving variograms in which previous specifications are used to weight the variogram at hand, constant pooled and mean variograms over the year, moving window variograms, a metric 3D variogram, a separable covariance functions, and a product-sum covariance function. It is argued that the last three distinguish themselves from the others , by incorporating both spatial and temporal semi-variance in the covariance structure. As such, they could be considered as single dual structure variograms rather than multiple specifications over time of a spatial variogram. Advantage of this is that it allows temporal information to interact. It is concluded that it is beneficial to include temporal correlation to extend spatial-only correlation. However, it is warned that improvements are deemed modest and that inclusion of temporal interactions comes at the cost of more complex modelling effort. At the same time, it is argued that the symmetric spatio-temporal covariance models discussed in this paper are the most simple methods available. More comprehensive models could address asymmetry in the spatio-temporal covariance and further improve results.

In a follow-up paper, Gräler et al. (2016) again consider PM-10 concentrations in Europe. However, this time they extend the covariance models of consideration with a sum-metric and simple sum-metric covariance function. Although the more advanced sum-metric and simple sum-metric methods have a lower MSE when considering the optimized value from the variogram estimation, they do not improve prediction results over the metric, separable and product-sum method in terms of RMSE, MAE, ME and COR. When compared to the pure spatial form of kriging, improvements are only marginal. It is argued that the limitations in performance increases can be explained by the weak temporal correlation. A lag of only a few days already leads to drastic variability compared to the spatial extend. As a solution they suggest investigating in temporal lags with a higher frequency. Additionally it is argued that the added value of spatio-temporal covariance structures, lies in the flexibility of the model. It allows for interpolation not only across space, but as well in time. Doing so enables the filling of gaps in time series, not solely based on time series, but also including spatial neighbours. This should be of specific benefit for irregular sampled data (like real estate sales), as sampling data is no longer limited to specific time slices.

At last we consider two returning topics that were found in the literature. First, while several space-time methods have been considered, a consensus seems to exists that advocates in favour of including some form of recent nearby-transaction variables. The importance of prior transactions by local neighbours is specifically stressed by Case et al. (2004). In a comparison of four models, they find that the two models that include some form of nearest-neighbour variables, outperform the other two significantly.

Second, almost all considered kriging methods that considers property prices incorporates the hedonic framework by inclusion of a trend in the interpolation process. Three main methods to do so are discussed in Hengl et al. (2007). Considered approaches are kriging with external drift, universal kriging and regression kriging. It is stressed that while modelling steps differ, all methods are mathematically equal. Hence, identical results should be obtained if parameters are held constant. For simplicity, we only use the term regression kriging.

Overall we can derive a number of conclusions from the existing literature. While several methods have been considered, they have in common that they account for dependence by considering neighbouring properties. Key differences are the specifications of how these neighbouring properties relate. Moreover, accounting for heterogeneity is commonly performed by means of data subsets. Again, a variety of methods can be used for the creation of such subsets. However, most papers consider subsets based on geography and a time component. Despite model innovations, inclusion of the hedonic framework remains common practice and is necessary to account for the heterogeneous property characteristics. Within geostatistics this is frequently done by a form of regression kriging. At last, the geostatistical literature clearly favours the use of space-time kriging methods over purely spatial kriging methods. Especially the use of joint space-time kriging is routed for.

Conceptual model

Based upon our introduction and literature review, we specify our conceptual model in figure 1. In here we combine the main concepts of the hedonic framework (Rosen 1974), regression kriging (Hengl et al. 2007) and spatio-temporal kriging (Heuvelink and Griffith 2010).

We start with a housing transaction dataset at the bottom left. As common in the real estate literature, transaction data should contain property prices, locations of the properties considered, times of sales and the structural characteristics of these properties. Ideally, these characteristics should contain all relevant determiners of property prices. However, in practise this is almost impossible. Therefore, key indicators will do. Similar to traditional hedonic methods, this transaction set is used for the specification of a ordinary least squares model. This model should predict property prices as adequately as possible while avoiding, to become unnecessarily complex. The usual OLS assumptions should be checked. These include: the regression should be linear in parameters, observations should be random sampled, the conditional mean of the error should be zero, no multi-collinearity, no heteroscedasticity, no autocorrelation and errors should be normally distributed. According to the literature it is expected that the homoscedasticity and non-autocorrelation assumptions are violated. To account for the heteroscedasticity, our dataset is divided into submarkets. Since we will account for autocorrelation in our next part, this assumption is relaxed for now. OLS specifications should be considered until all assumptions are met with exception of non-autocorrelation.

To relax the autocorrelation assumptions, we follow Heuvelink and Griffith (2010) and specify two additional assumptions: These are: the zero-mean residual is stochastic and multivariate normally distributed and the covariance between points only depends upon their distance in space and time. The next step consist of analysing the residuals obtained from the predictions made with the OLS model. Residuals can be obtained by predicting the sample data, just used to train our OLS model, and subtracting the actual prices. Since the autocorrelation assumption is violated, it is expected that the residuals are not independent from one-another. The relation of these residuals can be used to our advantage. However, to do so this relation first needs to be specified. Previous literature tells us that the autocorrelation is caused by spatial and temporal dependence. Hence, a relation is sought in the spatio-temporal domain. To do so, we consider a spatio-temporal variogram, as is common in the spatio-temporal geostatistical literature.

This variogram is a function that specifies the expected variance between residuals at different ranges in both space and time. To specify such a space-time variogram, the semi-variance between any combination of two residuals is considered. Every obtained residual is related to a transaction price, of which the location and time are known. Hence, we can obtain empirical data of semi-variance at all different ranges between points is space; all different ranges at points in time; and all different ranges of points separated both spatially and temporally. To smooth the semi-variance values and get an indication of more general spatio-temporal trends, obtained values are grouped into spatial bins and temporal lags. This provides us with an empirical spatio-temporal variogram. While this empirical variogram provides us with information about semi-variance at the exact specified ranges, it does not provide an indication about the values in between. Therefore, the empirical space-time variogram is used to fit a continuous theoretical space-time variogram that can be used for our semi-variance estimations at any point in space and time. The literature specifies five theoretical space-time variograms that are generally considered and each come with their own assumptions (see Heuvelink and Griffith 2010). These are: metric, separable, product-sum, simple sum-metric and sum-metric. Dependent upon the complexity our empirical space-time variogram outcomes, a theoretical variogram method can be chosen that best fits the data.

Once a method is chosen, it is fitted to the empirical data. A common method to do so is fitting by the eye. However, also various fitting procedures are available for this process. Once a theoretical space-time variogram is obtained, it can be used for the estimation of a covariance matrix of our data. This matrix is used to estimate a generalised least squares model with equal parameters as the considered OLS application. The GLS mode is used obtain in-sample residuals, which can be used for the estimation of a new empirical variogram. After this the fitting procedure is repeated and new coefficients are specified. Theoretically, this process should be repeated until the obtained parameters remain stationary. However, according to Hengl et al. (2007) a single iteration should be satisfactory.

Once satisfactory coefficients are obtained, these are used to make the out-of-sample predictions of the properties of interest. The last specified space-time variogram is used for simple kriging on the residual values of the sample data so that a trend surface is obtained. The two values can than be summed to obtain the final predictions.

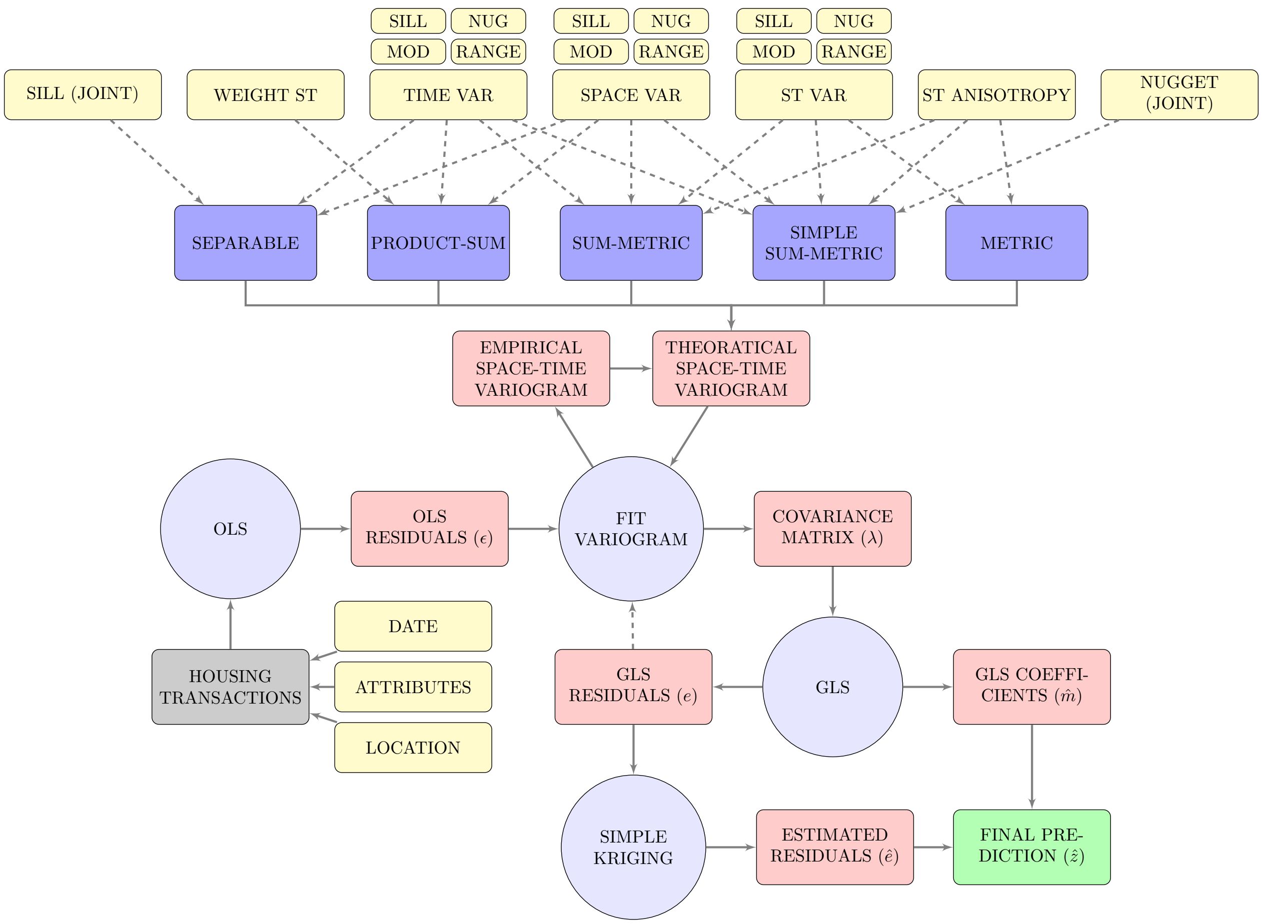


Figure 1: Conceptual model

Methodology

Model specification

We proceed to present the regression-kriging model, which was introduced by Odeh et al. (1995) and later brought to the spatio-temporal extent by Snepvangers et al. (2003). For our setup we assume a Gaussian spatio-temporal random field (Z) that is defined over a spatial domain $\mathcal{S}$ and a temporal domain $\mathcal{T}$. In-sample data contains log property transaction prices $\mathbf{z} = (z(s_1, t_1), ..., z(s_n, t_n))$ that are observed at a set of distinct spatio-temporal locations $(s_1, t_1), ..., (s_n, t_n) \in \mathcal{S} \times \mathcal{T} \subseteq \mathbb{R}^2 \times \mathbb{R}$ and may include simultaneous measures at multiple locations or repeated measures at the same location. We are interested in modelling Z from sample $\mathbf{z}$, in order to predict unobserved prices at any unobserved location in space and time (Gräler et al. 2016). We assume random field Z to be stationary and spatially isotropic across our domain of interest $\mathcal{S} \times \mathcal{T}$. Therefore, we can characterise the field (Z) as nothing but a deterministic trend $m(s, t)$ and a zero-mean stochastic residual $e(s, t)$, so that our basic model satisfies:

$$Z(s,t) = m(s,t) + e(s,t) \tag{1}$$

In a pure geostatistical approach the trend $m(s,t)$ is assumed to be constant in space and time (Hengl et al. 2007). However, this is unlikely for property prices, which are affected by various structural characteristics. Therefore, we consider the trend as a function of known explanatory variables (also known covariates). In line with the hedonic framework, we consider a linear regression approach to model the relationship between the dependent variable and the covariates, so that we get:

$$\hat{m}(s_0,t_0) = \alpha + \sum_{k=0}^{p} \beta_k Q_k(s_0,t_0) \tag{2}$$

in which p are predictors; k are auxiliary variables; α is a constant; $\hat{\beta}$ are to be estimated coefficients; and Q is the value of the auxiliary variables.

The zero-mean stochastic residual e from eq. (1) is assumed to be multivariate normally distributed. It can therefore be fully characterized by a covariance function C, that quantifies the covariance between $e(s,t)$ and $e(s',t')$ at any pair of points (s,t) and s',t' in $\mathcal{S} \times \mathcal{T}$. Given the stationarity assumption, the covariance solely depends upon separation distances across space $h \in \mathbb{R}_{\geq 0} = ||s-s'||$ and time $u \in \mathbb{R}_{\geq 0} = |t-t'|$ between points. Under the isotropy assumption h is considered to be a scalar two-dimensional Euclidean distance vector in space. From what follows we can estimate $e(s,t)$ at any place and time within our domain of interest applying ordinary kriging:

$$\hat{e}(s_0,t_0) = \sum_{i=1}^{n} \lambda_i e(s_i,t_j) \tag{3}$$

in which λ_i are kriging weights that are determined by the spatio-temporal dependence structure of the to be obtained residual(e).

Together, eqs. (1) to (3) provide the spatio-temporal regression kriging model:

$$\hat{z}(s_0,t_0) = \hat{m}(s_0,t_0) + \hat{e}(s_0,t_0) = \hat{\alpha} + \sum_{k=0}^{p} \hat{\beta}_k Q_k(s_0,t_0) + \sum_{i=1}^{n} \lambda_i e(s_i,t_j) \tag{4}$$

Several steps are required to solve our equation and obtain $\hat{m}$ and $\hat{e}$. For this, we require a specification of $\hat{\beta}_0$, $\hat{\beta}_k$, λ_i and e. We start with the estimation of the regression coefficients ($\hat{\beta}_k$). According to Hengl et al. (2007), these can be obtained by any fitting method. However, it is pointed out by Cressie (2015) that GLS should be preferred in case of correlation of the disturbances. Reason for this is that GLS takes correlation between individual observations into account. Hence, to obtain $\hat{\beta}$ we consider:

$$\hat{\beta}_{GLS} = \left(Q^T C^{-1} Q\right)^{-1} Q^T C^{-1} z, \tag{5}$$

where $\hat{\beta}_{GLS}$ is the vector of estimated regression coefficients; Q is a matrix of predictions at sampling locations; C is the covariance matrix of residuals e; and z is a vector of measured values of the target variable (Hengl et al. 2007).

A difficulty intrinsic to regression kriging is that both the regression coefficients (β) and the covariance structure (C) need to be estimated simultaneously. Reason for this is that in order to estimate GLS coefficients, we require the covariance function of residuals. However, in order to estimate the covariance function of residuals we need to know the GLS residuals, for which we require the GLS coefficients (Hengl et al. 2007). A commonly applied solution is to estimate the residuals by an iterative process: The drift model is estimated by means of OLS; then the covariance function of residuals is used to obtain GLS coefficients; the GLS coefficients are used to re-compute residuals; an updated covariance function is estimated, and-so-forth (Hengl et al. 2007). Ideally, multiple iterations should be executed until the predicted residuals stabilize. However, it is argued by Kitanidis (1993) that a single iteration often suffices as outcome differences are only minor.

We follow the single iteration approach and start with the following OLS specification:

$$y = \alpha + \sum_{k=1}^{p} \beta_{k_{OLS}} X_k + \epsilon \tag{6}$$

here y is the log price; α is a constant; k is the k^{th} relevant property characteristics ($k = 1, ..., p$); β is a coefficient to be estimated; X is a vector of the relevant property characteristic; and ϵ is the OLS residual. Using the transaction prices as dependent variable (y), and property characteristics as independent variables (X), we can fit eq. (6) and obtain the OLS residuals (ϵ).

We proceed with the estimation of a covariance matrix (C). This matrix is required for the estimation of the GLS coefficients ($\hat{\beta}_{GLS}$) and the specification of the kriging weights (λ). As discussed in the introduction, we apply a dual structure space-time variogram for our covariance specification. The spatio-temporal variogram indicates semi-variance between any pair of residuals separated by a spatial bin and temporal lag. It is defined as:

$$\gamma(h, u) = \frac{1}{2}\mathbb{E}\left(\epsilon(s_i, t_j) - \epsilon(s_i + h, t_j + u)\right)^2 \tag{7}$$

in which $\mathbb{E}$ denotes the mathematical expectation; h is a spatial bin; and u is a temporal lag. An empirical spatio-temporal sample variogram is created by averaging the semi-variance in regularly spatial bins and temporal lags that reflect the extent of the variance adequately (Cressie 2015). We can solve eq. (7) by specifying bins (h) and lags (u), and utilizing variance between residuals (ϵ) at all known locations. While this provides us with the semi-variance (γ) at the specified bins and lags, it does not offer a functional form that allows for the estimation of semi-variance in between these bings and lags. Therefore, we fit a theoretical spatio-temporal variogram model to the empirical variogram obtained in eq. (7). This empirical variograms takes a continuous functional form and can be used to predict the semi-variance of residuals at any point (Sherman 2011). Deriving flexible and permissible theoretical space-time variogram structures, has been of topic in a number of studies (e.g. De Cesare et al. 2001; Gneiting 2002; Stein 2005; Jost et al. 2005). This provides us with a rich variogram family, with models that rely on different assumptions and varying complexities (Heuvelink et al. 2017). We consider the sum-metric model, which is specified as follows:

$$\gamma_{\hat{e}}(h,u) = \gamma_s(h) + \gamma_t(u) + \gamma_{st}\left(\sqrt{h^2 + (k*u)^2}\right) \tag{8}$$

Here $\gamma_{\hat{e}}(h,u)$ is the space-time variogram; $\gamma_s(h)$ is the spatial variogram; $\gamma_t(u)$ is the temporal variogram; $\gamma_{st}\left(\sqrt{h^2+(k*u)^2}\right)$ is the joint variogram and k is a spatio-temporal anisotropy scaling parameter. An estimation of k is required, as well as a set of spatial variogram model parameters $\theta_s = \{\tau_s^2, \sigma_s^2, \phi_s\}$, a set of temporal variogram model parameters $\theta_t = \{\tau_t^2, \sigma_t^2, \phi_t\}$ and set of joint variogram model parameters $\theta_{joint} = \{\tau_{st}^2, \sigma_{st}^2, \phi_{st}\}$. The parameters contain a nugget (τ), a partial sill (σ) and a range (ϕ) of a specified variogram model. The nugget contains two components: the microscale variance and the variance induced by inaccuracies in the measurement device. The partial sill and range determine the shape of the variogram model.

The benefit of the sum-metric model is that it allows spatial-, temporal- and spatio-temporal parameters to be specified individually. As such, it can be fitted to specific situations and dependence structures. Moreover, since the components of the spatial-, temporal-, and spatio-temporal variograms are specified individually, they can be fairly easy interpreted in a physical sense (Snepvangers et al. 2003). The downside is that the offered flexibility comes with an increased complexity in terms of parameter specifications and assumptions. The sum-metric model contains thirteen interdependent variables that should be specified properly, to obtain reliable results (Heuvelink et al. 2017; Gräler et al. 2013). Moreover, the sum-metric model assumes that $\gamma_s(h)$, $\gamma_t(u)$ and $\gamma_{st}(h,u)$ are second-order stationary and mutually independent. Furthermore, the use of a joint spatio-temporal lag (γ_{st}) relies on assumption that distances in space and time can be reduced to a single entity, which might not be realistic in all situations (Snepvangers et al. 2003). Nevertheless, it is argued by Snepvangers et al. (2003) that model building is about providing sufficiently realistic descriptions of the real world, while still being applicable. Given these concerns, the level of complexity of the sum-metric model is deemed appropriate for our case.

To fit the sum-metric model to the empirical space-time variogram, we minimize the mean squared difference between the model and empirical variogram surfaces. This is done with the "L-BFGS-B" optimization procedure. This method is commonly used for efficient parameter estimates in machine learning (see Zhu et al. 1997). Regarding the fitting routine, we follow Gräler et al. (2016) and apply a method in which no weights are applied to squared residuals in the leasts squares estimation.

Although the fitting routine allows us to optimally fit a theoretical space-time variogram to our empirical data, it relies on a initial model that serves as a starting point. This requires the input of initial nuggets, sills, ranges and model types per one dimensional variogram and a anisotropy value that describes the relation between space and time in the joint dimension. Of particular interest is the specification of the individual variogram model types, as this is a fixed parameter. These model types explain the relation between semi-variance and the relevant separation measure. Earlier applications have predominantly applied spherical models for the specification of a functional form. However, these studies frequently fail to specify why the spherical function is chosen over others (Kuntz and Helbich 2014). More recently, Chica-Olmo et al. (2019) and Derdouri and Murayama (2020) consider the use of other model types. They argue that exponential models overall provide the best performance. Nevertheless, the functional forms discussed in these papers only focus on the spatial extend and do not consider temporal and spatio-temporal separation. Hence, the literature provides no consensus in which variogram model to use. For practical reasons, we limit ourself in this research to the exponential variogram models.

The considered exponential variogram model is adapted from Chica-Olmo et al. (2019) to fit any considered dimension and takes the following form:

$$\hat{\gamma}_{\hat{e}}(s|t|st) = \begin{cases} \tau + \sigma\left[1 - exp\left(-\frac{\omega}{\phi}\right)\right] & \omega > 0 \\ 0 & \omega = 0 \end{cases} \tag{9}$$

where $\hat{\gamma}_{\hat{e}}$ is the one-dimensional variogram of either space (s), time (t) or spacetime (st); and $\omega = h|u|h, u$, dependent upon the dimension of consideration.

The earlier specified isotropy and stationarity assumptions allow the same variogram to be used for directions and spatio-temporal locations (Kilibarda et al. 2014). Once all parameters of eq. (8) are specified, it is used to obtain a structured covariance matrix (C) of the residuals (ϵ) and to specify kriging weights (λ). Under the stationarity and isotropy, covariance is directly related to our variogram model by $\gamma(h, u) = C(0, 0) - C(h, u)$ where $C(0, 0)$ is the variance of Z (Heuvelink et al. 2017). It is to be noted that C is strictly positive definite, to ensure a unique solution for the to be estimated kriging weights (Chen et al. 2020). The vector of the kriging weights is calculated from the specified variogram model by $\lambda_i = \gamma(s, t)^T\Gamma^{-1}$, in which Γ is a matrix that contains semi-variances between possible combinations of space-time observations (Zoest et al. 2020). Now that C and λ can be estimated, we can solve eq. (5) and predict $\hat{z}(s, t)$ using our final model equation:

$$\hat{z}_{RK}(s_i, t_j) = q_0^T\hat{\beta}_{GLS} + \lambda_0^T(z - q\hat{\beta}_{GLS}) \tag{10}$$

Model validation

To validate the proposed methodology, we apply a prediction exercise. In this exercise we split our data into in-sample data and out-of-sample data. In-sample data is used to fit our models and out-of-sample data is used for validation. The advantage of using distinct subsets for model fitting and model validation is that it allows for more objective comparisons between different models (Refaeilzadeh et al. 2009). This is especially relevant for the interpolated residuals. These are known for the within-sampling locations, and should thus provide unrealistic nearly perfect predictions[2] (Zhang and Wang 2010). Moreover, results should be less prone misleading statistics caused by over-fitting (Hawkins 2004). Out-of-sample predictions are made based upon the following procedure.

Step 1, our whole dataset is divided into seven geographical- and five temporal zones so that a total of thirty-five subsets is obtained. The creation of these subsets serves four purposes. First, spatial and temporal submarkets allows us to account for spatio-temporal heterogeneity. Second, spatial and temporal submarkets allows us to identify different spatio-temporal dependence structures which helps us to meet the isotropy assumption. Third, spatio-temporal procedures are computationally heavy and splitting our data into subsets allows for more efficient computing. Fourth, spatial and temporal submarkets allow us to make out-of-sample predictions for varying market structures and allows establishment of model robustness. Spatial divisions are made based upon the k-means clustering algorithm (see Mouselimis 2022; Kaufman and Rousseeuw 2009) in which X and Y coordinates are considered. Although it is recognized that the use of experience-based submarkets might be preferred over deriving these statistically (Liu 2012), this is considered as improbable for our case as no such data was available. Moreover, the use of administrative borders, such as neighbourhoods, municipalities, zip-codes and cities, has been considered. However, these cause densities within certain subzones to be unfeasibly low, so that no model can be fitted. Temporal divisions are made based upon larger economic trends that affect the housing market. Considered periods are the years: 1980:1999, 2000:2007, 2008:2013, 2014:2018 and 2019:2021. These reflect two periods of stable price growths, the financial crisis, the post crisis period and the COVID-19 period respectively. Since these involve broad phenomena that do not affect all areas at a certain time equally hard, no specific transition dates are specified and the years are taken as a whole. While use of smaller time spans, such as single years, might be favoured to avoid any remaining heterogeneity this would deplete our subsamples to unworkable sizes. For the same reason earlier time spans involve larger time periods.

All spatial zones are deemed constant over time, and vice-versa. Although it is recognized that submarkets might be prone to changes (Kopczewska and Ćwiakowski 2021) (e.g. a macro-economic trend might not affect all places at the same time), it is argued that these are not likely to affect

[2] Predictions might not be entirely perfect as the predictions are not merely based upon interpolated residuals but also by a trend

Table 2: Sampling strategy robustness check: each value corresponds to the temporally ordered percentile distribution of individual dataset

In-Sample	Out-of-Sample
0-90	90-100
0-90	90-95
0-90	95-100
0-95	95-100

our outcomes in a significant way. Moreover, specifications of such varying subsets might hinder the interpretations of our results and make these unnecessarily difficult.

Step 2, for each subsample the transactions are sorted from earliest to latest, after which the first 90% is assigned as in-sample estimation data and the remaining 10 % as out-of-sample prediction data. It is argued by Liu (2012), that this procedure of sampling data should be favoured as it recognizes the intrinsic temporal structure of housing transactions. Namely, earlier transactions are relevant for predicting current prices and not the other way around. Moreover, this sampling strategy allows the validation of the type of out-of-sample predictions that are most likely to be used. That is to say, it validates the predictive power of temporally close forecasts, or in other words, contemporary prices. Since our main prediction exercise guaranties identical in-sample and out-of-sample information for predictions with both the OLS and LSTRK model, predictive power of both methods can be assessed on a level ground. Nevertheless, to see if prediction results are driven by particular out-of-sample specifications, a robustness check is applied in which different in-sample / out-of-sample distributions are specified. Considered temporally ordered percentiles are specified in table 2

Step 3, the in-sample datasets are used to fit a traditional OLS model (eqn. 6) and the earlier specified LSTRK model (eqn. 4). It is to note that spatio-temporal variogram parameters are estimated for each individual subset, so that the estimated covariance matrix is optimally specified for each submarket. Thereafter, the trained models are used to predict the property prices of the out-of sample datasets for the relevant submarket. Predicted prices are compared with actual sales prices. Regarding our validation statistics we follow Liu (2012) and consider mean absolute errors (MAE), median absolute errors (MedAE), mean squared errors (MSE) and root mean squared errors (RMSE). All these statistics penalize predictions errors in both directions of the prediction error distribution and are commonly used for model assessments (e.g. Palma et al. 2019; Chica-Olmo et al. 2019)

The MAE takes the average of all absolute errors and measures the difference between the true value and the expected value. Doing so it penalizes all errors with an equal weight. The MAE can be obtained with eqn. 11. Here $\hat{z}$ is the predicted value; z is the actual value; i is the property of consideration; and n is the number of observations (Chai and Draxler 2014).

$$\text{MAE} = \frac{\sum_{i=1}^{n} |\hat{z}_i - z_i|}{n} \tag{11}$$

The MedAE takes the median of all absolute differences between targets and predictions. Advantage is that it provides an indication of prediction accuracy that is robust to outliers. The trade-off of this robustness is that MedAE only considers the distribution of errors and not the actual values. As such it might be deemed insensitive (Armstrong and Collopy 1992). MedAE is obtained by:

$$\text{MedAE} = \text{median}(|z_1 - \hat{z}_1|, ..., |z_n - \hat{z}_n|) \tag{12}$$

The MSE is another metric to asses model predictions. Similar to the MAE it takes the average of errors. However, instead of using absolute errors, it makes errors non-negative by taking the squared value. Squaring the errors emphasizes large errors and cause relatively small errors to become even smaller. Therefore, this metric penalizes extreme errors (Chai and Draxler 2014). It is obtained by:

$$\text{MSE} = \frac{\sum_{i=1}^{n} (\hat{z}_i - z_i)^2}{n} \tag{13}$$

The RMSE is almost identical to MSE. However, it takes the root of the MSE outcome. Therefore, it is measured in the same unit as the response variable. Overall this is deemed more straight forward in terms of interpretation (Willmott 1982). RMSE is obtained by:

$$\text{RMSE} = \sqrt{\frac{\sum_{i=1}^{n} (\hat{z}_i - z_i)^2}{n}} \tag{14}$$

Step 4, to see how LSTRK performs compared to more extensive OLS models, out-of-sample performance of OLS with a trend surface analysis; OLS with interacting space-time dummies; and OLS with clustered errors, is analysed. To avoid the risk of over-fitting these models, they are estimated with the entirety of spatial zones. Hence, they could be considered as global models. Nevertheless, to assure that equal in-samples and out-samples are used for model fitting and for evaluating the price predictions, the temporal structure is kept. Hence, the procedure specified in step two is followed. However, the spatial clusters are merged both in-sample data and out-of-sample data.

Data

To validate our proposed model a dataset containing housing transactions in Mecklenburg County will be used. Mecklenburg County is located in the United States of America and belongs to the state of North Carolina (figure 2).

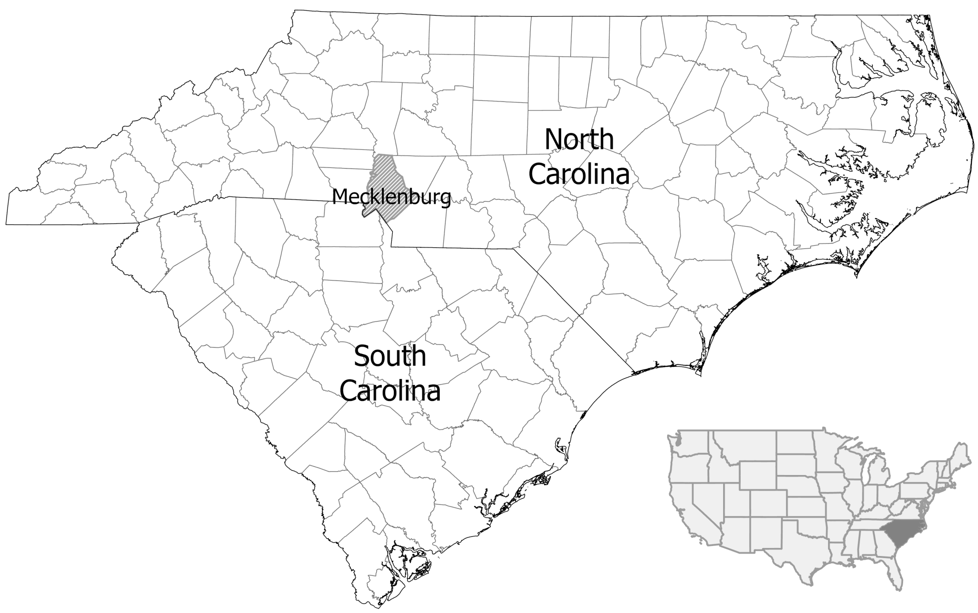


Figure 2: Location of Mecklenburg County; Created with ArcGIS

With an expected annual average growth rate of 1.99 % between 2010 and 2030, Mecklenburg County is considered among the ten fastest growing counties in the Unites States (Khabazi 2018). Mecklenburg County inhabited a population of 1.11 million in 2019. This makes it the second-most populated county in North Carolina (United States Census Bureau n.d.). Mecklenburg County covers a total surface of 1415 squared kilometres, of which 1363 squared kilometres are land. Furthermore, it contains seven municipalities, of which Charlotte is largest. The equally named city of Charlotte inhabits close to 700,000 residents and is the seventeenth largest city in the US. The city is home to a large banking industry and several fortunate 500 companies. This makes Charlotte the second largest financial center in the US after New York. Other important industries in Charlotte are insurance and real estate, causing the median wages in Charlotte to be the highest in the region (Khabazi 2018). Additional notable areas in Mecklenburg County are the towns of Cornelius, Davidson, Huntersville, Matthews, Mint Hill and Pineville. An overview of these towns, the city of Charlotte, and the main roads that connect them is provided in fig. 3. It can be noticed that Charlotte is surrounded by a ring-road that interconnects it with Pineville, Matthews and Mint Hill. Moreover, there is a highway that leads straight to Huntersville, Davidson and Cornelius. On the West, Mecklenburg County is bordered by Catawba River which flows through Lake Norman, Mountain Island Lake, Lake Wylie, Alison and Alison Creek from North to South respectfully. The Charlotte Douglas International Airport is located about 10 kilometres west from centre of Charlotte.

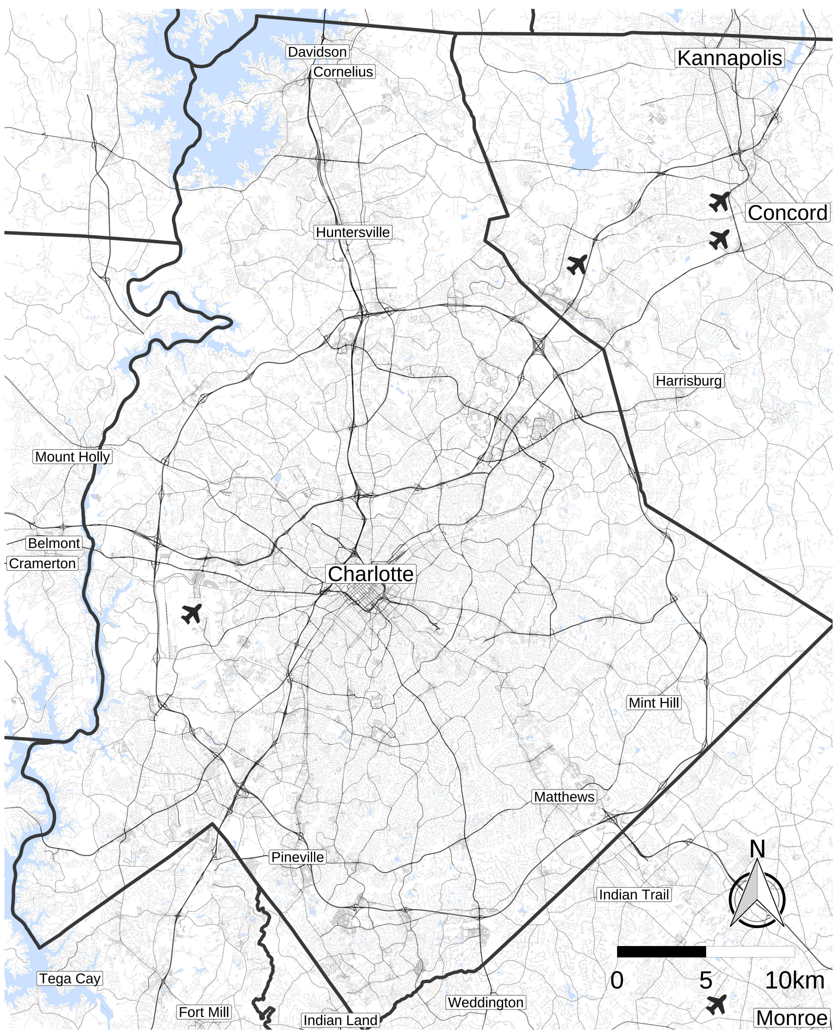


Figure 3: Mecklenburg County's main infrastructure, cities and towns. Created with data from Openstreetmap (Padgham et al. 2017)

Our housing transaction dataset has been developed by the Mecklenburg County Government and is publicly available from the Mecklenburg Open Data Portal (Charmeck n.d.). The dataset contains property related information such as ownership data, parcel boundaries and building characteristics. The data has been created to assist governmental agencies in resource management decisions through use of Geographic Information Systems (GIS). The data has been used for computer assisted mass appraisals with purpose of local tax administration and collection. Earlier editions of the Mecklenburg County transaction data have been used by Kane et al. (2003) to study the effects school accountability ratings have on housing values. Kane et al. (2006) use the same dataset to study the relationship between school characteristics and housing prices. Furthermore, Linden and Rockoff (2008) apply this data to estimate the impact of crime risk on property values. The raw dataset is a shapefile that contains sales information of 415,570 parcels. Multiple steps in R are taken to clean the dataset and filter out relevant transactions (see appendix). Cleaning steps include the removal of non-residential parcels, outliers and ambiguous transactions. Furthermore, the sample has been limited to sales of existing homes between January 1st, 1980 and January 14th 2022. After imposing these restrictions, 57,159 single family property parcels are left. The parcels have been transformed from polygons to points, by using the centroid of the polygon. In the case of large or particular formed parcels the point location might therefore slightly differ from the actual building. However, it is assumed that the centroid of the parcel reflects the building location rather accurately. The spatial data is projected in “NC State Plane NAD83 Feet”. Hence, distance is measured in feet.

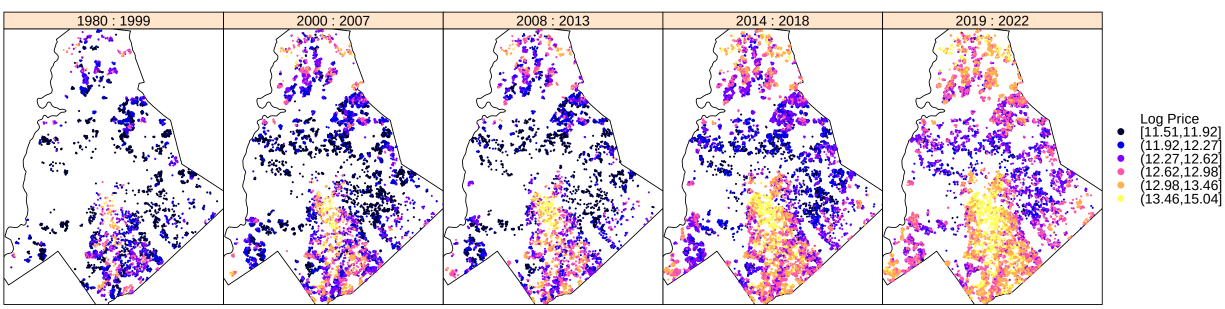


Figure 4: Property sale prices (log USD) per time period (years)

The cleaned dataset is plotted over space and time in figure 4. It can be seen that our sales data covers almost the entire area of interest. However, densities vary considerably per area. Moreover, clear differences in sampling sizes per period exist. Data density is relatively small for earlier periods and gradually increases for later periods. It is to be noted that earlier periods shown in this figure involve larger time spans, but the number of observations remains relatively similar between specified periods. In terms of prices, clear patterns are visible. Overall prices seem to be clustered in space and are increasing over time. Moreover, a trend can be spotted in which cluster areas of relatively high priced houses seem to grow over time.

As shown in table 3, our transaction data contains the main characteristics of individual properties. This includes sale prices, heated area, plots sizes, the number of bedrooms, the number of bathrooms, the amount of fireplaces, building ages, building quality, property conditions, external wall materials, story heights, times of sales and parcel locations. It can be noticed that sales prices vary considerably. This makes sense as our data covers a large area, over a considerable time span. As indicated in fig. 4 these factors explain a large part of the price variations. Nevertheless, properties vary notably in compositions. Also these differences are likely to affect property prices. Our numerical values are right tailed. Hence, larger standard deviations are predominantly caused by high values intrinsic to more luxury properties. Regarding the factor characteristics, it can be seen that certain values are more often present than others. One- and two story properties, made from face brick or aluminium/vinyl and with an average building condition, account for almost half of the properties in our data[3]. Nevertheless, considerable numbers of observations are present for other characteristics as well.

Table 3: Summary statistics of cleaned property dataset

				sales (N = 57,154)	
Sales price (USD)			**Building condition**		
mean (sd)	302,000	195,000	Average	38,271	(67%)
Heated area (ft)			Good	13,648	(24%)
mean (sd)	2,471	859	Very good	4,619	(8%)
Plot size (ac)			Excellent	616	(1%)
mean (sd)	0.12	0.13	**External wall**		
Bedrooms (#)			Face brick	11,244	(20%)
mean (sd)	3.57	0.68	Cedar/Redwood	529	(1%)
Half bathrooms (#)			Masonite	7,585	(13%)
mean (sd)	0.67	0.49	Plywood	1,744	(3%)
Full bathrooms (#)			Aluminium/Vinyl	28,211	(49%)
mean (sd)	2.31	0.63	Stucco	648	(1%)
Fireplaces (#)			Design vinyl	7,193	(13%)
mean (sd)	1.01	0.13	**Story height**		
Construction cohorts			1 Story	12,343	(22%)
1900,1930	350	(1%)	1.5 Story	4,727	(8%)
1930,1944	567	(1%)	2 Story	37,997	(66%)
1944,1960	1,953	(3%)	2.5 Story	1,499	(3%)
1960,1970	1,528	(3%)	Split level	588	(1%)
1970,1980	2,519	(4%)	**Season of sale**		
1980,1990	7,009	(12%)	Autumn	14,200	(25%)
1990,2000	16,586	(29%)	Spring	14,568	(25%)
2000,2010	16,736	(29%)	Summer	17,298	(30%)
2010,2020	9,906	(17%)	Winter	11,088	(19%)

[3]Cannot be derived from table

Cluster Statistics

An overview of the subsets, as specified in the second part of the methodology, is provided in table 4. In combination with figs. 3 to 5 it provides an indication of the kind of submarkets we are dealing with. While not comprehensive, these submarket statistics help us to interpret model outcomes.

Table 4: Summary statistics of the most relevant property characteristics per spatial cluster

Cluster	Price ($) (mean)	PSF ($) (mean)	Prop Size (SF) (mean)	Plot Size (SF) (mean)	Year Built (mean)	External Wall (mode)	Prop Grade (mode)
1	204000	101	2066	4736	1997	Alum/Vinyl	Average
2	267000	107	2507	7503	2006	Alum/Vinyl	Average
3	216000	95	2304	5896	2000	Alum/Vinyl	Average
4	332000	124	2688	6202	2003	Alum/Vinyl	Average
5	187000	96	1984	5877	2002	Alum/Vinyl	Average
6	501000	194	2515	2038	1973	Face Brick	Good
7	360000	127	2813	4542	1995	Face Brick	Good

Cluster 1 is located directly east towards the major city of Charlotte and can be classified as suburban. Properties are of average quality and around 25 years old overall. Compared to the other clusters, the heated area and plot sizes are relatively small. Prices are below average both in terms of price per square foot and in absolute terms.

Cluster 2 is located south-west towards Charlotte. However, despite its proximity, only few houses are located in proximity to the city center. At the North, this cluster is separated by the Charlotte Douglas International Airport. At the south-east side it is separated by industrial areas. Overall, properties are relatively new and of average size. However, plot sizes are considerably larger than those of other areas. Property prices seem to differ within the area. These are relatively low at the north east, and go up considerably at the far west. An explanation for this might be proximity to the lake, more natural areas and separation from the airport.

Cluster 3 is located north-west towards Charlotte and can be considered as another suburban region. Overall density is low and prices per square foot rank lowest of all considered clusters. Plot sizes are above average, while property sizes are below. Buildings are of average quality and age.

Cluster 4 involves the most northern part of our data extend. The main center is the town of Huntersville. While this area is more remote from Charlotte, it remains well connected and is still considered as a suburb. Property prices are above average and go up especially close to the lake in the most northern part. While prises per square feet and housing sizes are slightly below those of cluster 7, plot sizes are considerably larger.

Cluster 5 is located North West towards the center of Charlotte and north of the Charlotte Douglas International Airport. On average, properties within the cluster are the smallest and most affordable within our dataset. While properties are small, plot sizes rank second only slightly after the even less dense area of cluster 3.

Cluster 6 covers Charlotte's city center. Both total prices and prices per square foot are remarkably higher than for the other clusters. Buildings are considerably older than in the other regions. Nevertheless they are overall in good shape and made of more expensive bricks. Despite the high prices, property sizes remain average. Nevertheless, plot sizes are considerably smaller than for the other regions.

Cluster 7 comprehends well-of suburbs located south of Charlotte. Both total property prices and prices per square feet rank second (after cluster 6). However, property sizes and especially plot sizes are considerably larger. On average this cluster has the largest properties. Overall these are made of brick and in good condition. While houses are not as old as the properties in cluster 6, they remain considerably below average.

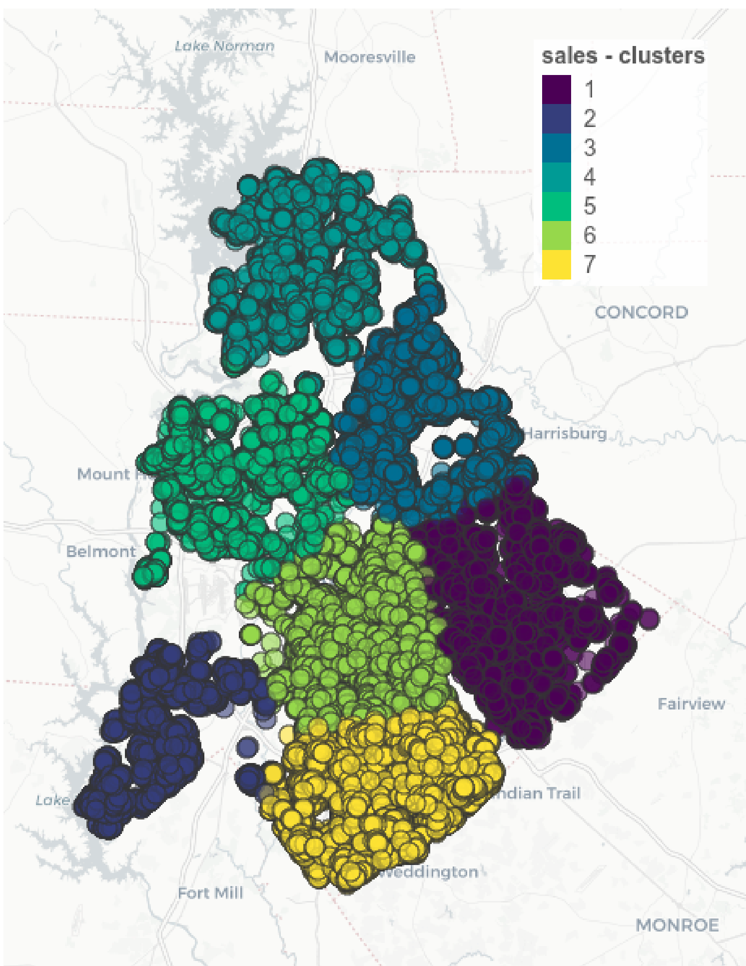


Figure 5: Property clusters based on k-means x and y coordinates

Results

Model estimates

Model estimates of the initial OLS specification (eq. (6)) and the deterministic trend part of our LSTRK specification (eq. (4)) are presented in table 5. For the sake of space, only key percentile values are specified as is common when a large ranges of coefficients are considered (e.g. Huang et al. 2010). While the summation of data omits some details, it offers a compact overview that can easily be used for comparisons. For the interpretation we expect that the heterogeneity in coefficients relates to the same subsamples for both models. In other words, minimum values are caused by the same subsample for both OLS and LSTRK.

An examination of the coefficients provide us with a couple of insights. First of all, it can be noticed that the coefficients across OLS and LSTRK models are rather equal in terms of strengths and directions. This suggests that our estimations are rather robust and not heavily affected by different model specifications. On the other hand, large differences can be found within model specifications. Also this is considered as a positive. It suggests that heterogeneity effects indeed occur and that not all parameters are valued equal within submarkets. This strengthens our believe that the choice to consider individual submarkets was right. Since the strengths and directions of the coefficient follow a similar pattern for both models we examine these simultaneously. Differences are discussed afterwards. Moreover, to keep our assessment concise, attention is paid to the median value unless specified else.

Intercept coefficients vary heavily within both models. These range from large negative values to high positive values. Given the large differences in prices between submarkets, these variations are expected. On the other hand, the negative signs for minimum values and lower quadrants are non-expected. This suggests that properties with no other characteristics are expected to have a negative value. Since this is practically impossible, no further action is taken.

The heated area, total acres, bedrooms, half bathroom, full bathroom, fireplace, year built and date of sale median coefficients all have a positive impact on property prices. This is in line with our expectations as these characteristics are indeed associated to add value to properties. While the 0 values for heated area and date of sale might suggest that these have no effects, it is stressed that this is due to rounding of coefficients and the large number these values typically take.

Coefficients of the building grades are in line with our expectations. These relate to properties of average quality and show increasingly higher values as the quality gets better. The external wall dummy variables relate to aluminium/vinyl. All material types have a positive median sign. This is also in line with our expectations as the aluminium/vinyl it the cheapest and easiest building option. Bricks, cedar/redwood and stucco have the highest strengths. Also this is expected as these could be considered as more luxury building materials.

Seasonal coefficients use the Autumn level as reference data. Coefficients have low values and thus have little impact on the estimated values. In theory we expect the spring and summer coefficient to be postive and the winter coefficient to be negative. This is the case for the median OLS coefficients but not for the LSTRK. However, given the low strength of these coefficients no further attention is paid to these.

Height effects are compared by the number of stores in which the 1.0 store is used as reference. The size of the coefficients indicate that one and one and a half store properties are valued over the others, which makes sense as condos are typically deemed more luxury.

Considering the differences between our OLS and LSTRK estimates, a first thing to note is that differences in intercept are less extreme for the latter. This suggests that values are indeed better explained by the specified characteristics. Moreover it can be noted that ranges of the numeric variables are less extreme for LSTRK, but that the medium stays constant. Also this could be considered as an indication that the LSTRK properly accounts for autocorrelation effects.

Table 5: OLS and GLS parameter estimate summary

	OLS					GLS				
	Min	LQ	Med	UQ	Max	Min	LQ	Med	UQ	Max
(Intercept)	-16.43	-1.64	1.99	6.78	19.03	-15.03	-1.46	1.78	4.82	13.65
Heatedarea	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00
Totalac	-0.21	0.00	0.07	0.17	0.46	-0.18	0.02	0.07	0.13	0.29
Bedrooms	-0.03	-0.00	0.01	0.02	0.05	-0.01	0.00	0.01	0.02	0.05
Halfbaths	-0.04	-0.01	0.01	0.03	0.10	-0.02	0.00	0.01	0.02	0.07
Fullbaths	-0.06	0.00	0.03	0.05	0.15	-0.04	0.01	0.02	0.03	0.06
Numfirepla	-0.18	-0.02	0.02	0.06	0.26	-0.16	-0.00	0.02	0.04	0.22
Yearbuilt	-0.00	0.00	0.00	0.01	0.01	-0.00	0.00	0.00	0.01	0.01
Dateofsale	-0.00	0.00	0.00	0.00	0.00	-0.00	0.00	0.00	0.00	0.00
Building grade										
Good	0.08	0.17	0.21	0.26	0.33	0.07	0.16	0.20	0.23	0.39
Very good	0.14	0.32	0.38	0.49	0.74	0.11	0.33	0.36	0.40	0.79
Excellent	0.30	0.48	0.56	0.74	0.78	0.33	0.44	0.54	0.64	0.95
External wall										
Face brick	0.06	0.11	0.14	0.16	0.25	0.06	0.09	0.10	0.12	0.21
Cedar/Redwood	-0.28	0.08	0.15	0.22	0.58	-0.15	0.05	0.08	0.14	0.72
Masonite	-0.02	0.03	0.05	0.08	0.15	-0.00	0.02	0.03	0.05	0.11
Plywood	-0.05	0.02	0.05	0.12	0.30	-0.13	0.01	0.04	0.08	0.22
Stucco	-0.49	0.11	0.17	0.27	0.52	-0.63	0.06	0.12	0.18	0.36
Design vinyl	-0.03	0.07	0.13	0.16	0.27	-0.03	0.04	0.06	0.08	0.14
Season										
Spring	-0.06	-0.00	0.01	0.02	0.04	-0.06	-0.01	-0.00	0.01	0.03
Summer	-0.02	0.00	0.01	0.03	0.06	-0.02	-0.00	0.00	0.01	0.04
Winter	-0.04	-0.01	-0.00	0.01	0.04	-0.04	-0.01	-0.00	0.00	0.03
Height										
1.5 store	-0.08	0.01	0.03	0.09	0.17	-0.04	0.00	0.02	0.03	0.10
2.0 store	-0.12	-0.03	-0.01	0.01	0.09	-0.05	-0.01	-0.00	0.01	0.09
2.5 store	-0.23	-0.08	-0.04	0.02	0.17	-0.20	-0.05	-0.02	0.00	0.12
Split level	-0.31	-0.14	-0.08	0.00	0.46	-0.17	-0.05	-0.02	0.03	0.37

This table summarizes covariate distributions to indicate the extent of variability. Values include minimum, median and maximum coefficient values and lower and upper quartiles.

Prediction results

Our main prediction results are presented in table 11. As argued we consider out-of-sample predictions of the temporally ordered top 90-100 percentiles, based upon in-sample data of the first temporally order 0-90 percentiles. While predictions are made for each distinct subset, we present the summary values grouped by space and time for the sake of space. To consider the amount of predictions within each group, we cannot merely take the mean statistics. Hence, the following procedure was used. First, subsets are created. Second, in-sample and out-of-sample groups are specified. Third, individual models are trained. Fourth, out-of-sample predictions are made for each subset using the corresponding model. Fifth, predictions are grouped back together. Sixth, statistics are calculated for each group.

A limitation of this approach is that values of subgroups with little values have little effect on the considered statistics. In particular these groups, more advanced models might not suffice. Hence, the summary statistics might be flavoured for the LSTRK approach. To account for this, a manual inspection was made for statistics per subgroup to see if there are in line with our data. No outliers were found.

When we compare the validation statistics for OLS and LSTRK at the corresponding groups, we can make a number of observations. For any statistic, LSTRK values are considerably lower than their OLS counterpart. Since our validation statistics are negatively-oriented scores, this indicates that predictions made by the LSTRK are more accurate than those made by the OLS model. Percentile differences between LSTRK and OLS have a negative sign for all considered groups and each validation statistic. This suggests the model improvements are consistent over different data subsets and validation statistics. Moreover, the magnitude of change is considerable, with total reductions up to 40%. The large reductions of LSTRK validation statistics compared to those of OLS suggests that accounting for spatio-dependence indeed improves price-forecasts accuracy substantially.

Nevertheless, considerable differences in terms of performance improvements can be found per statistic and for individual subgroups. Zooming in to individual statistics, we find that on average LSTRK reduces the MAE by 28.87%. However, this number goes up to 34.11% for zone 2 and drops to 19.86% at temporal cluster 1980:1999. While this might suggests that improvements are not consistent, this does not seem to be the case. MAE values are dependent upon the property values of consideration. Therefore, larger values in more recent time spans and at more expensive spatial clusters is expected. Moreover, MAE value differences between subgroups seem to be smaller for LSTRK. Between groups MAE values of LSTRK range from 0.0774 to 0.1369. This is considerably less than the OLS values that range from 0.1057 to 1.942. In this regard, model improvement differences seem to be predominantly caused by the large variety in the MSE values of the OLS model. The same pattern seems to occur for the other validation statistics. Whereas values between groups vary considerably for the OLS model, they remain relatively stable for LSTRK. This suggests that LSTRK predictions are relatively little affected by variations in property market types. These findings are in line with our theory, that argues in favour of space-time geostatistics, as it is able to fit dynamic dependence structures (Heuvelink et al. 2017).

Overall, improvements made by accounting for spatio-temporal dependence seem to be in line with the literature. E.g.the STAR method used by Liu (2012) reduces the average MAE by 27.6%

and the average RMSE by 24.2%, relative to OLS. For the spatio-temporal regression cokriging method applied by Chica-Olmo et al. (2019), reductions go up to 18% and 30% for MAE and MSE respectively. However, since different datasets, structural characteristics and sampling strategies have been used, these statistics only provide a rough comparison. In this regard, additional research is required to see how these models compare in prediction accuracies.

As an additional robustness check, we consider if prediction results might be driven by different in-sample and out-of-sample specifications. Therefore, we perform two forecasts exercises with different out-of-sample data, and one forecasts exercise with both different in-sample and out-of-sample data. Using the same in-sample data, we consider the temporally ordered 90-95 percentiles and the temporally ordered 95-100 percentiles in table 7 and table 8 respectfully. Results of the temporally ordered 0-95 percentile in-sample data, and the temporally ordered 95-100 percentile out-of-sample data are presented in table 9.

Outcomes of tables 7 to 9 are in line with our previous results, and thus argue in favour for the robustness of LSTRK. Upon closer examination, validation statistics seem to be slightly higher for prediction of the temporally far away data (8) compared to the spatially closer data (7). However, these effect are small and expected, as it is easier to make temporally close forecasts than far away forecasts. On the other hand, a decrease in the magnitude of prediction improvements can be noticed. This suggests that LSTRK works best for temporally close predictions and is in line with our theory (Gräler et al. 2016). Regarding table 9, we expect the lowest validation statistics for both models, as this sampling distribution utilizes relatively large in-sample datasets to make temporally close predictions. This is indeed the case. It is notable that model improvements of LSTRK over OLS are even larger in this scenario. This might be another indication that LSTRK performs best at predicting prices of properties that are spatially and temporally close to the in-sample data. Moreover, it could indicate that LSTRK profits more from the increased sampling sizes than OLS does. Both tables 7 and 9 consider temporally close out-of-sample predictions. However, LSTRK predictions seem to be better for the latter, both absolutely and relatively to OLS. While this effect might be caused by the different in out-of-sample datasets, it is more likely that improvements are caused by larger in-sample data sizes. When examining model improvements per individual zone, it can be seen that especially zones with relatively little data profit from the additional 5% in sample data. This suggests, that initial in-sample datasets might have been to small to accurately fit the LSTRK model. Hence, increasing in-sampling data sizes leads to diminishing returns.

Table 6: Prediction results obtained using the first 90% of temporally ordered transactions as in-sample and 90-100% of temporally order transactions as out-of-sample

	Zone 1 (N = 673)	Zone 2 (N = 652)	Zone 3 (N = 1053)	Zone 4 (N = 938)	Zone 5 (N = 484)	Zone 6 (N = 660)	Zone 7 (N = 1252)	Total (N = 5,712)
OLS								
Mean‖error‖	0.1297	0.1175	0.1255	0.1395	0.1381	0.1942	0.1194	0.1350
Median‖error‖	0.1056	0.0972	0.1014	0.1145	0.1092	0.1565	0.0950	0.1078
MSE	0.0283	0.0231	0.0258	0.0332	0.0331	0.0631	0.0244	0.0316
RMSE	0.1682	0.1520	0.1606	0.1821	0.1818	0.2512	0.1562	0.1778
LSTRK								
Mean‖error‖	0.1022	0.0774	0.0842	0.0957	0.1065	0.1369	0.0880	0.0962
Median‖error‖	0.0802	0.0537	0.0655	0.0746	0.0874	0.1059	0.0667	0.0721
MSE	0.0201	0.0143	0.0129	0.0168	0.0210	0.0342	0.0152	0.0182
RMSE	0.1419	0.1198	0.1134	0.1296	0.1450	0.1848	0.1234	0.1349
Difference (%)								
Mean‖error‖	-21.22	-34.11	-32.95	-31.43	-22.90	-29.47	-26.28	-28.74
Median‖error‖	-24.06	-44.82	-35.40	-34.85	-19.97	-32.31	-29.72	-33.10
MSE	-28.81	-37.93	-50.15	-49.37	-36.45	-45.87	-37.60	-42.42
RMSE	-15.63	-21.22	-29.39	-28.85	-20.28	-26.42	-21.01	-24.12

(a) Predictions summarized by spatial submarket

	1980:1999 (N = 527)	2000:2007 (N = 1208)	2008:2013 (N = 754)	2014:2018 (N = 1742)	2019:2022 (N = 1481)	Total (N = 5,712)
OLS						
Mean‖error‖	0.1057	0.1310	0.1653	0.1182	0.1531	0.1350
Median‖error‖	0.0743	0.1034	0.1402	0.0879	0.1323	0.1078
MSE	0.0211	0.0300	0.0424	0.0265	0.0372	0.0316
RMSE	0.1452	0.1732	0.2058	0.1628	0.1929	0.1778
LSTRK						
Mean‖error‖	0.0847	0.0981	0.1130	0.0841	0.1046	0.0962
Median‖error‖	0.0624	0.0742	0.0929	0.0594	0.0814	0.0721
MSE	0.0141	0.0172	0.0221	0.0152	0.0220	0.0182
RMSE	0.1186	0.1311	0.1488	0.1234	0.1484	0.1349
Difference (%)						
Mean‖error‖	-19.86	-25.15	-31.67	-28.85	-31.71	-28.74
Median‖error‖	-15.96	-28.27	-33.74	-32.36	-38.45	-33.10
MSE	-33.29	-42.77	-47.72	-42.55	-40.85	-42.42
RMSE	-18.32	-24.35	-27.70	-24.21	-23.09	-24.12

(b) Predictions summarized by temporal submarket

Table 7: Prediction results obtained using the first 90% of temporally ordered transactions as in-sample and 90-95% of temporally order transactions as out-of-sample

	Zone 1 (N = 338)	Zone 2 (N = 328)	Zone 3 (N = 527)	Zone 4 (N = 469)	Zone 5 (N = 243)	Zone 6 (N = 330)	Zone 7 (N = 626)	Total (N = 2,861)
OLS								
Mean‖error‖	0.1276	0.1125	0.1253	0.1358	0.1382	0.1999	0.1185	0.1340
Median‖error‖	0.1032	0.0926	0.1027	0.1086	0.1135	0.1604	0.0947	0.1060
MSE	0.0280	0.0212	0.0253	0.0319	0.0326	0.0687	0.0240	0.0316
RMSE	0.1672	0.1457	0.1592	0.1785	0.1805	0.2621	0.1548	0.1777
LSTRK								
Mean‖error‖	0.1029	0.0751	0.0854	0.0902	0.1033	0.1383	0.0832	0.0942
Median‖error‖	0.0863	0.0512	0.0647	0.0651	0.0869	0.1014	0.0615	0.0688
MSE	0.0202	0.0133	0.0136	0.0156	0.0184	0.0377	0.0133	0.0178
RMSE	0.1423	0.1152	0.1166	0.1248	0.1357	0.1942	0.1152	0.1334
Difference (%)								
Mean‖error‖	-19.31	-33.23	-31.82	-33.61	-25.25	-30.81	-29.78	-29.70
Median‖error‖	-16.46	-44.65	-36.99	-39.99	-23.42	-36.78	-35.06	-35.10
MSE	-27.62	-37.50	-46.31	-51.10	-43.48	-45.10	-44.60	-43.63
RMSE	-14.92	-20.94	-26.73	-30.07	-24.82	-25.91	-25.57	-24.92

(a) Predictions summarized by spatial submarket

	1980:1999 (N = 265)	2000:2007 (N = 605)	2008:2013 (N = 379)	2014:2018 (N = 872)	2019:2022 (N = 740)	Total (N = 2,861)
OLS						
Mean‖error‖	0.1125	0.1370	0.1569	0.1172	0.1474	0.1340
Median‖error‖	0.0751	0.1109	0.1330	0.0871	0.1259	0.1060
MSE	0.0251	0.0333	0.0374	0.0262	0.0357	0.0316
RMSE	0.1585	0.1826	0.1933	0.1620	0.1890	0.1777
LSTRK						
Mean‖error‖	0.0898	0.0991	0.1073	0.0814	0.1002	0.0942
Median‖error‖	0.0613	0.0718	0.0878	0.0576	0.0753	0.0688
MSE	0.0155	0.0182	0.0187	0.0144	0.0218	0.0178
RMSE	0.1245	0.1350	0.1369	0.1200	0.1476	0.1334
Difference (%)						
Mean‖error‖	-20.23	-27.68	-31.62	-30.54	-32.00	-29.70
Median‖error‖	-18.31	-35.31	-34.01	-33.89	-40.15	-35.10
MSE	-38.28	-45.36	-49.88	-45.12	-39.03	-43.63
RMSE	-21.44	-26.08	-29.20	-25.92	-21.91	-24.92

(b) Predictions summarized by temporal submarket

Table 8: Prediction results obtained using the first 90% of temporally ordered transactions as in-sample and 95-100% of temporally order transactions as out-of-sample

	Zone 1 (N = 335)	Zone 2 (N = 324)	Zone 3 (N = 526)	Zone 4 (N = 469)	Zone 5 (N = 241)	Zone 6 (N = 330)	Zone 7 (N = 626)	Total (N = 2,851)
OLS								
Mean‖error‖	0.1318	0.1226	0.1258	0.1432	0.1380	0.1884	0.1202	0.1361
Median‖error‖	0.1086	0.1082	0.1009	0.1212	0.1073	0.1522	0.0952	0.1099
MSE	0.0286	0.0250	0.0263	0.0345	0.0335	0.0575	0.0248	0.0317
RMSE	0.1692	0.1582	0.1620	0.1856	0.1831	0.2399	0.1576	0.1780
LSTRK								
Mean‖error‖	0.1014	0.0798	0.0829	0.1012	0.1096	0.1355	0.0928	0.0983
Median‖error‖	0.0756	0.0562	0.0662	0.0825	0.0884	0.1116	0.0722	0.0750
MSE	0.0200	0.0154	0.0121	0.0180	0.0236	0.0306	0.0172	0.0186
RMSE	0.1416	0.1243	0.1101	0.1341	0.1537	0.1750	0.1311	0.1365
Difference (%)								
Mean‖error‖	-23.08	-34.94	-34.08	-29.35	-20.53	-28.04	-22.83	-27.78
Median‖error‖	-30.36	-48.06	-34.4	-31.95	-17.63	-26.64	-24.16	-31.74
MSE	-29.98	-38.3	-53.86	-47.78	-29.56	-46.78	-30.84	-41.21
RMSE	-16.32	-21.45	-32.07	-27.74	-16.07	-27.05	-16.84	-23.32

(a) Predictions summarized by spatial submarket

	1980:1999 (N = 262)	2000:2007 (N = 603)	2008:2013 (N = 375)	2014:2018 (N = 870)	2019:2022 (N = 741)	Total (N = 2,851)
OLS						
Mean‖error‖	0.0989	0.1250	0.1738	0.1192	0.1589	0.1361
Median‖error‖	0.0741	0.0973	0.1434	0.0882	0.1430	0.1099
MSE	0.0170	0.0267	0.0474	0.0268	0.0387	0.0317
RMSE	0.1304	0.1634	0.2177	0.1636	0.1967	0.1780
LSTRK						
Mean‖error‖	0.0797	0.0970	0.1187	0.0868	0.1089	0.0983
Median‖error‖	0.0645	0.0776	0.0989	0.0626	0.0875	0.0750
MSE	0.0126	0.0161	0.0256	0.0161	0.0222	0.0186
RMSE	0.1123	0.1270	0.1600	0.1267	0.1491	0.1365
Difference (%)						
Mean‖error‖	-19.43	-22.36	-31.71	-27.19	-31.45	-27.78
Median‖error‖	-12.91	-20.25	-31.00	-29.00	-38.84	-31.74
MSE	-25.83	-39.53	-46.00	-40.03	-42.53	-41.21
RMSE	-13.88	-22.24	-26.52	-22.56	-24.19	-23.32

(b) Predictions summarized by temporal submarket

Table 9: Prediction results obtained using the first 95% of temporally ordered transactions as in-sample and 95-100% of temporally order transactions as out-of-sample

	Zone 1 (N = 335)	Zone 2 (N = 326)	Zone 3 (N = 526)	Zone 4 (N = 468)	Zone 5 (N = 242)	Zone 6 (N = 331)	Zone 7 (N = 627)	Total (N = 2855)
OLS								
Mean‖error‖	0.1286	0.1152	0.1204	0.1359	0.1304	0.1880	0.1161	0.1310
Median‖error‖	0.1079	0.0973	0.0959	0.1105	0.0973	0.1535	0.0909	0.1040
MSE	0.0272	0.0228	0.0242	0.0316	0.0313	0.0577	0.0232	0.0299
RMSE	0.1648	0.1511	0.1557	0.1777	0.1770	0.2402	0.1523	0.1729
LSTRK								
Mean‖error‖	0.0973	0.0735	0.0773	0.0833	0.0930	0.1399	0.0864	0.0908
Median‖error‖	0.0667	0.0518	0.0607	0.0589	0.0638	0.1103	0.0678	0.0663
MSE	0.0201	0.0141	0.0108	0.0140	0.0207	0.0328	0.0154	0.0172
RMSE	0.1416	0.1186	0.1038	0.1183	0.1438	0.1812	0.1239	0.1310
Difference (%)								
Mean‖error‖	-24.29	-36.20	-35.79	-38.67	-28.72	-25.59	-25.54	-30.71
Median‖error‖	-38.16	-46.80	-36.67	-46.73	-34.42	-28.17	-25.36	-36.26
MSE	-26.12	-38.40	-55.56	-55.67	-34.04	-43.11	-33.85	-42.54
RMSE	-14.05	-21.51	-33.34	-33.42	-18.78	-24.58	-18.67	-24.20

(a) Predictions summarized by spatial submarket

	1980:1999 (N = 262)	2000:2007 (N = 605)	2008:2013 (N = 377)	2014:2018 (N = 870)	2019:2022 (N = 741)	Total (N = 2,855)
OLS						
Mean‖error‖	0.1001	0.1244	0.1620	0.1177	0.1473	0.1310
Median‖error‖	0.0778	0.0978	0.1330	0.0881	0.1255	0.1040
MSE	0.0173	0.0265	0.0426	0.0263	0.0347	0.0299
RMSE	0.1315	0.1629	0.2065	0.1623	0.1864	0.1729
LSTRK						
Mean‖error‖	0.0774	0.0913	0.1048	0.0832	0.0970	0.0908
Median‖error‖	0.0643	0.0700	0.0808	0.0553	0.0699	0.0663
MSE	0.0123	0.0147	0.0223	0.0156	0.0201	0.0172
RMSE	0.1109	0.1212	0.1494	0.1250	0.1418	0.1310
Difference (%)						
Mean‖error‖	-22.68	-26.60	-35.34	-29.36	-34.16	-30.71
Median‖error‖	-17.39	-28.46	-39.21	-37.19	-44.35	-36.26
MSE	-28.84	-44.60	-47.68	-40.67	-42.12	-42.54
RMSE	-15.64	-25.57	-27.66	-22.98	-23.92	-24.20

(b) Predictions summarized by temporal submarket

Global model comparisons

Our final empirical results compare predictive power of LSTRK to global OLS models. In here we consider OLS with trend surface analysis; OLS with interacting space-time dummies; and OLS with inverse distance weighting (IDW) residuals. Multiple model specifications have been tested to obtain optimal predictions with the specified method. Considered variations include the introduction of polynomials and the specification of several administrative boundaries, including municipalities, cities, neighbourhood and zip-code districts and the specified k-means clusters. Eventually the seven specified k-means clusters with no polynomials provided best results for all cases.

The trend surface is created through inclusion of individual x and y coordinates, the squared coordinates and a multiplication of x coordinates times y coordinates. Spatial cluster dummies were added to further improve the trend surface results. The interacting space time dummies consider the k-means spatial clusters and sale dates grouped per whole year. The IDW model uses the same covariates as applied in our LSTRK model, but does include k-means cluster dummies to account for heterogeneity. As the considered LSTRK model remains unaltered and the exact same dataset is used, these statistics remain the same.

Since we no longer focus on spatial subgroups, results are only shown for the entire area. Temporal distinctions remain present and prediction results are shown for each specified temporal subgroup. The MAE and RMSE are considered as criteria statistics. The results can be seen in table 10.

Prediction results of the global OLS models are in line with our previous findings. The trend surface model and interacting space-time dummies model are relatively similar in terms of MAE and RMSE, with slightly better results for the firmer. Compared to the local OLS model, these models also seem to be more accurate, although prediction improvements are not consistent. Despite these marginal prediction improvements, both MAE and RMSE values remain considerably higher than those of LSTRK. On the other hand, the IDW model considerably improves upon earlier obtained OLS results. In terms of MAE and RMSE, prediction results of this model are more similar to those of LSTRK than those of the other OLS models. Considering MAE improvements, it can be seen that LSTRK is able to reduce these values from 21.06% (1980:1999) up to 44.27% (2014:2018) compared to the trend surface model. Compared to the IDW model, MAE improvements only range from 5.51% (2008:2013) up to % 34.79 (2014:2018). Regarding RMSE values, relative imrovements made by LSTRK range from 20.77% (1980:1999) up to 35.76% (2014:2018) and from 7.12% (2008:2013) up to 27.53% (2014:2018) for the trend- and IDW model respectfully.

While the improvement rates of LSTRK over the IDW model are considerably lower than for the other OLS models, they remain considerable. Hence, these results again argue in favour for the use of LSTRK. Noteworthy is that the highest improvement rates are achieved in the 2014:2018 time period. One might recall from table 11, that this is the period with the most observations. The opposite is true for 1980:1999 and 2008:2013. In these periods the smallest improvements are

Table 10: Out-of-sample prediction statistics for global OLS models

	OLS+Trend		OLS+STdum		OLS+IDWres		LSTRK	
	MAE	RMSE	MAE	RMSE	MAE	RMSE	MAE	RMSE
1980:1999	0.1073	0.1497	0.1071	0.1487	0.0896	0.1305	0.0847	0.1186
2000:2007	0.1287	0.1737	0.1394	0.1854	0.1070	0.1482	0.0981	0.1311
2008:2013	0.1443	0.1908	0.1674	0.2093	0.1196	0.1602	0.1130	0.1488
2014:2018	0.1509	0.1921	0.1286	0.1755	0.1331	0.1703	0.0841	0.1234
2019:2022	0.1604	0.2017	0.1575	0.2001	0.1394	0.1768	0.1046	0.1484

This table includes MAE and RMSE error statistics for out-of-sample forecasts made by OLS with trend surface analysis, OLS with interacting space-time dummies, OLS with clustered errors and LSTRK

made. Table 11 indicates that these are the periods with the smallest sampling sizes. Hence, these results again suggests that LSTRK seems to benefit relatively more from increased sampling sizes than OLS does.

Conclusion

In order to answer our main research question we start by discussing our two sub-questions. For our first sub-question we ask how LSTRK performs compares to traditional hedonic models in terms of out-of-sample property price forecast accuracy. Our empirical analyses provide level ground prediction exercises between LSTRK and several OLS specifications. For all validation parameters LSTRK clearly outperform their OLS equivalents in a consisting manner. This suggests that LSTRK indeed has the potential to outperform traditional hedonic OLS models. Moreover, several robustness checks are made to find if LSTRK might somehow be favoured by the sampling strategy. This does not seem to be the case. For any considered submarket and at any percentile distribution of the in-sample data and out-of-sample data, LSTRK provides the best property price forecasts. This suggest that LSTRK does not only has the potential to outperform traditional OLS models, but also seems to do so in a consisting manner. Nevertheless, our empirical analysis identifies two phenomena that limit LSTRK prediction accuracy relatively much compared to OLS. First, LSTRK should only be used for predictions that are close to the space-time domain of the fitting data. While this is the case for most prediction models, results suggest that it hampers LSTRK more than OLS. Second, LSTRK seems to require larger sampling sizes to adequately fit the model than is the case for OLS. While still outperforming OLS models, improvements are considerably lower in cases where little data is used to fit our model. Despite these drawbacks, LSTRK still performs best in all prediction accuracy comparisons. Although the considered OLS methods might not cover all traditional hedonic methods, they do cover some of the best known and applied techniques. Therefore, it is concluded that LSTRK indeed improves property price forecast accuracy over traditonal hedonic models.

Our second sub-question considers the practicality of LSTRK. In here we ask how performable LSTRK is in terms of model complexity, parameter specification and computing power. As indicated in our methodology, LSTRK could be considered as quite complex, especially compared to traditional hedonic OLS methods. To perform LSTRK requires the assumptions of stationarity, second-order stationarity, isotropy, a linear relation, absence of multicollinearity, a spatio-temporal random field, zero-mean stochastic residuals and mutually independent variograms. While space-time kriging models are heavily adaptable and can thus be modified to meet or relax one or more of these assumptions, it remains unlikely that all these simplifications can hold in reality. Moreover, extensions also increase the complexity of the model. Therefore, a trade-off has to be made in terms of providing sufficiently realistic descriptions of real phenomena and being applicable (Snepvangers et al. 2003). In this regard, also the proposed LSTRK model can be considered as a compromise. Some of the made assumptions are almost certainly violated. E.g. it is doubtful if the obtained residuals are fully dependent upon separation distances in space and time, and dependence structures are unlikely to be fully scalar by euclidean distances (H. Crosby et al. 2018). Accounting for this, would make our model considerably more complicated. Hence, in order to limit the effects of these violations, LSTRK is fitted by individual submarkets. Considering our prediction results, this trade-off seems to work.

Despite these compromises, LSTRK can still be considered as complex. Apart from the specification of the initial OLS model, LSTRK requires various inputs for the specification of the covariance structure. Especially, the specification of empirical variogram parameters can be a time consuming process as it requires knowledge of de dependence structure that is to be captured. As such, several trial and errors are necessary in order to get an understanding of the data at hand. The disadvantage of using non-sparse covariance matrices, becomes present here. Computing times go up drastically when larger sampling sizes are used. As such, an additional trade-off needs to be made. Ideally, one uses all the relevant information to obtain a high level detail empirical variogram. However, the trade-off is longer computing times. Therefore, one might favour the use of smaller sampling sizes. Again, the use of local submarkets provide a solution here, as it speeds up the fitting process considerably while all in-sample data is utilized.

Given these considerations, there is no one-way answer to our second sub-question. The proposed methodology requires considerably more modelling effort than the considered OLS models, and might therefore be deemed unusable. However, it also provides notable improvements of the prediction accuracy. Moreover, the proposed method offers huge flexibility and can adapt itself to fit any kind of situation. If there is no temporal dependence our model transforms in a purely spatial regression kriging. If there is no spatial dependence our model transforms in a purely temporal regression model. Once neither is present, all covariances are considered to be equal and a traditional OLS model remains. It is up to the user to decide if this flexibility and the prediction improvements are worth the extra trouble. In this regard, LSTRK provides a trade-off between accuracy and applicability.

To conclude we answer our main research-question in which we ask if LSTRK provides adequate property values and if its theoretically and practically warranted. Overall our findings suggest that this seems to be the case indeed. Compared to traditional OLS models, property price predictions constantly more accurate. Moreover, the accounting for spatio-temporal dependence and spatio-temporal heterogeneity that this model does is recommended in several studies (e.g. Liu 2012; B. Wu et al. 2014; Hayunga and Kolovos 2015). In terms of practicality, the model can be marked as complex. However, in return it offers high flexibility and accurate predictions. Given these constraints, the considered methodology is not deemed fit for quick analyses in high precision is no necessity. Hence, easier to use purely spatial methods could provide more suitable outcomes in these scenarios. However, if optimal predictions should be made and time is no scarcity, use of spatio-temporal regression kriging could provide a valuable option. The proposed method offers huge flexibility and can adapt itself to fit any kind of situation.

Discussion

Our research results argue in favour of further exploration of space-time geostatistics in the field of property price valuations. Although our empirical analysis provides favourable results for this method, additional validation is required to establish its robustness. Hence, prediction accuracy should be assessed in various situations and compared to additional models. In this regard, especially performance comparisons between LSTRK and other space-time models are of interest. For future research one might for example consider a comparison with the STAR model (Liu 2012) and the GTWR model (Soltani et al. 2021). Moreover, it is of interest to see how space-time geostatistical methods can be further extended. Several opportunities are available to do so (Gräler et al. 2016). For future research one might consider the use of more advanced approaches for the determination of the model trend. In this regard, the use of machine learning methods could provide an interesting option (Heuvelink et al. 2017). Moreover, inclusion of alternative covariates might effect the dependence structure. A systematic exploration of dependence structures might provide a key to obtain better covariance structures. Also the use of alternative submarkets could further improve prediction accuracies. While this research used fixed spatial and temporal zones, one could for example consider the use of flexible submarkets.

At last, we address some limitations of thesis, that might be addressed in future research. As argued, some of the considered LSTRK assumptions are unlikely to hold in reality. In this regard it would have been of interest to see how predictions are affected by alternative model extensions. Hence, we suggest an in-depth analysis of LSTRK performance base upon different assumptions. Moreover, it is noted that our model specifications fail to include accessibility measures. These are considered highly important indicator for property prices according to Chica-Olmo et al. (2019). Therefore, we recommend inclusion of accessibility covariates, such as distance to the CBD, for future analyses.

Furthermore, we address that our specifications of the theoretical variograms only considers the sum-metric space-time variogram. Several other options are available that should be considered (Gräler et al. 2016). The same goes for the model selection of the one-dimensional variograms. For these, we limited our options to the exponential variable. While providing overall accurate fits, it remains unknown if alternative model specifications might have provided even better results. Therefore, we suggest the exploration of alternative variogram models.

Literature

Armstrong, J Scott and Fred Collopy (1992). "Error measures for generalizing about forecasting methods: Empirical comparisons". In: *International journal of forecasting* 8.1, pp. 69–80.

Bourassa, Steven C, Eva Cantoni, and Martin Hoesli (2007). "Spatial dependence, housing submarkets, and house price prediction". In: *The Journal of Real Estate Finance and Economics* 35.2, pp. 143–160.

— (2008). "Predicting house prices with spatial dependence: impacts of alternative submarket definitions". In: *Swiss Finance Institute Research Paper* 08-01.

Case, Bradford et al. (2004). "Modeling spatial and temporal house price patterns: A comparison of four models". In: *The Journal of Real Estate Finance and Economics* 29.2, pp. 167–191.

Chai, T. and R. R. Draxler (2014). "Root mean square error (RMSE) or mean absolute error (MAE)? – Arguments against avoiding RMSE in the literature". In: *Geoscientific Model Development* 7.3, pp. 1247–1250. DOI: `10.5194/gmd-7-1247-2014`. URL: `https://gmd.copernicus.org/articles/7/1247/2014/`.

Charmeck (n.d.). *openmapping*. Online. Accessed[2021-01-14]. URL: `http://maps.co.mecklenburg.nc.us/openmapping/index.html`.

Chen, Wanfang, Marc G Genton, and Ying Sun (2020). "Space-time covariance structures and models". In:

Chica-Olmo, Jorge (2007). "Prediction of housing location price by a multivariate spatial method: Cokriging". In: *Journal of Real Estate Research* 29.1, pp. 91–114.

Chica-Olmo, Jorge, Rafael Cano-Guervos, and Mario Chica-Olmo (2013). "A Coregionalized Model to Predict Housing Prices". In: *Urban Geography* 34.3, pp. 395–412. DOI: `10.1080/02723638.2013.778662`. eprint: `https://doi.org/10.1080/02723638.2013.778662`. URL: `https://doi.org/10.1080/02723638.2013.778662`.

Chica-Olmo, Jorge, Rafael Cano-Guervos, and Mario Chica-Rivas (2019). "Estimation of housing price variations using spatio-temporal data". In: *Sustainability* 11.6, p. 1551.

Clapp, John (Feb. 2004). "A Semiparametric Method for Estimating Local House Price Indices". In: *Real Estate Economics* 32, pp. 127–160. DOI: `10.1111/j.1080-8620.2004.00086.x`.

Crespo, Ricardo (2009). "Statistical extensions of GWR: Spatial interpolation and a spatiotemporal approach". PhD thesis. National University of Ireland, Maynooth.

Cressie, Noel (2015). *Statistics for spatial data*. John Wiley & Sons.

Crosby, Henry et al. (2016). "A spatio-temporal, Gaussian process regression, real-estate price predictor". In: *Proceedings of the 24th ACM SIGSPATIAL International Conference on Advances in Geographic Information Systems*, pp. 1–4.

Crosby, Henry et al. (2018). "Road distance and travel time for an improved house price Kriging predictor". In: *Geo-Spatial Information Science* 21.3, pp. 185–194.

Crosby, Neil et al. (2018). "Valuations and their importance for real estate investments". In: *Routledge Companion to Real Estate Investment*. Routledge, pp. 143–171.

De Cesare, Luigi, DE Myers, and D Posa (2001). "Product-sum covariance for space-time modeling: an environmental application". In: *Environmetrics: The official journal of the International Environmetrics Society* 12.1, pp. 11–23.

Derdouri, Ahmed and Yuji Murayama (2020). "A comparative study of land price estimation and mapping using regression kriging and machine learning algorithms across Fukushima prefecture, Japan". In: *Journal of Geographical Sciences* 30, pp. 794–822.

Dubé, Jean and Diègo Legros (2013). "A spatio-temporal measure of spatial dependence: An example using real estate data". In: *Papers in Regional Science* 92.1, pp. 19–30.

— (2014). "Spatial econometrics and the hedonic pricing model: what about the temporal dimension?" In: *Journal of Property Research* 31.4, pp. 333–359. DOI: `10.1080/09599916.2014.913655`. eprint: `https://doi.org/10.1080/09599916.2014.913655`. URL: `https://doi.org/10.1080/09599916.2014.913655`.

Dubin, Robin A. (1998a). "Predicting House Prices Using Multiple Listings Data". In: *The Journal of Real Estate Finance and Economics* 17.1, pp. 35–59. URL: `https://EconPapers.repec.org/RePEc:kap:jrefec:v:17:y:1998:i:1:p:35-59`.

— (1998b). "Spatial Autocorrelation: A Primer". In: *Journal of Housing Economics* 7.4, pp. 304–327. ISSN: 1051-1377. DOI: `https://doi.org/10.1006/jhec.1998.0236`. URL: `https://www.sciencedirect.com/science/article/pii/S1051137798902364`.

Elhorst, J Paul (2014). *Spatial econometrics from cross-sectional data to spatial panels*. Springer.

Fischer, Manfred M and Arthur Getis (2010). *Handbook of applied spatial analysis: software tools, methods and applications*. Springer.

Francke, Marc K and Gerjan A Vos (2004). "The hierarchical trend model for property valuation and local price indices". In: *The Journal of Real Estate Finance and Economics* 28.2, pp. 179–208.

Gatzlaff, Dean and Doğan Tirtiroğlu (1995). "Real Estate Market Efficiency: Issues and Evidence". In: *Journal of Real Estate Literature* 3.2, pp. 157–189. DOI: `10.1080/10835547.1995.12090046`. eprint: `https://doi.org/10.1080/10835547.1995.12090046`. URL: `https://doi.org/10.1080/10835547.1995.12090046`.

Gibbons, Stephen and Henry G Overman (2012). "Mostly pointless spatial econometrics?" In: *Journal of Regional Science* 52.2, pp. 172–191.

Gneiting, Tilmann (2002). "Nonseparable, stationary covariance functions for space–time data". In: *Journal of the American Statistical Association* 97.458, pp. 590–600.

Gräler, Benedikt, Lydia Gerharz, and Edzer Pebesma (Mar. 2013). "Spatio-temporal analysis and interpolation of PM 10 measurements in Europe". In:

Gräler, Benedikt, Edzer Pebesma, and Gerard Heuvelink (Jan. 2016). "Spatio-Temporal Interpolation using gstat". In: *The R Journal* 8, pp. 204–218. DOI: `10.32614/RJ-2016-014`.

Harville, David (1976). “Extension of the Gauss-Markov theorem to include the estimation of random effects”. In: *The Annals of Statistics* 4.2, pp. 384–395.

Hawkins, Douglas M (2004). “The problem of overfitting”. In: *Journal of chemical information and computer sciences* 44.1, pp. 1–12.

Hayunga, Darren K and Alexander Kolovos (2015). “Advanced space–time analysis: constructing a real estate price index”. In: *American Economic Association Conference.*

Helbich, Marco and Daniel A Griffith (2016). “Spatially varying coefficient models in real estate: Eigenvector spatial filtering and alternative approaches”. In: *Computers, Environment and Urban Systems* 57, pp. 1–11.

Hengl, Tomislav (Jan. 2009). *A Practical Guide to Geostatistical Mapping.*

Hengl, Tomislav, Gerard B. M. Heuvelink, and David G. Rossiter (2007). “About regression-kriging: From equations to case studies”. In: *Computers & Geosciences* 33.10. Spatial Analysis, pp. 1301–1315. ISSN: 0098-3004. DOI: https://doi.org/10.1016/j.cageo.2007.05.001. URL: http://www.sciencedirect.com/science/article/pii/S0098300407001008.

Heuvelink, Gerard B. M. and Daniel A Griffith (2010). “Space-Time Geostatistics for Geography: A Case Study of Radiation Monitoring Across Parts of Germany”. In: *Geographical Analysis* 42.2, pp. 161–179.

Heuvelink, Gerard B. M., Edzer J. Pebesma, and Benedikt Gräler (2017). “Space-Time Geostatistics”. In: *Encyclopedia of GIS.* Ed. by Shashi Shekhar, Hui Xiong, and Xun Zhou. Cham: Springer International Publishing, pp. 1919–1926. ISBN: 978-3-319-17885-1. DOI: 10.1007/978-3-319-17885-1_1647. URL: https://doi.org/10.1007/978-3-319-17885-1_1647.

Huang, Bo, Bo Wu, and Michael Barry (2010). “Geographically and temporally weighted regression for modeling spatio-temporal variation in house prices”. In: *International Journal of Geographical Information Science* 24.3, pp. 383–401. DOI: 10.1080/13658810802672469. eprint: https://doi.org/10.1080/13658810802672469. URL: https://doi.org/10.1080/13658810802672469.

Jost, G, GBM Heuvelink, and A Papritz (2005). “Analysing the space–time distribution of soil water storage of a forest ecosystem using spatio-temporal kriging”. In: *Geoderma* 128.3-4, pp. 258–273.

Kane, Thomas J., Stephanie K. Riegg, and Douglas O. Staiger (Aug. 2006). “School Quality, Neighborhoods, and Housing Prices”. In: *American Law and Economics Review* 8.2, pp. 183–212. ISSN: 1465-7252. DOI: 10.1093/aler/ahl007. eprint: https://academic.oup.com/aler/article-pdf/8/2/183/19439414/ahl007.pdf. URL: https://doi.org/10.1093/aler/ahl007.

Kane, Thomas J. et al. (2003). “School Accountability Ratings and Housing Values [with Comments]”. In: *Brookings-Wharton Papers on Urban Affairs*, pp. 83–137. ISSN: 15287084. URL: http://www.jstor.org/stable/25067396.

Kaufman, Leonard and Peter J Rousseeuw (2009). *Finding groups in data: an introduction to cluster analysis*. John Wiley & Sons.

Khabazi, Maryam (2018). "Regional Economic Analysis of Mecklenburg County, NC". In:

Kilibarda, Milan et al. (Mar. 2014). "Spatio-temporal interpolation of daily temperatures for global land areas at 1 km resolution". In: *Journal of Geophysical Research: Atmospheres* 119, n/a–n/a. DOI: 10.1002/2013JD020803.

Kitanidis, Peter K (1993). "Generalized covariance functions in estimation". In: *Mathematical Geology* 25.5, pp. 525–540.

Kobylińska, Katarzyna and Radoslaw Cellmer (Dec. 2016). "The Use of Indicator Kriging for Analyzing Prices in the Real Estate Market". In: *Real Estate Management and Valuation* 24. DOI: 10.1515/remav-2016-0025.

Kopczewska, Katarzyna and Piotr Ćwiakowski (2021). "Spatio-temporal stability of housing submarkets. Tracking spatial location of clusters of geographically weighted regression estimates of price determinants". In: *Land Use Policy* 103, p. 105292.

Kuntz, Michael and Marco Helbich (2014). "Geostatistical mapping of real estate prices: an empirical comparison of kriging and cokriging". In: *International Journal of Geographical Information Science* 28.9, pp. 1904–1921.

Kyriakidis, Phaedon C. and André G. Journel (1999). "Geostatistical Space-Time Models: A Review". In: *Mathematical Geology* 31.6, pp. 651–684. ISSN: 1573-8868. DOI: 10.1023/A:1007528426688. URL: https://doi.org/10.1023/A:1007528426688.

Leung, Charles K et al. (2010). "House market in Chinese cities: Dynamic modeling, in-sample fitting and out-of-sample forecasting". In: *Sample Fitting and Out-of-Sample Forecasting (August 22, 2010)*.

Li, Lixin and Peter Revesz (2004). "Interpolation methods for spatio-temporal geographic data". In: *Computers, Environment and Urban Systems* 28.3, pp. 201–227.

Linden, Leigh and Jonah E. Rockoff (June 2008). "Estimates of the Impact of Crime Risk on Property Values from Megan's Laws". In: *American Economic Review* 98.3, pp. 1103–27. DOI: 10.1257/aer.98.3.1103. URL: https://www.aeaweb.org/articles?id=10.1257/aer.98.3.1103.

Liu, Xiaolong (2012). "Spatial and temporal dependence in house price prediction". In: *The Journal of Real Estate Finance and Economics* 47.2, pp. 341–369.

Mayer, Michael et al. (2019). "Estimation and updating methods for hedonic valuation". In: *Journal of European real estate research*.

McMillen, Daniel P (2012). "Perspectives on spatial econometrics: linear smoothing with structured models". In: *Journal of Regional Science* 52.2, pp. 192–209.

Miller, Eric J (1997). "interpolation utilizing kriging". In: *Innovations in GIS* 4, p. 181.

Montero, Jose and Beatriz Larraz (2011). "Interpolation methods for geographical data: housing and commercial establishment markets". In: *Journal of Real Estate Research* 33.2, pp. 233–244.

Montero-Lorenzo, José-María and Beatriz Larraz-Iribas (2012). "Space-time approach to commercial property prices valuation". In: *Applied Economics* 44.28, pp. 3705–3715. DOI: 10.1080/00036846.2011.581212. eprint: https://doi.org/10.1080/00036846.2011.581212. URL: https://doi.org/10.1080/00036846.2011.581212.

Mooya, Manya M. (2016). *Real Estate Valuation Theory: a critical appraisal.* 20169330492016933049. Springer. ISBN: 978-3-662-49163-8.

Mouselimis, Lampros (2022). *ClusterR: Gaussian Mixture Models, K-Means, Mini-Batch-Kmeans, K-Medoids and Affinity Propagation Clustering.* R package version 1.2.6. URL: https://CRAN.R-project.org/package=ClusterR.

Nappi-Choulet, Ingrid and Tristan-Pierre Maury (2009). "A spatiotemporal autoregressive price index for the Paris office property market". In: *Real Estate Economics* 37.2, pp. 305–340.

Odeh, I.O.A., A.B. McBratney, and D.J. Chittleborough (1995). "Further results on prediction of soil properties from terrain attributes: heterotopic cokriging and regression-kriging". In: *Geoderma* 67.3, pp. 215–226. ISSN: 0016-7061. DOI: https://doi.org/10.1016/0016-7061(95)00007-B. URL: https://www.sciencedirect.com/science/article/pii/001670619500007B.

Otto, Philipp and Wolfgang Schmid (2018). "Spatiotemporal analysis of German real-estate prices". In: *The Annals of Regional Science* 60.1, pp. 41–72.

Pace, Kelley et al. (Jan. 1998). "Spatio-Temporal Autoregressive Models of Neighborhood Effects". In: *The Journal of Real Estate Finance and Economics* 17, pp. 15–33. DOI: 10.1023/A:1007799028599.

Pace, Kelley et al. (2000). "A method for spatial–temporal forecasting with an application to real estate prices". In: *International Journal of Forecasting* 16.2, pp. 229–246. ISSN: 0169-2070. DOI: https://doi.org/10.1016/S0169-2070(99)00047-3. URL: https://www.sciencedirect.com/science/article/pii/S0169207099000473.

Padgham, Mark et al. (June 2017). "osmdata". In: *The Journal of Open Source Software* 2.14. DOI: 10.21105/joss.00305. URL: https://doi.org/10.21105/joss.00305.

Palma, Monica et al. (2019). "The residential real estate market in Italy: A spatio-temporal analysis". In: *Quality & Quantity* 53.5, pp. 2451–2472.

Pinkse, Joris and Margaret E Slade (2010). "The future of spatial econometrics". In: *Journal of Regional Science* 50.1, pp. 103–117.

Refaeilzadeh, Payam, Lei Tang, and Huan Liu (2009). "Cross-Validation". In: *Encyclopedia of Database Systems.* Ed. by LING LIU and M. TAMER ÖZSU. Boston, MA: Springer US, pp. 532–538. ISBN: 978-0-387-39940-9. DOI: 10.1007/978-0-387-39940-9_565. URL: https://doi.org/10.1007/978-0-387-39940-9_565.

Rosen, Sherwin (1974). "Hedonic prices and implicit markets: product differentiation in pure competition". In: *Journal of political economy* 82.1, pp. 34–55.

Sherman, Michael (2011). *Spatial statistics and spatio-temporal data: covariance functions and directional properties.* John Wiley & Sons.

Smith, Tony and Peggy Wu (Dec. 2009). “A Spatio-Temporal Model of Housing Prices Based on Individual Sales Transactions Over Time”. In: *Journal of Geographical Systems* 11, pp. 333–355. DOI: 10.1007/s10109-009-0085-9.

Snepvangers, J.J.J.C, G.B.M Heuvelink, and J.A Huisman (2003). “Soil water content interpolation using spatio-temporal kriging with external drift”. In: *Geoderma* 112.3. Pedometrics 2001, pp. 253–271. ISSN: 0016-7061. DOI: https://doi.org/10.1016/S0016-7061(02)00310-5. URL: https://www.sciencedirect.com/science/article/pii/S0016706102003105.

Soltani, Ali et al. (2021). “Housing price variations using spatio-temporal data mining techniques”. In: *Journal of Housing and the Built Environment* 36.3, pp. 1199–1227.

Stein, Michael L (2005). “Space–time covariance functions”. In: *Journal of the American Statistical Association* 100.469, pp. 310–321.

Sun, Hua, Yong Tu, and Shi-Ming Yu (Feb. 2005). “A Spatio-Temporal Autoregressive Model for Multi-Unit Residential Market Analysis*”. In: *The Journal of Real Estate Finance and Economics* 31, pp. 155–187. DOI: 10.1007/s11146-005-1370-0.

United States Census Bureau (n.d.). *QuickFacts Mecklenburg County, North Carolina.* Online. [Accessed 2020-04-10]. URL: https://www.census.gov/quickfacts/mecklenburgcountynorthcarolina.

Venkatachalam, Parvatham and Manoj Kumar (2017). “Space-Time Variogram Modeling”. In: *Encyclopedia of GIS*. Ed. by Shashi Shekhar, Hui Xiong, and Xun Zhou. Cham: Springer International Publishing, pp. 1932–1938. ISBN: 978-3-319-17885-1. DOI: 10.1007/978-3-319-17885-1_1643. URL: https://doi.org/10.1007/978-3-319-17885-1_1643.

Willmott, Cort J (1982). “Some comments on the evaluation of model performance”. In: *Bulletin of the American Meteorological Society* 63.11, pp. 1309–1313.

Wu, Bo, Rongrong Li, and Bo Huang (May 2014). “A Geographically and Temporally Weighted Autoregressive Model with Application to Housing Prices”. In: *International Journal of Geographical Information Science* 28, pp. 1186–1204. DOI: 10.1080/13658816.2013.878463.

Xiao, Hui (2020). “Essays on Model Uncertainty and Real Estate Markets”. PhD thesis. University of Guelph.

Yao, Jing and A. Fotheringham (June 2015). “Local Spatiotemporal Modeling of House Prices: A Mixed Model Approach”. In: *The Professional Geographer* 68, pp. 1–13. DOI: 10.1080/00330124.2015.1033671.

Yoo, E.-H. and Phaedon C. Kyriakidis (2009). “Area-to-point Kriging in spatial hedonic pricing models”. In: *Journal of Geographical Systems* 11.4, p. 381. ISSN: 1435-5949. DOI: 10.1007/s10109-009-0090-z. URL: https://doi.org/10.1007/s10109-009-0090-z.

Zhang, Hao and Yong Wang (2010). “Kriging and cross-validation for massive spatial data”. In: *Environmetrics: The official journal of the International Environmetrics Society* 21.3-4, pp. 290–304.

Zhu, Ciyou et al. (1997). “Algorithm 778: L-BFGS-B: Fortran subroutines for large-scale bound-constrained optimization”. In: *ACM Transactions on mathematical software (TOMS)* 23.4, pp. 550–560.

Zoest, Vera van et al. (2020). “Spatio-temporal regression kriging for modelling urban NO2 concentrations”. In: *International Journal of Geographical Information Science* 34.5, pp. 851–865. DOI: 10.1080/13658816.2019.1667501. eprint: https://doi.org/10.1080/13658816.2019.1667501. URL: https://doi.org/10.1080/13658816.2019.1667501.

Appendix

Data preparation

```
### PACKAGES
library(sf)
library(dplyr)
library(lubridate)
library(mapview)
library(sp)
library(spacetime)

### DATA PROCESSING
select <- "pid, totalac, totalvalue, dateofsale, price, yearbuilt, heatedarea,
↪  storyheigh, aheatingty, actype, extwall, foundation, numfirepla, fireplaces,
↪  bldggrade, fullbaths, halfbaths, bedrooms, units, neighbourh, landvalue,
↪  municipali, city, zipcode"
from <- "\"Parcel_TaxData\""

where <- "descbuildi='RES' AND price >= 100000 AND price <= 5000000 AND
↪  extravalue = 0 AND yearbuilt <= 2022 AND price >= 0.3*totalvalue AND price
↪  <= 1.8*totalvalue AND accounttyp ='INDIVIDUAL' AND heatedarea >= 50 AND
↪  yearbuilt >= 1800 AND state='NC' AND vacantorim <> 'VAC' AND ownertype IS
↪  NULL AND bedrooms >= 1 AND bedrooms <= 5 AND fullbaths >= 1 AND fullbaths
↪  <=5 AND halfbaths <= 2 AND numfirepla <= 3"

query <- paste("SELECT", select, "FROM", from, "WHERE", where, sep = " ")

RawDat <- "/home/philibert/Documents/Thesis/Real Estate
↪  Studies/Parcel_TaxData.shp"
data <- read_sf(RawDat, query = query)
rm(list = c("from","query","select","where"))

sales <- data %>%
  mutate(
    dateofsale = as.Date(dateofsale),
    BuiltCohort = case_when(
      yearbuilt >= 1900 & yearbuilt <= 1930 ~ "[1900,1930]",
```

```
    yearbuilt >  1930 & yearbuilt <= 1944 ~ "(1930,1944]",
    yearbuilt >  1944 & yearbuilt <= 1960 ~ "(1944,1960]",
    yearbuilt >  1960 & yearbuilt <= 1970 ~ "(1960,1970]",
    yearbuilt >  1970 & yearbuilt <= 1980 ~ "(1970,1980]",
    yearbuilt >  1980 & yearbuilt <= 1990 ~ "(1980,1990]",
    yearbuilt >  1990 & yearbuilt <= 2000 ~ "(1990,2000]",
    yearbuilt >  2000 & yearbuilt <= 2010 ~ "(2000,2010]",
    yearbuilt >  2010 & yearbuilt <= 2020 ~ "(2010,2020]"),
  BuiltCohort = factor(BuiltCohort, levels=c("[1900,1930]", "(1930,1944]",
  ↪  "(1944,1960]", "(1960,1970]", "(1970,1980]", "(1980,1990]",
  ↪  "(1990,2000]", "(2000,2010]", "(2010,2020]")),
  city = case_when(
    city == "CHAROTTE" ~ "CHARLOTTE",
    city == "MARIN" ~ "MARVIN",
    city == "MOORSEVILLE" ~ "MOORESVILLE",
    city == "MORISVILLE" ~ "MOORESVILLE",
    city == "MT HOLLY" ~ "MOUNT HOLLY",
    city == "SHERRILL FORD" ~ "SHERRILS FORD",
    city == "SHERRILLSFORD" ~ "SHERRILS FORD",
    city == "WINSTON SALEM" ~ "WINSTON-SALEM",
    city == "WINTSON SALEM" ~ "WINSTON-SALEM",
    TRUE ~ as.character(city)),

  Season = case_when(
    month(dateofsale) %in% c(3,4,5) ~ "Spring",
    month(dateofsale) %in% c(6,7,8) ~ "Summer",
    month(dateofsale) %in% c(9,10,11) ~ "Autumn",
    month(dateofsale) %in% c(12,1,2) ~ "Winter"),
  Season = as.factor(Season)) %>%

distinct(pid, dateofsale, price, .keep_all=TRUE) %>%

distinct(dateofsale, price, neighbourh, .keep_all=TRUE) %>%

filter(
  year(dateofsale) >= yearbuilt,
  totalac <= quantile(totalac, 0.95)) %>%
```

```
  mutate(
    saleyear = year(dateofsale),
    storyheigh = as.factor(storyheigh),
    aheatingty = as.factor(aheatingty),
    actype = as.factor(actype),
    extwall = as.factor(extwall),
    foundation = as.factor(foundation),
    #fireplaces = as.factor(fireplaces),
    bldggrade = as.factor(bldggrade),
    psf = price/heatedarea) %>%
  na.omit(sales) %>%
  filter(extwall %in% names(which(table(extwall)>= 500)),
         bldggrade %in% names(which(table(bldggrade)>= 500)),
         storyheigh %in% names(which(table(storyheigh)>= 500)),
         psf <= 550) %>%
  mutate(extwall = droplevels(extwall),
         bldggrade = droplevels(bldggrade),
         storyheigh = droplevels(storyheigh)) %>%
  dplyr::select(-c("fireplaces", "actype", "totalvalue", "aheatingty",
   ↪  "foundation", "units"))

sales <- st_point_on_surface(sales)

### CREATE CLUSTERS
x <- st_coordinates(sales)[,1]
y <- st_coordinates(sales)[,2]
sf <- sales$heatedarea
pz <- sales$totalac

ClusDat1 <- data.frame(x,y)
ClusDat2 <- data.frame(sf, pz)
set.seed(27)
Clus1 <- kmeans(ClusDat1, 7, nstart = 50, iter.max = 50)
Clus2 <- kmeans(ClusDat2, 4, nstart = 50, iter.max = 50)

sales$Clus1 <- as.factor(Clus1$cluster)
sales$Clus2 <- as.factor(Clus2$cluster)
mapview(sales,zcol="Clus1")
```

```
mapview(sales,zcol="Clus2")

### SPACE TIME IRREGULAR DATA FRAME
sales <- as(sales, "Spatial")
SalesSP <- SpatialPoints(sales@coords, proj4string = CRS("+proj=lcc +lat_0=33.75
↪  +lon_0=-79 +lat_1=34.3333333333333 +lat_2=36.1666666666667
↪  +x_0=609601.220000001 +y_0=0 +datum=NAD83 +units=us-ft +no_defs"))
SalesDF <- data.frame(sales)
SalesTM <- as.Date(sales$dateofsale)
SalesSTIDF <- STIDF(SalesSP,SalesTM,data=SalesDF)
saveRDS(SalesSTIDF, "SampleData")
```

Fitting and predicting

```
### PACKAGES
library(spacetime)
library(sp)
library(dplyr)
library(Metrics)
library(nlme)
library(lattice)
library(gstat)

### PREPARATION
#To keep coefficients add below rule 62: LST <<- lst;COEF <<- beta
#In terminal use insert to alter text; ctrl+:; w to save; ctrl+: q to quit
trace("krigeST.df", edit=T, where = krigeST)

setwd("/home/philibert/Dropbox/Masters Thesis Real Estate Studies/Thesis Real
↪  Estate Studies")

theme.novpadding <- list(layout.heights = list(top.padding = 0, main.key.padding
↪  = 0, key.axis.padding = 0, axis.xlab.padding = 0, xlab.key.padding = 0,
↪  key.sub.padding = 0, bottom.padding = 0), layout.widths = list(left.padding
↪  = 1, key.ylab.padding = 0, ylab.axis.padding = 0, axis.key.padding = 0,
↪  right.padding = 1))

SampleData <- readRDS("SampleData")

it <- 0
predlist <- list()
varlist <- list()
coeflist <- list()
tclus <- c("1980::1999","2000::2007","2008::2013","2014::2018","2019::2022")
sclus <- levels(SampleData@data$Clus1)

### FITTING AND PREDICTING
for(TiEx in tclus){
  for(GeEx in sclus){
    it <- it+1
```

```
xy <- SampleData@data$Clus1==GeEx
SubSamp <- SampleData[xy,TiEx]
print(length(SubSamp))
SubSamp@data <- droplevels(SubSamp@data)
top9 <- round(length(SubSamp)*0.90)
TrainSamp <- SubSamp[0:top9]
TestSamp <- SubSamp[(top9+1):length(SubSamp)]

#fit0 <- lm(log(price) ~ heatedarea + totalac + numfirepla + fullbaths +
↪  halfbaths + bedrooms + Season + dateofsale + coords.x1 + coords.x2 +
↪  I(coords.x1^2) + I(coords.x2^2) + I(coords.x1*coords.x2), data =
↪  TrainSamp@data)

#fit0 <- lm(log(price) ~ heatedarea + totalac + extwall + storyheigh +
↪  numfirepla + bldggrade + fullbaths + halfbaths + bedrooms + Season +
↪  dateofsale + yearbuilt, data = TrainSamp@data)

fit0 <- lm(log(price) ~ heatedarea + totalac + extwall + storyheigh +
↪  numfirepla + bldggrade + fullbaths + halfbaths + bedrooms + Season +
↪  dateofsale + BuiltCohort, data = TrainSamp@data)

TestSamp@data <- TestSamp@data %>%
  mutate(
    extwall = replace(extwall, !extwall %in% unlist(fit0$xlevels[1]),NA),
    storyheigh = replace(storyheigh, !storyheigh %in%
    ↪  unlist(fit0$xlevels[2]),NA),
    bldggrade = replace(bldggrade, !bldggrade %in%
    ↪  unlist(fit0$xlevels[3]),NA),
    Season = replace(Season, !Season %in% unlist(fit0$xlevels[4]),NA),
    BuiltCohort = replace(BuiltCohort, !BuiltCohort %in%
    ↪  unlist(fit0$xlevels[5]), NA)
  )

for(i in c("extwall","storyheigh","bldggrade","Season","BuiltCohort")){
  index <- which(is.na(TestSamp@data[[i]]))
  if(length(index) > 0){
    TestSamp <- TestSamp[-index,]
  }
```

```
}

TrainSamp$res <- fit0$residuals

if(length(TrainSamp) >= 3000){
  x <- sample(nrow(TrainSamp@data), size = 3000, replace = FALSE)
  VarSamp <- TrainSamp[x,]
} else {VarSamp <- TrainSamp}

time <-
↪  as.numeric(max(VarSamp@data$dateofsale)-min(VarSamp@data$dateofsale))
ll = !is.na(is.projected(VarSamp@sp)) && !is.projected(VarSamp@sp)
dist <- spDists(t(VarSamp@sp@bbox), longlat = ll)[1, 2]/8



vgram <- variogramST(formula = res ~ 0, data = VarSamp, tunit = "days",
↪  tlags = seq(0,time/2,time/8), cutoff = dist, width = dist/20, boundaries
↪  = , progress = TRUE, pseudo = 0, assumeRegular = F, na.omit = T, cores =
↪  7)

vgram$gamma <- vgram$gamma*1000 #re-scaling required for fit.stvariogram to
↪  work

f = function(x) attr(m.fit <<- fit.StVariogram(
  vgram,
  vgmST("sumMetric",
        space=vgm(psill=14,
                  model="Exp",
                  range=x,
                  nugget=0),
        time =vgm(psill=4,
                  model="Exp",
                  range=time/2,
                  nugget=0),
        joint=vgm(psill=4,
                  model="Exp",
                  range=x,
```

```
                    nugget=0),
          stAni=1,
          temporalUnit = "days"),
    lower = c(0, 10, 0, 0, 10, 0, 0, 10, 0, 1),
    upper = c(50, 30000, 10, 20, 30000, 50, 50, 30000, 20, 100)),
    "optim.output")$value

optimize(f, c(3000,12000))

m.fit$space$psill <- m.fit$space$psill/1000
m.fit$time$psill <- m.fit$time$psill/1000
m.fit$joint$psill <- m.fit$joint$psill/1000
vgram$gamma <- vgram$gamma/1000

i <- 13
UKrige <- NA
while(suppressWarnings(is.na(UKrige))){

  Ind <- "log(price)"

  if(i>0){
    Dep <-c("heatedarea", "totalac", "extwall", "storyheigh", "numfirepla",
    ↪   "bldggrade", "fullbaths", "halfbaths", "bedrooms", "Season",
    ↪   "dateofsale", "yearbuilt")
    f <- as.formula(paste(Ind, paste(Dep[-i], collapse = " + "), sep = " ~
    ↪   "))
  } else {
    Dep <- c("heatedarea", "totalac"," storyheigh", "fullbaths",
    ↪   "halfbaths", "bedrooms", "Season", "dateofsale", "yearbuilt")
    f <- as.formula(paste(Ind, paste(Dep, collapse = " + "), sep = " ~ "))
    print("SHORT")
  }

  UKrige <- tryCatch(
    expr = {krigeST(f,
                    data = TrainSamp,
                    newdata = TestSamp,
                    modelList = m.fit,
```

```
                        computeVar = FALSE,
                        fullCovariance = FALSE,
                        nmax= )},
        error = function(cnd){
          print(cnd)
          return(NA)}
      )
      i <- i-1
    }




    pred.ols <- exp(predict(fit0, TestSamp@data))
    pred.rk <- exp(UKrige$var1.pred)
    act <- TestSamp@data$price

    preddat <- data.frame(act,pred.ols,pred.rk,TiEx,GeEx)
    predlist[[it]] <-preddat

    nug_s <- m.fit$space$psill[1]
    sill_s <- m.fit$space$psill[2]
    range_s <- m.fit$space$range[2]
    nug_t <- m.fit$time$psill[1]
    sill_t <- m.fit$time$psill[2]
    range_t <- m.fit$time$range[2]
    nug_j <- m.fit$joint$psill[1]
    sill_j <- m.fit$joint$psill[2]
    range_j <- m.fit$joint$range[2]
    st_ani <- as.numeric(m.fit$stAni)

    var <- data.frame(nug_s,sill_s, range_s, nug_t, sill_t, range_t, nug_j,
    ↪  sill_j, range_j, st_ani)
    var$TiEx <- TiEx;var$GeEx<-GeEx
    varlist[[it]] <- var

    CHAR <- as.character(colnames(LST$X))
    df <- data.frame(CHAR,COEF)
    df$TiEx <- TiEx;df$GeEx<-GeEx;df$Mod<-"RK"
```

```
    coeflist[[it]] <- df

    pdf(file=paste("wireframe",TiEx,"_", GeEx,".pdf", sep=""), width = 15,
     ↪  height = 7.5)
    print(plot(vgram,m.fit,all=T,wireframe=T,
               xlab=list("distance (ft)", rot=30), ylab=list("time lag (days)",
                ↪  rot=-35), scales=list(arrows=F, z = list(distance = 5)),
               par.settings=theme.novpadding))
    dev.off()
  }
}

predstat <- bind_rows(predlist)

varstat <- bind_rows(varlist)

coefstat <- bind_rows(coeflist)

save.image(file = "LSTRK_90-10")
```

Tables and figures

```
### PACKAGES

library(spacetime)
library(BAMMtools)
library(USAboundaries)
library(USAboundariesData)
library(qwraps2)
library(xtable)
library(Metrics)
library(dplyr)

### FIGURE 4
SampleData <- readRDS("SampleData")
MBC <- as_Spatial(us_counties(states = "NC")[16,])

SampleData@data$logprice <- log(SampleData@data$price)
k <- getJenksBreaks(SampleData@data$logprice,7)
tcuts <- as.Date(c("1980-01-01", "2000-01-01",
↪ "2008-01-01","2014-01-01","2019-01-01","2022-01-01"))

SalePrice <- stplot(SampleData[,,"logprice"], sp.layout = MBC, cuts = k, tcuts =
↪ tcuts, key.space = "right", cex = 0.3, names.attr = c("1980 : 1999","2000 :
↪ 2007","2008 : 2013","2014 : 2018","2019 : 2022"))

# MODIFY LEGEND
SalePrice <- SalePrice
SalePrice$legend$right$args$key$points$pch <- c(16, 16, 16, 16, 16, 16, 16)
SalePrice$legend$right$args$key$points$col <- c("black","#000033FF",
↪ "#0000ECFF", "#8200FFFF", "#FF55AAFF", "#FFB24DFF", "#FFFF60FF")
SalePrice$legend$right$args$key$points$cex <- c(0,1,1,1,1,1,1)
SalePrice$legend$right$args$key$text[[1]] <- c("Log Price","[11.51,11.92]",
↪ "(11.92,12.27]", "(12.27,12.62]", "(12.62,12.98]", "(12.98,13.46]",
↪ "(13.46,15.04]")

#SAVE FIGURE
pdf(file = "SalePrice.pdf", width = 15, height = 4.2)
```

```
SalePrice
dev.off()

### TABLE 6 (9 can be obtained by loading predictions LSTRK_95-05)
load("~/Dropbox/Masters Thesis Real Estate Studies/Thesis Real Estate
↪  Studies/LSTRK_90-10")
options(qwraps2_markup = "latex")

predstat$act <- log(predstat$act)
predstat$pred.ols <- log(predstat$pred.ols)
predstat$pred.rk  <- log(predstat$pred.rk)

predsum <-
  list("OLS" =
         list("Mean|error|"          = ~ round(mae(act,pred.ols),4),
              "Median|error|"        = ~ round(mdae(act,pred.ols),4),
              "MSE"                  = ~ round(mse(act,pred.ols),4),
              "RMSE"                 = ~ round(rmse(act,pred.ols),4)),
       "LSTRK" =
         list("Mean|error|"          = ~ round(mae(act,pred.rk),4),
              "Median|error|"        = ~ round(mdae(act,pred.rk),4),
              "MSE"                  = ~ round(mse(act,pred.rk),4),
              "RMSE"                 = ~ round(rmse(act,pred.rk),4)),
       "DIF" =
         list("Mean|error|"          = ~ round((mae(act,pred.rk)-mae(act,
         ↪  pred.ols))/mae(act, pred.ols)*100,2),
              "Median|error|"        = ~ round((mdae(act,pred.rk)-mdae(act,
              ↪  pred.ols))/mdae(act, pred.ols)*100,2),
              "MSE"                  = ~ round((mse(act,pred.rk)-mse(act,
              ↪  pred.ols))/mse(act, pred.ols)*100,2),
              "RMSE"                 = ~ round((rmse(act,pred.rk)-rmse(act,
              ↪  pred.ols))/rmse(act, pred.ols)*100,2)))

sumtab1 <- cbind(summary_table(group_by(predstat,GeEx),predsum),
                 summary_table(predstat, predsum))

sumtab2 <- cbind(summary_table(group_by(predstat,TiEx),predsum),
                 summary_table(predstat, predsum))
```

```
sumtab3 <- cbind(summary_table(group_by(predstat,GeEx,TiEx),predsum),
                 summary_table(predstat, predsum))

print(sumtab1, caption = "", markup = "latex")

print(sumtab2, caption = "", markup = "latex")

print(sumtab3, caption = "", markup = "latex")

### TABLE 7 0-90-90:95
predlist_l <- list()

for(j in 1:35){
  bot5 <- round(as.numeric(nrow(predlist[[j]]))/2)
  predlist_l[[j]] <- predlist[[j]][0:bot5,]
}

predstatl <- bind_rows(predlist_l)
predstatl$act <- log(predstatl$act)
predstatl$pred.ols <- log(predstatl$pred.ols)
predstatl$pred.rk  <- log(predstatl$pred.rk)


sumtab1 <- cbind(summary_table(group_by(predstatl,GeEx),predsum),
                 summary_table(predstatl, predsum))

sumtab2 <- cbind(summary_table(group_by(predstatl,TiEx),predsum),
                 summary_table(predstatl, predsum))

print(sumtab1, caption = "", markup = "latex")

print(sumtab2, caption = "", markup = "latex")



### TABLE 8 0-90-95:100
predlisth <- list()
```

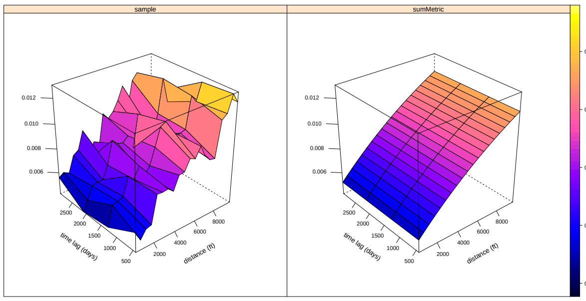
```
for(j in 1:35){
  bot5 <- round(as.numeric(nrow(predlist[[j]]))/2)
  tot <- as.numeric(nrow(predlist[[j]]))
  predlisth[[j]] <- predlist[[j]][(bot5+1):tot,]
}

predstath <- bind_rows(predlisth)
predstath$act <- log(predstath$act)
predstath$pred.ols <- log(predstath$pred.ols)
predstath$pred.rk  <- log(predstath$pred.rk)


sumtab1 <- cbind(summary_table(group_by(predstath,GeEx),predsum),
                 summary_table(predstath, predsum))




sumtab2 <- cbind(summary_table(group_by(predstath,TiEx),predsum),
                 summary_table(predstath, predsum))
print(sumtab1, caption = "", markup = "latex")
print(sumtab2, caption = "", markup = "latex")


### TABLE 6



### VARIOGRAM PARAMETERS
for(t in unique(varstat$TiEx)){
  temp <- varstat %>%
    filter(TiEx == t) %>%
    mutate(
      nug_s = nug_s * 1000,
      sill_s = sill_s * 1000,
      nug_t = nug_t * 1000,
      sill_t = sill_t * 1000,
      nug_j = nug_j * 1000,
      sill_j = sill_j * 1000
```

```
  )

  clname <- paste("Zone",temp$GeEx,sep = " ")
  temp <- as.table(t(temp[1:10]))

  rownames(temp) <- c("$\\tau_s^2$",
                      "$\\sigma_s^2$",
                      "$\\phi_s$",
                      "$\\tau_t^2$",
                      "$\\sigma_t^2$",
                      "$\\phi_t$",
                      "$\\tau_{j}^2$",
                      "$\\sigma_{j}^2$",
                      "$\\phi_{j}$",
                      "$\\kappa$")

  colnames(temp) <- clname

  dig <- matrix(rep(rbind(2,2,0,2,2,0,2,2,0,2),8), ncol = 8)
  x.width <- xtable(temp, digits = dig)
  align(x.width) <- "XRRRRRRR"
  print(x.width, tabular.environment = "tabularx", width = "\\textwidth",
   ↪  include.rownames=T,
        include.colnames=T, floating=F, sanitize.rownames.function = identity,
         ↪  booktab=TRUE, file=paste("VarPar_90-10", t))
}

### TABLE 4
getmode <- function(v) {
  uniqv <- unique(v)
  uniqv[which.max(tabulate(match(v, uniqv)))]
}

ClusSum <- as.data.frame(
  SampleData@data %>%
    group_by(Clus1) %>%
    summarise(Price = round(mean(price)),
              PSF = round(mean(psf)),
```

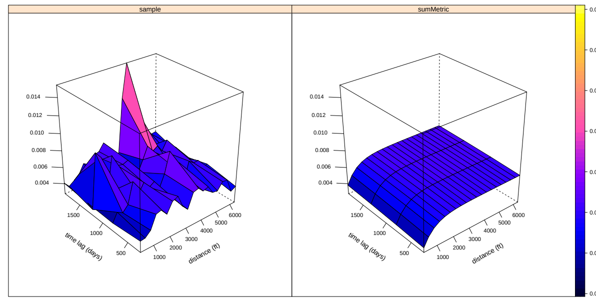
```
              SF = round(mean(heatedarea)),
              PlotSize = round(mean(totalac*43560)),
              YearBuilt = round(mean(yearbuilt)),
              Wall = getmode(extwall),
              Grade = getmode(bldggrade))
)
print(xtable::xtable(ClusSum, digits = 0), auto=TRUE, include.rownames =
↪  F,include.colnames = T, floating = T, booktabs = T)


### TABLE 10
library(tidyr)

SampleData <- readRDS("SampleData")

it <- 0
trainlist <- list()
testlist <- list()
tclus <- c("1980::1999","2000::2007","2008::2013","2014::2018","2019::2022")
sclus <- levels(SampleData@data$Clus1)
for(TiEx in tclus){
  for(GeEx in sclus){
    it <- it+1
    xy <- SampleData@data$Clus1==GeEx
    SubSamp <- SampleData[xy,TiEx]
    print(length(SubSamp))
    SubSamp@data <- droplevels(SubSamp@data)
    top9 <- round(length(SubSamp)*0.90)
    TrainSamp <- SubSamp[0:top9]
    TestSamp <- SubSamp[(top9+1):length(SubSamp)]
    trainlist[[it]] <-TrainSamp@data
    trainlist[[it]]$G <- GeEx; trainlist[[it]]$T <- TiEx

    testlist[[it]] <- TestSamp@data
    testlist[[it]]$G <- GeEx; testlist[[it]]$T <- TiEx
  }
}
```

```
traindat <- bind_rows(trainlist)
testdat <- bind_rows(testlist)
traindat <-st_as_sf(traindat, coords = c("coords.x1", "coords.x2"))
testdat <- st_as_sf(testdat, coords = c("coords.x1", "coords.x2"))


traindat$coords.x1 <- st_coordinates(traindat)[1]
traindat$coords.x2 <- st_coordinates(traindat)[2]
testdat$coords.x1 <- st_coordinates(testdat)[1]
testdat$coords.x2 <- st_coordinates(testdat)[2]

predlist <- list()
it <- 0
for(TiEx in tclus){
  it <- it + 1
  tdata <- traindat %>% filter(T == TiEx)
  vdata <- testdat %>% filter(T == TiEx)

  fit1.1 <- lm(log(price) ~ heatedarea + totalac + extwall + storyheigh +
   ↪ numfirepla + bldggrade + fullbaths + halfbaths + bedrooms + Season +
   ↪ dateofsale + BuiltCohort + coords.x1 + coords.x2 + I(coords.x1^2) +
   ↪ I(coords.x2^2) + I(coords.x1*coords.x2) + municipali , data = tdata)
  fit1.2 <- lm(log(price) ~ heatedarea + totalac + extwall + storyheigh +
   ↪ numfirepla + bldggrade + fullbaths + halfbaths + bedrooms + Season +
   ↪ dateofsale + BuiltCohort + coords.x1 + coords.x2 + I(coords.x1^2) +
   ↪ I(coords.x2^2) + I(coords.x1*coords.x2) + Clus1 , data = tdata)

  fit2.1 <- lm(log(price) ~ heatedarea + totalac + extwall + storyheigh +
   ↪ numfirepla + bldggrade + fullbaths + halfbaths + bedrooms + Season +
   ↪ dateofsale + BuiltCohort + municipali*saleyear , data = tdata)
  fit2.2 <- lm(log(price) ~ heatedarea + totalac + extwall + storyheigh +
   ↪ numfirepla + bldggrade + fullbaths + halfbaths + bedrooms + Season +
   ↪ dateofsale + BuiltCohort + Clus1*saleyear , data = tdata)

  fit3.1 <- lm(log(price) ~ heatedarea + totalac + extwall + storyheigh +
   ↪ numfirepla + bldggrade + fullbaths + halfbaths + bedrooms + Season +
   ↪ dateofsale + BuiltCohort + municipali , data = tdata)
```

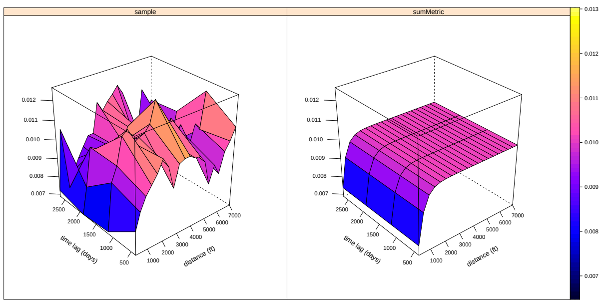
```
  fit3.2 <- lm(log(price) ~ heatedarea + totalac + extwall + storyheigh +
   ↪  numfirepla + bldggrade + fullbaths + halfbaths + bedrooms + Season +
   ↪  dateofsale + BuiltCohort + Clus1 , data = tdata)


  idw(fit3.2$residuals~1,tdata, vdata)$var1.pred

  pred1.1 <- predict(fit1.1, vdata)
  pred1.2 <- predict(fit1.2, vdata)
  pred2.1 <- predict(fit2.1, vdata)
  pred2.2 <- predict(fit2.2, vdata)
  pred3.1 <- predict(fit3.1, vdata) + idw(fit3.1$residuals~1,tdata,
   ↪  vdata)$var1.pred
  pred3.2 <- predict(fit3.2, vdata) + idw(fit3.2$residuals~1,tdata,
   ↪  vdata)$var1.pred

  pred <- data.frame(log(vdata$price),
   ↪  pred1.1,pred1.2,pred2.1,pred2.2,pred3.1,pred3.2)
  pred$TiEx <- TiEx
  predlist[[it]] <- pred
}
preds <- bind_rows(predlist) %>%
    pivot_longer(., cols = c(pred1.1, pred1.2, pred2.1, pred2.2, pred3.1,
     ↪  pred3.2))
colnames(preds) <- c("price", "TiEx","mod", "pred")

GlobModSum <-
  preds %>%
    group_by(mod, TiEx) %>%
    summarise("MAE" =  round(mae(price,pred),4),
              "MdAE" = round(mdae(price,pred),4),
              "MSE" = round(mse(price,pred),4),
              "RMSE" = round(rmse(price,pred),4))
```



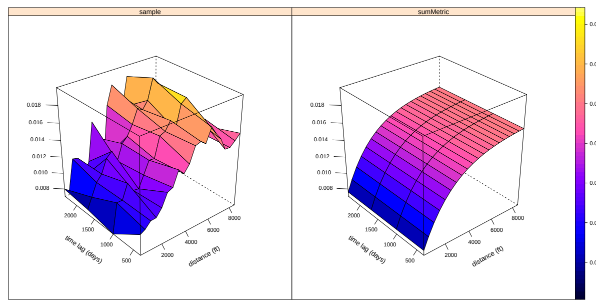
(a) Zone 1



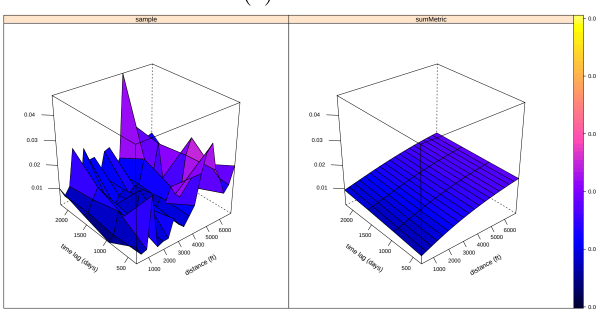
(b) Zone 2



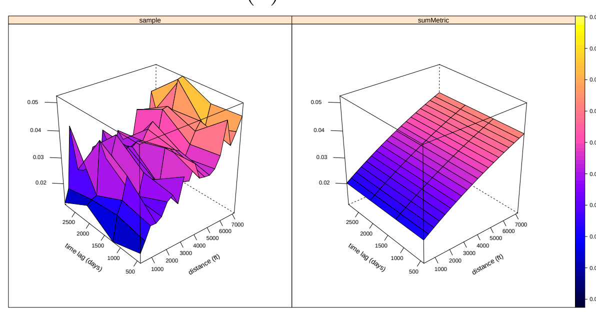
(c) Zone 3



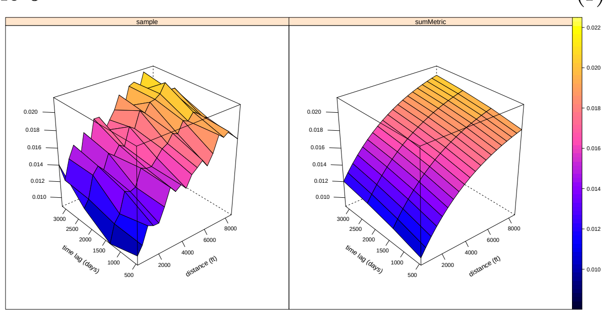
(d) Zone 4



(e) Zone 5

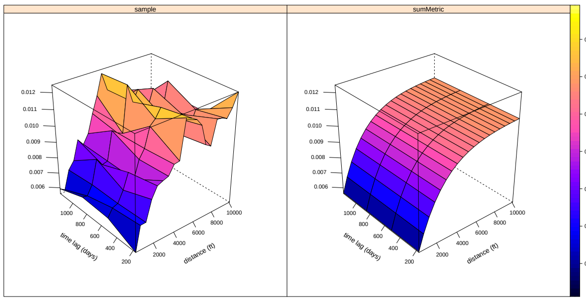


(f) Zone 6

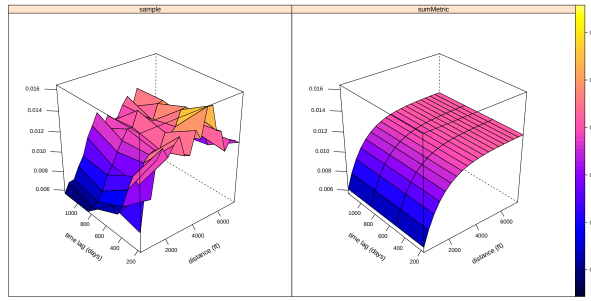


(g) Zone 7

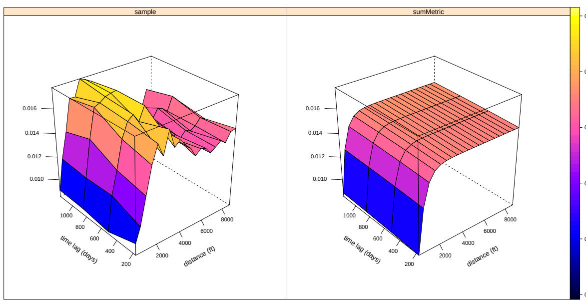
Figure 6: Empirical and fitted space-time variograms 1980:1999



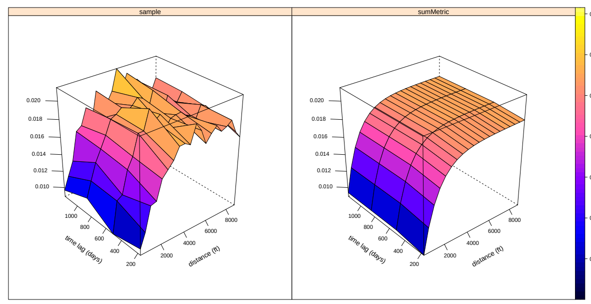
(a) Zone 1



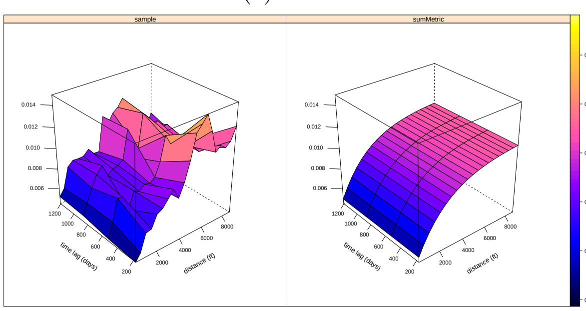
(b) Zone 2



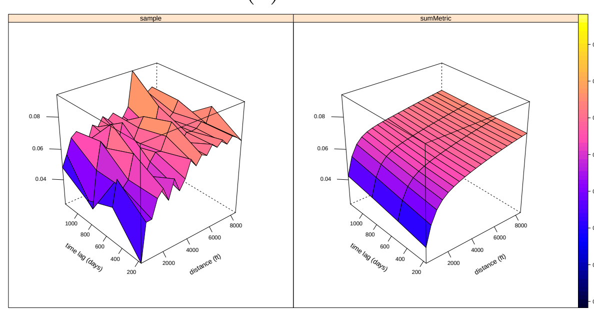
(c) Zone 3



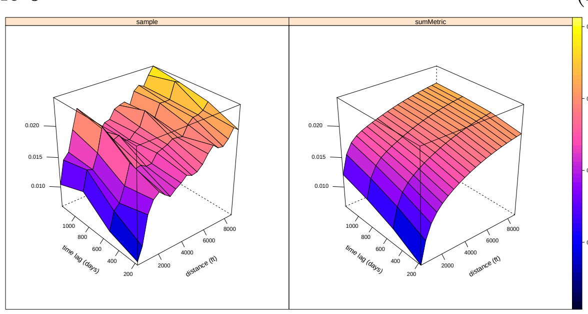
(d) Zone 4



(e) Zone 5

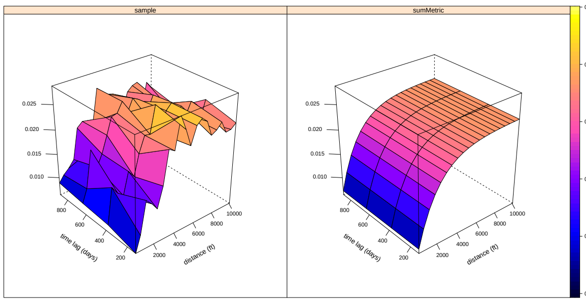


(f) Zone 6

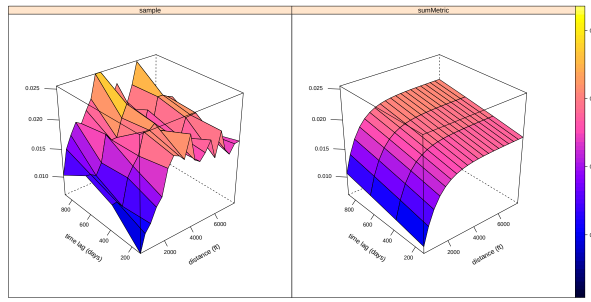


(g) Zone 7

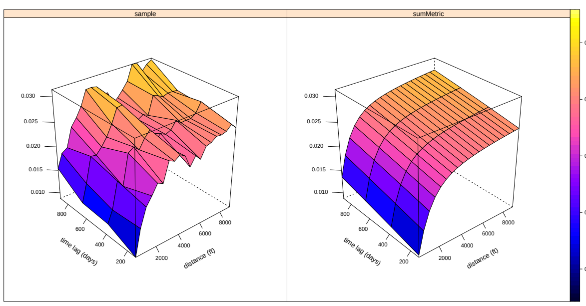
Figure 7: Empirical and fitted space-time variograms 2000:2007



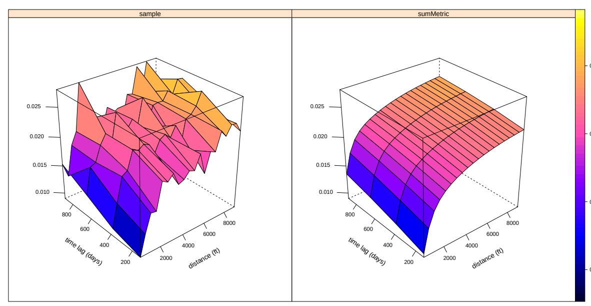
(a) Zone 1



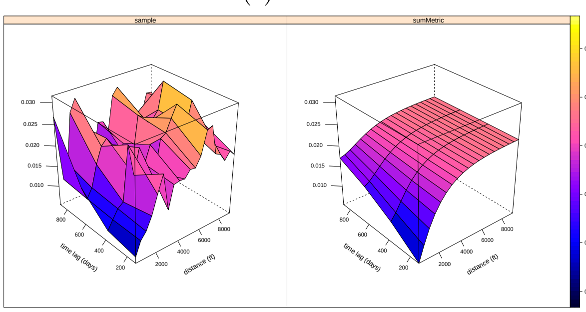
(b) Zone 2



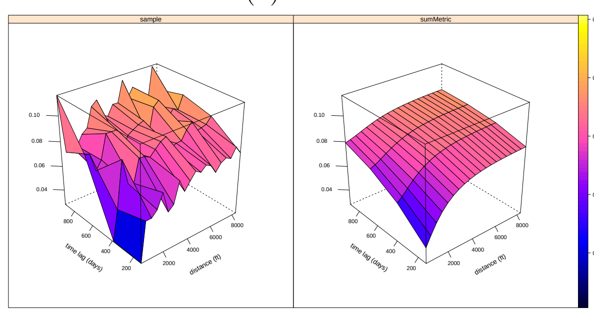
(c) Zone 3



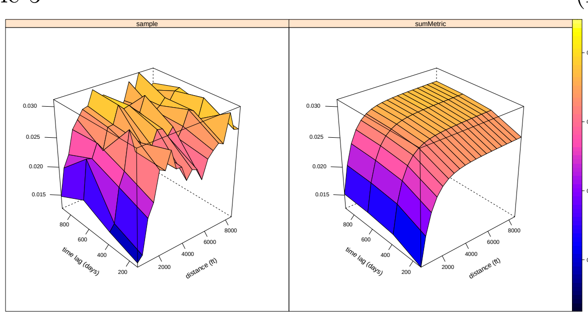
(d) Zone 4



(e) Zone 5

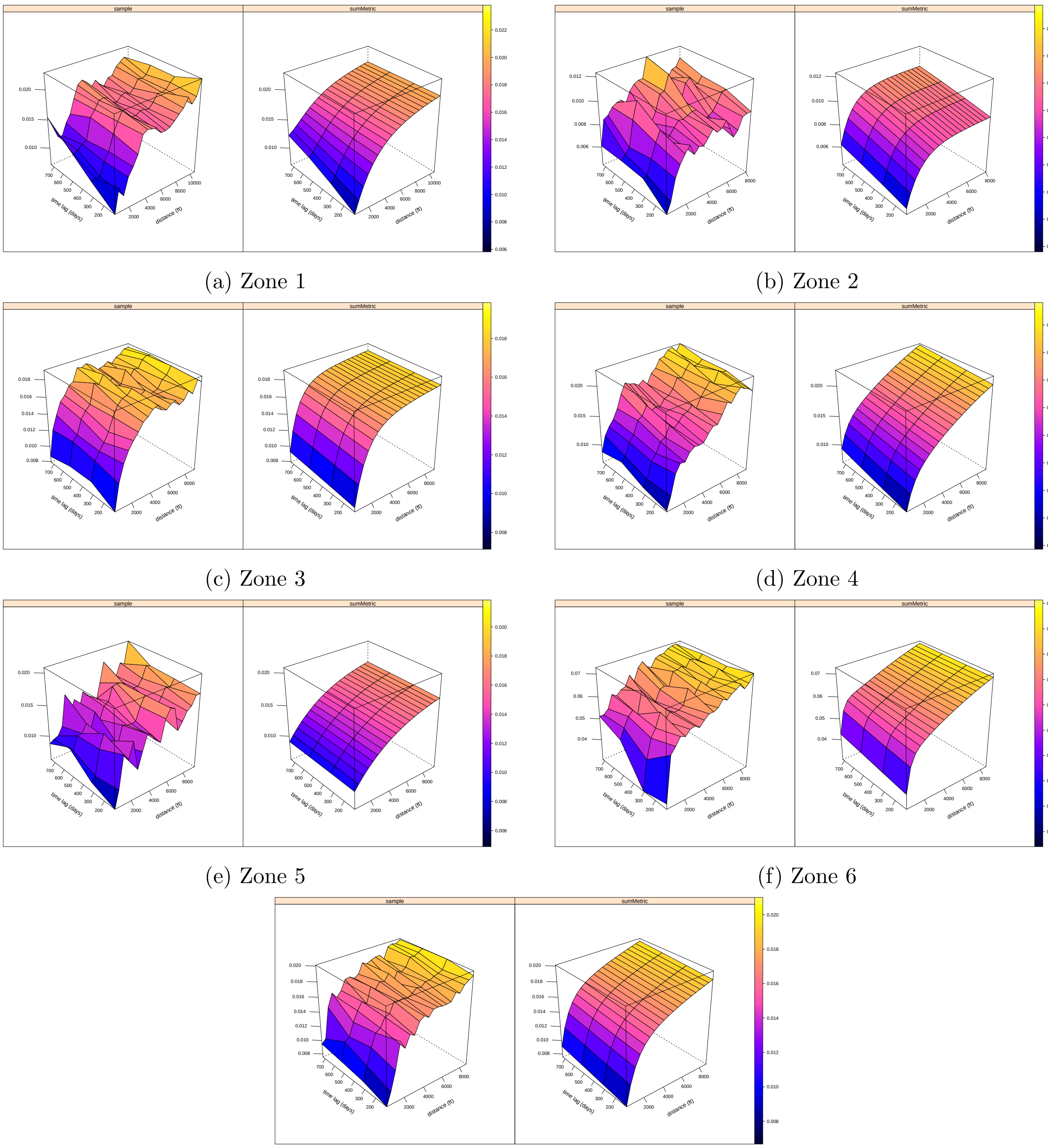


(f) Zone 6



(g) Zone 7

Figure 8: Empirical and fitted space-time variograms 2008:2013



(a) Zone 1

(b) Zone 2

(c) Zone 3

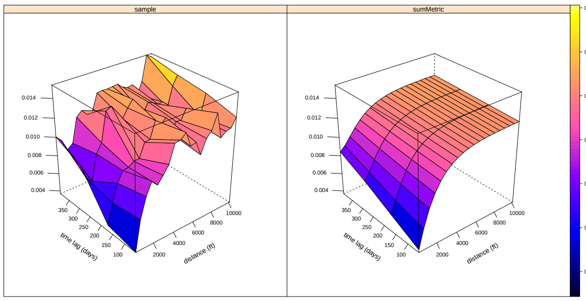
(d) Zone 4

(e) Zone 5

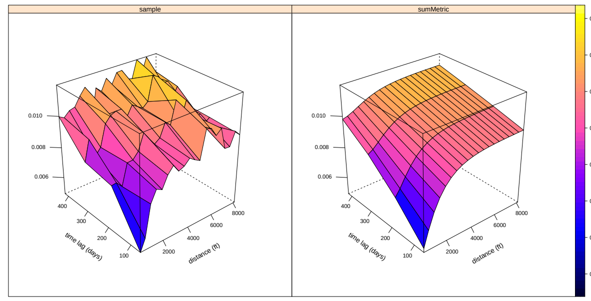
(f) Zone 6

(g) Zone 7

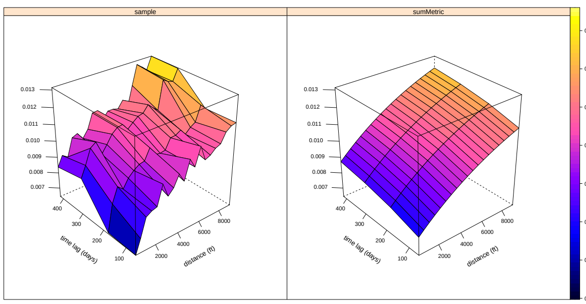
Figure 9: Empirical and fitted space-time variograms 2014:2018



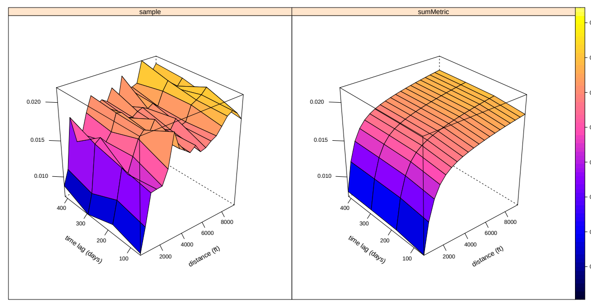
(a) Zone 1



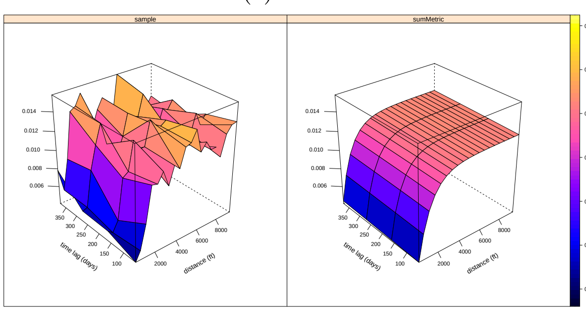
(b) Zone 2



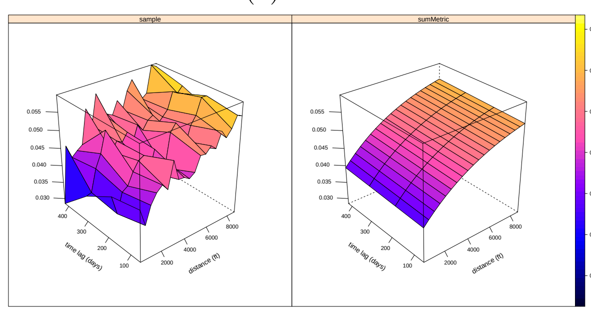
(c) Zone 3



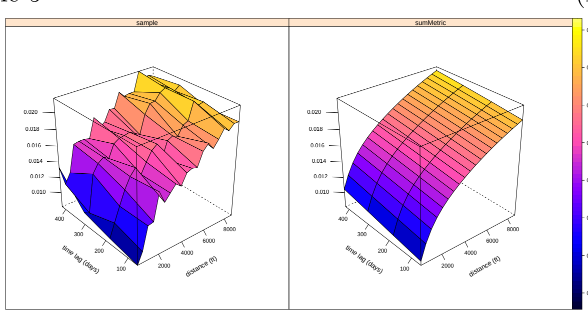
(d) Zone 4



(e) Zone 5



(f) Zone 6



(g) Zone 7

Figure 10: Empirical and fitted space-time variograms 2019:2022

Alternative Clusters

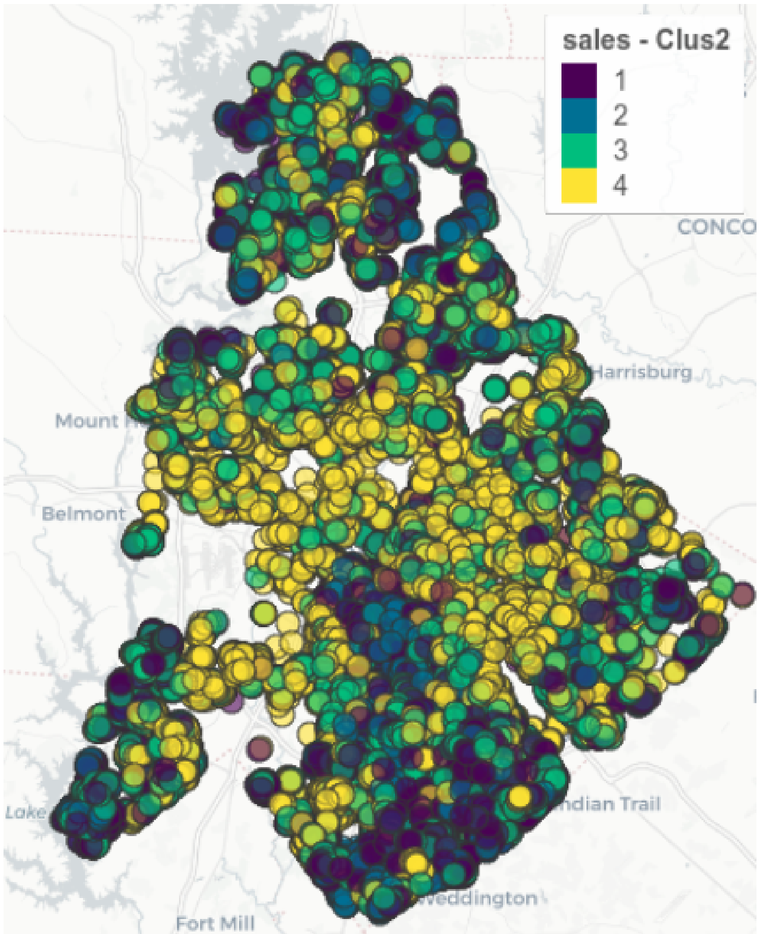


Figure 11: Property clusters based upon heated area and plot size

Table 11: Prediction results obtained using the first 90% of temporally ordered transactions as in-sample and 90-100% of temporally order transactions as out-of-sample, using alternative submarkets based upon property size and plot size

	Zone 1 (N = 408)	Zone 2 (N = 1395)	Zone 3 (N = 2024)	Zone 4 (N = 1885)	Zone 5 (N = 5712)
OLS					
Mean‖error‖	0.16	0.15	0.1514	0.1524	0.152
Median‖error‖	0.1235	0.1214	0.1185	0.1143	0.1187
MSE	0.046	0.0385	0.0418	0.0422	0.0414
RMSE	0.2144	0.1963	0.2044	0.2053	0.2035
LSTRK					
Mean‖error‖	0.1367	0.0942	0.0931	0.1032	0.0998
Median‖error‖	0.0964	0.0661	0.0692	0.0764	0.072
MSE	0.0363	0.0186	0.0183	0.0207	0.0204
RMSE	0.1904	0.1364	0.1352	0.1437	0.1429

(a) Predictions summarized by spatial submarket

	1980:1999 (N = 527)	2000:2007 (N = 1208)	2008:2013 (N = 754)	2014:2018 (N = 1742)	2019:2022 (N = 1481)	Total (N = 5,712)
OLS						
Mean‖error‖	0.1053	0.1528	0.1858	0.1372	0.1681	0.152
Median‖error‖	0.0765	0.1237	0.1512	0.1006	0.1385	0.1187
MSE	0.0218	0.04	0.0572	0.0358	0.048	0.0414
RMSE	0.1476	0.2001	0.2392	0.1892	0.2192	0.2035
LSTRK						
Mean‖error‖	0.0873	0.1079	0.1204	0.0869	0.1024	0.0998
Median‖error‖	0.0605	0.0807	0.093	0.0615	0.0735	0.072
MSE	0.0141	0.0216	0.0268	0.0164	0.0225	0.0204
RMSE	0.1269	0.1471	0.1638	0.1279	0.15	0.1429

(b) Predictions summarized by temporal submarket