UNIVERSITY OF ECONOMICS IN BRATISLAVA

Faculty of Economic Informatics

Evidence number: 103003/D/2025/36124048430084356

Functional data in spatial econometric models

Dissertation thesis

2025 **Peter Knížat, MSc.**

UNIVERSITY OF ECONOMICS IN BRATISLAVA

Faculty of Economic Informatics

Functional data in spatial econometric models

Dissertation thesis

Study program: Data Science in Economics

Field of study: Economics and Management

Study department: Department of Operations Research and Econometrics

Thesis Supervisor: assoc. Professor Ing. Andrea Furková, PhD.

Bratislava 2025 **Peter Knížat, MSc.**

THESIS ASSIGNMENT

Name and Surname:	Peter Knížat, MSc.
Study programme:	Data Science in Economics (Single degree study, Ph.D. (philosophy) III. deg., part time form)
Field of Study:	8. - Economics and Management
Type of Thesis:	Dissertation
Language of Thesis:	English
Secondary language:	Slovak
Thesis's title:	Functional data in spatial econometric models

Literature:

Ramsay, J. O., & Silverman, B. W. (2005). Functional Data Analysis. Second Edition. Springer Series in Statistics, Springer.

Ramsay, J. O., Hooker G., & Graves, S. (2009). Functional Data Analysis with R and MATLAB. Use R!. Springer.

Kopczewska, K (2022). Spatial machine learning: new opportunities for regional science. The Annals of Regional Science, 68, 713–755. https:// doi.org/10.1007/s00168-021-01101-x.

De Brouwer, P. J. S. (2020). The Big R-Book: From Data Science to Learning Machines and Big Data. 1st Edition. Wiley.

Kopczewska, K. (2020). Applied Spatial Statistics and Econometrics: Data Analysis in R. 1st edition. Routledge.

Pineda-Ríos, W., Giraldo, R., & Porcu, E. (2019). Functional SAR models: With application to spatial econometrics. Spatial Statistics, 29, 145-159. https:// doi.org/10.1016/j.spasta.2018.12.002

Goal:

The aim of the dissertation thesis will be the analysis of functional data in the context of econometric models and spatial econometric models, with an expected use of a new data source derived from highway toll system data.

Annotation:

Technological advancements in information technology enable the recording of large sets of economic data that can be described using mathematical functions. These mathematical objects can subsequently be directly incorporated into econometric regression models. Empirical studies have shown that regional economic processes are usually not spatially isolated, but are influenced by economic activities in neighbouring regions. The dissertation thesis will focus on statistical methodologies that allow for the analysis of functional data in the context of econometric models and spatial econometric models. The work will present theoretical concepts and computational strategies for both parametric and non-parametric regression models with functional data, including their extensions to models with spatial effects. The primary analytical tool is expected to be a new class of models known as Functional Spatial

Autoregressive (Functional SAR) models, along with the use of a new data source originating from the highway toll system to enable “rapid” estimates of selected macroeconomic indicators.

Tutor: doc. Ing. Andrea Furková, PhD.
Department: KOVE FHI - Department of Operations Research and Econometrics

Assigned: 30.03.2022

Approved: 01.09.2022 prof. Mgr. Juraj Pekár, PhD.

osoba zodpovedná za realizáciu študijného programu

Declaration of honour

I hereby declare that I made this dissertation thesis by myself and that I mentioned and cited properly all the sources and literature.

Date: …….. ..……

Acknowledgements

I express my sincere gratitude to my dissertation thesis supervisor, assoc. Professor Ing. Andrea Furková, PhD. for her invaluable support, expert guidance, and insightful feedback throughout the process of writing and refining my thesis. The valuable feedback was also provided by assoc. Professor Ing. Michaela Chocholatá, PhD. that lead to a significant improvement of the content of the dissertation thesis. I am also deeply appreciative to Ing. Helena Glaser-Opitzová, PhD. and PhDr. Ľudmila Ivančíková, PhD. for their support and professional feedback at Statistical Office of the Slovak Republic. It is through their collective collaboration and mentorship that I was able to develop and complete my dissertation thesis.

ABSTRACT

KNÍŽAT, Peter: Functional data in spatial econometric models. – University of Economics in Bratislava. Faculty of Economic Informatics; Department of Operations Research and Econometrics. – Thesis supervisor: assoc. Professor Ing. Andrea Furková, PhD. – Bratislava: FHI EU, 2025, 133 pages

The objective of the dissertation thesis is the contribution to both methodological and empirical areas of the researched field. In the methodological study, our goal is to demonstrate the extension of spatial regression models to functional space. We commence by outlining a statistical methodology for functional data analysis, i.e. when observed data are generated as mathematical functions. It follows that a theoretical framework of spatial regression models must be adjusted to functional space. The estimation procedure consists of integrations since we deal with mathematical functions in the least squares criterion. Our key achievement is an improvement of the iterative algorithm which is used for estimation of the spatial autoregressive parameter. Specifically, in the iterative algorithm, we propose to minimise the generalised *cross-validation* score of estimated models that leads to a robust solution. The robustness of the solution is verified in the empirical analysis. To the best of our knowledge, there is no previous study of our proposed algorithm. In the empirical application, we use a new data source to propose an innovative econometric model that can explain the dynamics of the industrial production of districts in Slovakia. Data are electronic records of vehicles that pass through monitored sections of the road network, collected by the digital toll system, in Slovakia. The empirical results show that the proposed econometric model has a high potential to provide accurate estimates of the industrial production of districts in Slovakia.

Key words:

Functional data analysis, spatial regression models, generalised cross-validation score, iterative algorithm, digital toll system

ABSTRAKT

KNÍŽAT, Peter: Funkcionálne dáta v priestorových ekonometrických modeloch. - Ekonomická univerzita v Bratislave. Fakulta hospodárskej informatiky; Katedra operačného výskumu a ekonometrie. – Vedúci záverečnej práce: doc. Ing. Andrea Furková, PhD. – Bratislava: FHI EU, 2025, 133 strán

Cieľom dizertačnej práce je prínos do metodologickej a empirickej oblasti skúmanej tematiky. V metodologickej štúdii je naším cieľom demonštrovať rozšírenie priestorových regresných modelov do funkcionálneho priestoru. Začneme predstavením štatistickej metodológie pre analýzu funkcionálnych údajov, t. j. keď sú pozorované údaje generované ako matematické funkcie. Z toho vyplýva, že teoretický rámec priestorového regresného modelu musí byť prispôsobený funkcionálnemu priestoru. Postup odhadu modelu pozostáva z integrácií, pretože sa zaoberáme matematickými funkciami v kritériu najmenších štvorcov. Naším kľúčovým úspechom je vylepšenie iteratívneho algoritmu, ktorý sa používa pri odhade priestorového autoregresného parametra. Konkrétne v iteratívnom algoritme navrhujeme minimalizovať zovšeobecnené skóre *krížovej-validácie* odhadovaných modelov, ktorý vedie k robustnému riešeniu. Robustnosť riešenia je overená v empirickej analýze. Pokiaľ je nám známe, neexistuje žiadna predchádzajúca štúdia nášho navrhovaného algoritmu. V empirickej aplikácii použijeme nový dátový zdroj pre navrhnutie inovatívneho ekonometrického modelu, ktorý dokáže vysvetliť dynamiku priemyselnej produkcie v okresoch na Slovensku. Dáta sú elektronické záznamy o vozidlách, ktoré prechádzajú cez monitorované úseky cestnej siete, zbierané digitálnym mýtnym systémom na Slovensku. Empirické výsledky ukazujú, že navrhovaný ekonometrický model má vysoký potenciál poskytnúť presné odhady priemyselnej produkcie v okresoch na Slovensku.

Kľúčové slová:

Analýza funkcionálnych dát, priestorové regresné modely, zovšeobecnené skóre krížovej-validácie, iteratívny algoritmus, digitálny mýtny systém

Contents

List of abbreviations

ANOVA	Analysis of variance
BND	Boundary
CPI	Consumer Price Index
FDA	Functional Data Analysis
GCV	Generalised *cross-validation*
GLS	Generalised least squares
MSE	Mean squared error
NSI	National Statistical Institute
SAR	Spatial Autoregressive Model
SARAR	Spatial Autoregressive Model with SAR Disturbances
SD	Standard deviation
SDM	Spatial Durbin Model
SEM	Spatial Error Model
SLX	Spatial Linear Model with Lagged X
SOSR	Statistical Office of the Slovak Republic
SSE	Sum of squared errors
VAR	Variance

List of Figures

List of Tables

Introduction

The dissertation thesis conducts research of two main problems of the statistical application in econometrics. A methodological problem makes a study of the theoretical framework of spatial regression models and its estimation procedures expanded to functional spaces. Here, the main contribution is the improvement of the iterative procedure that is used for estimation of the model. Specifically, within the iterative procedure, we propose to minimise the generalised *cross-validation* score of estimated models, which ensures that a robust estimate of the spatial autoregressive parameter can be obtained. To the best of our knowledge, there is no previous study of our proposed algorithm. An empirical problem makes a study of new data source in official statistics. Data are electronic records of vehicles that pass through monitored sections of the road network, which are automatically collected by the digital toll system, in Slovakia. This new data source has a potential to provide analytical information on economic performance, which can be used to shape a decision making by policymakers. Specifically, we propose a spatial econometric model expanded to functional space, which uses a weekly freight intensity to explain the annual industrial production of districts in Slovakia.

Nowadays, the advancements in information technology allows for collecting a large volume of *high-frequency* data. In some cases, data generation processes can be described by mathematical functions. A statistical theory which can be used for analysis of such data is called functional data analysis. The main difference to the classical *real-space* framework is that the observations are now in the form of mathematical objects. Initially, a *point-wise* observations are transformed onto mathematical functions, i.e. curves. If the curves are associated with spatial locations, a geographical dependence between observations must be taken into account. In the past decade, researchers and academics dealt with a problem of extending the theory of functional regressions for spatially dependent data. In this dissertation thesis, our aim is to study a theoretical framework of the spatial regression model, where the covariate matrix contains functional observations, i.e. curves, and the response is a vector of *real-valued* observations. The model assumes a spatially autocorrelated error term that requires an application of the spatial autoregressive model on the observed residuals. The main contribution of our research is that we

propose an improvement of the iterative procedure for estimation of the spatial autoregressive parameter. The proposed algorithm minimises the generalised *cross-validation* score of the estimated model that leads to a more robust solution than previously proposed algorithm by Schabenberger and Gotway (2005). In the empirical analysis, we demonstrate that the proposed iterative algorithm leads to a robust estimate of the spatial autoregressive parameter.

The empirical analysis commences with a description of data generation processes that amass electronic records from the digital toll system in Slovakia. A large volume of daily electronic records are of very *high-frequency*, with a size of monthly files approximately 6-8 gigabytes. The digital toll system is based on the satellite technology, which records a time stamp of vehicles that pass through monitored sections of roads. We propose to study an economic relationship between the intensity of the freight and the industrial production of districts in Slovakia. In the initial step, we aggregate the daily electronic records into a weekly intensity of the freight, i.e. a count of passing vehicles through monitored sections of roads, for each district. These are then transformed to mathematical functions, which form the functional observations, i.e. curves, of the functional covariate matrix. The response variable is a vector of *real-valued* observations of the industrial production for each district. The functional spatial econometric model assumes a spatially autocorrelated error term structure. The empirical application of the proposed functional spatial econometric model is carried out in R.

We believe that the dissertation thesis stems further research in the field of functional spatial regression models and its applicability to *real-world* problems.

The dissertation thesis is organised as follows. Section 1 outlines the current state of the research problem, both abroad and domestically. Section 2 provides an overview of the main and partial objectives of the dissertation thesis. Section 3 discusses a theory of functional data analysis. It shows a methodological framework of functional regression models and its estimation procedures, which is followed with an extension to include spatial dependencies between observations. Section 4 gives details of the new data source and its aggregation. The empirical analysis demonstrates the application of functional regression models shown in Section 3. Sections Discussion and Conclusion summarise the results and achievements of the dissertation thesis, discuss the shortcomings of data and the proposed model, and outline further research in the field.

1 The-state-of-the-art

The problem of analysing functions in statistics is a relatively recent phenomenon that has been extensively investigated over the past few decades. The functionality in observed data, i.e., when data are generated through a process naturally described as functional, occurs in many areas of sciences. In the field of applied economics, the *real-valued* observations that stem from the functional process are usually generated across time periods.

In the first step the *real-valued* observations are transformed to mathematical functions, i.e. curves. The mathematical functions are imposed on the *real-valued* observations by the process of smoothing or interpolation. The choice depends on the initial strategy of data representations. If the observation values are assumed to be errorless, then it is an interpolation process, but if they contain some observational error that requires removing, then the transformation from *real-valued* data to mathematical functions may involve smoothing.

Various types of mathematical functions can be used for the expansion of the original *real-valued* data. For example, a Fourier basis system is most suitable for the periodic data that do not exhibit too rapid fluctuation in any given interval. The other data fitting techniques that are computationally simple to implement are the kernel smoothing and the local polynomial fitting. However, these methods could lead to suboptimal solutions with respect to minimising the total sum of squared errors (Ramsay and Silverman, 2005). A more flexible basis system that is used for the functional transformation of noisy data is the expansion by basis splines, also called B-splines. The basis spline system combines piece-wise polynomials that are joined at the predefined breakpoints. The breakpoints are used to divide the original data into usually equally spaced intervals. A linear combination of these polynomials forms a basis spline function. The code for working with basis splines is widely available in various programming languages, including R. Theoretical properties of basis splines are discussed by de Boor (2001). Its disadvantage is that it is not straightforward how to implicitly determine the degree of smoothing, which is controlled by a number of breakpoints, i.e. a number of basis splines. A large number of basis splines could cause the instable estimates at boundary intervals and, subsequently, in their derivatives that are needed in further statistical analysis of functional data. Nevertheless, these limitations can be circumvented by adding a roughness penalty within the method of least squares for estimation of parameters of basis splines (Ramsay and Silverman,

2005). The other advantages of using the roughness penalty in the least squares estimator for data approximation is provided by (Green and Silverman, 1994). In the empirical analysis of the dissertation thesis, we use the expansion by basis splines for transformation of the original *real-valued* observations to mathematical functions.

Let's assume that we observe a matrix of real numbers $\mathbf{X} = x_{ij}$ representing a p-dimensional *column-wise* matrix, where $j = 1, \dots, p$ refers to a number of observed variables, which are recorded for observations, $i = 1, \dots, n$. A notation of the transformed matrix of *real-valued* observations into functional observations can be defined as

$$\mathbf{X} \rightarrow \chi_i(t) \tag{1.1}$$

where $\chi_i(t)$ is the matrix of mathematical functions, i.e. observed curves for $i = 1, \dots, n$, which are now observable at any argument t on a continuous domain $t \in T$, where T refers to an interval of functions[1]. The observed curves can now be discretised at any point t. In Section 3, we describe an analytical procedure for transformation of the *real-valued* observations to functional observations. The theoretical properties of functional spaces are provided by Ramsay and Silverman (2005).

It follows that the matrix of observed curves serves as an input for further statistical analyses. In the following section, we describe further statistical analyses that tackle the problem of identifying variations and dependencies within and between functional data.

1.1 Functional data analysis

The basic descriptive statistics of functional data, which explore the inherent variability in the observed curves, are means, standard deviations and covariance structures. The main difference to their classical counterparts is that all of these are functional, i.e. observed on a continuous domain $t \in T$, rather than *point-wise* estimates. The functional mean of the observed curves is estimated as

[1] The interval of functions is bounded by the number of observed variables 1 and p, respectively.

$$\bar{\chi}(t) = \frac{1}{n}\sum_{i=1}^{n} \chi_i(t) \tag{1.2}$$

where the summation is *point-wise* across the observed curves. It follows that the functional variance of observed curves[2] is estimated as

$$var_{\chi}(t) = \frac{1}{n-1}\sum_{i=1}^{n} [\chi_i(t) - \bar{\chi}(t)]^2 \tag{1.3}$$

The square root of eq. (1.3) is the functional standard deviation. Similarly, we can estimate functional covariance and correlation structures, which can be either within or between the observed curves (Ramsay and Silverman, 2005). The functional observations that are used in the empirical analysis of the dissertation thesis are naturally correlated, both within and between observed curves, since the original observations are weekly time series across geographical locations.

The preliminary statistics of functional data play a pivotal role as starting points to further applications of classical statistical techniques of regression models and analysis of variance which both investigate the way in which variability in observed functional data can be accounted for by other known or observed functions. The *state-of-the-art* manuscript for functional data analysis, which covers parametric regression models, is Ramsay and Silverman (2005).

The analysis of variance extended to functional space is a statistical method for identifying the variability within and between the categorised observed curves, i.e. each observed curve falls within the specific category $c = 1, \dots, C$. It takes the equivalent regression formulation (Ramsay and Silverman, 2005)

$$\chi_{ic}(t) = \sum_{j=1}^{C+1} \beta_j(t) z_{(ic)j} + \varepsilon_{ic}(t) \tag{1.4}$$

where $\mathbf{Z} = z_{(ic)j}$ are the binary values of the $n \times (C+1)$ design matrix, i.e. the values of the design matrix are 0's or 1's that code the category of the observed curves. The first column

[2] Note that we deal with a sample of observed curves.

takes the values of 1 that identifies the overall mean function across all observed curves. $\beta_j(t)$ are the unknown functional regression parameters that need to be estimated. Since the *left-hand* side of eq. (1.4) is functional, the components of the error term $\varepsilon_{ic}(t)$ are also of the functional form. The model in eq. (1.4) assumes that $\varepsilon_{ic}(t)$ are independently and identically distributed.

The regression analysis is a set of statistical methods for estimating the relationships between the response and the covariate(s). In the context of functional data analysis, either one or both of the response and the covariate can take the functional form. The following notation defines a functional regression model with a *real-valued* response and the functional covariate (Ramsay and Silverman, 2005)

$$y_i = \alpha + \int \beta(t)\chi_i(t)dt + \varepsilon_i \tag{1.5}$$

where $\mathbf{y} = y_i$ is the vector of *real-valued* responses corresponding to each curve $\chi_i(t)$ and the integral $\int dt$ is computed across the continuous domain $t \in T$. The error term $\boldsymbol{\varepsilon} = \varepsilon_i$ is a *real-valued* vector of residuals, which are assumed independently and identically distributed. The unknown functional regression parameter $\beta(t)$, which needs to be estimated, is observable on the continuous domain $t \in T$. Its advantage is that it uncovers local variations and patterns on the entire domain of the observed functional covariate. The statistical significance of $\beta(t)$ is inspected visually, i.e. if the values of $\beta(t)$ are close to the *x*-axis at $y = 0$, the values of $\chi_i(t)$ have a small effect on the values of y_i. The computational procedure for estimation of both functional regression models in eq. (1.4) and eq. (1.5) is shown in Section 3.

The other interests may lay in predictions of new functions based on a fitted functional model where the nonparametric techniques commonly provide the most accurate predictive options. The *state-of-the-art* book for functional data analysis, which covers nonparametric regression models, is Ferraty and Vieu (2006). The other possibility to explore the variability in observed functional data, or its prediction, is to use a *non-frequentist* perspective which is based on the distributional assumption, i.e. a Bayesian regression approach which can be deployed in both functional parametric and nonparametric modelling. A theoretical concept of these methods is discussed in the book by Shi and Choi (2011).

1.2 Spatial regression models

A theory of spatial dependencies between observations in the context of the classical regression analysis is a well-developed statistical field. In the spatial scenario, we assume that the observations of economic phenomena are associated with geographical locations. It follows that in the *open-market* economy, the economic *spill-over* effect causes a different degree of dependence between locations, i.e. the observations with closer proximity of the geographical contiguity have higher spatial dependency. If the geographical location is measured by a lattice structure, i.e. the longitude and the latitude are the centroid of the geographical location, the spatial dependency can be obtained by measuring a distance or a connectivity between neighbourhoods. The simplest way is to assign a binary weight to neighbourhoods that share the boundaries (LeSage, 2009)

$$w_{ij} = \begin{cases} 1, & bnd(i) \cap bnd(j) \neq \emptyset \\ 0, & bnd(i) \cap bnd(j) = \emptyset \end{cases} \tag{1.6}$$

where w_{ij} are the values of the $n \times n$ weight matrix $\mathbf{W} = w_{ij}$, $bnd(\cdot)$ refers to a boundary between different geographical locations i and j, i.e. if the locations do not share the boundary, w_{ij} takes a value of 0, otherwise it is assigned a value of 1. We note that if $i = j$, w_{ij} also takes a value of 0. The method in eq. (1.6) consists of a set of estimation methods of spatial weights between geographical locations. Figure 1 shows the possible scenarios of neighbourhood contiguity when using the method in eq. (1.6)

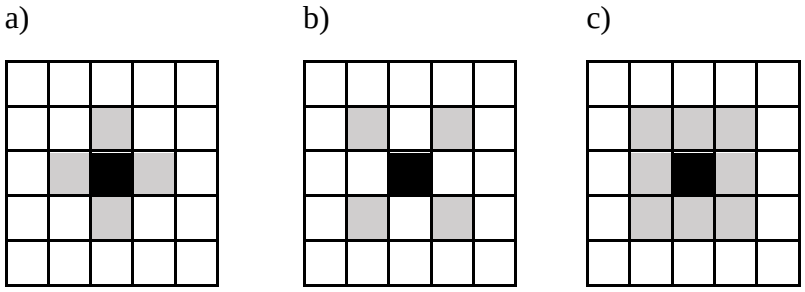

Figure 1. Possible scenarios of geographical contiguity
Source: Author's construction

Figure 1 shows that we can obtain three possible scenarios of neighbourhood contiguity, which have the following definition: a) rook, b) bishop and c) queen. The most commonly used for economic phenomena is the *queen-type* estimation method. We note that Figure 1 shows a

symmetrical distribution of weights. For geographical locations, the symmetry of weights is not a requirement, i.e. geographical locations can have a different number of neighbours. The other possibility is to use a *distance-based* measure approach, where the contiguity is determined based on the distance between centroids rather than just shared boundaries. Section 3 shows the estimation of w_{ij} by the generalised *queen-type* method, which we use for estimation of spatial weights w_{ij} in the empirical analysis.

Moreover, in the empirical analysis, the spatial weights are not used in their binary form, but are subject to a standardisation, i.e. most commonly a *row-standardisation* (Anselin and Rey, 2014)

$$w_{ij} = \frac{w_{ij}}{\sum_{j=1}^{n} w_{ij}} \tag{1.7}$$

As a result of eq. (1.7), each row sum of w_{ij} equals 1. The further exploratory analysis of spatial observations must take an account of spatial dependencies. In the context of the regression analysis, the spatial dependencies are incorporated within the theoretical framework of the regression model by using the estimated values of the matrix $\boldsymbol{W} = w_{ij}$. The spatial regression models differ in the way to which the spatial dependence is incorporated within its theoretical framework.

We commence with defining a theoretical framework of the classical regression model (Anderson, 2003)

$$y_i = \alpha + \sum_{j=1}^{p} \beta_j x_{ij} + \varepsilon_i \tag{1.8}$$

where $\mathbf{y} = y_i$ is the $1 \times n$ vector of the response values and $\mathbf{X} = x_{ij}$ are the values of the $n \times p$ matrix of independent variables, where $i = 1, \dots, n$ refers to the number of observations and $j = 1, \dots, p$ is the number of independent variables. The unknown regression parameters α and $\boldsymbol{\beta} = \beta_j$ remain to be estimated. In eq. (1.8), we assume that the values in the error term $\boldsymbol{\varepsilon} = \varepsilon_i$ are identically and independently distributed.

In time series of the economic phenomenon, the presence of autocorrelations is modelled by autoregressive models, which are constructed as a linear combination of past values. In the framework of spatial regressions, the spatial autocorrelation can be imposed on the vector of residual errors by using the spatial dependencies defined by the values w_{ij} (Schabenberger and Gotway, 2005)

$$\varepsilon_i = \rho \sum_{j=1}^{n} w_{ij}\varepsilon_j + v_i \tag{1.9}$$

where $\mathbf{W\varepsilon} = \sum_{j=1}^{n} w_{ij}\varepsilon_j$ is called the spatially lagged error term and ρ refers to the spatial autoregressive parameter, which shows the strength of the spatial dependency between observations. In eq. (1.9), we impose the spatial autocorrelation structure by regressing ε_i on all other values ε_j, where the values of w_{ij} are *non-zero*. We also note that the diagonal values of w_{ij} are zero so ε_i are not regressed on themselves. The values in the error term $\mathbf{v} = v_i$ are assumed to be independently and identically distributed. Rearranging eq. (1.9) and substituting into eq. (1.8), we obtain the following spatial autoregressive model formalised in the matrix notation

$$\mathbf{y} = \alpha + \mathbf{X\beta} + (\mathbf{I} - \rho\mathbf{W})^{-1}\mathbf{v} \tag{1.10}$$

where $\mathbf{I}$ is the $n \times n$ identity matrix. After some rearrangements of eq. (1.10), we can obtain different spatial autoregressive models (Schabenberger and Gotway, 2005)

$$\mathbf{y} = \alpha + \mathbf{X\beta} - \rho\mathbf{WX\beta} + \rho\mathbf{Wy} + \mathbf{v} \tag{1.11}$$

where $\mathbf{WX\beta} = \sum_{j=1}^{p} \beta_j \sum_{j=1}^{n} w_{ij}x_{ij}$ and $\mathbf{Wy} = \sum_{j=1}^{n} w_{ij}y_j$, including $\mathbf{W\varepsilon} = \sum_{j=1}^{n} w_{ij}\varepsilon_j$ in eq. (1.9), are called spatially lagged variables. The model in eq. (1.11) assumes that the spatial autoregressive parameter ρ exists and that it can take a different value for each lagged variable. The model in eq. (1.11) is also called spatial autoregressive model and it was first introduced by Whittle (1954). The simplest version of the spatial autoregressive model is to assume a one-parameter model for ρ, i.e. where we only keep the lagged term of $\sum_{j=1}^{n} w_{ij}y_j$ on the *right-hand* side of eq. (1.11). Section 3 introduces a different lagged structure of the spatial autoregressive model.

For a *well-defined* model, the term $(\mathbf{I} - \rho\mathbf{W})$ in eq. (1.10) must be invertible, which implies that some conditions on $\mathbf{W}$ must be imposed. One of the conditions is that $\mathbf{W}$ must be *row-standardised* as shown in eq. (1.7). Additionally, the restriction is placed on the spatial parameter ρ, Haining (1993)

$$\frac{1}{\lambda_{min}} \leq \rho < \frac{1}{\lambda_{max}} \tag{1.12}$$

where λ_{min} and λ_{max} are the smallest and the largest eigenvalues of $\mathbf{W} = w_{ij}$, which can be obtained through the familiar technique of the singular value decomposition. Nevertheless, for the *row-standardised* $\mathbf{W}$, $\lambda_{min} \leq -1$ and $\lambda_{max} = 1$, so that $\rho < 1$ but could be less than -1 (Haining, 1993). The estimation technique of the spatial autoregressive model is shown in Section 3.

A conceptual framework of the spatial econometrics, and its corresponding methodology, has been introduced and extensively researched since the second half of the 20th century. It has gradually progressed from methods concerned with measuring similarity of data objects and its proximity of distances (e.g., Whittle, 1954) to introducing a spatial autocorrelation in the methodological framework (e.g., Cliff and Ord, 1973). A need to address the economical and statistical problem of dependence between locations in the econometric models lead to advances in the spatial regression framework. A new phenomenon of the spatial econometrics has been introduced by Paelnick (1974). A monography by Paelinck and Klaasen (1979), was a first publication concerned with introducing new concepts for spatial econometric modelling. A modelling approach, in the complex form that defined a spatial dependence and correlation, has been formalised in the manuscript by Anselin (1988). In the current research, the spatial econometrics is adapted within both a theoretical framework of the econometrics and in the area of the empirical analysis. An important breakthrough in the history of the spatial econometrics (Furková, 2017) has been an introduction of the theory called a new economic geography (NEG), which has begun with works of, e.g., Krugman (1991). The NEG models define a framework for spatial analysis of economic data within research problems such as regional convergence, regional concentration of economic activities and dynamics of adaptation. Moreover, a lot of research attention is given to the spatial autoregressive models, i.e. models that take into account

a spatial dependence by including lagged dependent variables in the model. The Spatial Autoregressive (SAR) model falls into this category (Cliff and Ord, 1973; Cliff and Ord, 1981).

The textbook by LeSage and Pace (2009) provides a *state-of-the-art* introduction to econometric analysis of spatially dependent data. It is intended for students and researchers who are interested to learn the application of econometric modelling to spatial data. The book by Schabenberger and Gotway (2022) covers the most recent statistical methods in the analysis of spatial data structures. Its intention is to provide the *model-based* applied research tools that can serve to solve *real-world* problems. Its frequentist approach leads to showcasing different procedures for estimation of spatial parameters of linear and generalised linear spatial models. It also covers a theory of *spatio-temporal* models that are used to estimate parameters across both space and time domain, which would be mainly of interest to PhD students and academic researchers.

Domestically, Furková (2017) provides a comprehensive review of the econometric models and its estimation techniques for spatial regressions in her habilitation thesis. Her habilitation thesis is the main source of the review of the current state of research, literature sources and methodological advances of the spatial regressions in the econometrics for this dissertation thesis. Furková (2017) mentions that, historically, an econometric theory had been dominated by researching problems of a time dependence, and there was not enough focus dedicated to a spatial dependence. Moreover, Furková (2013), in her separate two publications, provides a methodological analysis of the SAR model and analysis of the spatial dependence in econometric models, respectively. The paper by Furková and Chocholatá (2016) provides options of the software used for spatial data analysis. Additionally, the paper by Furková and Knížat (2024), of which the abridged version was presented at the 17th International Symposium on Operations Research in Bled (Slovenia) 2023, considers the spatial effects and nonlinearities on semiparametric spatial autoregressive models. It focuses on the empirical application of semiparametric spatial autoregressive econometric models to explain the dynamics of the European Union's regional unemployment.

1.3 Extension to functional space

The extension of spatial regression models to functional spaces has also been widely studied in the recent decade. Its main challenges are caused by coding computational strategies for estimation procedures of too many unknown parameters that are defined in the theoretical framework of the models. No textbook has been written so far, which would provide a comprehensive overview of the theory and the application of functional spatial regression models.

The textbook by Horváth and Kokoszka (2012) provides an introduction to analysis of functional data in the context of spatially dependent observations. It concludes that in the presence of these dependencies the mean function is not simply an unweighted arithmetic average of the observed curves. It proposes three different methods for estimation of the mean function for functional data observed at spatial locations. The fundamental principle of the functional spatial statistics is to consider that the observed functions, or curves, at *close-by* locations contribute similar information, and hence these get smaller weights than observed curves at sparse locations in the estimation of the overall mean function. The most recent advances in functional spatial statistics, accompanied by the corresponding papers, are briefly described by Aneiros et al. (2019). The paper states that developing statistical methodology for spatial functional data is a challenging topic both from a mathematical and conceptual framework. The paper is included in the special issue of Journal of Multivariate Analysis in 2019 that is dedicated to spatial statistical methodologies, which also contains two additional papers that concern spatial functional regressions: one on modelling spatially dependent functional data with partial differential equation structure (Arnone et al., 2019) and the other on spatial functional principal component analysis with an empirical application to imaging data (Lui et al., 2017). The extension of tests for the spatial autocorrelation and clustering based on a nonparametric variogram to the functional space is proposed by Giraldo et al. (2018) and Romano et al. (2011), respectively.

The spatial autoregressive model with the spatially autocorrelated errors, extended to functional space, is studied by Pineda-Ríos et al. (2009). The authors propose a theoretical framework and the estimation procedure for functional spatial autoregressive model, where the response is *real-valued* and the covariate matrix contains curves. One of its proposed estimation

technique is to use the generalised method of least squares. The drawback of the method is that it uses an iterative algorithm for estimation of the spatial autoregressive parameter, which could lead to relatively large estimates of ρ, or to a *non-convergence* issues due to the *over-parametrisation* of the model. One of the main achievements of this dissertation thesis is that we propose an adjustment of the iterative algorithm such that it leads to a robust estimate of the spatial autoregressive parameter. We discuss the details of the estimation in Section 3.

The focus of the empirical research in the dissertation thesis is the extension, and its application, of the spatial autoregressive model with spatially autocorrelated errors to functional space. It can be formalised as

$$y_i = \alpha + \int \beta(t)\chi_i(t)dt + \varepsilon_i \tag{1.13}$$

where the integration $\int dt$ is taken across the continuous domain $t \in T$, the unknown regression parameter $\beta(t)$ and the values in the covariate matrix $\chi_i(t)$ are functional. The values in the response y_i and the residual values ε_i are both *real-valued*, with $\varepsilon_i = \rho \sum_{j=1}^{n} w_{ij}\varepsilon_j + v_i$ as defined in eq. (1.9). The solution of the model in eq. (1.13) for unknown parameters $\beta(t)$ and ρ leads to a system of equations. The system of equations is then solved iteratively. The estimation procedure is shown in Section 3.

In the recent publication, Římalová et al. (2022), shows an estimation of the fully functional spatial regression model with spatially correlated errors through the iterative procedure using the generalised method of least squares. We can formalise a fully functional spatial autoregressive model as

$$\Upsilon_i(t) = \alpha(t) + \int \beta(t)\chi_i(t)dt + \varepsilon_i(t) \tag{1.14}$$

where the response $\Upsilon_i(t)$, the intercept $\alpha(t)$ and the error term $\varepsilon_i(t)$ are now of the functional form, noting that $\varepsilon_i(t)$ are spatially autocorrelated. We note that the *right-hand* side of eq. (1.14) can also include other functional or *real-valued* covariates (Římalová et al., 2022). It is evident that the estimation procedure of the model in eq. (1.14) contains too many unknown parameters, which bring both analytical and computational complexities.

Moreover, the paper by Římalová et al. (2022) proposes a permutation method for testing the statistical significance of the functional regression parameters. The paper concludes that the performance of the estimation procedure on the simulated dataset performs well when the range of spatial correlation is not too high. For larger ranges the procedure requires a larger sample size. The empirical application is carried out on a dataset that includes yearly measurements of per capita amount of the municipal waste from 1997 to 2011 in the 49 cities of the Venice province that are regressed on spatial coordinates of town centres and the additional functional covariate of tourism indicator. The results of the model show that the tourism effect is significant in explaining the municipal waste by regions, unlike the spatial covariates (longitude and latitude).

1.4 New data source: toll data

In the empirical study of the dissertation thesis, the focus is on analysis of the new data source for use in official statistics. We analyse data that are collected by the digital toll system in Slovakia. The research on using toll data for nowcasting[3] macroeconomic indicators, transport statistics or identifying economic cycles is usually conducted by National Statistical Institutes (NSIs) since its unavailability to public. It is one of the projects on the agenda of the innovation portfolio 2023-2025 for development of *short-term* business statistics of the European Statistical System Committee[4]. NSIs usually cooperate with research communities and academics that can tackle the complex statistical problems, which crop up when dealing with various data structures.

Statistical Office of the Slovak Republic organised and chaired a panel discussion on the international statistical conference in Mexico City, which was held under the auspices of International Association for Official Statistics and International Statistical Institute[5]. The presidents of NSIs of Austria, Hungary, Ireland and Slovakia presented the most recent innovative projects that are related to new data sources in official statistics. The presidents agreed that one of the main challenges is that if NSIs do not follow the trend in innovations and

[3] The word nowcasting emerged as a combination of "now" and "forecasting" and it is used in macroeconomics for predicting the present, the recent past and the near future (Stock et al., 2016).

[4] ESS innovation agenda 2023 portfolio, June 2023 (Eurostat internal document).

[5] https://www.isi-next.org/proposals/proposal/563/detail/.

the use of new data sources in official statistics, the private institutions could take a role of the leader in producing new reliable statistics. Hence, NSIs must undergo transformations and become more agile by researching and implementing new data sources into official and experimental statistics.

Moreover, in the separate panel session, Statistical Office of the Slovak Republic presented the most recent research on using toll data for nowcasting the Industrial Production Index[6]. The research study is approved for publication in the special issue of Statistical Journal of the IAOS (Knížat et al., 2025). The paper shows a comparison of the *truck-passage* index, which is estimated by using different price index formulae, with the Industrial Production Index in Slovakia. The use of price index formulae for estimation of the *truck-passage* index is inspired by the author's previous research on theory and applications of price indices (Knížat, 2023a; Knížat, 2023b; Knížat, P. and Glaser-Opitzová, H., 2023 and Knížat et al., 2024). Additionally, the paper shows a decomposition of time series of indices through empirical mode decomposition to study its cyclicality component. It concludes that the *truck-passage* index has a potential to serve for early nowcasting and the identification of economic cycles of the industrial production output. In general, the authors state that "*a new data source offers an opportunity for further research in the area of transport statistics, short-term business statistics and road control management that can be useful for policymakers in various government sectors*".

In recent years, other European countries conducted various research on using toll data. For example, the German federal statistical office has been widely analysing toll data for the use in nowcasting *short-term* statistics or for the identification of economic cycles. The corresponding paper by Cox et al. (2018) states that there is a close relationship between the economic activity and the freight traffic by trucks since the economic activity requires transport services. The authors propose an index, called *truck-toll-mileage* index, which represents a monthly change of the freight transport through German highways. It describes a collection of toll data, its processing and a construction of the *fixed-base truck-toll-mileage* index. A strong relationship is shown between the *truck-toll-mileage* index and individual components of the German industrial production output. It further decomposes time series of the *truck-toll-mileage* index into trend, cycle and seasonal components and compare the results with the German

[6] https://www.isi-next.org/abstracts/submission/1956/view/

production indices. It concludes that toll data hold high information value and a clear correlation with domestic trade, certain service sectors and the overall economy is evident. Further research is required whether toll data can be used for official statistics that would lead to an improvement in timeliness and lower the burden on respondents when collecting data through sample surveys.

Domestically, the first study of processing and analysing toll data and its application for nowcasting the gross domestic product that uses the multivariate *seasonally-adjusted* integrated autoregressive and moving average model is carried out by Knížat et al. (2024). The paper shows that toll data have a high potential to capture the overall macroeconomic trend, but time series are too short in order to draw any robust conclusions. The authors suggest implementing the model for experimental statistics and to monitor its performance. The other research study by Knížat (2024), which is published in the conference proceedings, shows the application of the functional analysis of variance of the regional freight. The main objective of the paper is to explore the co-variability of the intensity of the freight between and within regions in Slovakia across time. The paper states that the advantage of the functional analysis of variance compared to its classical counterpart is that the estimated parameters are continuous that allows for identification of the intensity of the freight across time. The author concludes that there are no common seasonal patterns of the intensity of the freight between regions and that the overall mean function is different from regional profiles. The results can be used to shape transport policies for more efficient management of the road transportation between regions.

Section 4 discusses the data generation process and the aggregation of electronic records from the digital toll system, which is consequently used in the empirical analysis of the dissertation thesis.

Table 1 shows a summary of the main literature, including a brief description of its content, which is used for research in the dissertation thesis. The author’s publications are also listed.

Name	Author(s)	Methodology
Functional Data Analysis. Second Edition. Springer Series in Statistics, Springer, 2005.	Ramsay, J. O. and Silverman, B. W.	The *state-of-the-art* textbook provides a *ground-breaking* foundation to functional data analysis.
Functional Data Analysis with R and MATLAB. Use R!, Springer, 2009.	Ramsay, J. O., Hooker G., Graves, S.	The textbook provides an application of functional data analysis in R and MATLAB.
Introduction to Spatial Econometrics. Taylor and Francis Group LLC, CRC Press, 2009.	LeSage, J. and Pace, K. R.	The textbook provides a *state-of-the-art* foundation to spatial econometrics along with numerous applied examples. It is intended as a text for students and researchers interested in learning about spatial regression models.
Statistical Methods for Spatial Data Analysis. Taylor and Francis Group LLC, CRC Press, 2005.	Schabenberger, O. and Gotway, C. A.	The textbook provides a comprehensive overview of statistical methods used for analysis of spatially dependent data, including computational strategies for different estimation procedures.
Modely a metódy priestorovej ekonometrie. Habilitačná práca, Ekonomická univerzita v Bratislave, Fakulta hospodárskej informatiky, 2017.	Furková, A.	The habilitation thesis provide a thorough overview of the theory and application of spatial regression models, including different estimation methods (in Slovak).

New data sources in statistics: the effect of freight intensity on macroeconomic indicators. Slovak statistics and demography, Vol. 34, no. 1, 2024, p. 7–26.	Knížat, P., Ceľuchová Bošanská, D., Janík, M., Nguyen, F.	The paper shows a first attempt to process and analyse toll data in Slovakia. It applies a multivariate time series model to assess a relationship between the freight and the gross domestic product (in Slovak).
Regional Intensity of the Freight: Functional Analysis of Variance. In: Mathematical Methods in Economics 2024. The 42nd International Conference on Mathematical Methods in Economics, Proceedings, Praha. The Czech Society for Operations Research, 2024.	Knížat, P.	The conference paper that was presented at the international conference on Mathematical Methods in Economics in Ústí nad Labem, Czechia, 2024. The content of the paper demonstrates a study of the application of functional analysis of variance of the regional freight in Slovakia.
Nowcasting industrial production index with high-frequency toll data. Statistical Journal of the IAOS. Vol. 41, no. 1, 2025, p. 102–111. https://doi.org/10.1177/18747655241307540	Knížat, P., Furková, A., Glaser-Opitzová, H., Peťko, P.	The paper proposes to use a theory of price indices for estimation of different *truck-passage* indices. The *truck-passage* index is then used for nowcasting the Industrial Production Index in Slovakia.

Multilateral indices in official price statistics and a new additive splicing method. Quality and Quantity Journal, Vol. 58, no. 5, 2024, p. 4207–4222. https://doi.org/10.1007/s11135-024-01848-3.	Knížat, P., Glaser-Opitzová, H., Furková, A., Vojtková, M.	The paper demonstrates a *deep-dive* analysis of multilateral price indices, particularly, it shows an application of the GEKS method. It proposes a new method of splicing, which is used for extension of multilateral price indices. For empirical analysis, it uses *web-scraped* online prices of particular products in Slovakia.
Web scraped data in consumer price indices. Statistical Journal of the IAOS, Vol. 39, no. 1, 2023, p. 203–212. https://doi.org/10.3233/SJI-220115.	Knížat, P.	The paper shows the application of bilateral and multilateral price indices on *web-scraped* online prices of particular products in Slovakia.
Consumer price index from web-scraped data: analysis of specific product category. Slovak Statistics and Demography, Vol. 33, no. 1, 2023, p. 37–49.	Knížat, P. and Glaser-Opitzová, H.	The paper shows the empirical analysis of different price index formulae on web scraped online prices in the context of the Slovak market (in Slovak).
Hedonic consumer price index: diagnostics and analysis of variance. Slovak Statistics and Demography, Vol. 33, no. 3, 2023, p. 21–38.	Knížat, P.	The paper conducts a *deep-dive* analysis of the hedonic consumer price index.

Beyond Parametric Bounds: Exploring Regional Unemployment Patterns Using Semiparametric Spatial Autoregression. Business Systems Research, Vol. 15, no. 2, 2024, p. 48-66. https://doi.org/10.2478/bsrj-2024-0017.	Furková, A. and Knížat, P.	The paper demonstrates the application of the *semi-parametric* nonlinear approach to spatial autoregressive models. The empirical analysis attempts to explain the dynamics of the EU's regional unemployment.

Table 1. The main literature used for research in the dissertation thesis, including the author's publications
Source: Author's construction

2 Objectives

In this section we outline the main and partial objectives of the dissertation thesis as well as its achievements. Generally, we set two key goals of the research: methodological and empirical, respectively.

The *methodological objective* aims to demonstrate the extension of spatial regression models to functional space. The main difference of functional spatial regression models compared to its classical counterparts is that the regression parameters are now functional, i.e. they are observed on the continuous domain. Its advantage is that the user can evaluate the estimated functional regression parameter on the entire domain of the original *real-valued* variables, not only as a *point-wise* estimate. The computational strategy of the method of least squares is modified since we deal with mathematical functions in the least squares criterion. Our main focus is on the design of the functional spatial regression framework with the spatially correlated error term. We consider the generalised method of least squares for estimation of the functional regression parameter and the spatial autoregressive parameter. Our main methodological aim is the enhancement of the iterative algorithm for estimation of the spatial autoregressive parameter, which is used to solve the system of equations obtained by the generalised method of least squares. We propose its adjustment such that it leads to a more robust solution than its previous version, which is currently provided in the scientific literature. To the best of our knowledge, there is no previous study of our proposed algorithm. The details are discussed throughout the dissertation thesis.

The *objective of empirical research* is to conduct an analysis of new data source, which can be used for production of new official and experimental statistics. The new data source contains electronic records of vehicles that pass through monitored sections of the road network in Slovakia. It is automatically collected by the digital toll system. Our main focus is to develop an econometric model that assess the relationship between the intensity of the freight across time and the annual industrial production by districts in Slovakia. The econometric model uses a spatial regression framework with spatially correlated error term that is extended to functional space. The functional form of the model allows a study of the intensity of the freight across the entire time domain. We intend to address the economic hypothesis, which are set by the

empirical research, whether toll data have a potential to capture the dynamics of the industrial production of districts in Slovakia.

Moreover, we formulate a list of partial objectives of the dissertation thesis. The list also contains study objectives that are deemed to be a standard part of the doctoral studies:

- Literature and scientific research: we conduct a comprehensive study of the most recent literature and scientific research that is relevant for researched field of the dissertation thesis. We intend to identify the areas for potential improvement.
- Functional data analysis: we commence the analysis by introducing the foundation to applied functional data analysis, i.e. a transformation of *real-valued* observations to mathematical functions.
- Spatial regression models and extension to functional space: we commence the analysis by introducing the classical regression framework ,which is followed by its extension to functional space. Furthermore, a theoretical framework of functional regression models is modified to accommodate dependencies between spatial observations. We show a set of models that falls under functional spatial autoregressive models.
- Empirical analysis and interpretation of results: one of the main challenges of empirical analysis is the interpretation of estimated functional parameters. In the results, we attempt to explain the estimated functional parameters in the context of applied economics and its usefulness for policymakers in various government sectors.

The dissertation thesis opens up the opportunities for further research in the field of functional spatial regression models and the use of new data sources for official and experimental statistics. Particularly, there are ample opportunities to develop functional extensions of more complex spatial regression models, including the computational strategies for its estimation, and the use of electronically generated data to produce more timely statistics that can be used for early decision making to shape economic policies, respectively.

3 Methodology and methods

This section describes a theoretical framework and its corresponding estimation procedure of functional regression models. The focus of our research is on regression models, where observations are recorded at geographical locations. We commence by outlining the methodology of transforming *real-valued* observations to mathematical functions. We then follow with a definition of functional regression models and its extension to spatially dependent data.

Section 3 is organised as follows. In 3.1, we show a transformation procedure of the *real-valued* space onto functional space, which is carried out by using the expansion by basis splines. In 3.2, we define a theoretical framework of functional parametric regression models and its computational strategies by the modified method of least squares. In 3.3, we extend the model to spatially dependent observations, where our main achievement is that we propose to modify the iterative algorithm for estimation of the spatial autoregressive parameter by minimising the *cross-validation* score of estimated models. Our proposed procedure leads to a more robust solution than the previously proposed algorithm, which is verified in Section 4 Results.

3.1 Transformation from real space to functional

We recall that we observe a matrix of real numbers $\mathbf{X} = x_{ij}$ that represents a p-dimensional *column-wise* matrix, $j = 1, \dots, p$, recorded for sample observations, $i = 1, \dots, n$. We assume that the observed dataset comes from a functional process. Here, we demonstrate the transformation procedure of the *real-valued* observations to mathematical functions. We note that a function is fitted onto a matrix of real numbers x_{ij} across each sample observation i.

Without a loss of generality, we consider the first sample observation $i = 1$. It follows that the first observation curve $\chi_1(t)$ can be expressed in terms of the expansion by basis splines (Ramsay and Silverman, 2005)

$$\chi_1(t) = \sum_{k=1}^{K} c_{1k} B_{k,m}(t, \tau_l) \tag{3.1}$$

where $B_{k,m}(t,\tau_l)$ is a $k \times 1$ vector of piecewise polynomials of order m that are automatically tied together at breakpoints τ_l. The order m refers to a number of constants that define the polynomial, i.e. one more than its degree. The breakpoints τ_l, $l = 1, \dots, L-1$ are values at L subintervals over which a function is approximated. The parameter $k = 1, \dots, K$ refers to a number of polynomials, where $K = m + L - 1$, which are used in the expansion. The parameter K determines a degree of the smoothness of original data. The $\mathbf{c}_1 = c_{1k}$ is a vector of unknown parameters that needs to be estimated. To simplify the notation, we set $B_{k,m}(t,\tau_l) = B_k(t)$.

In the empirical analysis, we select the parameter K at its maximum, i.e. a breakpoint is set at each observed variable, which ensures a minimal loss of information of original data. Hence, a degree of the smoothness in eq. (3.1) is determined and controlled by adding the following roughness penalty (Ramsay and Silverman, 2005)

$$RP = \lambda \int D^j B_k(t) D^j B_k(t) dt \tag{3.2}$$

where D^j is j^{th} the derivative of the polynomial $B_k(t)$ and λ is the smoothing parameter that controls the level of the smoothness. The order of the derivative depends on the observed data structure and various orders should be empirically tested for the optimal degree of the smoothness. The smoothing parameter λ can be estimated through a *cross-validation* method, or it can be set at a relatively low value to minimise a loss of information of original data (Ramsay and Silverman, 2005).

The application of smoothing by basis splines to real data can be done through the familiar technique of fitting statistical models to data by minimizing the sum of squared errors, which leads to an estimation of unknown parameters c_{1k}. The sum of squared errors, including the roughness penalty in eq. (3.2), can be derived

$$SSE = \left(x_{1j} - B_k(t)c_{1k}\right)^{\mathrm{T}} \left(x_{1j} - B_k(t)c_{1k}\right) + \lambda {c_{1k}}^{\mathrm{T}} RP c_{1k} \tag{3.3}$$

where $\mathbf{x}_1 = x_{1j}$ is a vector of values of the first sample observation across all observed variables $j = 1, \dots, p$ and a subscript T refers to a transpose. If we take the derivative of eq. (3.3)

with respect to a vector of parameters $\mathbf{c}_1 = c_{1k}$ and solve it equal to 0, we get the following solution

$$\hat{c}_{1k} = \left(B_k(t)^{\mathrm{T}} B_k(t) + \lambda R P^j\right)^{-1} B_k(t)^{\mathrm{T}} x_{1t} \tag{3.4}$$

Naturally, the estimation of curves for each sample observation can be extended for each i. The estimation method now leads to a matrix of estimated parameters $\hat{\mathbf{C}} = \hat{c}_{ik}$. We note that the same $B_k(t)$ is used across all curves.

3.2 Functional parametric regression models

The aim of this subsection is to describe a methodology which combines linear parametric modelling with mathematical functions observed as either the response or the covariate. The assumption of independence between observed curves must be satisfied, i.e. $cov\left(\chi_i(t), \chi_j(t)\right) = 0$, for all $i \neq j$.

The parametric functional models are similar to its *real-valued* counterparts in the sense that the estimation of the unknown parameters is carried out by minimising the least squares criterion, although with a modified computational strategy. The main difference is that the regression coefficients now become the regression coefficient function(s), $\beta(t)$, which are observable on a continuous domain t.

3.2.1 The response is functional and the covariate is binary

Here, we show a theoretical concept for estimation of functional analysis of variance. The features of the functional matrix of observed curves $\chi_i(t)$ are used to fit a regression model by constructing a set of binary independent variables in the covariate matrix $\mathbf{Z} = z_{(ic)j}$, with a dimension of $n \times (C + 1)$. The number of curves in each category, $c = 1, \dots, C$, is denoted by n_c, i.e. the i^{th} sample curve falls in the specific c^{th} group.

The columns $j = 1, \dots, C + 1$ of the covariate matrix $\mathbf{Z}$ are constructed as a binary indicator that represents a category of each observed curve, with ones in the first column that serves for estimation of the mean function across all curves. In this case, the statistical technique of analysis of variance, which is extended to functional space, addresses the problem because the values of the independent variables are 0's or 1's coding the group of the observed curves.

We commence with a formalised model of functional analysis of variance (Ramsay and Silverman, 2005)

$$\chi_{ic}(t) = \mu(t) + \alpha_c(t) + \varepsilon_{ic}(t) \tag{3.5}$$

where the subscript c refers to the category of each observed curve i and the function $\mu(t)$ is the overall mean across all n sample curves. The terms $\alpha_c(t)$ are the specific effect functions which represent departures from the overall mean function specific for each category c. The functional error term $\varepsilon_{ic}(t)$ is the unexplained variation specific to the i^{th} sample curve. $\varepsilon_{ic}(t)$ are assumed to be identically and independently distributed. Similarly, as in the classical analysis of variance framework, the following constraint has to be satisfied in order to identify the specific effects of categories

$$\sum_c \alpha_c(t) = 0 \quad for\ all\ t \tag{3.6}$$

There are a number of ways of imposing the constraint of eq. (3.6). We do it by adding the additional $n + 1$ curve into $\chi_i(t)$, i.e. $\chi_{n+1}(t) = 0$.

The regression model in eq. (3.5) can be re-expressed by defining a set of regression functions $\beta_j(t)$ by setting $\beta_1(t) = \mu(t)$, $\beta_2(t) = \alpha_1(t)$, and so on to $\beta_{C+1}(t) = \alpha_C(t)$, so that the functional vector $\beta_j(t) = \big(\mu(t), \alpha_1(t), \dots, \alpha_C(t)\big)^{\mathrm{T}}$. In these terms, eq. (3.5) can be reformulated to eq. (1.4).

In order to estimate the model in eq. (1.4), the regression functions $\beta_j(t)$ are also expressed in terms of basis splines, $\beta_j(t) = b_{jk_\beta}\theta_{k_\beta}(t)$, where $\theta_{k_\beta}(t)$ is a $k_\beta \times 1$ vector of basis splines[7], and b_{jk_β} are the values of the unknown $(C + 1) \times k_\beta$ matrix of coefficients that need to be estimated.

It follows that if we use the modified method of least squares, we get the following computational technique for estimation of coefficients b_{jk_β} (Ramsay and Silverman, 2005)

[7] We can keep the number k_β and the order m of basis splines as for $B_k(t)$.

$$vec\left(\hat{b}_{jk_\beta}\right) = \left(\mathbf{J}_{\theta_{k_\beta}\theta_{k_\beta}} \otimes (\mathbf{Z}^{\mathrm{T}}\mathbf{Z})\right)^{-1} vec\left(\mathbf{Z}^{\mathrm{T}}\mathbf{C}\mathbf{J}_{B_k(t)\theta_{k_\beta}}\right) \tag{3.7}$$

where $vec(\cdot)$ denotes a vectorised matrix, $\otimes$ is the Kronecker product, $\mathbf{J}_{\theta_{k_\beta}\theta_{k_\beta}} = \int \theta_{k_\beta}\theta_{k_\beta}{}^{\mathrm{T}}(t)dt$ and $\mathbf{J}_{B_k(t)\theta_{k_\beta}} = \int B_k(t)\theta_{k_\beta}{}^{\mathrm{T}}(t)dt$ are the derivatives of spline functions across the domain $t \in T$. A detailed derivation is provided by Ramsay and Silverman (2005).

3.2.2 The response is real-valued and the covariate is functional

Here, we show a computational technique for the functional regression model, where the response is *real-valued* and the covariate is a functional matrix that contains the observed curves. We recall that such a model is defined by eq. (1.5). The model in eq. (1.5) is a functional extension of a linear regression where the usual summation of $\sum_{j=1}^{p}\beta_j x_{ij}$ is replaced by integration $\int \chi_i(t)\beta(t)dt$ over a continuous domain $t \in T$, where $\beta(t)$ is a regression coefficient function observable on a continuous domain $t \in T$.

Similarly, as in the previous section, we first decompose the regression coefficient function $\beta(t)$ in terms of basis splines, $\beta(t) = b_{k_\beta}\theta_{k_\beta}(t)$, where $\theta_{k_\beta}(t)$ is a vector of basis splines[8] of length $k_\beta \times 1$ and b_{k_β} are the corresponding values of unknown coefficients. It follows that the model in eq. (1.5) can be re-expressed in terms of the matrix notation

$$\mathbf{y} = \int \mathbf{C}B_k(t)\theta_{k_\beta}(t)^{\mathrm{T}}\mathbf{b}dt + \boldsymbol{\varepsilon} \tag{3.8}$$

where the *point-wise* parameter α is absorbed into $\theta_{k_\beta}(t)^{\mathrm{T}}\mathbf{b} = b_{k_\beta}\theta_{k_\beta}(t)$ by adding a first column of ones for its estimation. The following sum of squared errors can be obtained (Ramsay and Silverman, 2005)

$$SSE = \sum_{i=1}^{n}\left(\mathbf{y} - \int \mathbf{C}B_k(t)\theta_{k_\beta}(t)^{\mathrm{T}}\mathbf{b}dt\right)^2 \tag{3.9}$$

[8] Again, we can keep the number k_β and the order m of basis splines as in $B_k(t)$.

Similarly, to make the functional parameter $\beta(t)$ smoother, we can add the roughness penalty into eq. (3.9), $RP^{\beta(t)} = \lambda \int D^j \theta_k(t) D^j \theta_k(t) dt$. The smoothing parameter λ controls the level of the smoothness of $\beta(t)$, specifically, when the number of basis splines is too large.

It follows that if we use the modified method of least squares, we get the following computational technique for estimation of the unknown coefficients b_{k_β} (Ramsay and Silverman, 2005)

$$\left(\hat{\alpha}, \hat{b}_1, \dots, \hat{b}_{k_\beta}\right)^{\mathrm{T}} = (\mathbf{Z}^{\mathrm{T}}\mathbf{Z})^{-1}\mathbf{Z}^{\mathrm{T}}\mathbf{y} \tag{3.10}$$

where $\mathbf{Z} = \left[\mathbf{1C} \int B_k(t) \theta_{k_\beta}(t)^{\mathrm{T}} dt\right]$ is a $n \times \left(k_\beta + 1\right)$ matrix, where $\mathbf{1}$ refers to a constant basis for estimation of α. A detailed derivation is provided by Ramsay and Silverman (2005).

3.3 Spatial econometric models

Here, we discuss the estimation procedure of linear regression models with spatial dependencies, which occur when sample observations are associated with geographical locations identifiable through a set of spatial coordinates, i.e. the longitude and the latitude.

We commence by recalling a theoretical framework of the classical regression model in eq. (1.8), where the values of the random error term are identically and independently distributed with 0 mean and a constant variance σ^2. One of the estimation procedures of the unknown parameters α and β_j of the model is the ordinary method of least squares (Anderson, 2003).

In the following section, we describe an adjustment of the classical regression model, including the modification of the method of least squares, through the addition of the spatially lagged structure to the model. The main focus of our research is the model with spatially autocorrelated error term structure.

3.3.1 Spatial autoregressive models

The observations across different geographical locations are known to be spatially autocorrelated when the observation from the location i depends on observations from other location j, where $i \neq j$. The spatial autocorrelation between observations can be formally defined

in terms of the corresponding nonzero covariance between observed responses $cov(y_i, y_j) = E(y_i y_j) - E(y_i)E(y_j) \neq 0,\ i \neq j$, where $E(\cdot)$ is an expected value (LeSage and Pace, 2009).

We can incorporate the spatial autocorrelation within the theoretical framework of the spatial autoregressive model in different ways, which is shown in eq. (1.11). Table 2 shows a list of models with different spatially lagged structures (Furková, 2017).

Name	**Model**	**Spatially lagged structure**
SEM	$\mathbf{y} = \alpha + \mathbf{X\beta} + \boldsymbol{\varepsilon}$ $\boldsymbol{\varepsilon} = \rho\mathbf{W\varepsilon} + \mathbf{v}, \quad where\ \mathbf{v} \sim N(0, \sigma^2\mathbf{I})$	Spatially lagged error term
SAR	$\mathbf{y} = \alpha + \mathbf{X\beta} + \rho\mathbf{Wy} + \boldsymbol{\varepsilon},$ $where\ \boldsymbol{\varepsilon} \sim N(0, \sigma^2\mathbf{I})$	Spatially lagged response
SARAR	$\mathbf{y} = \alpha + \mathbf{X\beta} + \rho\mathbf{Wy} + \boldsymbol{\varepsilon}$ $\boldsymbol{\varepsilon} = \rho\mathbf{W\varepsilon} + \mathbf{v}, \quad where\ \boldsymbol{\varepsilon} \sim N(0, \sigma^2\mathbf{I})$	Both response and error term are spatially lagged
SDM	$\mathbf{y} = \alpha + \mathbf{X\beta} - \rho\mathbf{WX\beta} + \rho\mathbf{Wy} + \boldsymbol{\varepsilon}, \quad where\ \boldsymbol{\varepsilon} \sim N(0, \sigma^2\mathbf{I})$	Both response and covariate are spatially lagged
SLX	$\mathbf{y} = \alpha + \mathbf{X\beta} - \rho\mathbf{WX\beta} + \boldsymbol{\varepsilon},$ $where\ \boldsymbol{\varepsilon} \sim N(0, \sigma^2\mathbf{I})$	Spatially lagged covariate

Table 2. The spatial autoregressive model with different spatially lagged structures
Source: Author's construction

In the following text, we show two different versions of the method of least squares for estimation of the spatial error model (SEM) and the spatial autoregressive model (SAR), respectively.

In the literature, the notation of SEM is commonly used for spatial error model. SEM represents a spatial autoregressive model, where the model's spatial autocorrelation is embedded in its error term.

To remove the spatial autocorrelation from the error term, we multiply the model in eq. (1.10), or refer to Table 2 SEM, by $(\mathbf{I} - \rho\mathbf{W})$, which leads to the following

$$(\mathbf{I} - \rho\mathbf{W})\mathbf{y} = (\mathbf{I} - \rho\mathbf{W})\mathbf{X}\boldsymbol{\beta} + \mathbf{v} \tag{3.11}$$

where, as usual, the unknown parameter α is absorbed into $\mathbf{X}$ by adding ones in the first column. The estimation of the unknown parameters $\boldsymbol{\beta}$ and ρ can be carried by the generalised method of least squares. It follows that the following sum of squared errors can be obtained

$$SSE = [(\mathbf{I} - \rho\mathbf{W})\mathbf{y} - (\mathbf{I} - \rho\mathbf{W})\mathbf{X}\boldsymbol{\beta}]^2 \tag{3.12}$$

By applying the usual technique of minimising the sum of squared errors, i.e. taking the derivative of eq. (3.12) with respect to $\boldsymbol{\beta}$ and ρ, respectively, and solving it equal to zero, we get the following set of equations

$$\begin{aligned} \widehat{\boldsymbol{\beta}} &= [(\mathbf{X} - \hat{\rho}\mathbf{W}\mathbf{X})^{\mathrm{T}}(\mathbf{X} - \hat{\rho}\mathbf{W}\mathbf{X})]^{-1}(\mathbf{X} - \hat{\rho}\mathbf{W}\mathbf{X})^{\mathrm{T}}(\boldsymbol{y} - \hat{\rho}\mathbf{W}\mathbf{y}) \\ \hat{\rho} &= \left[\left(\boldsymbol{y} - \mathbf{W}\mathbf{X}\widehat{\boldsymbol{\beta}}\right)^{\mathrm{T}}\left(\boldsymbol{y} - \mathbf{W}\mathbf{X}\widehat{\boldsymbol{\beta}}\right)\right]^{-1}\left(\boldsymbol{y} - \mathbf{W}\mathbf{X}\widehat{\boldsymbol{\beta}}\right)^{\mathrm{T}}\left(\boldsymbol{y} - \mathbf{W}\mathbf{X}\widehat{\boldsymbol{\beta}}\right) \end{aligned} \tag{3.13}$$

The solution to the set of equations in (3.13) can be calculated iteratively by setting first $\hat{\rho} = 0$, which implies that, in the first step, $\widehat{\boldsymbol{\beta}}$ are estimated through the ordinary method of least squares. We then plug $\widehat{\boldsymbol{\beta}}$ into the second equation and recalculate $\hat{\rho}$ and so on until the desired convergence of $\hat{\rho}$ is attained, i.e. a specified difference between the most recent and the previous estimate of $\hat{\rho}$. This iterative procedure is shown by Schabenberger and Gotway (2005).

In the following text, we show the estimation of the spatial autoregressive model with the spatially lagged response structure, which is also called spatial autoregressive model (SAR). We recall the expression of the model

$$\mathbf{y} = \alpha + \mathbf{X\beta} + \rho\mathbf{Wy} + \boldsymbol{\varepsilon} \tag{3.14}$$

The term $\mathbf{Wy} = \sum_{j=1}^{n} w_{ij} y_j$ represents a linear combination of the values y_i, which is constructed from response values in locations j that are neighbours of the location i, as defined by the values in w_{ij}. In eq. (3.14), the correlation occurs between the random errors ε_i and the values in the term $\sum_{j=1}^{n} w_{ij} y_j$, which implies that the estimation of the regression parameters β_j by the method of ordinary least squares would lead to inconsistent estimates (Green, 2003). To eliminate this correlation, also called endogeneity, we can apply the estimation strategy by the method of *two-stage* least squares, which first requires an identification of the instrumental variables.

We commence with the rearrangement of the term **Wy**

$$\mathbf{Wy} = \mathbf{W}(\mathbf{I} - \rho\mathbf{W})^{-1}\mathbf{X\beta} + (\mathbf{I} - \rho\mathbf{W})\boldsymbol{\varepsilon} \tag{3.15}$$

where **I** is an $n \times n$ identity matrix. Applying the constraint on $|\rho| < 1$, we can rewrite the term $(\mathbf{I} - \rho\mathbf{W})^{-1}$ in terms of power series (Anselin and Rey, 2014)

$$(\mathbf{I} - \rho\mathbf{W})^{-1} = \sum_{p=0}^{\infty} \rho^p \mathbf{W}^p \tag{3.16}$$

where the subscript p refers to the infinite expansion of power series. Substituting eq. (3.16) into eq. (3.15) gives

$$\mathbf{Wy} = \mathbf{W} \sum_{p=0}^{\infty} \rho^p \mathbf{W}^p \, \mathbf{X\beta} + \mathbf{W} \sum_{p=0}^{\infty} \rho^p \mathbf{W}^p \boldsymbol{\varepsilon} \tag{3.17}$$

Substituting eq. (3.17) into eq. (3.14), the model in eq. (3.14) can be rewritten as

$$\mathbf{y} = \sum_{p=0}^{\infty} \rho^p \mathbf{W}^p \, \mathbf{X\beta} + \mathbf{X\beta} + \boldsymbol{\varepsilon} \tag{3.18}$$

where the term $\sum_{p=0}^{\infty} \rho^p \mathbf{W}^p \mathbf{X}\boldsymbol{\beta}$ consists of instrumental variables, which is a linear combination of the spatially lagged covariate variables of several power series, i.e. $\mathbf{WX}$, $\mathbf{W}^2\mathbf{X}$, $\mathbf{W}^3\mathbf{X}$ etc. In practice, an expansion of power series needs to be cut off at some finite point.

The literature offers various options of constructing a set of instrumental variable. For example, Kelejian et al. (2004) proposes the following matrix of instrumental variables

$$\mathbf{Q} = \left[\mathbf{X}, \mathbf{W}\sum_{p=0}^{\infty} \rho^p \mathbf{W}^p \mathbf{X}\boldsymbol{\beta}\right] \tag{3.19}$$

where eq. (3.16) is substituted into eq. (3.15). It follows that by applying the method of *two-stage* least squares, we obtain the following estimated parameters

$$\widehat{\boldsymbol{\beta}} = [(\mathbf{Q}(\mathbf{Q}^{\mathrm{T}}\mathbf{Q})^{-1}\mathbf{Q}^{\mathrm{T}}\mathbf{X})^{\mathrm{T}}\mathbf{X}]^{-1}(\mathbf{Q}(\mathbf{Q}^{\mathrm{T}}\mathbf{Q})^{-1}\mathbf{Q}^{\mathrm{T}}\mathbf{X})^{\mathrm{T}}\mathbf{y} \tag{3.20}$$

A detailed derivation of $\widehat{\boldsymbol{\beta}}$ is given by Lee (2003), albeit with slightly different instrumental variables.

There exists other *well-known* estimation methods for spatial autoregressive models, e.g. a maximum likelihood estimation. The habilitation thesis by Furková (2017) provides a comprehensive overview of other estimation methods.

3.3.2 Extension to functional space

Here, we extend the spatial autoregressive model, including its estimation procedure, to functional space. Our focus is on the functional extension of the spatial error model (SEM).

The paper by Kelejian and Prucha (1997) shows that the *two-stage* least square estimator for spatial regression models with spatially autocorrelated errors leads to inconsistent estimates. The authors propose to use the generalised least square estimator, where the spatial autoregressive parameter is estimated by the iterative procedure, similarly as in eq. (3.13). They demonstrate that the generalized least squares estimator is consistent and asymptotically normal.

The model is constructed such that the covariate matrix $\chi_i(t)$ is observed in the functional space and the response $\mathbf{y} = y_i$ is a *real-valued* variable. We recall that the functional extension of the spatial error model can be formalised

$$\mathbf{y} = \alpha + \int \beta(t)\chi_i(t)dt + \boldsymbol{\varepsilon}$$
$$\boldsymbol{\varepsilon} = \rho\mathbf{W}\boldsymbol{\varepsilon} + \mathbf{v} \tag{3.21}$$

where $\boldsymbol{\varepsilon}$ are spatially autocorrelated but $\mathbf{v}$ are assumed independently and identically distributed. We note that the regression parameter $\beta(t)$ is functional, the intercept α and the spatial autoregressive parameter ρ are *point-wise*.

From Section 3.2.2, we recall that both $\beta(t)$ and $\chi_i(t)$ can be defined in terms of basis splines. It follows that the model in eq. (3.21) can be rewritten

$$\mathbf{y} = \int \mathbf{C}B_k(t)^{\mathrm{T}}\theta_{k_\beta}(t)^{\mathrm{T}}\mathbf{b}dt + (\mathbf{I} - \rho\mathbf{W})^{-1}\mathbf{v} \tag{3.22}$$

We note that $\beta(t) = \theta_{k_\beta}(t)^{\mathrm{T}}\mathbf{b}$, $\chi_i(t) = \mathbf{C}B_k(t)^{\mathrm{T}}$ and $\boldsymbol{\varepsilon}$ is replaced by rearranging $\boldsymbol{\varepsilon} = (\mathbf{I} - \rho\mathbf{W})^{-1}\mathbf{v}$. The vector of unknown parameters $\mathbf{b}$ remains to be estimated. The unknown parameter α is absorbed into $\theta_{k_\beta}(t)$ by adding a constant basis.

We remove the spatial autocorrelation from the error term in eq. (3.22) by multiplying with $(\mathbf{I} - \rho\mathbf{W})$, which leads to the following

$$(\mathbf{I} - \rho\mathbf{W})\mathbf{y} = (\mathbf{I} - \rho\mathbf{W}) \int \mathbf{C}B_k(t)^{\mathrm{T}}\theta_{k_\beta}(t)^{\mathrm{T}}\mathbf{b}dt + \mathbf{v} \tag{3.23}$$

The estimation of the model in eq. (3.23) is proposed by Pineda-Riós et al. (2018) through the generalised method of least squares. In the first step, the authors eliminate the basis functions $B_k(t)$ and $\theta_{k_\beta}(t)$ from the integral part by using the assumption of orthonormal basis. In our case, the basis splines are not orthonormal so the integral part can only be rearranged

$$\begin{aligned}\int \mathbf{C}B_k(t)^{\mathrm{T}}\theta_{k_\beta}(t)^{\mathrm{T}}\mathbf{b}dt &= \int (\mathbf{C}B_k(t)^{\mathrm{T}})\left(\theta_{k_\beta}(t)^{\mathrm{T}}\mathbf{b}\right)dt \\ &= \mathbf{C}\mathbf{b}^{\mathrm{T}}\left(\int B_k(t)\theta_{k_\beta}(t)^{\mathrm{T}}dt\right)\end{aligned} \tag{3.24}$$

It follows that the sum of squared errors can be obtained

$$SSE = \left(\boldsymbol{y} - \mathbf{C}\mathbf{b}^{\mathrm{T}} \int B_k(t)\theta_{k_\beta}(t)^{\mathrm{T}} dt\right)^{\mathrm{T}} (\mathbf{I} - \rho\mathbf{W})^2 \left(\boldsymbol{y} - \mathbf{C}\mathbf{b}^{\mathrm{T}} \int B_k(t)\theta_{k_\beta}(t)^{\mathrm{T}} dt\right) \quad (3.25)$$

By taking the derivative of eq. (3.25) with respect to $\mathbf{b}$ and ρ, respectively, and solving it equal to 0, we obtain the following set of equations

$$\begin{aligned} \hat{\mathbf{b}} &= (\mathbf{Z}^{\mathrm{T}}(\mathbf{I} - \hat{\rho}\mathbf{W})^2\mathbf{Z})^{-1}\mathbf{Z}^{\mathrm{T}}(\mathbf{I} - \hat{\rho}\mathbf{W})^2\mathbf{y} \\ \hat{\rho} &= \left(\left(\mathbf{y} - \hat{\mathbf{b}}\mathbf{Z}\right)^{\mathrm{T}}\mathbf{W}^2\left(\mathbf{y} - \hat{\mathbf{b}}\mathbf{Z}\right)\right)^{-1}\left(\mathbf{y} - \hat{\mathbf{b}}\mathbf{Z}\right)^{\mathrm{T}}\mathbf{W}\left(\mathbf{y} - \hat{\mathbf{b}}\mathbf{Z}\right) \end{aligned} \quad (3.26)$$

where $\mathbf{Z} = \left[\mathbf{1}\mathbf{C}\int B_k(t)\theta_{k_\beta}(t)^{\mathrm{T}} dt\right]$ is a $n \times \left(k_\beta + 1\right)$ matrix. The term $\mathbf{1}$ refers to a constant basis for estimation of α. The system of equations in eq. (3.26) cannot be solved analytically. The solution can be obtained through an iterative procedure as shown by Schabenberger and Gotway (2005), refer to section 3.3.1. However, the solution for ρ can lead to suboptimal results, i.e. the estimated parameter $\hat{\rho}$ can be relatively large (Waller and Gotway, 2004), or the iterative algorithm can suffer from *non-convergence* issues due to *over-parameterisation* in $\chi_i(t)$ and $\beta(t)$.

Hence, we propose to modify the iterative procedure shown by Schabenberger and Gotway (2005), which minimises the generalised *cross-validation* score of estimated models. The modified algorithm leads to a more optimal solution for $\hat{\rho}$. To the best of our knowledge, there is no previous study of our proposed algorithm.

The *step-wise* iterative procedure that finds the solution for $\hat{\mathbf{b}}$ and $\hat{\rho}$

1. Estimate $\hat{\mathbf{b}}$ through eq. (3.26) for specified range of $\hat{\rho}$;
2. For each estimated model in step 1, calculate the generalised *cross-validation* score;
3. Select the estimated model, i.e. $\hat{\mathbf{b}}$ and $\hat{\rho}$, of which the generalised *cross-validation* score is at its minima.

The range of $\hat{\rho}$ can be selected empirically, i.e. for large values of $\hat{\rho}$, the generalised *cross-validation* score of estimated models converges to infinity. In the empirical analysis, Section 4, we demonstrate that our proposed iterative procedure leads to a robust solution for spatial autoregressive parameter $\hat{\rho}$.

The generalised *cross-validation* score of the estimated model can be defined (Ramsay and Silverman, 2005)

$$GCV = \frac{(\mathbf{y} - \hat{\mathbf{y}})^{\mathrm{T}}(\mathbf{y} - \hat{\mathbf{y}})}{\left(1 - \frac{trace(\mathbf{S})}{n}\right)^2} \tag{3.27}$$

where $\mathbf{S} = (\mathbf{Z}^{\mathrm{T}}(\mathbf{I} - \hat{\rho}\mathbf{W})^2\mathbf{Z})^{-1}\mathbf{Z}^{\mathrm{T}}(\mathbf{I} - \hat{\rho}\mathbf{W})^2$ is the $n \times n$ matrix, also called the *hat* matrix, and the $trace(.)$ function returns the sum of its diagonal values. The nominator of eq. (3.27) is the sum of squared errors of the estimated model.

Moreover, the spatial error model (SEM) can be extended to the fully functional form, i.e. both the response and the covariate are functional. We recall that it can be formalised

$$\begin{gathered} \Upsilon_i(t) = \alpha(t) + \int \beta(t)\chi_i(t)dt + \varepsilon_i(t) \\ \varepsilon_i(t) = \rho\mathbf{W}\varepsilon_i(t) + v_i(t) \end{gathered} \tag{3.28}$$

where $\Upsilon_i(t)$ is the $i \times k_\Upsilon$ matrix of observed response curves. The parameter k_Υ refers to the number of basis splines used in the expansion of $\Upsilon_i(t)$. A difference to the previous form is that $\alpha(t)$, $\varepsilon_i(t)$ and $v_i(t)$ are now all functional. It is evident that a derivation of the analytical solution for model in eq. (3.28) can be very complex.

The paper by Římalová et al. (2022) shows the estimation procedure of the model in eq. (3.28), albeit with no *spatially-weighted* correlation structure, by using the generalised method of least squares. The estimator function for $\mathbf{b} = b_{k_\beta}$ is weighted by the unknown spatial *variance-covariance* matrix (Římalová et al., 2022)

$$\hat{\mathbf{b}} = (\mathbf{Z}^{\mathrm{T}}\boldsymbol{\Sigma}^{-1}\mathbf{Z})^{-1}\mathbf{Z}^{\mathrm{T}}\boldsymbol{\Sigma}^{-1}\Upsilon_i(t) \tag{3.29}$$

where $\mathbf{Z} = \left[\mathbf{1C}\int B_k(t)\theta_{k_\beta}(t)^{\mathrm{T}}dt\right]$ is a $n \times \left(k_\beta + 1\right)$ matrix. The term $\mathbf{1}$ refers to a constant basis for estimation of α, and $\boldsymbol{\Sigma}$ is the unknown $n \times n$ *variance-covariance* matrix. We note that the functional response $\Upsilon_i(t)$ is also decomposed in terms of basis splines.

The estimator in eq. (3.29) can be computed iteratively. In the first step, the parameters $\hat{\mathbf{b}}$ are estimated through the method of ordinary least squares. In the next step, the unknown

variance-covariance matrix **Σ** is estimated from residuals and *plugged-in* back to eq. (3.29). The procedure is then iterated until a specified difference between the most recent and the previously estimated residuals is attained. Moreover, the authors propose a significance test of the estimated model, which is based on the permutation testing procedure of spatially filtered residuals; refer to Římalová et al. (2022) for details.

4 Results

In this section, we show the empirical analysis of new data source, i.e. electronic records from toll system in Slovakia, which is briefly introduced in Section 1. We describe the detailed data generation process and the aggregation of daily electronic records. We then follow with an application of functional regression models discussed in the previous section.

The main objective of the empirical analysis of the dissertation thesis is to propose an econometric model, which captures the dynamics of the regional industry production through the intensity of the freight. The econometric model is based on the functional version of the spatial error model (SEM). We show the estimation of the model for five consecutive years by using our proposed iterative procedure to ascertain its robustness, specifically, for estimation of the spatial autoregressive parameter. Prior to the application of the econometric model, we carry out an exploratory analysis of the regional intensity of the freight across time by using the functional analysis of variance.

Section 4 is organised as follows. In Section 4.1, we describe a process of data collection and data aggregation. Moreover, the transformation of *real-valued* aggregated data to functional observations, i.e. a set of observed curves, is shown. In Sections 4.2, we estimate functional descriptive statistics and discuss its features. In Section 4.3, we continue with further exploratory analysis by using the functional analysis of variances. In Section 4.4, we discuss the proposed econometric model and show the results of estimated parameters of the functional SEM. The measure of the mean squared error (MSE) is used for evaluation of the estimated econometric model.

4.1 Data generation process

In the following text, we describe all data sources that are used in the empirical analysis: electronic records from toll system and industrial production in Slovakia. Moreover, we show the transformed functional observations for 2018-2023, separately, and discuss its features.

4.1.1 Electronic records from toll system

A digital system to collect a toll on highways was put in operation in Slovakia in January 2010, which after being extended to all classes of roads, became the longest network of tollable

roads in the European Union – the satellite technology covers 17,600 km of defined sections of roads in Slovakia[9]. Hence, the daily digital records of passing vehicles through sections of roads generated by the toll system create an opportunity to utilise these data for nowcasting macroeconomic variables since it captures a large part of the road transport of goods and commodities in Slovakia. In general, data are collected by the electronic toll system in an automated way as an inherent function of on-board units, installed in trucks[10], which uses satellite-covered sections of the Slovak road network. The on-board units record current geographic data on satellite-covered sections of roads. If the *on-board* unit's algorithm detects that the vehicle passed through a defined section of the road, the record is created and stored in the system, also called a toll event. It is important to note that not all sections of roads are chargeable. The record of the toll event contains date and time, identification of the road section, vehicle identification and information related to the vehicle. A section of the road is generally defined from the border of the intersection that forms its starting point, with an end point defined by the subsequent intersection. Each section of the road is marked with a unique identifier, the beginning of the section and the end of the section[11].

Generated data are records of the movement of vehicles through sections of the road network observed across the calendar years 2018-2023. Data are available by the 10th day of the following month. The dataset of records includes:

- unique identifier of the vehicle, so that it is possible to correctly record its trips made through highways,
- date and time of vehicle location recording, so that records are counted in the correct time window,
- vehicle category, so that only vehicles relevant to freight transport are taken into account, and not, for example, buses,
- identifier of the defined road section, which determines the position of the vehicle for the given record and its direction of travel,

[9] Skytoll, Electronic collection of toll, Slovak Republic. [2023-09-11]. Available at: https://www.skytoll.com/elektronicky-mytny-system-sr/.
[10] In the Slovak Republic, all vehicles with the total weight of over 3.5 tons or vehicles with the total weight of over 3.5 tons listed in § 4 par. 2 letters b) and c) of Act no. 106/2018 Coll. on the operation of vehicles in road traffic (vehicles of category M and N) except for motor vehicles of category M1 and except for vehicle sets consisting of motor vehicles of category M1 and N1, are obligated to install the on-board unit.
[11] The Supreme Audit Office of the Slovak Republic, 2019. Electronic toll collection: Final report.

- name of the region that locates the end point of each section of the road.

The limit for the dataset of records is that the distance is only provided for the chargeable sections of roads. The electronic system for other sections of roads only records a passing vehicle, with no record of the distance travelled. The collected dataset from the digital system represents a set of data points in time that can be considered as time series of *high-frequency* electronic records. The raw data records are subject to the following data aggregation.

4.1.2 Data aggregation

The electronic records, which represent digital time stamps of vehicles that pass through a point of entry, or exit, of monitored sections of the road network must be aggregated to a meaningful dataset for further statistical analysis. There are many possibilities for data aggregation. We decide to take a simple count of passing vehicles through entry / exit points per week between 2018-2023. This aggregation reflects the weekly intensity of the freight over time. It allows studying the intensity of the freight over time and the effect of its dynamics on the industrial production.

At the onset, we filter out all vehicles that are not relevant for transport of goods to manufacturers, e.g. buses. We then count a number of vehicles, which pass through each entry / exit point of monitored sections of roads, for each district. The *passage-count* is done on weekly basis between 2018-2023. Table 3, in Appendix, shows the weekly *passage-count* for each district between 2018-2023 (313 weeks). Each calendar year contains 52 weeks, except 2022 that contains 53 weeks.

Table 3 reveals that there are large differences in the volume of the intensity of the freight between districts, e.g. the district with the largest and the lowest volume of the intensity of the freight are Gelnica (in hundreds per week) and Košice (in hundreds of thousands per week), respectively. This phenomenon can cause a distortion of the estimated statistics, which would lead to incorrect statistical inferences for both exploratory statistics and regression analysis. Hence, we scale the original values (counts) by the logarithmic function with a base equal to ten. Table 4, in Appendix, shows the scaled values.

Moreover, we note that there are only 71 districts out of total 79 in Slovakia since we merge the districts of Bratislava I-V into Bratislava and Košice I-IV, including Košice-Šaca, into Košice.

4.1.3 From real space to functional

In this subsection, we display the weekly *passage-count* for each district transformed to functional observations (refer to Section 3.1, the estimation of eq. (3.1)), i.e. we now have 71 observed curves for 2018-2023, respectively. The features and patterns of the observed curves within and between years are also discussed.

Figures 2-7 display the observed curves for 71 districts, $\chi_i(t)$, for 2018-2023, respectively.

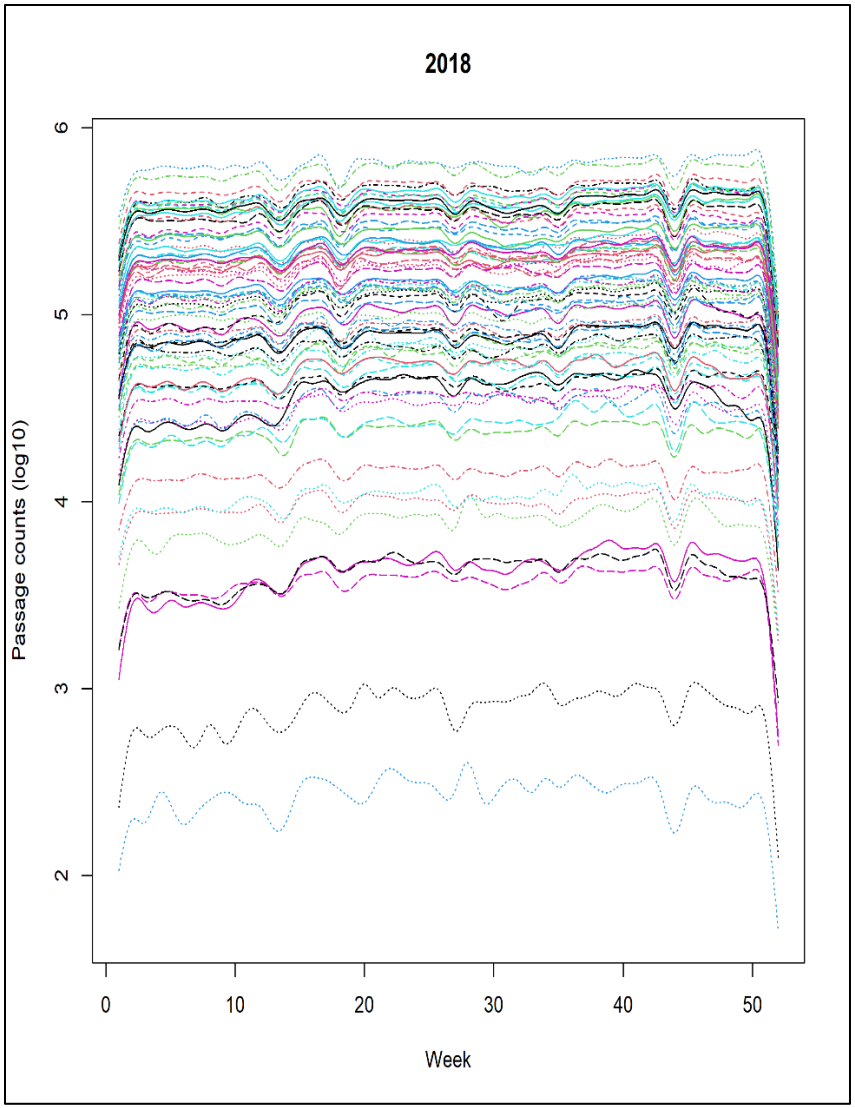

Figure 2. The observed curves (log_{10}) for 71 districts (2018)
Source: Author's construction

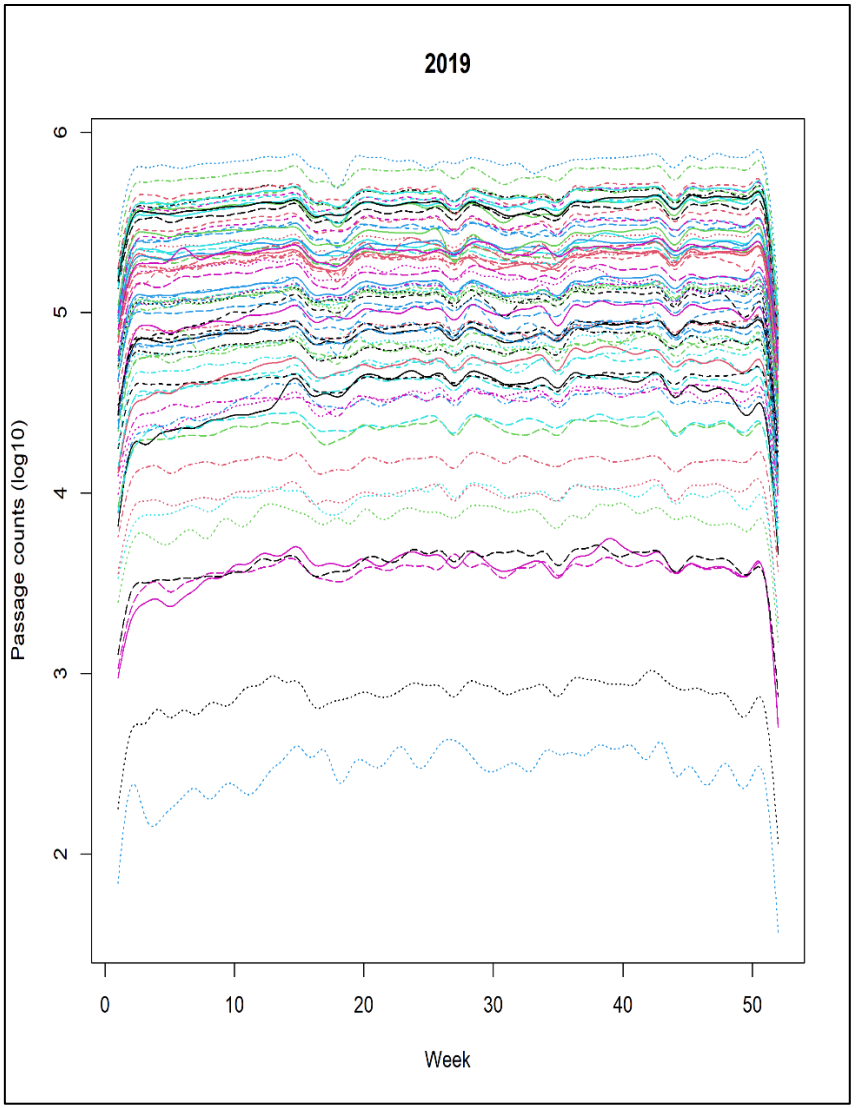

Figure 3. The observed curves (log_{10}) for 71 districts (2019)
Source: Author's construction

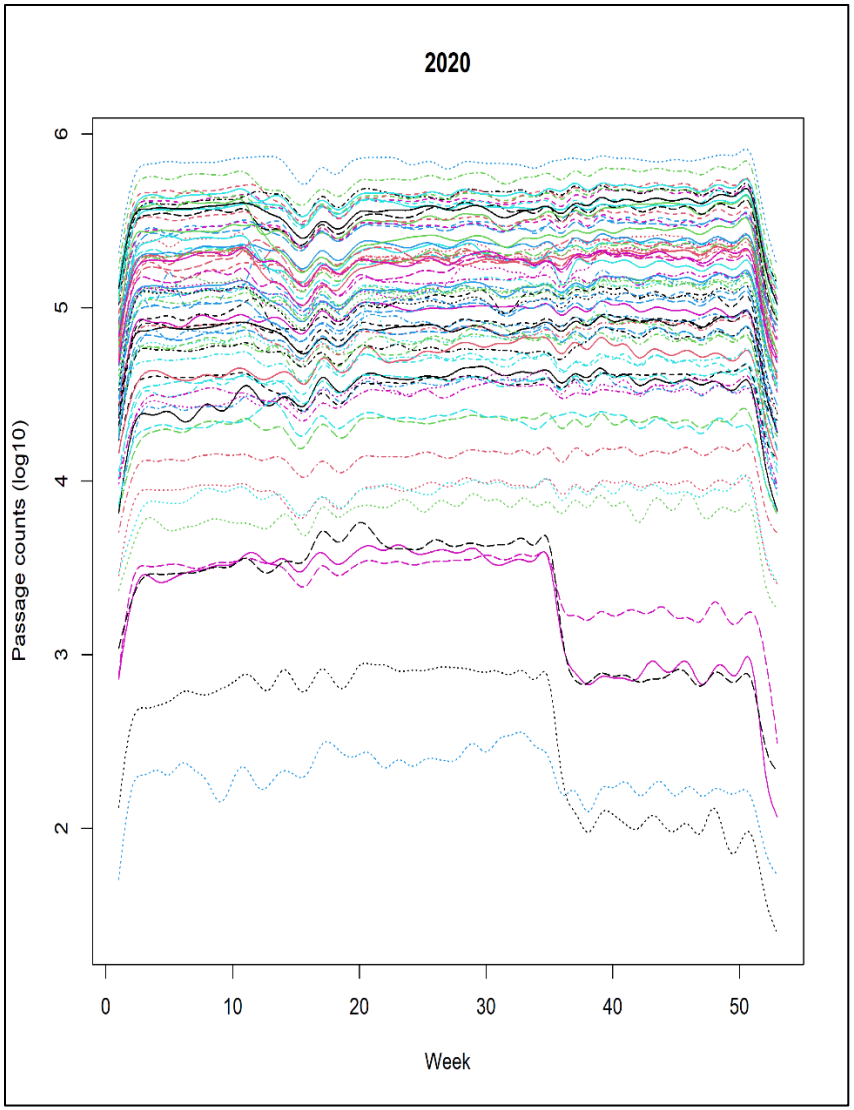

Figure 4. The observed curves (log_{10}) for 71 districts (2020)
Source: Author's construction

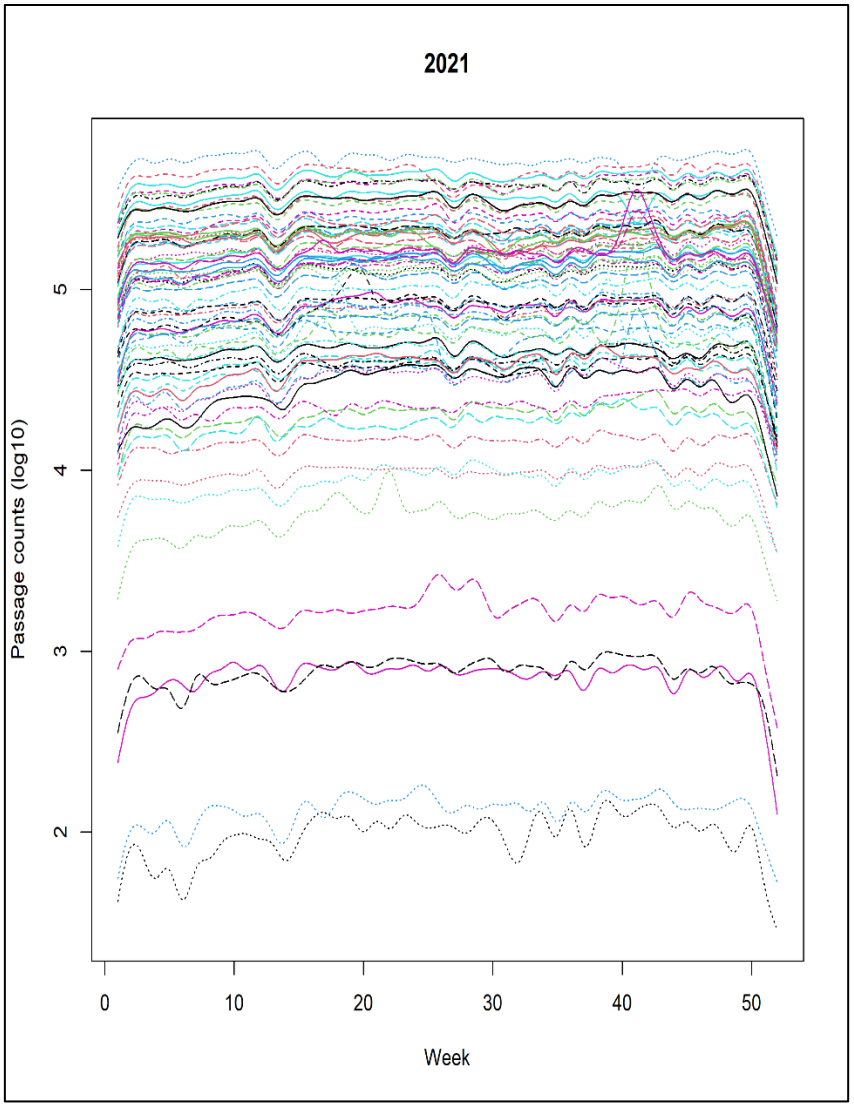

Figure 5. The observed curves (log_{10}) for 71 districts (2021)
Source: Author's construction

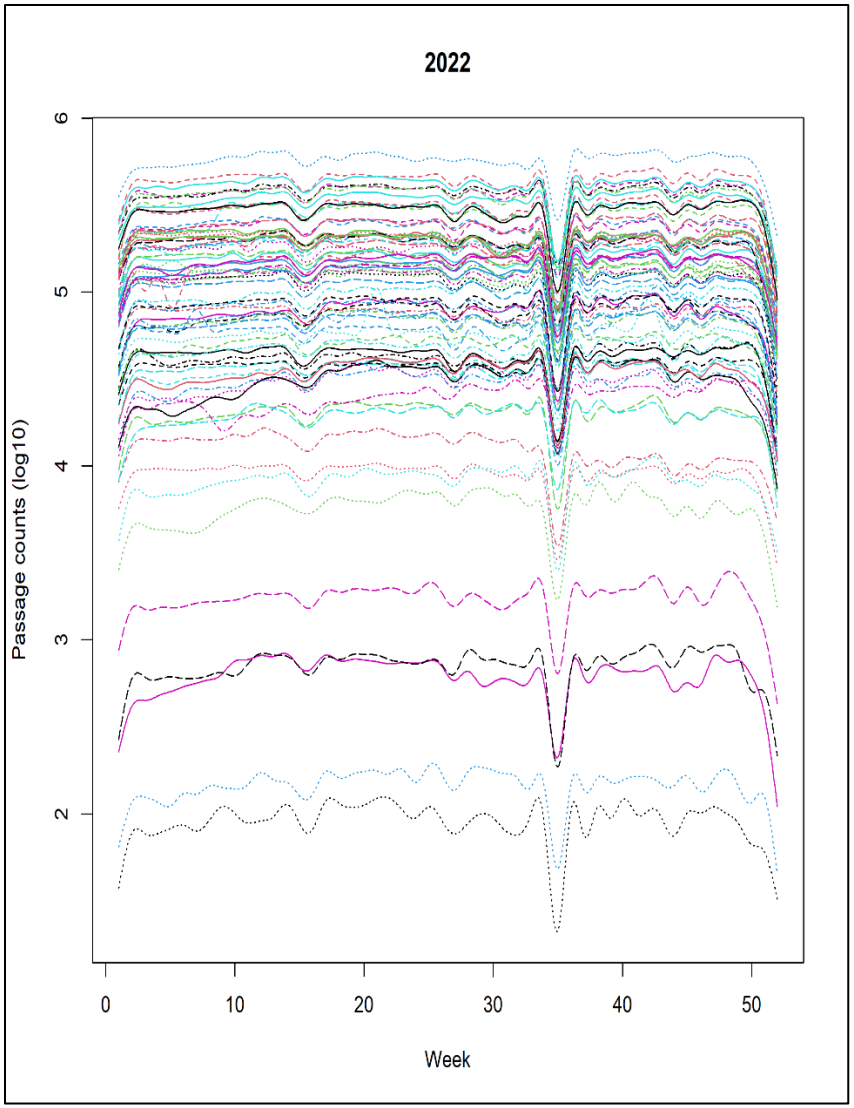

Figure 6. The observed curves (log_{10}) for 71 districts (2022)
Source: Author's construction

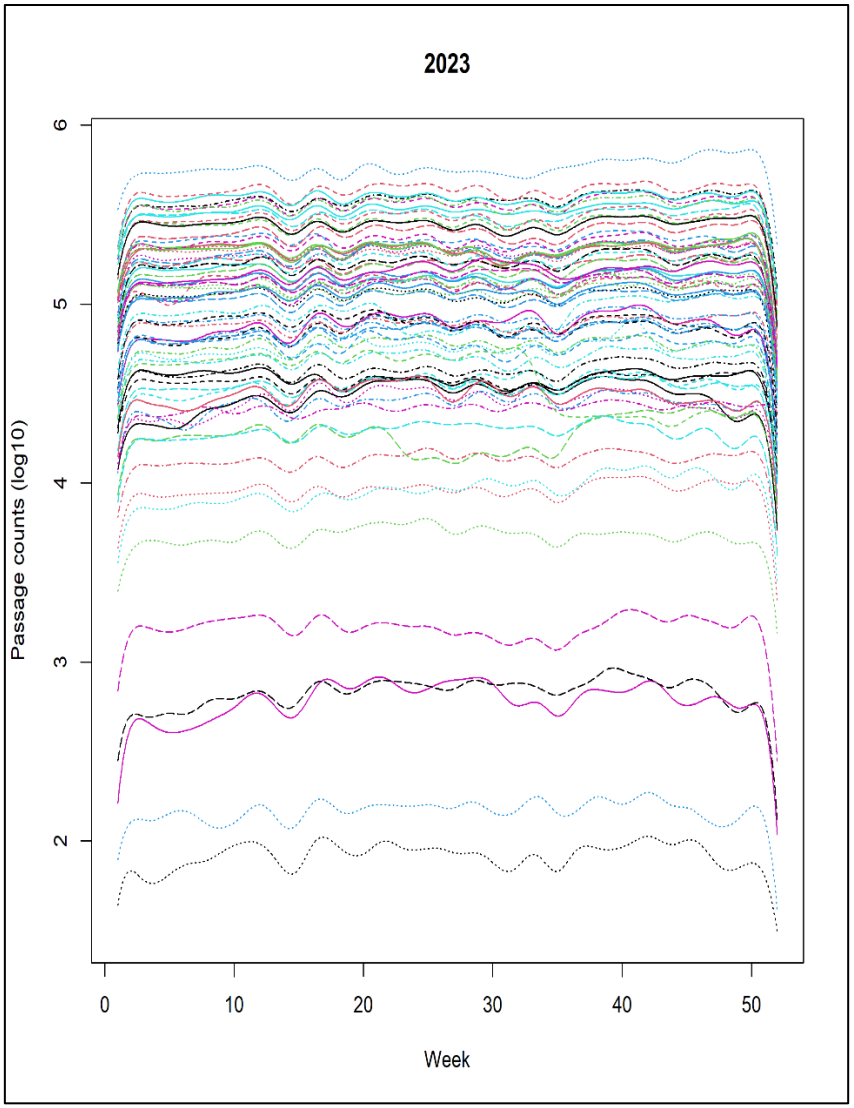

Figure 7. The observed curves (log_{10}) for 71 districts (2023)
Source: Author's construction

Figures 2-7 show that differences between the observed curves are significantly reduced after the logarithmic transformation. We can draw the following general conclusions:

- Districts with the lowest transport intensity are Gelnica and Medzilaborce, where the weekly counts of passages of vehicles is only in few hundreds, whereas the districts with the highest transport intensity are Žilina and Svidník, where the weekly counts of passages of vehicles is above five hundred thousand;
- An occurrence of frequent local fluctuations within each year, which can be similar between some districts, albeit no common pattern is detected for all districts;
- It is not easy to assess differences between years based on the smoothness of individual curves. An overall drop of the intensity of the freight can be seen in 2021 that could be caused by an economic slowdown due to the pandemic (Covid-19) months;
- A sudden drop of the intensity of the freight occur in 2020 for 5 districts with the lowest volume of the road transport, which could be caused by the legislation change that took an effect on 1st September 2020[12]. However, other districts do not seem to be affected by the legislation's changes;
- A sharp downward spike ate around 35th week in 2022 is likely caused by the *3-day* working week.

We note that the weekly intensity of the freight is affected by the number of working days within particular weeks, i.e. there is a legislation in Slovakia that only allows a restricted heavy vehicle transportation during weekends and bank holidays[13]. The other undesired cause of changes of the freight's intensity could be legislation changes related to the restructuring of monitored sections of the road network. However, the legislation changes are usually implemented from the following year, i.e. 1st January, except in 2020.

The drawback of data on road transportation is that it also include vehicles that are only transitory through Slovakia, i.e. the transportation of goods by these freight does not contribute to local economies in districts. It is not straightforward how to exclude a transitory freight from original data. It would require a complex algorithm, which could lead to some inaccuracies due

[12] Source: § 35 ods. 2 písm. a) zákona č. 474/2013 Z. z. o cestnej premávke (in Slovak).
[13] Source: § 39 zákona č. 8/2009 Z. z. o cestnej premávke (in Slovak).

to the complex European legislation that requires mandatory breaks for drivers of heavy vehicles[14]. However, it is subject to further research.

A further exploratory analysis of the observed curves, which can uncover patterns of variations between years, is carried out through the familiar technique of functional analysis of variance, refer to Section 4.4 below.

4.1.4 Industrial production

In this subsection, we describe the data source for industrial production, which is used as a *real-valued* response in the econometric model. The main focus of the econometric model is to assess the dynamics between the intensity of the freight and the industrial production by districts in Slovakia. The econometric model is estimated for each year, respectively.

The source of data is the annual report on Yearbook of Industry of the Slovak Republic, which is published by Statistical Office of the Slovak Republic[15]. Generally, from the methodological perspective, Statistical Yearbook of Industry is one component part of publication system of Statistical Office of the Slovak Republic that provides indicators on the whole aggregate of industrial outputs by economic activities, by districts and regions in Slovakia. The annual statistics in the report are published with a *two-year* lag, i.e. the industrial production for 2023 is published in the 2025 annual report.

Table 5, in Appendix, shows the industrial production by districts in Slovakia for 2018-2022. We note that for districts of Levoča and Medzilaborce, the industrial production is not provided due to data privacy obligations, i.e. if provided, it would be possible to identify the industrial production of specific manufacturers since their major contributions to the whole economy of these districts. Hence, the econometric model only includes observations of 69 districts.

[14] Source: Employing road transport workers: driving time and rest periods.
[15] Source: *Ročenka priemyslu SR*; to see the list of the annually published Yearbooks of Industry of the Slovak Republic by Statistical Office of the Slovak Republic click here (in Slovak).

4.2 Spatial weights

In this section, we describe the data source that is used to obtain the spatial coordinates, i.e. the latitude and the longitude, which locate the centroid of districts in Slovakia. Moreover, the estimation of spatial weights using the generalised *queen-type* contiguity is also shown.

We collect the spatial coordinates from the geographical database *GeoNames*[16], which is available free of charge. The database contains geographical data such as names of places in various languages, elevation and population. The latitude and the longitude, which are the geographical measures of the spatial location, are in the format WGS84 (World Geodetic System 1984). We use the R software, the package *geonames*[17], to connect and download the longitude and the latitude from the *GeoNames* database for districts in Slovakia.

Table 6, in Appendix, displays the spatial coordinates of districts and regions in Slovakia. There are 79 districts which falls within 8 regions in Slovakia. Note that we keep the original *place-names* as shown in the *GeoNames* database. The other information[18] is related to the population, size and density, which provide some additional information on districts.

For estimation of spatial weights, we use a method that is proposed by Anselin and Morrison (2019). The authors create a website that contains a description of the methodology and its application in the R software. It is based on the generalised concept of contiguity between spatial points, i.e. the latitude and the longitude. In the following text, we briefly describe the methodological steps and its application on spatial coordinates of districts in Slovakia to compute the spatial weights.

Initially, the spatial coordinates are converted to Thiessen polygons, which create a connected map of neighbourhoods[19]. The algorithm to create the Thiessen polygons from the spatial coordinates is proposed by Lee et al. (1980), which is coded in the package *deldir*[20] in R. In the next phase, the estimation procedure computes the *queen-type* contiguity weights between

[16] https://www.geonames.org/.
[17] https://cran.r-project.org/web/packages/geonames/geonames.pdf.
[18] Source: https://sk.wikipedia.org/wiki/Zoznam_okresov_na_Slovensku.
[19] We note that the map of neighbourhoods created by Thiessen polygons is an approximation of the actual map of districts in Slovakia.
[20] The function *deldir* from the R package *deldir*; https://cran.r-project.org/web/packages/deldir/deldir.pdf.

these polygons[21]. The advantage of this approach is that it deals with the problem of the uneven distribution of neighbours.

Figure 8 shows the estimated contiguity between districts in Slovakia.

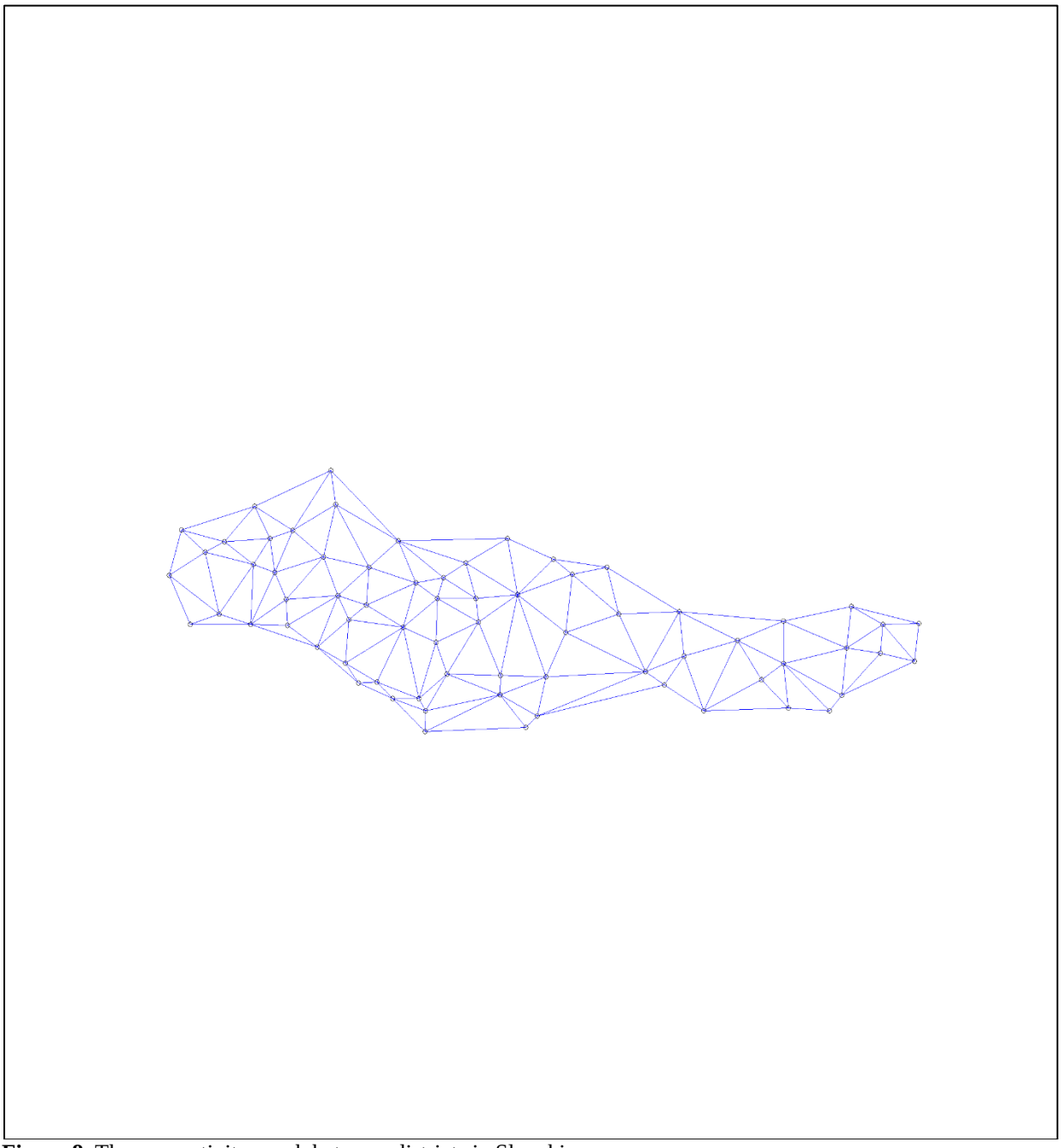

Figure 8. The connectivity graph between districts in Slovakia
Source: Author's construction

[21] We use the function *st_relate* from the R package *sf*; https://cran.r-project.org/web/packages/sf/sf.pdf.

Figure 8 shows that we have a fully connected graph, with no isolated (unconnected) subgroups of districts. The main reason to use the generalised version of the *queen-type* contiguity, i.e. the centroids of districts for estimation of neighbourhoods, is that we merge 5 districts of Bratislava and 5 districts of Košice, for which we use the centroids of Bratislava I and Košice I, respectively. Moreover, its advantage is that districts that are close to each other, i.e. based on its distance between centroids, but share no borders due to uneven borderlines are also connected.

Table 7 shows the distribution of the number of neighbourhoods associated with the number of districts.

No. of neighbours	3	4	5	6	7	8	9
No. of districts	3	16	29	14	7	1	1

Table 7. Distribution of the number of neighbourhoods associated with the number of districts
Source: Author's construction

Table 7 shows that the generalised *queen-type* contiguity approach creates 3 least connected districts, i.e. Námestovo, Skalica and Sobrance, with 3 links and 1 most connected region, i.e. Detva, with 9 links.

It follows that the spatial weights are then computed for estimated neighbourhoods as in eq. (1.7).

4.3 Descriptive statistics

In this section, we compute the basic descriptive statistics of functional observations $\chi_i(t)$ for 2018-2023. Figures 9-10 display the overall mean function and the standard deviation function for 2018-2023, respectively.

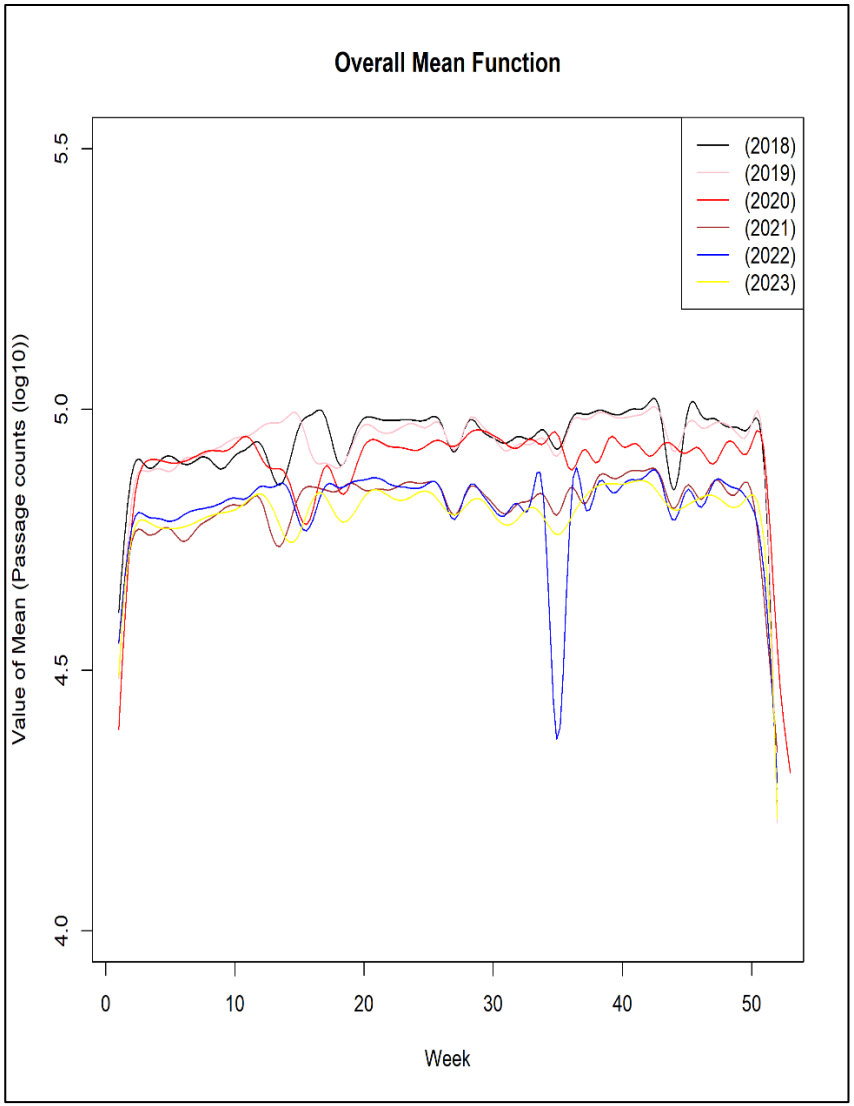

Figure 9. The mean function $\overline{\chi}(t)$ for 2018-2023
Source: Author's construction

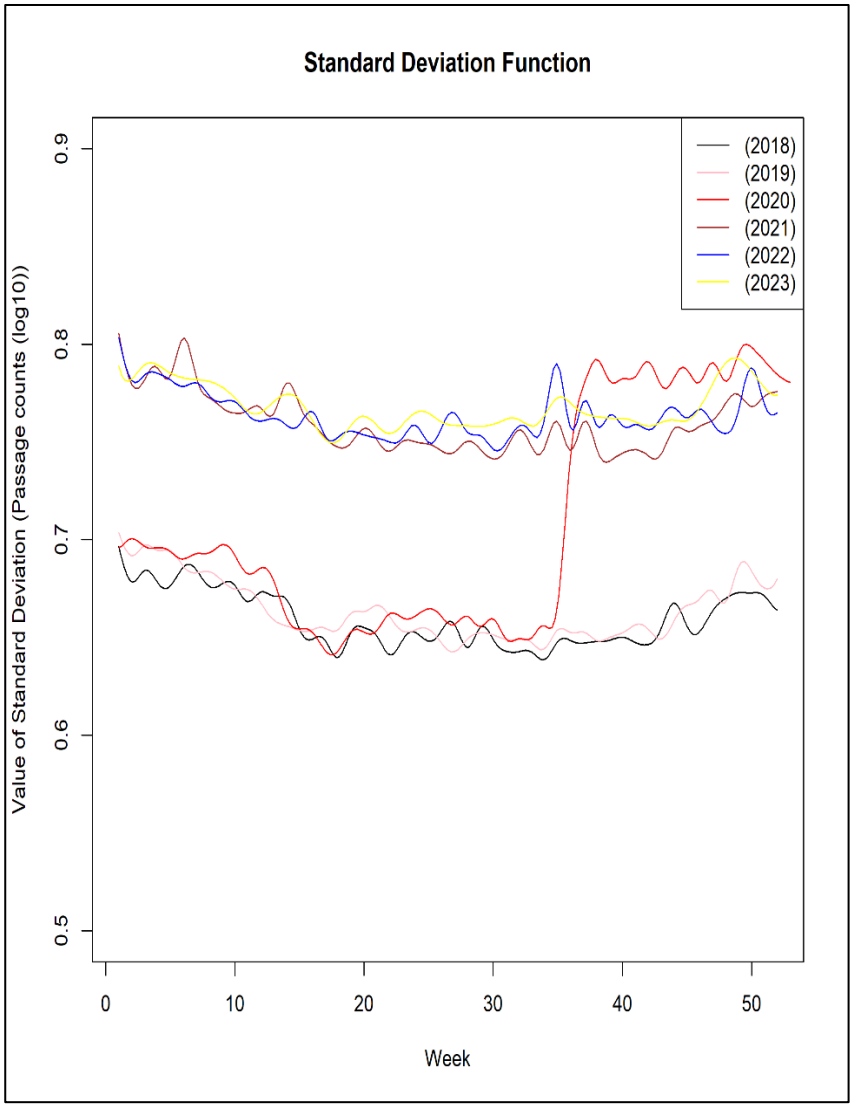

Figure 10. The standard deviation function $sd(\chi_i(t))$ for 2018-2023
Source: Author's construction

Figure 9 shows that the overall mean function of the intensity of the freight is much lower for 2021-2023 compared to the previous years, which is likely caused by the economic crisis of the Covid19 pandemic. Moreover, Figure 10 shows that the standard deviation functions are much higher for 2021-2023, including the second half 2020, i.e. the spread between the functional observations significantly increases for these years.

4.4 Functional analysis of variance

In the further exploratory analysis, we show the application of the methodology, of which the estimation is outlined in Section 3.2.1, of the functional analysis of variance. We estimate the parameters that are formalised in eq. (3.5), i.e. the overall mean function $\mu(t)$ across all 71 curves (districts) and the specific effect functions $\alpha_c(t)$ for 8 regions, where $c = 1, \dots, 8$, for 2018-2023, respectively. The terms $\alpha_c(t)$ can be interpreted as departures from the overall mean function $\mu(t)$ specific for each region c. We note that each district falls within specific region.

Table 8 shows the categorisation of observed curves, i.e. a number of districts that falls within specific regions.

Category	Region	Number of curves (districts)
1	Banská Bystrica	13
2	Bratislava	4
3	Košice	7
4	Nitra	7
5	Prešov	13
6	Trenčín	9
7	Trnava	7
8	Žilina	11

Table 8. Regions in Slovakia with a number of corresponding observed curves (districts)
Source: Author's construction

We note that eq. (3.5) is redefined into eq. (1.4), where the columns of the design matrix $\mathbf{Z} = z_{i(c)j}$ are constructed as a binary indicator, which represent specific regions with its corresponding districts. The terms $\alpha_c(t)$ are estimated as the regression coefficient functions for each region. A vector of ones in the first column of $\mathbf{Z}$ serves for estimation of the overall mean function $\mu(t)$.

Figures 11-16 show the estimated regression coefficient functions $\hat{\alpha}_c(t)$ for 8 regions and the estimated overall mean function $\hat{\mu}(t)$ for 2018-2023, respectively.

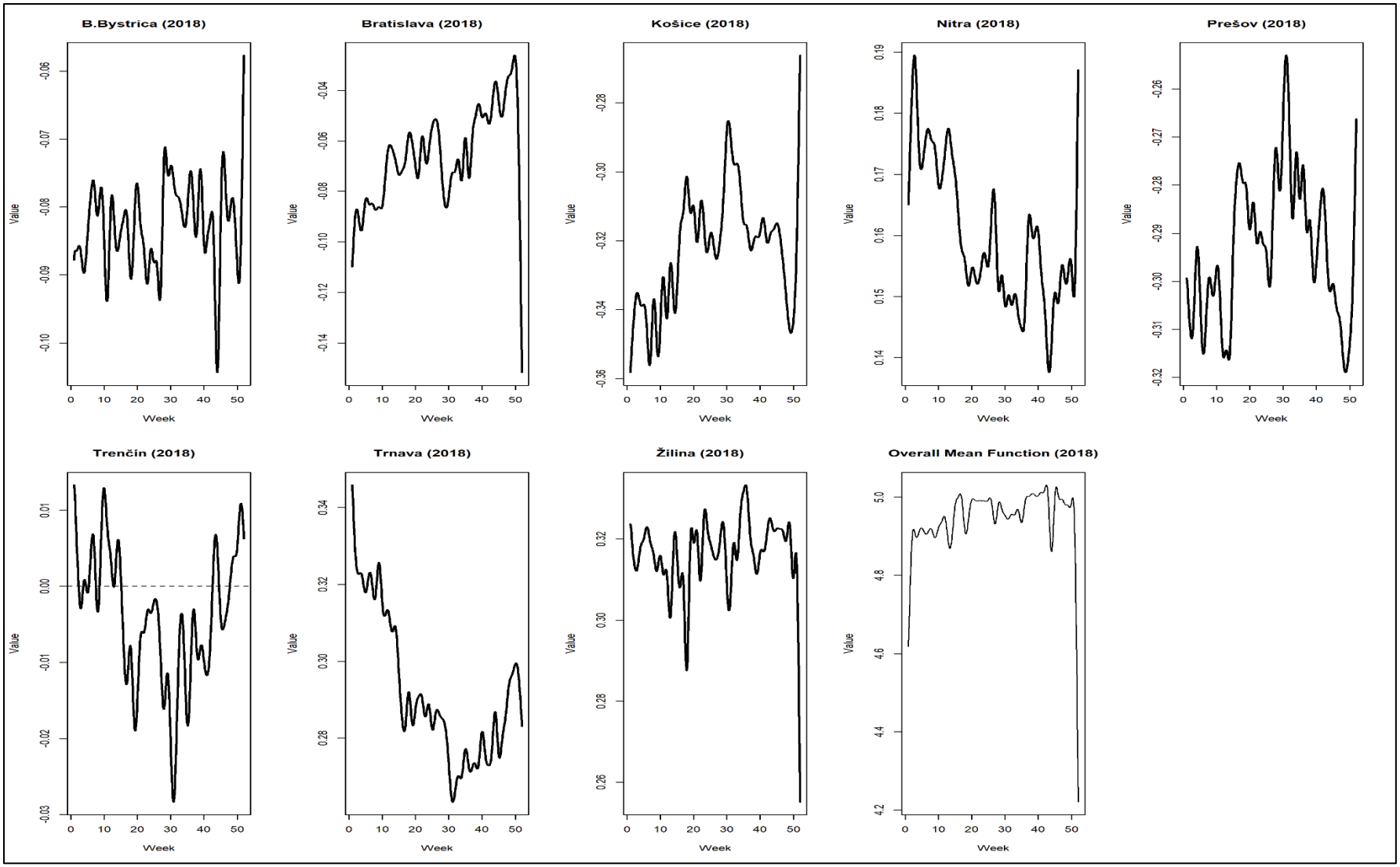

Figure 11. The estimated regression coefficient functions $\hat{\alpha}_c(t)$ and the estimated overall mean function $\hat{\mu}(t)$ (2018)
Source: Author's construction

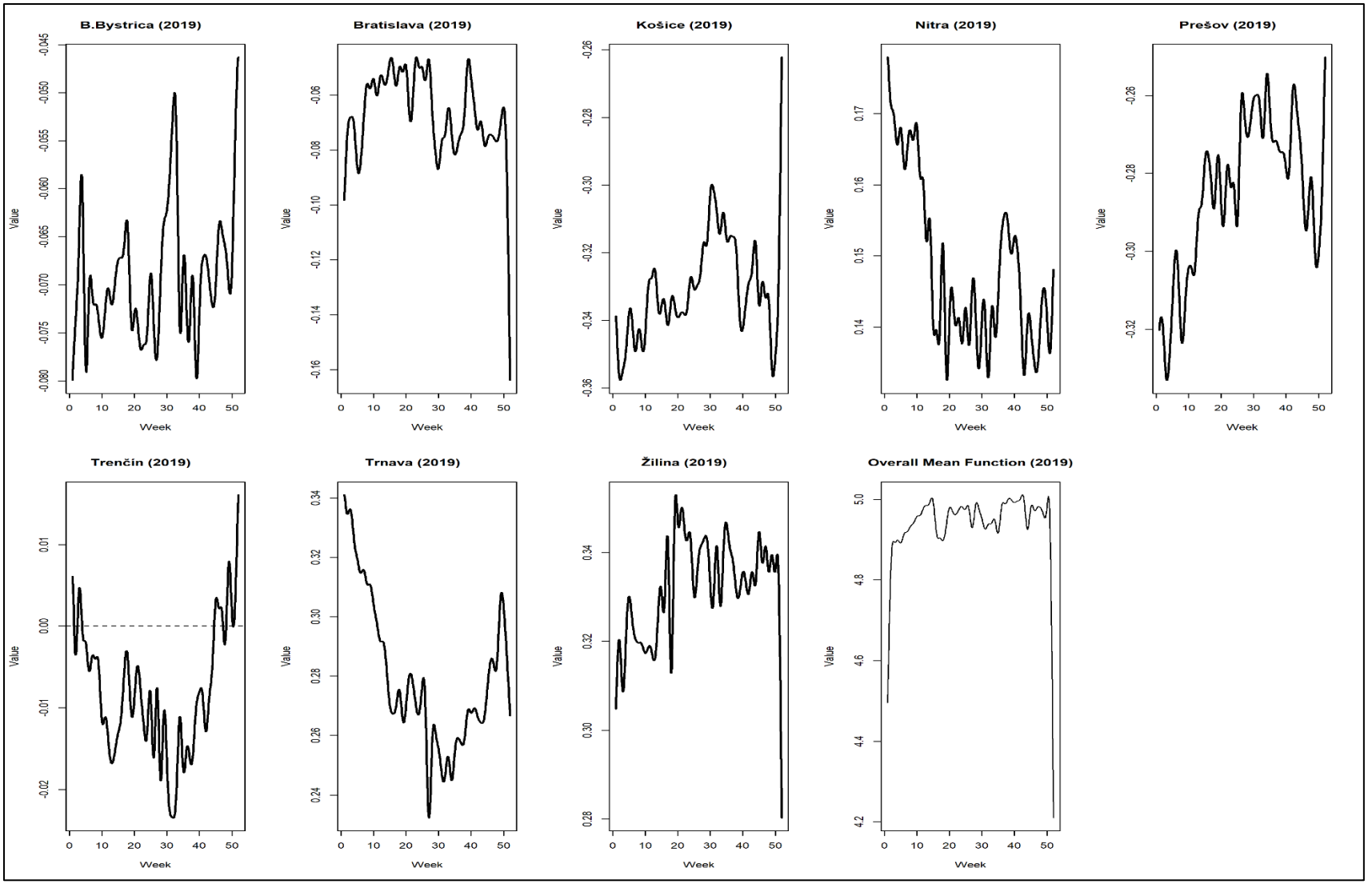

Figure 12. The estimated regression coefficient functions $\hat{\alpha}_c(t)$ and the estimated overall mean function $\hat{\mu}(t)$ (2019)
Source: Author's construction

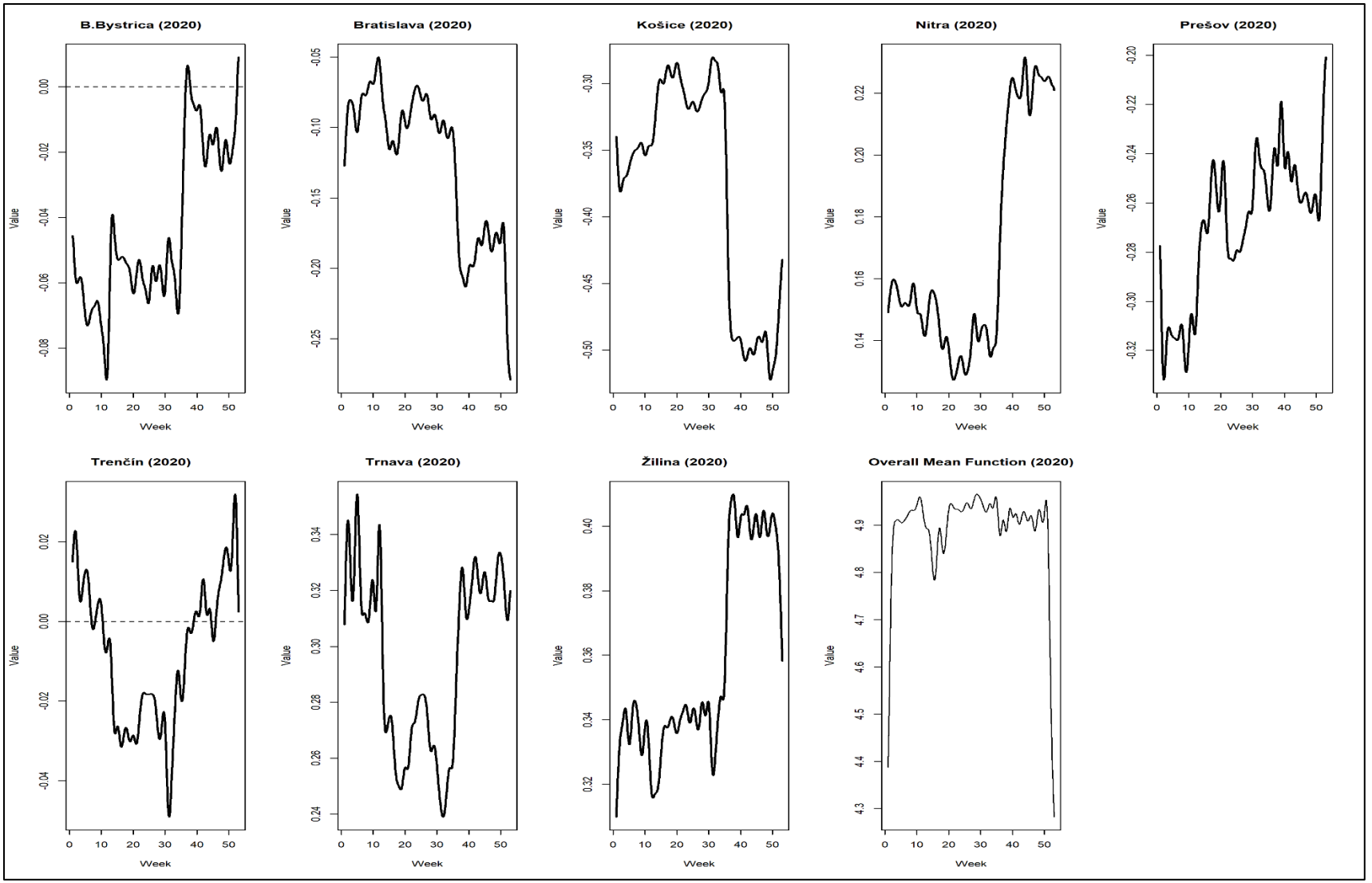

Figure 13. The estimated regression coefficient functions $\hat{\alpha}_c(t)$ and the estimated overall mean function $\hat{\mu}(t)$ (2020)
Source: Author's construction

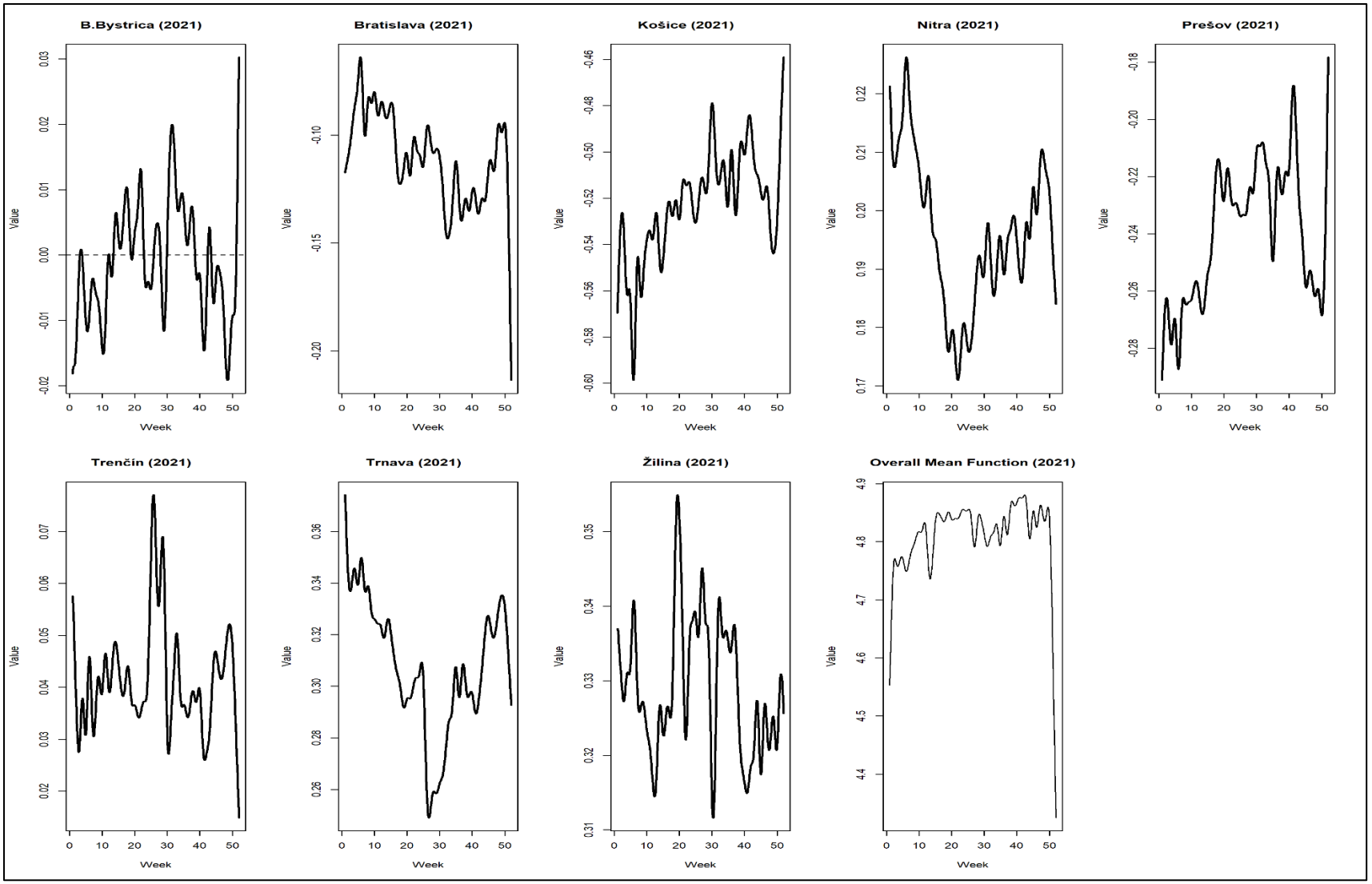

Figure 14. The estimated regression coefficient functions $\hat{\alpha}_c(t)$ and the estimated overall mean function $\hat{\mu}(t)$ (2021)
Source: Author's construction

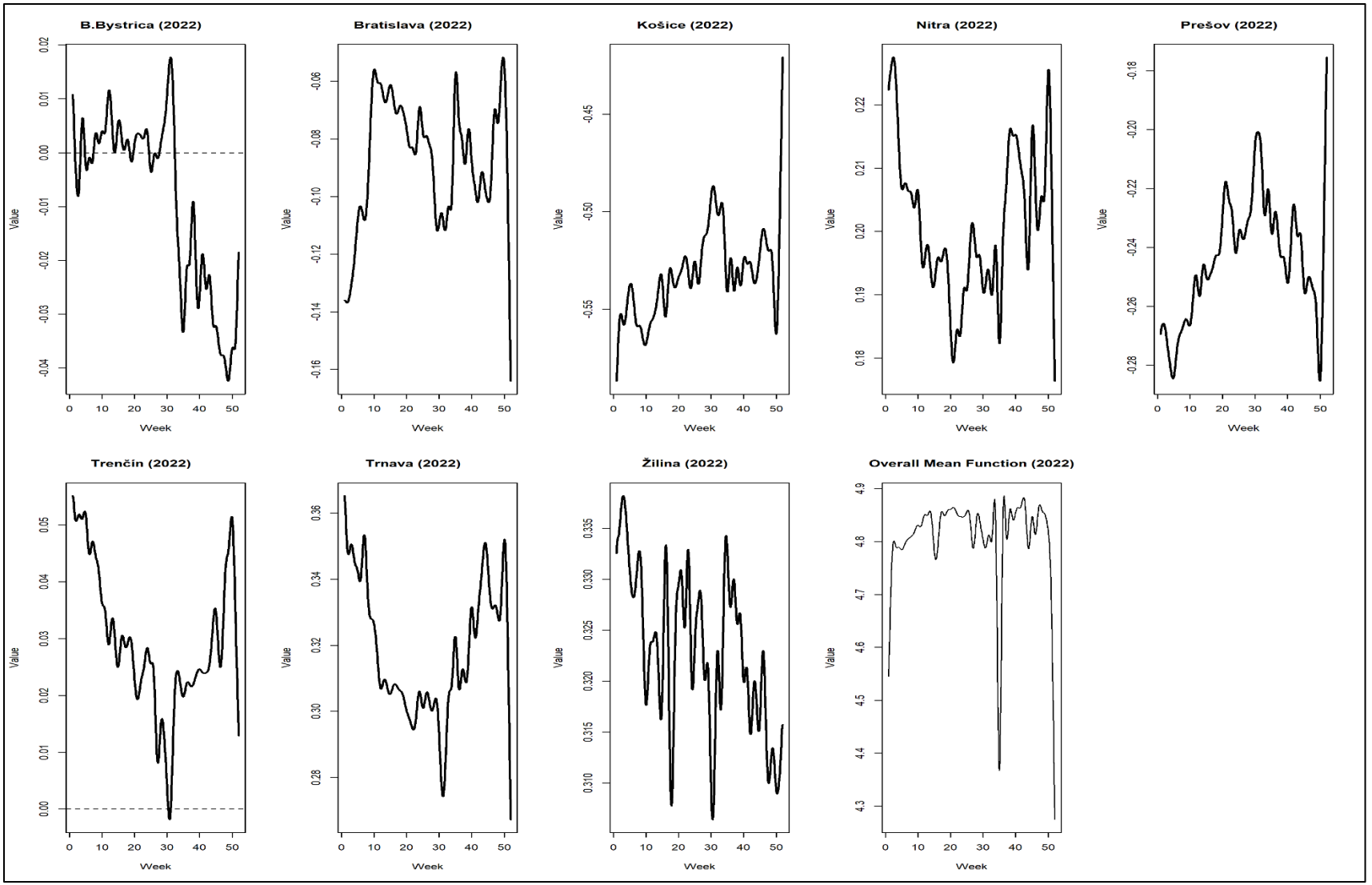

Figure 15. The estimated regression coefficient functions $\hat{\alpha}_c(t)$ and the estimated overall mean function $\hat{\mu}(t)$ (2022)
Source: Author's construction

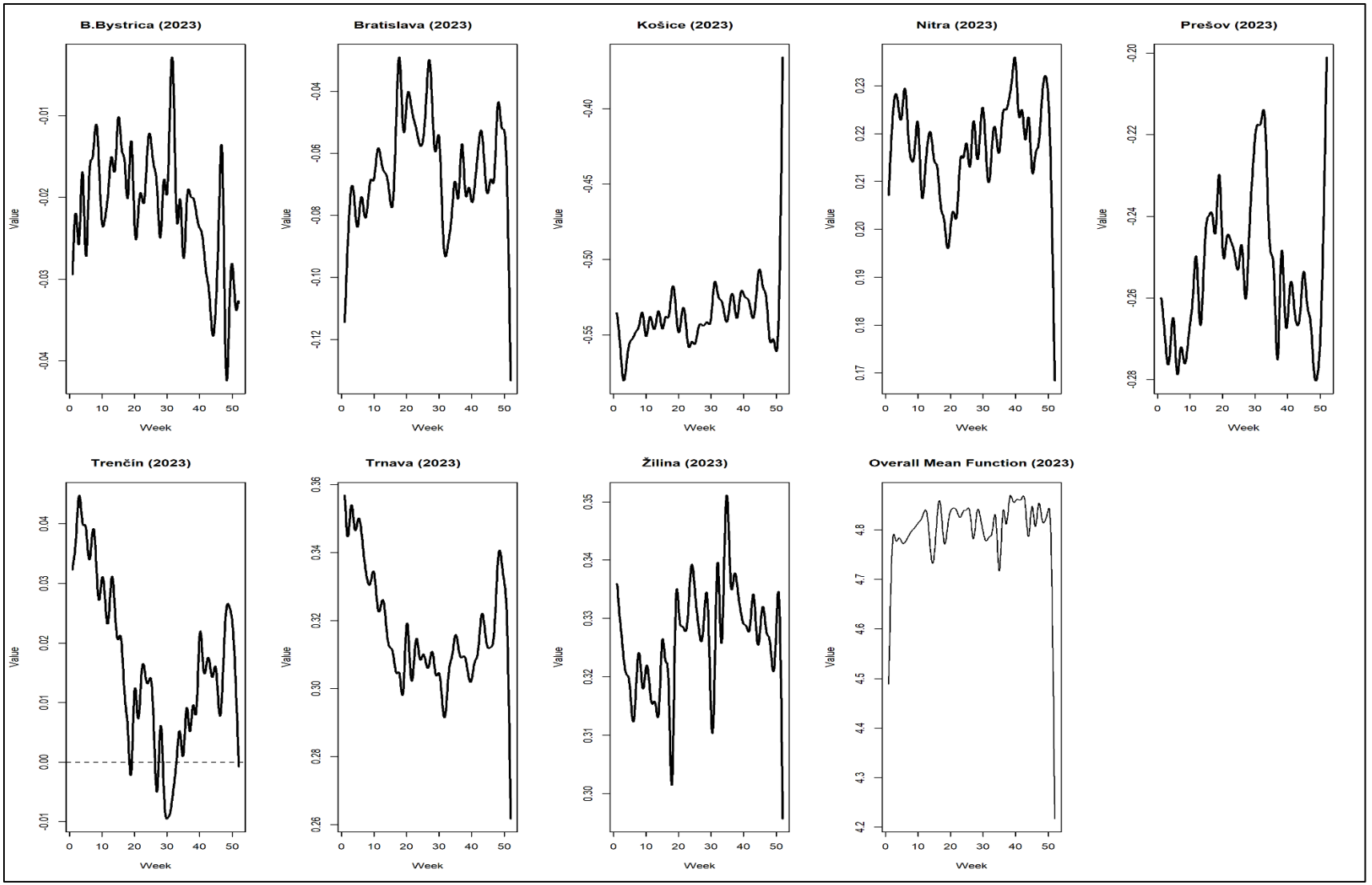

Figure 16. The estimated regression coefficient functions $\hat{\alpha}_c(t)$ and the estimated overall mean function $\hat{\mu}(t)$ (2023)
Source: Author's construction

In Figures 11-16, respectively, the plots 1-8 are the estimated regression coefficient functions $\hat{\alpha}_c(t)$, which correspond to the specific effect functions, for regions $c = 1, \dots, 8$ and the last plot is the estimated overall mean function $\hat{\mu}(t)$. The statistical significance of the estimated regression coefficient functions can be evaluated visually. If the value of the estimated functional parameter $\hat{\alpha}_c(t)$ crosses, or is close to, zero (dashed line), it can be deemed as statistically not significant from zero, i.e. there is no difference between the estimated overall mean function $\hat{\mu}(t)$ and the corresponding estimated regional effect function $\hat{\alpha}_c(t)$.

Based on Figures 11-16, we can draw the following conclusions

- The shape and the magnitude of the estimated functional parameters $\hat{\alpha}_c(t)$ are significantly different between regions, with no clear occurrence of seasonal patterns within regions;
- Figure 11 shows that in 2018 Trnava and Žilina have the largest volume of the intensity of the freight (values are above zero), whereas Košice and Prešov have the smallest volume (values are below zero). This also remains in 2019-2023 (Figures 12-16);
- The regions which have a similar intensity of the freight to the overall mean function (values are very close to zero), all through 2018-2023 (Figures 11-16), are Banská Bystrica, Bratislava and Trenčín;
- Figure 10 shows that there is a drop in the volume of the freight at around week 30 for Bratislava, Trenčín, Nitra and Trnava, whereas Banská Bystrica, Košice and Prešov show and spike. After the 30th week, there is a reversal of the trend;
- The volume of the freight significantly increases in the region Nitra over time (values are rising well above zero).

To emphasise one of the advantages of having a continuous estimated parameters $\hat{\alpha}_c(t)$ and $\hat{\mu}(t)$, instead of *point-wise* estimates as in the classical analysis of variance, is that we can identify and assess local variabilities of the regional intensity of the freight throughout time.

To gain a clearer picture of the deviation between the specific effect functions of regions and the overall mean function, it is worthwhile examining the estimated profiles $\hat{\mu}(t) + \hat{\alpha}_c(t)$. Figures 17-22 show the estimated profiles $\hat{\mu}(t) + \hat{\alpha}_c(t)$ for 8 regions, which are displayed as the

solid lines, and the estimated overall mean functions $\hat{\mu}(t)$, which are displayed as the dashed lines, for 2018-2023, respectively.

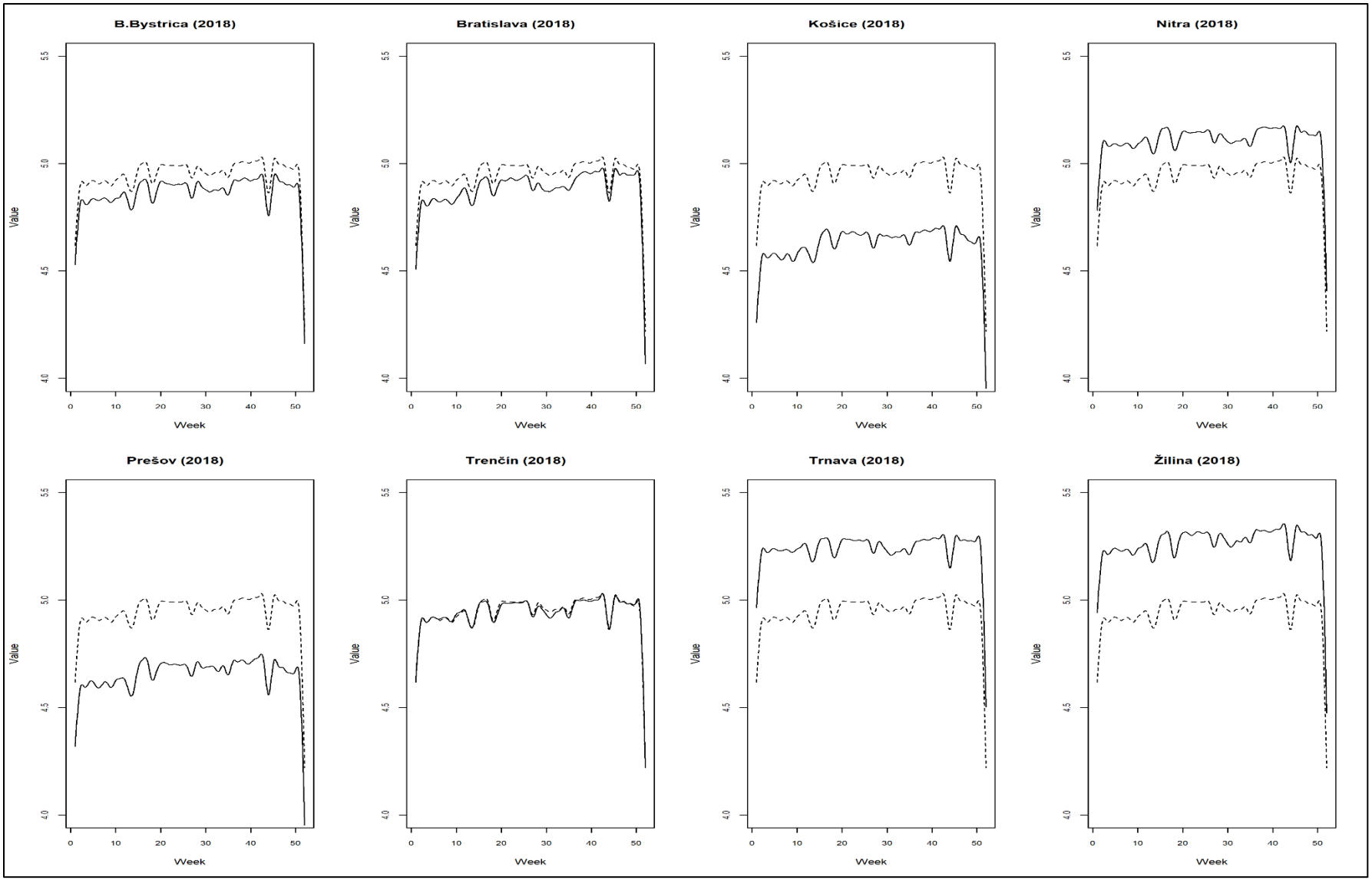

Figure 17. The estimated profiles $\hat{\mu}(t) + \hat{\alpha}_c(t)$ and the estimated overall mean function $\hat{\mu}(t)$ (dashed lines) (2018)
Source: Author's construction

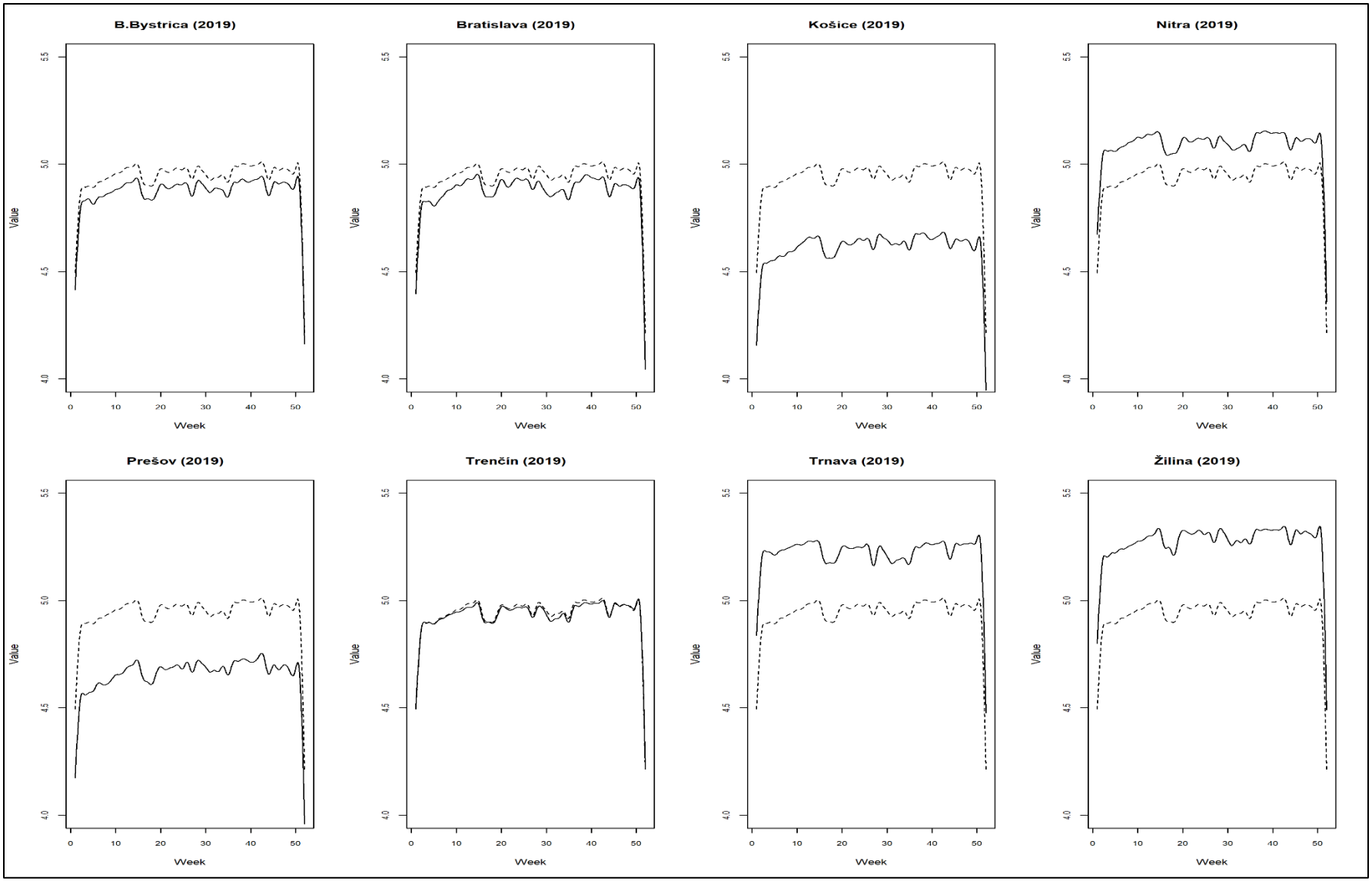

Figure 18. The estimated profiles $\hat{\mu}(t) + \hat{\alpha}_c(t)$ and the estimated overall mean function $\hat{\mu}(t)$ (dashed lines) (2019)
Source: Author's construction

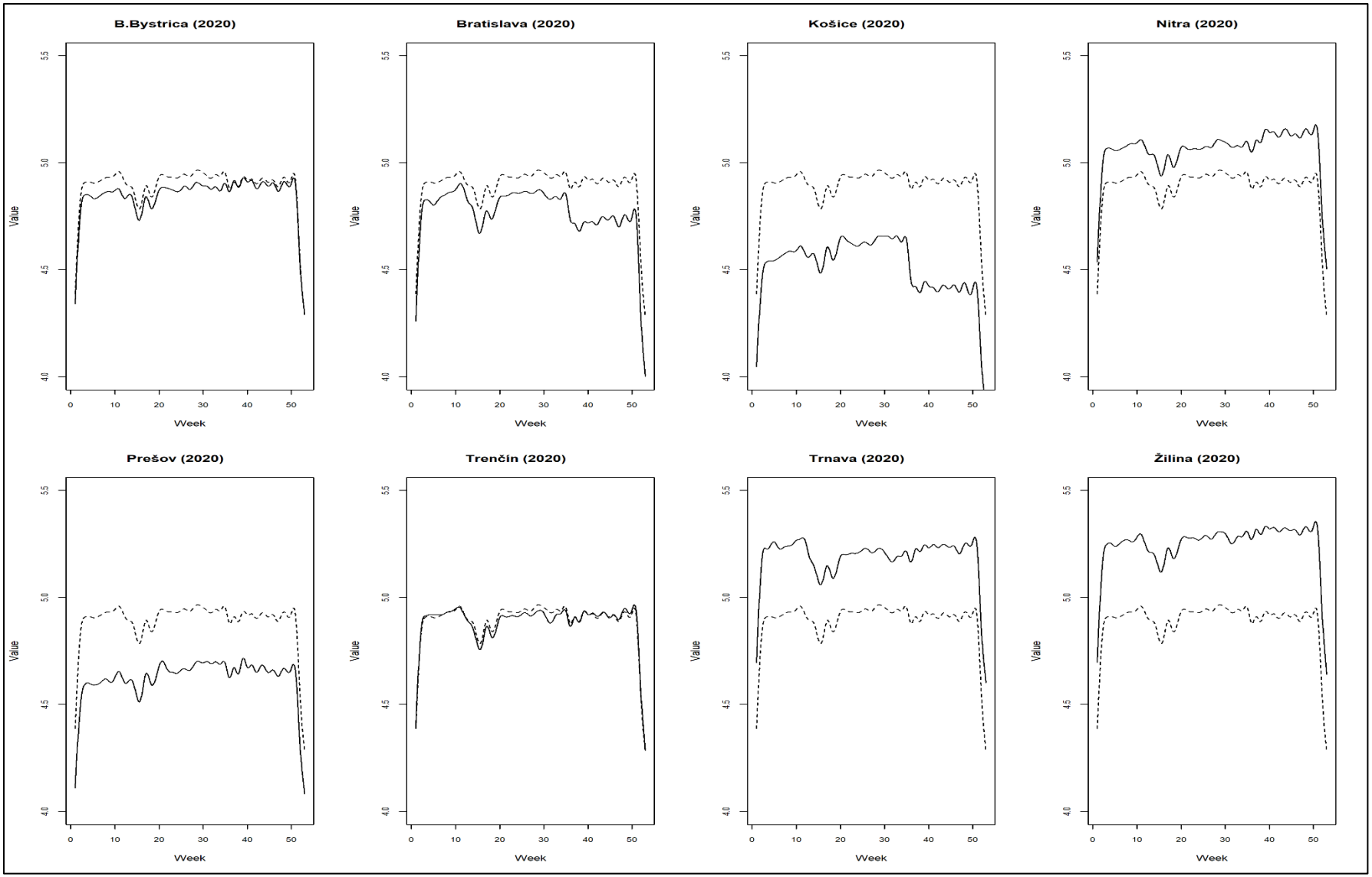

Figure 19. The estimated profiles $\hat{\mu}(t) + \hat{\alpha}_c(t)$ and the estimated overall mean function $\hat{\mu}(t)$ (dashed lines) (2020)
Source: Author's construction

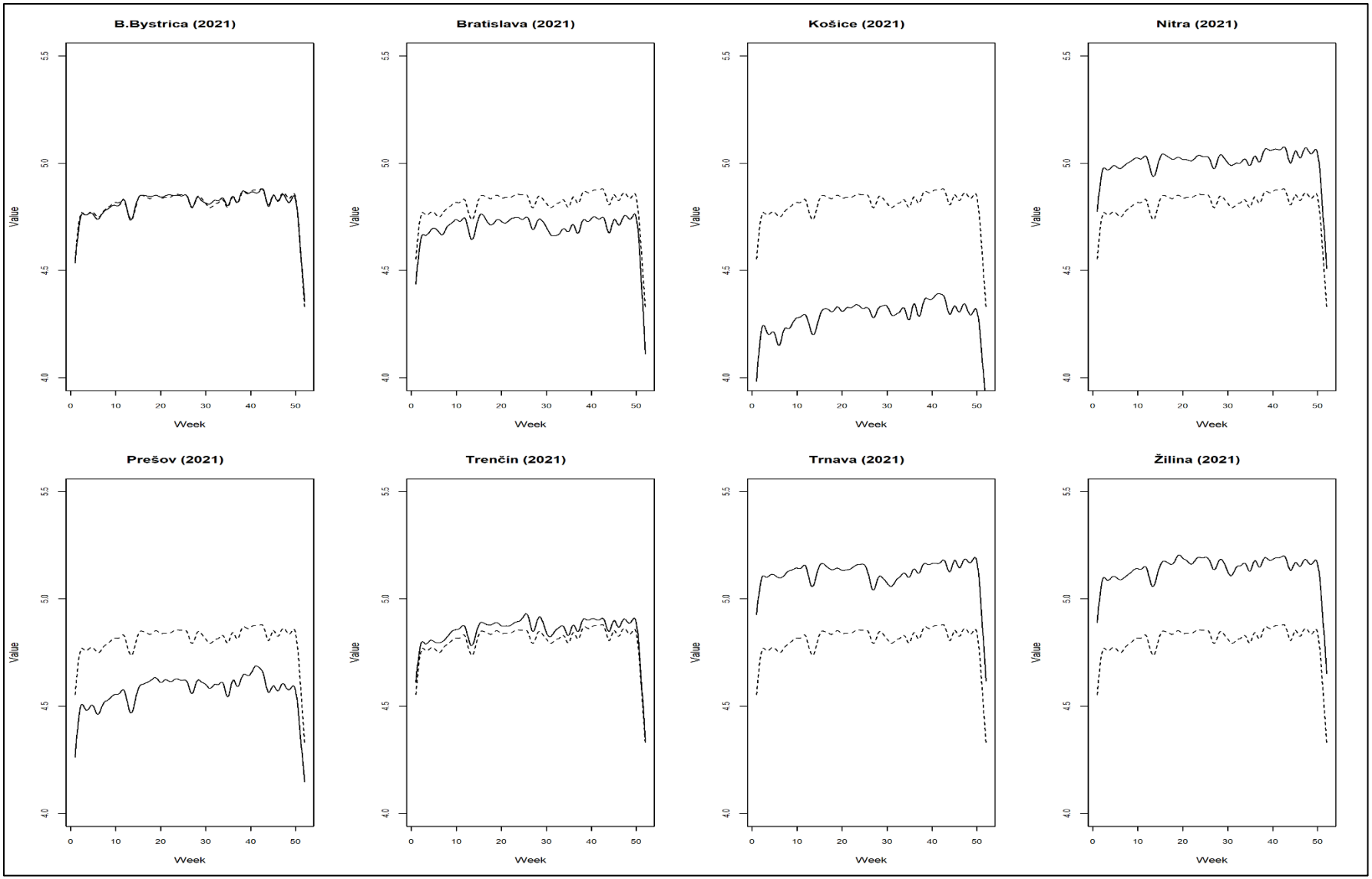

Figure 20. The estimated profiles $\hat{\mu}(t) + \hat{\alpha}_c(t)$ and the estimated overall mean function $\hat{\mu}(t)$ (dashed lines) (2021)
Source: Author's construction

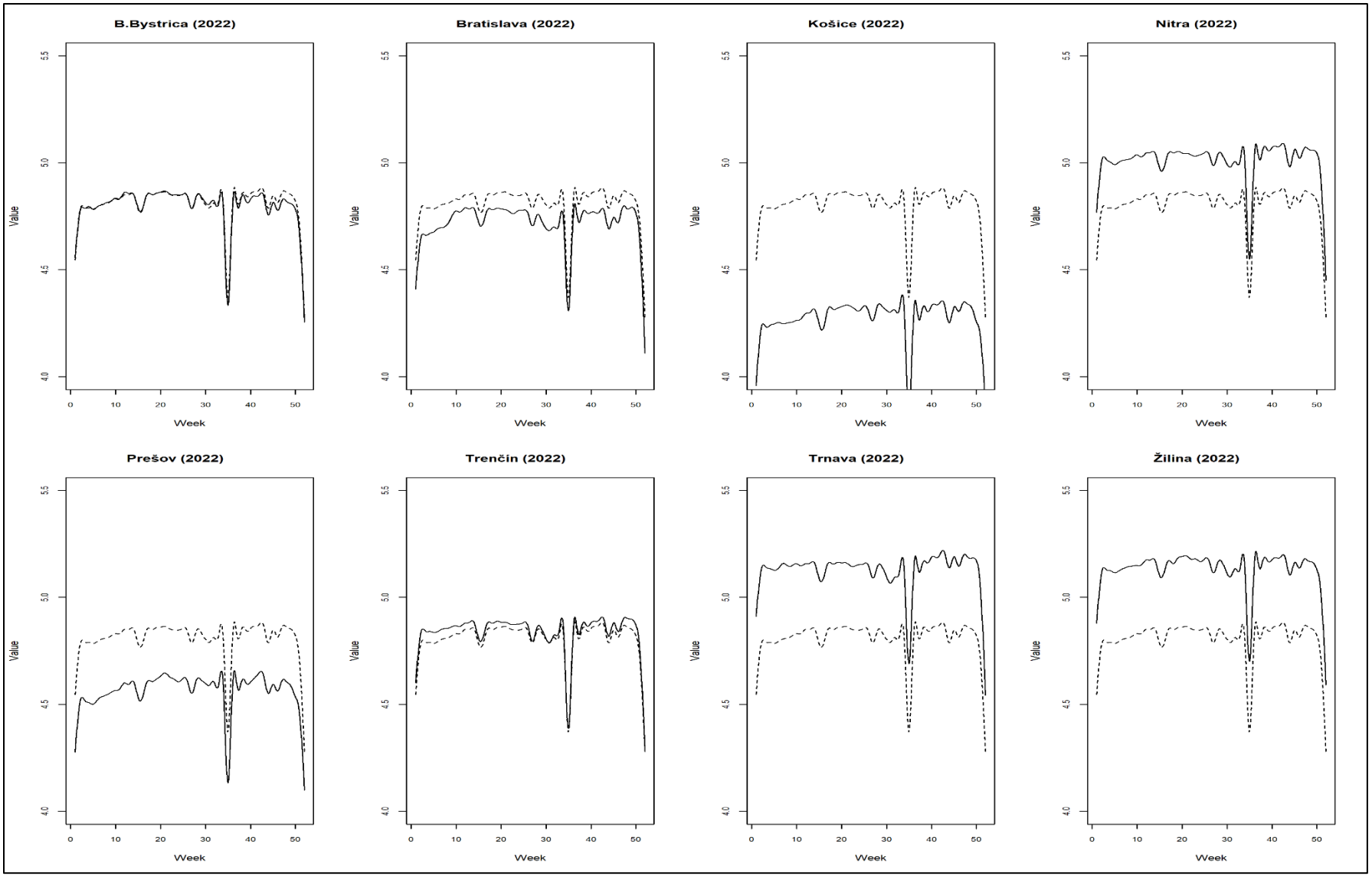

Figure 21. The estimated profiles $\hat{\mu}(t) + \hat{\alpha}_c(t)$ and the estimated overall mean function $\hat{\mu}(t)$ (dashed lines) (2022)
Source: Author's construction

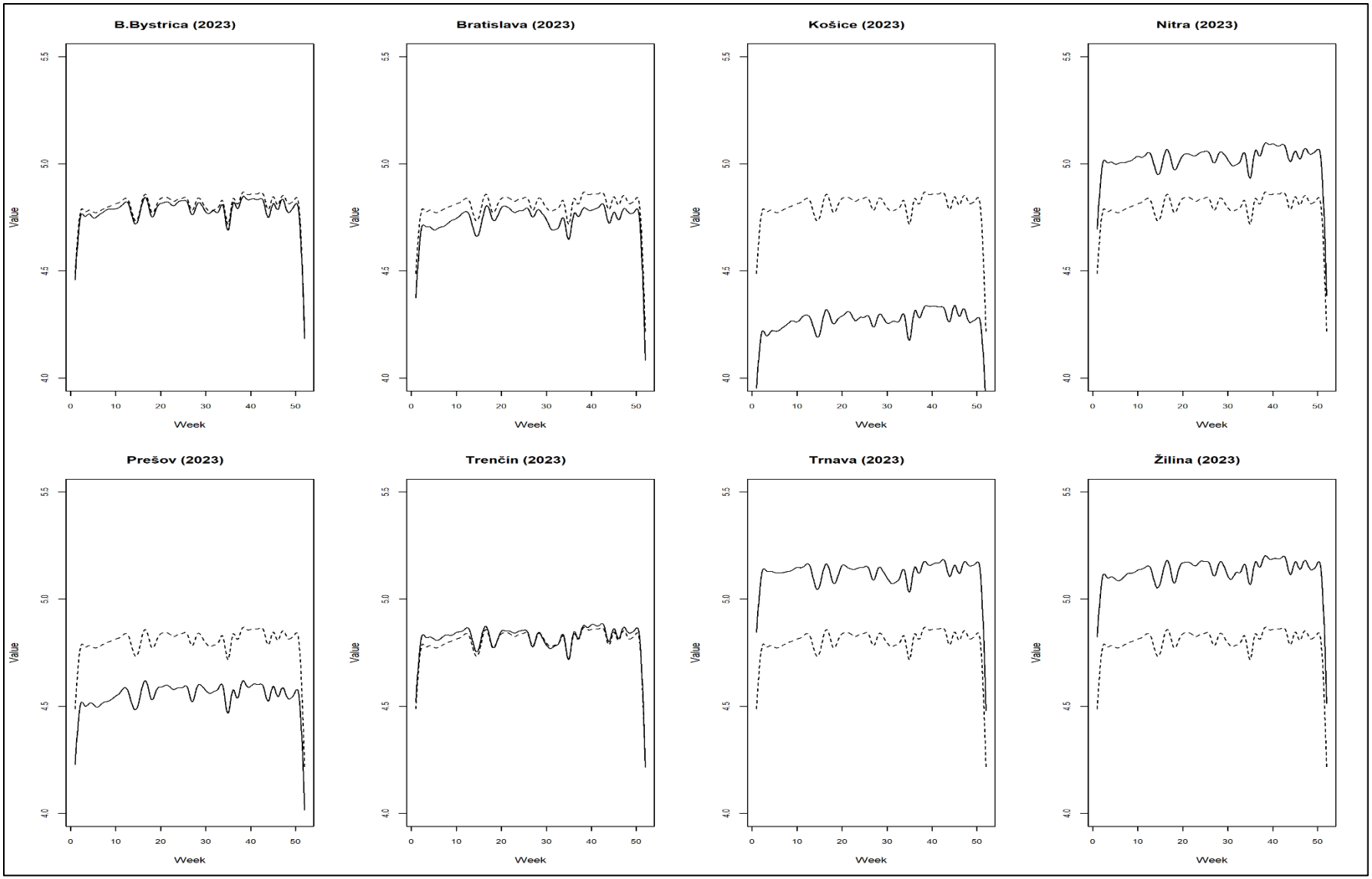

Figure 22. The estimated profiles $\hat{\mu}(t) + \hat{\alpha}_c(t)$ and the estimated overall mean function $\hat{\mu}(t)$ (dashed lines) (2023)
Source: Author's construction

Figures 17-22 show a clearer picture of differences between the regional intensity of the freight and the overall mean. Additionally, we can draw the following conclusions

- There are noticeable differences between some regional profile functions and the overall mean function, i.e. the regions with the above average intensity of the freight are Nitra, Trnava and Žilina, whereas the regions below average intensity of the freight are Košice and Prešov. No significant differences between the profile functions and the overall mean function is shown in Banská Bystrica, Bratislava and Trenčín;
- The other interesting feature that can be identified is if we compare Figures 17-22 with Table 5 (in Appendix; Industrial production by districts and regions in Slovakia for 2018-2022). The volume of the freight does not reflect the volume of the region's industrial production. For example, Bratislava has the largest industrial production but has a below average volume of the freight across 2018-2022.

In general, we can conclude that the empirical results show very different and complex seasonal patterns of variations and significant differences of the freight intensity between and within regions.

We note that the theoretical framework of functional analysis of variance does not take an account of spatial dependencies between districts. The phenomenon of spatial dependencies can have a significant effect on the estimated functional parameters, which is briefly discussed in Horváth and Kokoszka (2012). Further research is required in the area of functional analysis of variance, which incorporate spatial effects within its theoretical framework.

4.5 Econometric model: functional SEM

In this section, we show the estimated parameters of the econometric model, which is defined within the theoretical framework of the functional spatial error model (SEM). The main focus of the econometric model is to capture the dynamics of the annual industrial production by the intensity of the freight, where the observation units are districts in Slovakia. The intensity of the freight is expressed as a mathematical function and the annual industrial production is a *real-*

valued observation, which are both observed for each district. A theoretical framework of the functional SEM is defined in Section 3.3.2.

The functional regression parameter $\beta(t)$ and the spatial autoregressive parameter ρ are estimated by the generalised method of least squares, which uses the iterative procedure to find the optimal solution for $\beta(t)$ and ρ, refer to Section 3.3.2.

The *real-valued* response variable y_i is the annual industrial production and the functional covariate contains the observed curves $\chi_i(t)$, where $i = 1, \dots, 69$[22] refers to the number of districts. The intensity of the freight, i.e. a functional covariate $\chi_i(t)$, is observed on the continuous domain t, refer to Section 4.1.3 for details. The spatial weights between districts are computed as the $n \times n$ *row-standardised* $\mathbf{W}$ matrix, refer to Section 4.2. The econometric model is estimated for 2018-2022, respectively.

It follows that we estimate eq. (3.26) for $\hat{\mathbf{b}}$ through the iterative procedure, which is shown in Section 3.3.2. In the iterative procedure, we require to select a range of $\hat{\rho}$, for which eq. (3.26) is computed. We select the range of $\hat{\rho} = \langle -3,3 \rangle$ and estimate a set of models with corresponding generalised *cross-validation* scores, refer to eq. (3.27).

Figures 23-27 show the estimated spatial autoregressive parameters $\hat{\rho}$ against the generalised *cross-validation* scores of the estimated models for 2018-2022, respectively.

[22] Note that the observations for districts Levoča and Medzilaborce are excluded since its industrial productions are not published due to data privacy restrictions, see Section 4.1.4.

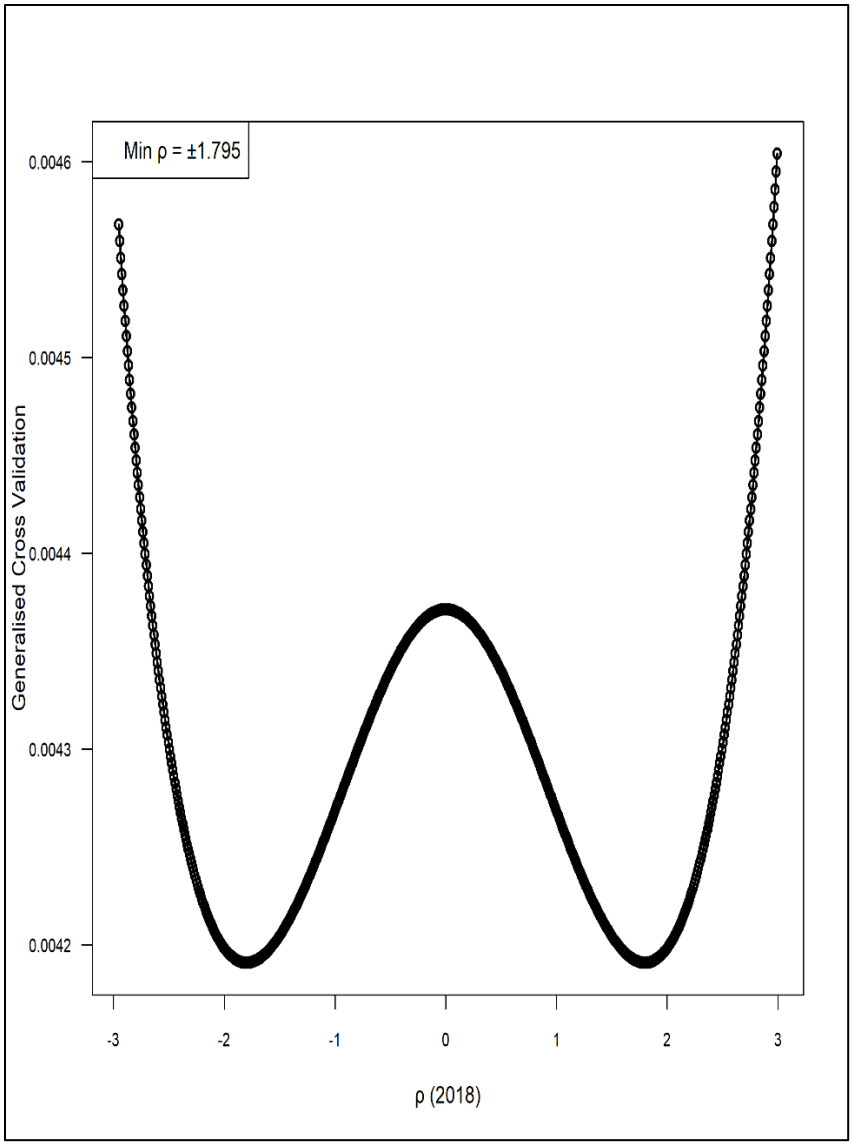

Figure 23. The estimated spatial autoregressive parameter $\hat{\rho}$ (2018)
Source: Author's construction

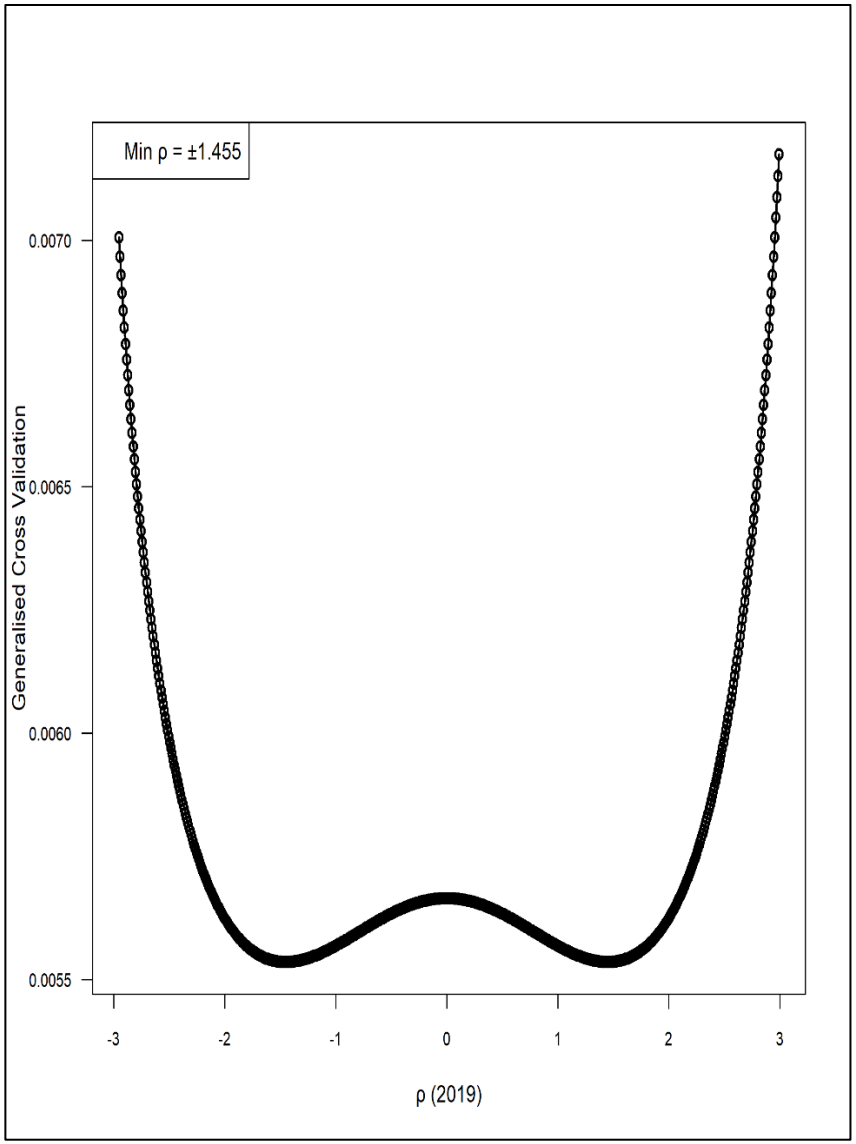

Figure 24. The estimated spatial autoregressive parameter $\hat{\rho}$ (2019)
Source: Author's construction

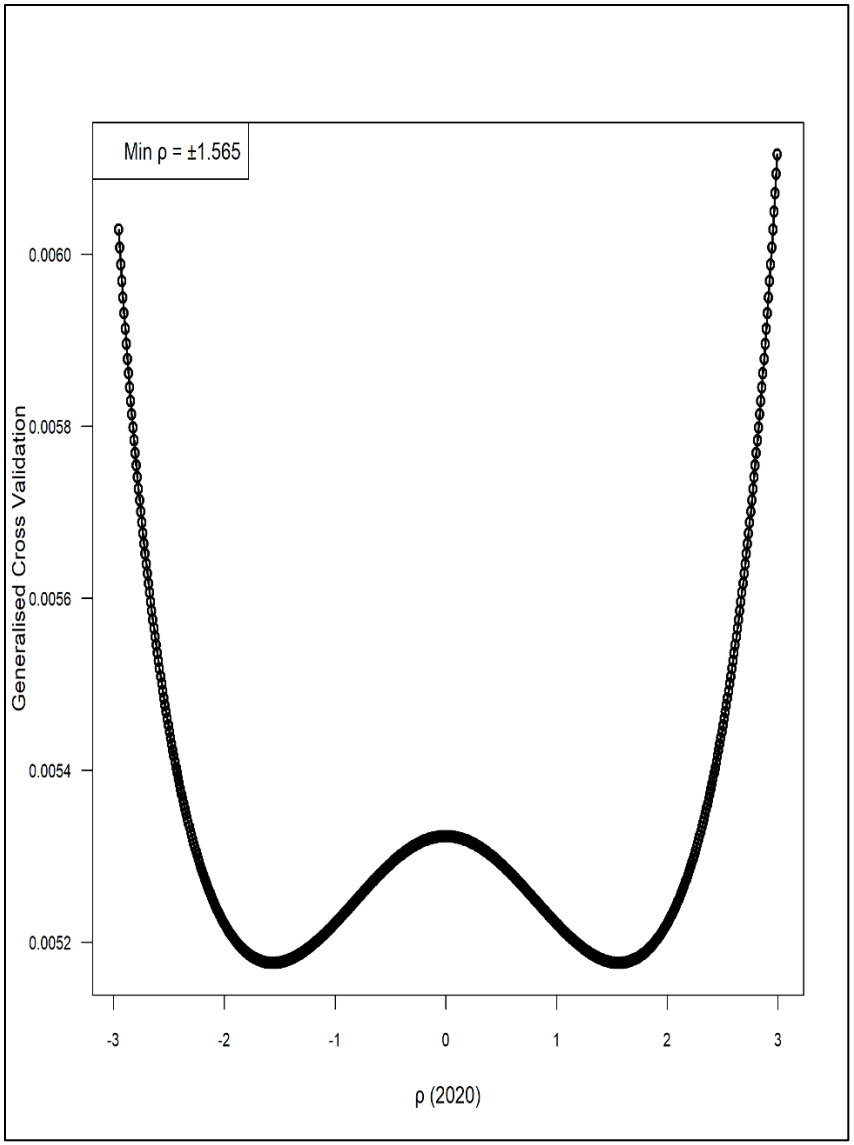

Figure 25. The estimated spatial autoregressive parameter $\hat{\rho}$ (2020)
Source: Author's construction

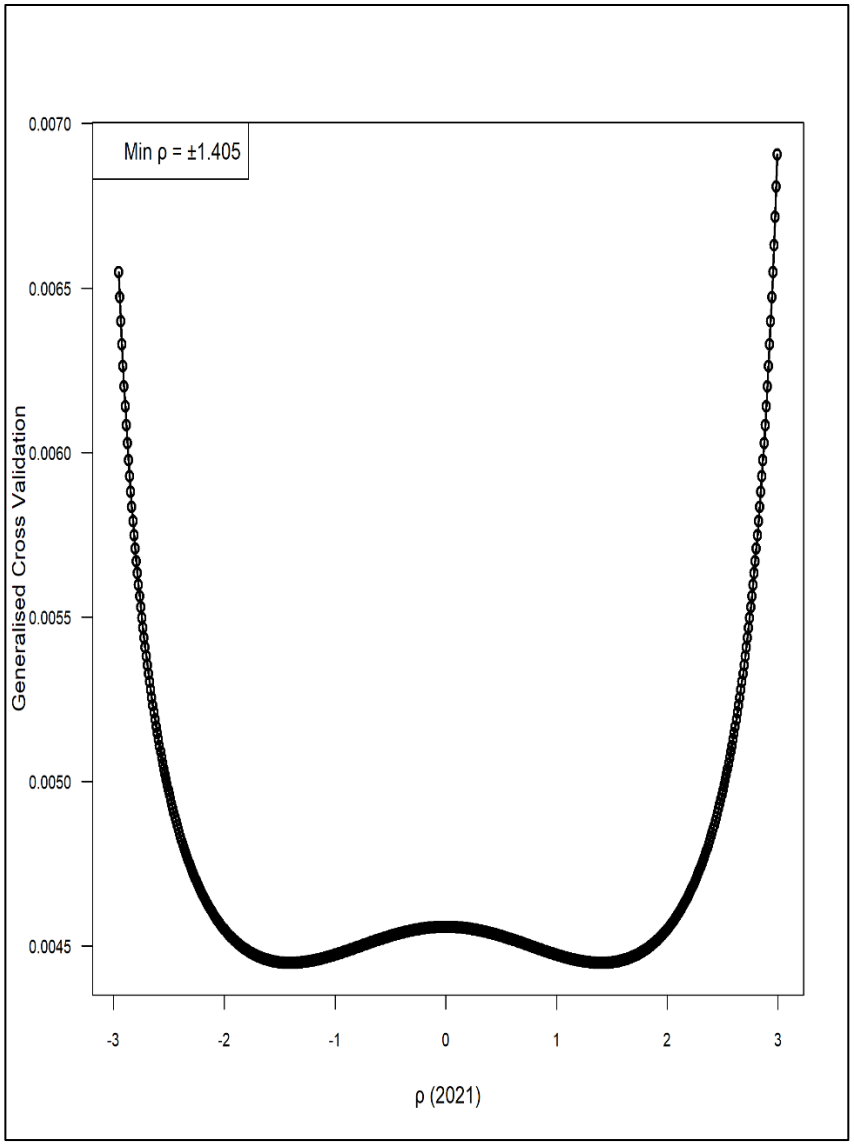

Figure 26. The estimated spatial autoregressive parameter $\hat{\rho}$ (2021)
Source: Author's construction

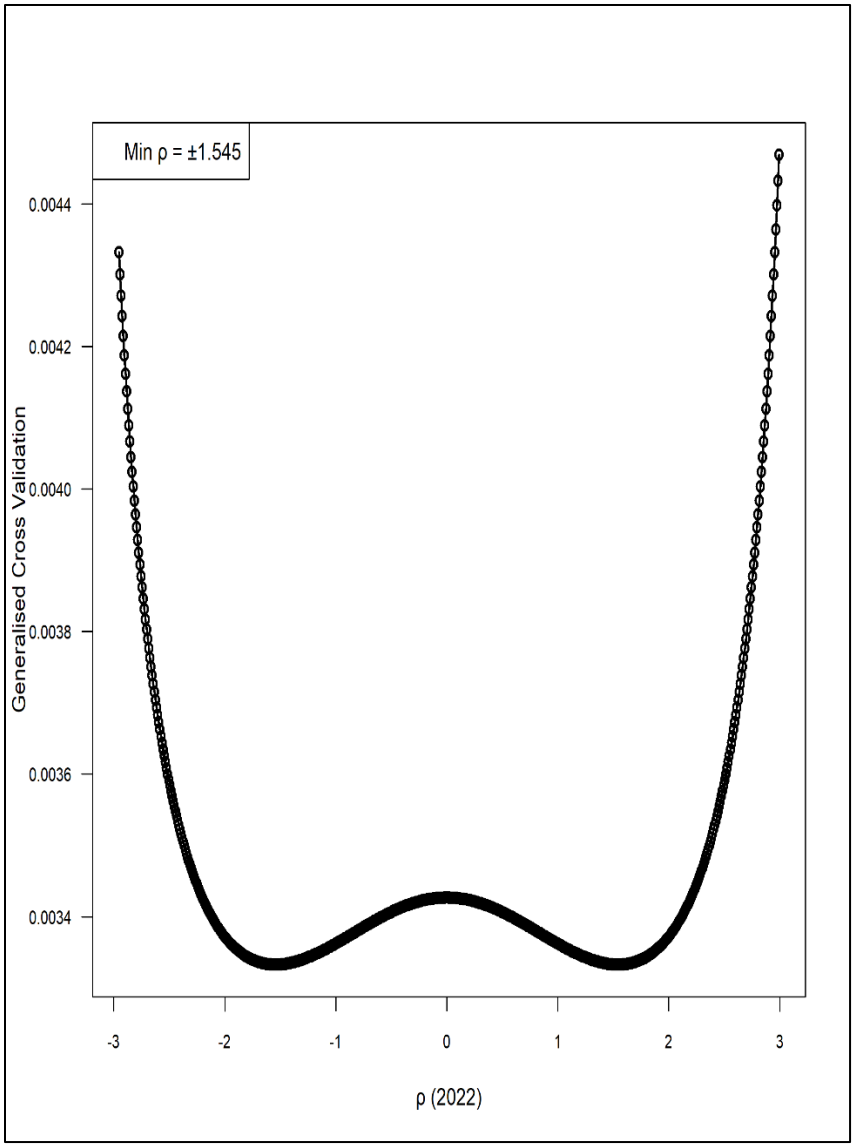

Figure 27. The estimated spatial autoregressive parameter $\hat{\rho}$ (2022)
Source: Author's construction

Figures 23-27 show that the estimated spatial autoregressive parameters $\hat{\rho}$ have similar shapes across 2018-2022. The magnitude of the generalised *cross-validation* (GCV) score is slightly above zero for values of $\hat{\rho}$ between $\langle -3{,}3 \rangle$, after which it increases sharply[23]. The GCV score is always minimised for two values of $\hat{\rho}$, the same negative and positive. The magnitude of the estimated (minimum) values of $\hat{\rho}$ is acceptable, i.e. always moderately below -1 and above 1, respectively[24]. We can infer that the estimation procedure leads to a robust solution for spatial autoregressive parameter $\hat{\rho}$.

Moreover, to assess the direction of the estimated $\hat{\rho}$, we examine the pattern of the observed values against the residuals. Figures 28-32, in Appendix, show the observed values against the residuals, which are obtained for minimum $\hat{\rho}$, for 2018-2022, respectively. Note that both positive and negative minimum $\hat{\rho}$ lead to the same estimated model, refer to eq. (3.26).

Figures 28-32, in Appendix, show that there is a positive correlation between the observed values and the residuals. It follows that we assume that an increase in the intensity of the freight leads to an increase in the industrial production. Hence, we select a positive minimum for spatial autoregressive parameter $\hat{\rho}$.

Moreover, we compare the results of the estimated spatial autoregressive parameter $\hat{\rho}$ between the iterative procedure proposed in this dissertation thesis and the iterative procedure shown by Schabenberger and Gotway (2005)[25]. Table 9 shows the results of the estimated autoregressive parameters $\hat{\rho}$ for 2018-2023, respectively, for two different versions of the iterative procedure.

[23] We select the cut off values for estimation of ρ, in the iterative procedure, for range $\langle -3{,}3 \rangle$ by steps of 0.01. Beyond this range, the generalised *cross-validation* score increases sharply.
[24] The most important condition for estimation of the *well-defined* model is that $(\mathbf{I} - \hat{\rho}\mathbf{W})$ in eq. (1.10) must be invertible. This condition is met for all estimated models. Refer to Section 1.2 for details.
[25] We set the convergence criterion of the iterative procedure at 0.1 (absolute value), i.e. a difference between the previous and the most recent estimate of $\hat{\rho}$, refer to Section 3.3.1. If we set a smaller convergence criterion, the iterative procedure does not converge for most of periods.

Period	Iterative procedure (dissertation thesis)	Iterative procedure (Schabenberger and Gotway, 2005)
2018	$\hat{\rho} = 1.795$	$\hat{\rho} = 4.371$
2019	$\hat{\rho} = 1.455$	$\hat{\rho} = 3.734$
2020	$\hat{\rho} = 1.565$	$\hat{\rho} = 4.469$
2021	$\hat{\rho} = 1.405$	$\hat{\rho} = 4.242$
2022	$\hat{\rho} = 1.545$	$\hat{\rho} = 4.332$

Table 9. Estimated spatial autoregressive parameter $\hat{\rho}$ for two different versions of the iterative procedure
Source: Author's construction

Table 9 shows that the iterative algorithm shown by Schabenberger and Gotway (2005) leads to relative high estimates of $\hat{\rho}$ across all periods. Based on the empirical results, we can conclude that the proposed algorithm in this dissertation thesis leads to a more robust solution for estimation of the spatial autoregressive parameter $\hat{\rho}$.

It follows that for estimated $\hat{\rho}$ (positive minimum), we obtain a set of the estimated functional regression parameters $\hat{\beta}(t)$ for 2018-2022. Note that the solution of the estimator function in eq. (3.26) leads to $\hat{\mathbf{b}}$, which is used to construct $\hat{\beta}(t) = \theta_{k_\beta}\hat{\mathbf{b}}$, refer to eqs. (3.21) and (3.22). Figures 33-37 show the estimated functional regression parameter $\hat{\beta}(t)$, including the *point-wise* intercept $\hat{\alpha}$, for 2018-2022, respectively.

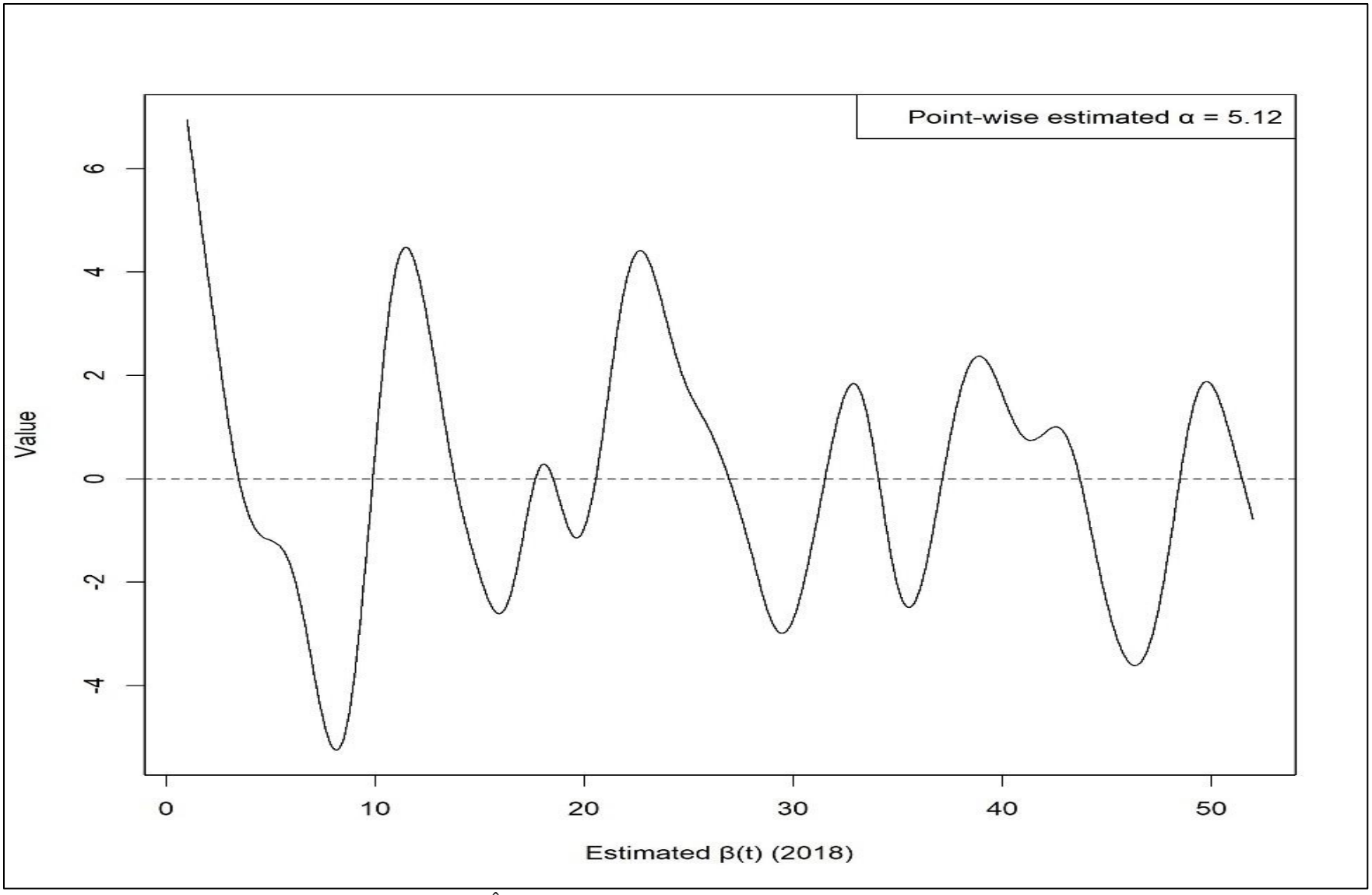

Figure 33. The estimated functional regression parameter $\hat{\beta}(t)$ and the *point-wise* intercept $\hat{\alpha}$ (2018)
Source: Author's construction

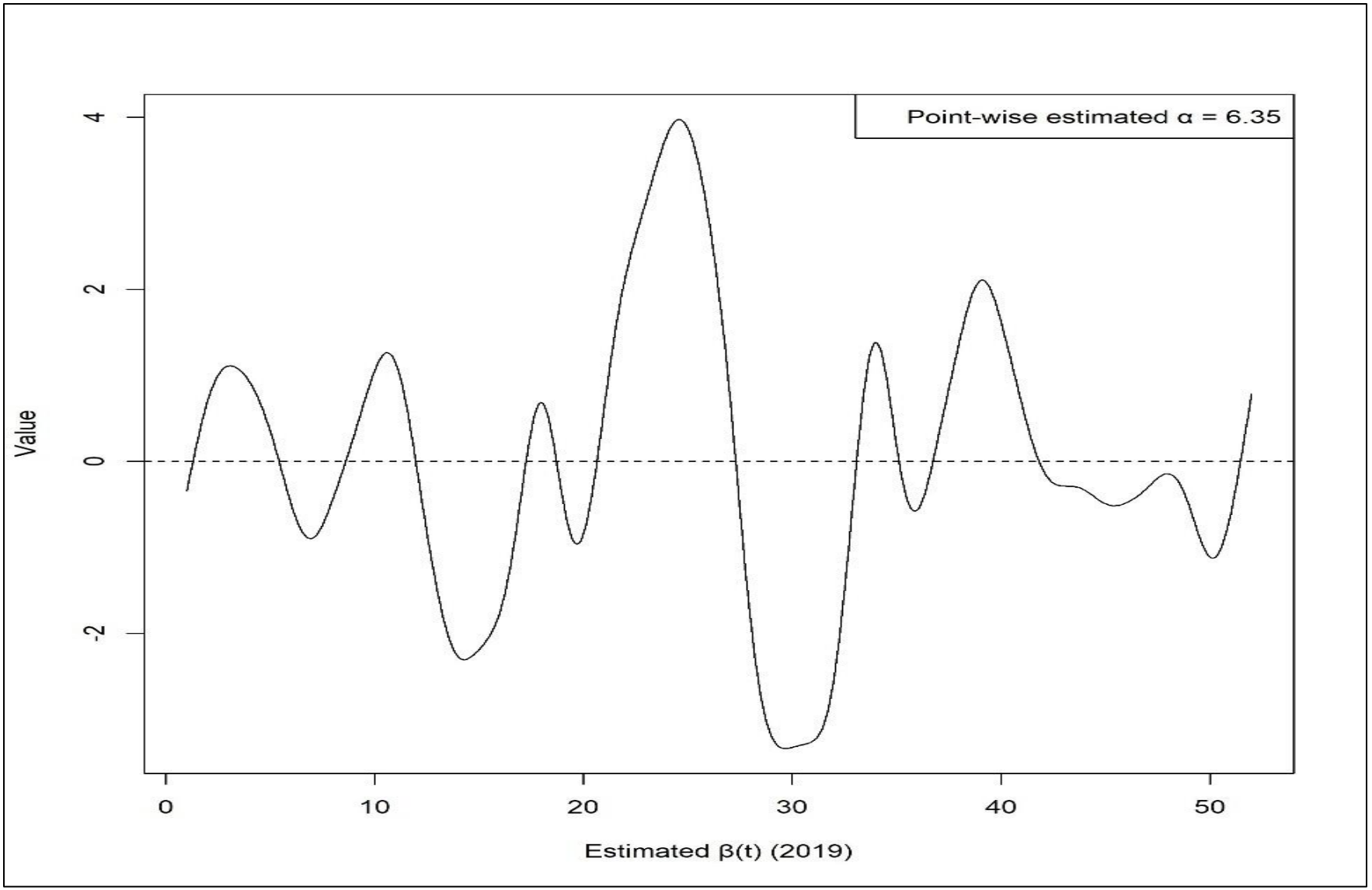

Figure 34. The estimated functional regression parameter $\hat{\beta}(t)$ and the *point-wise* intercept $\hat{\alpha}$ (2019)
Source: Author's construction

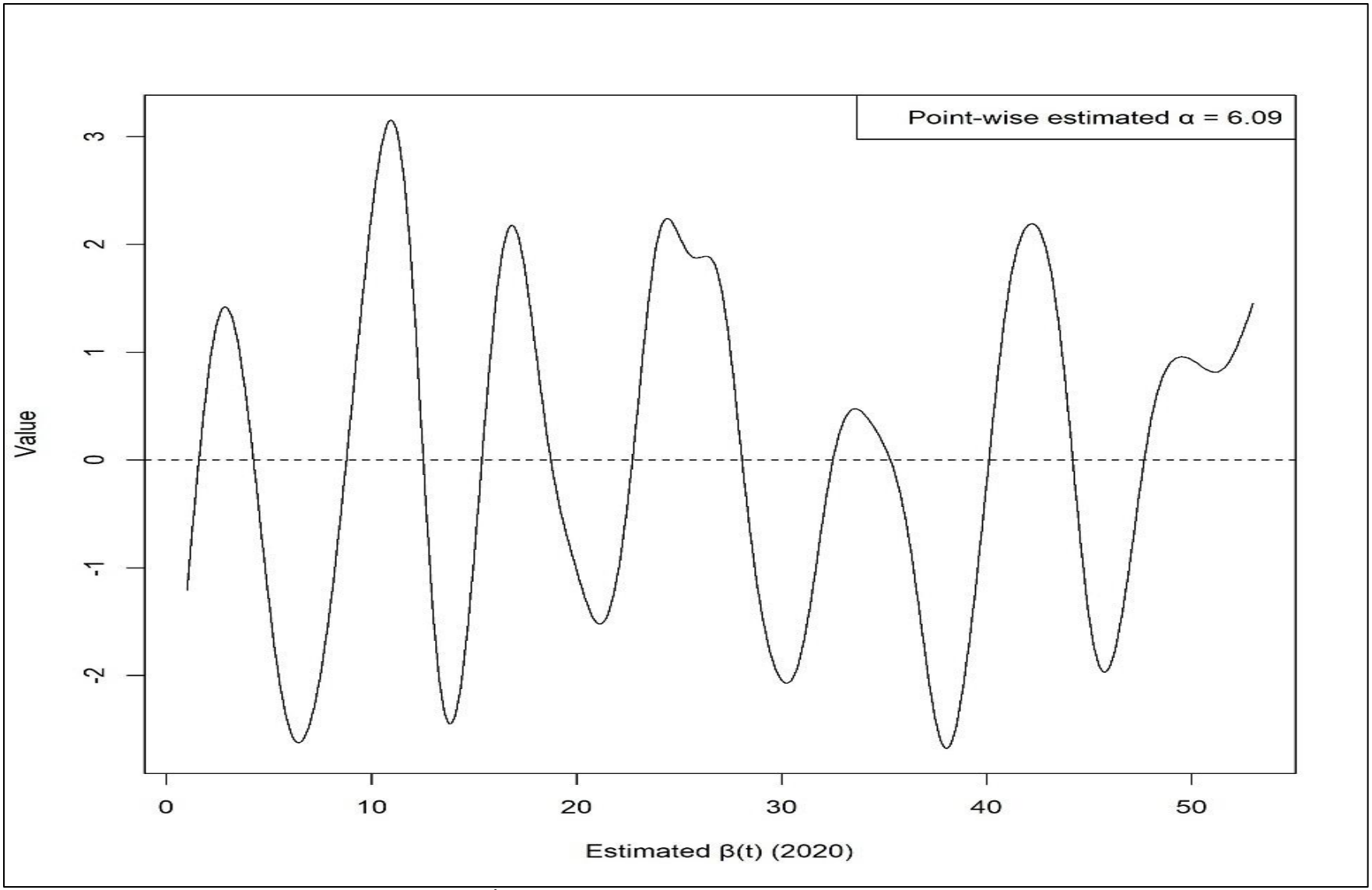

Figure 35. The estimated functional regression parameter $\hat{\beta}(t)$ and the *point-wise* intercept $\hat{\alpha}$ (2020)
Source: Author's construction

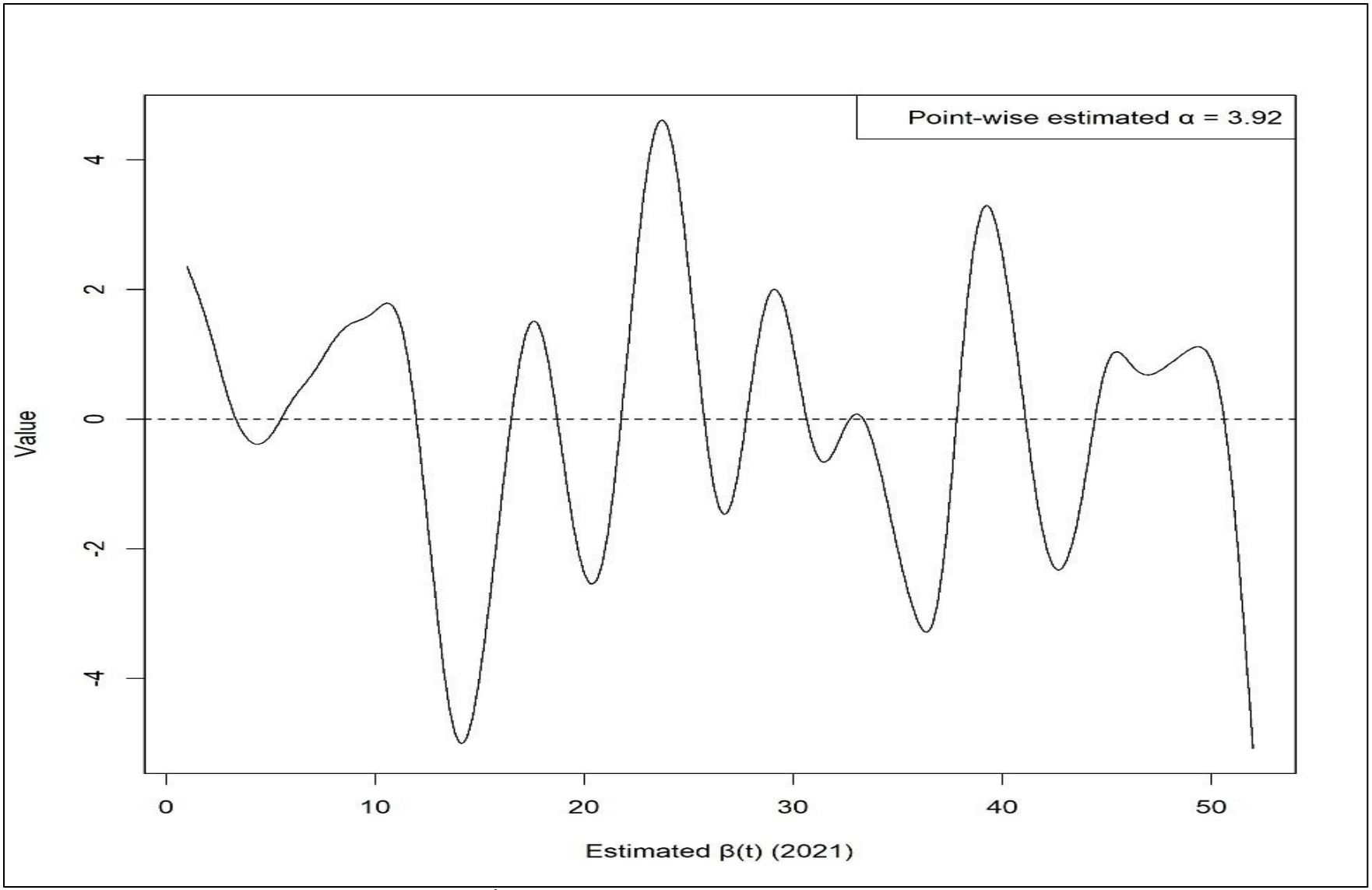

Figure 36. The estimated functional regression parameter $\hat{\beta}(t)$ and the *point-wise* intercept $\hat{\alpha}$ (2021)
Source: Author's construction

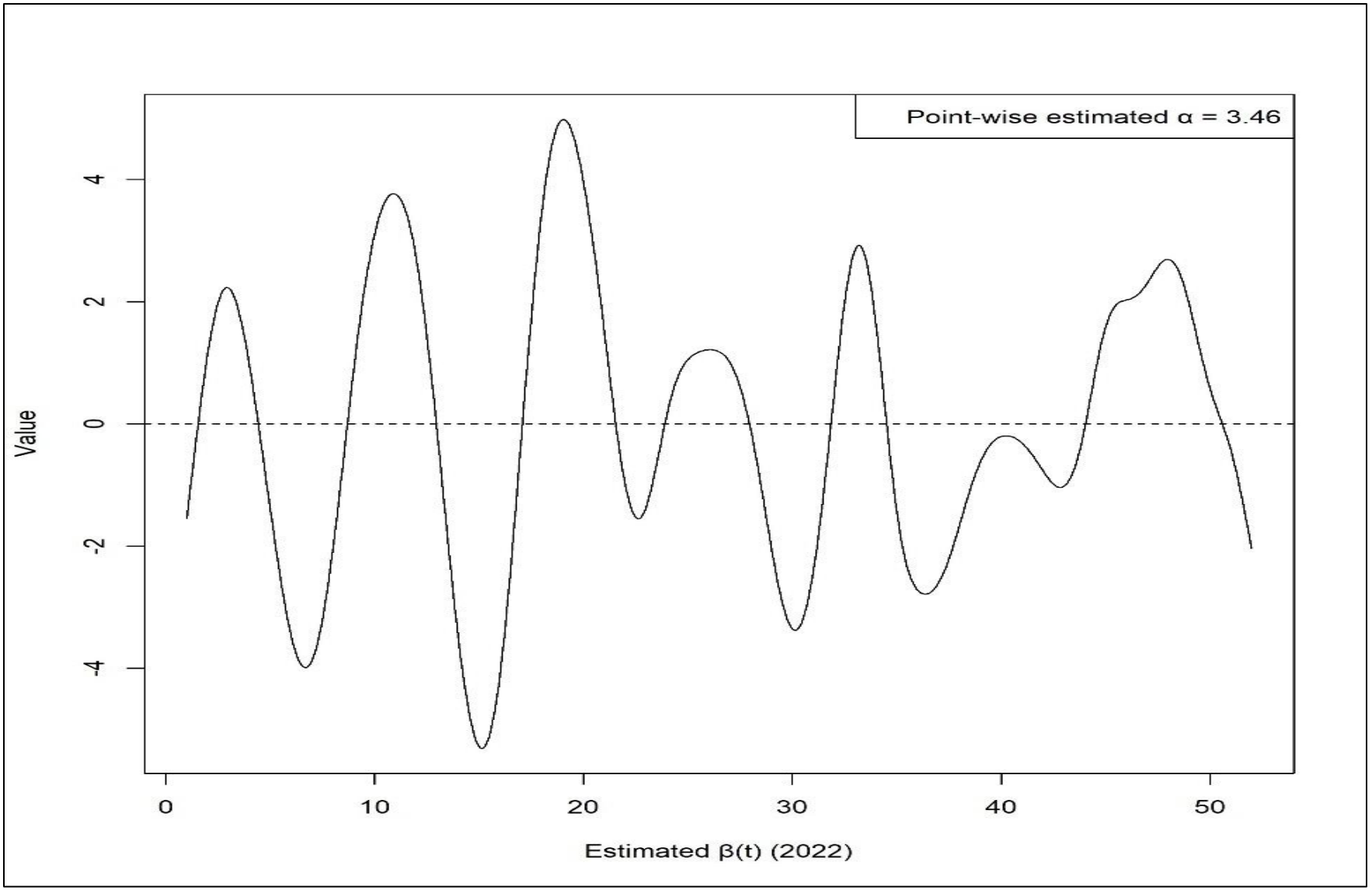

Figure 37. The estimated functional regression parameter $\hat{\beta}(t)$ and the *point-wise* intercept $\hat{\alpha}$ (2022)
Source: Author's construction

Figures 33-37 show that the estimated functional regression parameters $\hat{\beta}(t)$ oscillate around zero in each year. The frequent oscillations within short time periods can be caused by a large number basis splines used for expansion of $\beta(t)$ and $\chi_i(t)$. However, the advantage of using a large number of basis splines is that we can uncover local variations in the original data, which can be useful for studying the burden of the freight on the road traffic or the need of manufacturers to use the freight for transport of goods within specific time periods. We note that when the value of $\hat{\beta}(t)$ crosses, or is close to, zero, it can be considered statistically not significant.

Based on Figures 33-37, we can draw the following conclusions:

- The estimated parameters $\hat{\beta}(t)$ are statistically significant for all 2018-2022, i.e. the values of the estimated parameters $\hat{\beta}(t)$ oscillate well above or below zero. The industrial production is affected by the intensity of the freight;
- The values of the estimated parameter $\hat{\beta}(t)$ are very volatile within each year. This can be caused by frequent changes of the volume of the freight on roads within short time periods;
- It is not straightforward how to interpret the pattern of the estimated functional parameters $\hat{\beta}(t)$. Perhaps a study of the monthly industrial production could uncover more details of the freight volatility (no data is published on monthly industrial production).

To evaluate the estimated econometric model, we assess the difference between the observed and the fitted values. Figures 38-42 show the fitted values (black dots) alongside the observed values (red dots) of the response for districts, including the mean squared error[26] (MSE), for 2018-2022, respectively.

[26] The mean squared error (MSE) is calculated (Anderson, 2003) as $MSE = \frac{(\mathbf{y}-\hat{\mathbf{y}})^2}{n}$, where $\mathbf{y}$ is a vector of observed values, $\hat{\mathbf{y}}$ is a vector of fitted values and n is the number of observations.

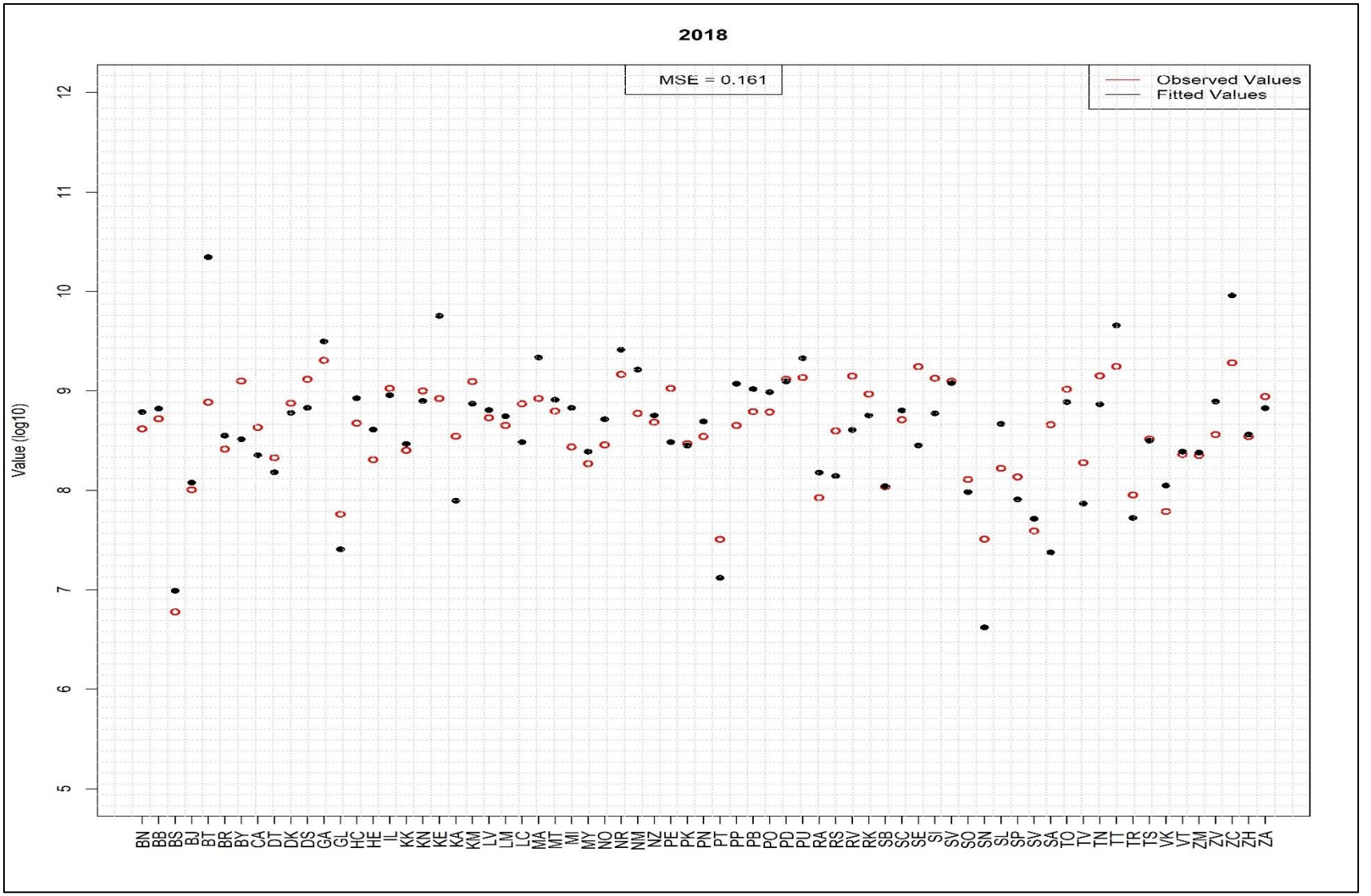

Figure 38. The fitted values (black dots) alongside the observed values (red dots) of the response for districts (2018), including the mean squared error (MSE)
Source: Author's construction

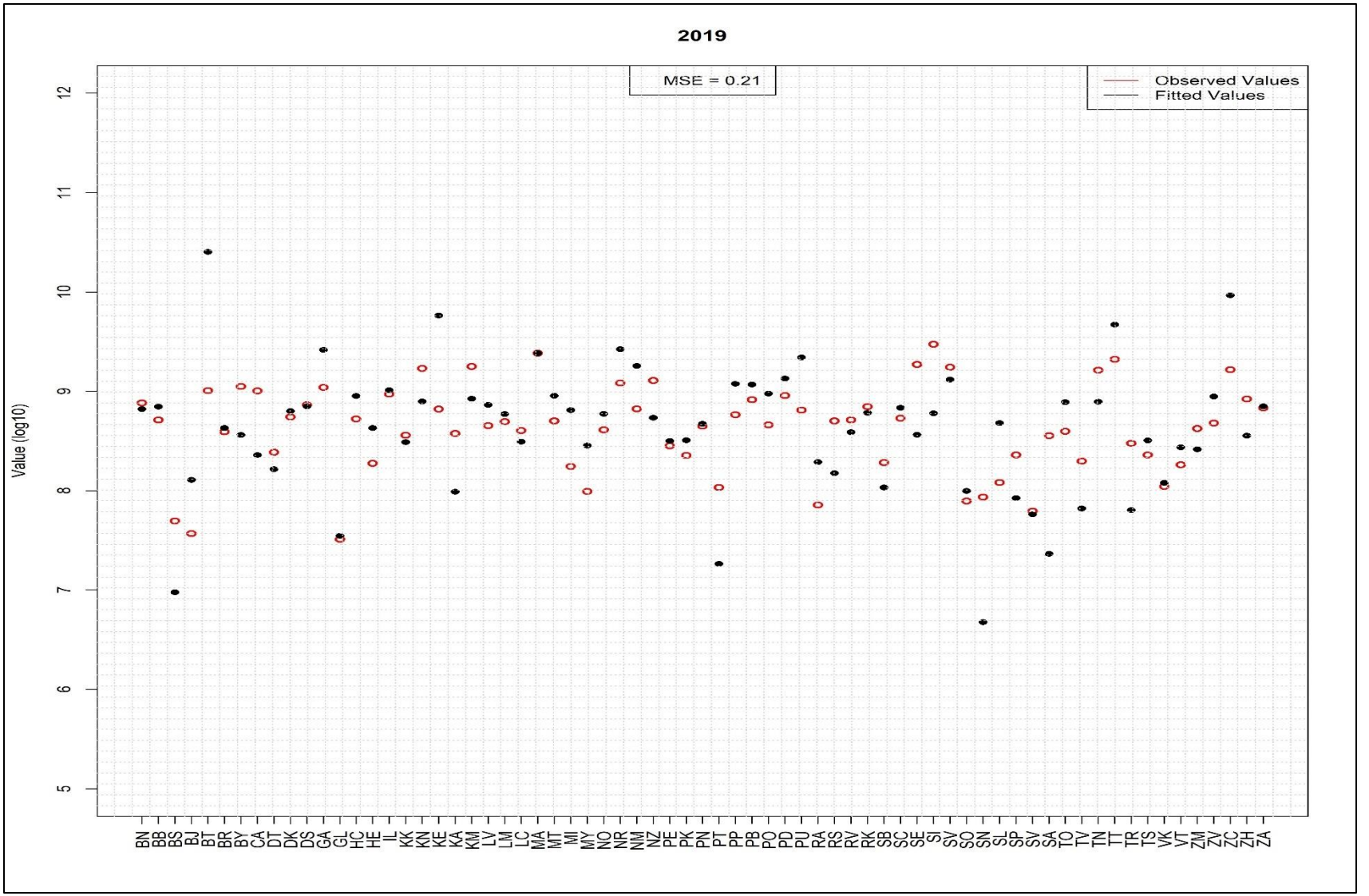

Figure 39. The fitted values (black dots) alongside the observed values (red dots) of the response for districts (2019), including the mean squared error (MSE)
Source: Author's construction

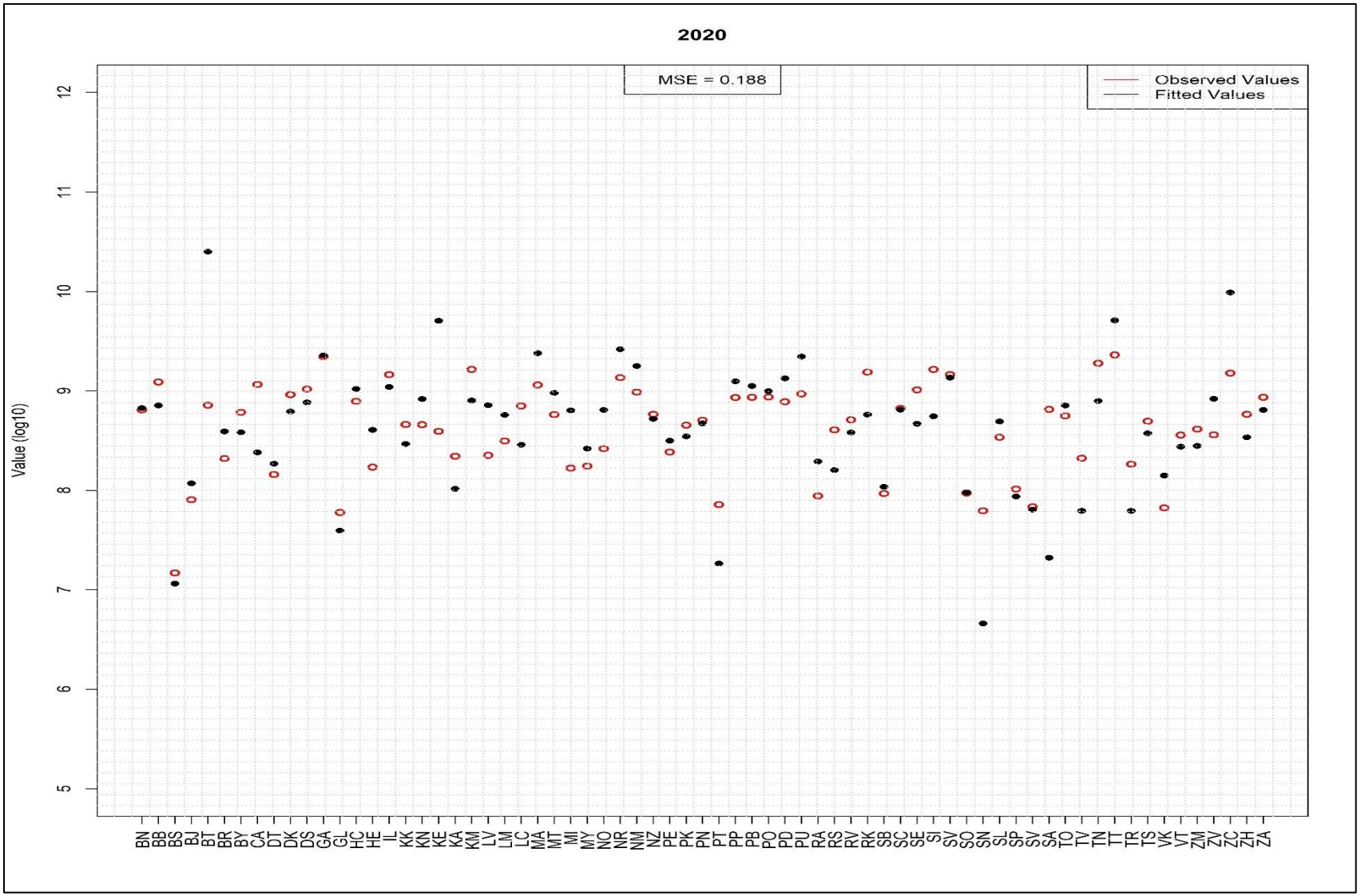

Figure 40. The fitted values (black dots) alongside the observed values (red dots) of the response for districts (2020), including the mean squared error (MSE)
Source: Author's construction

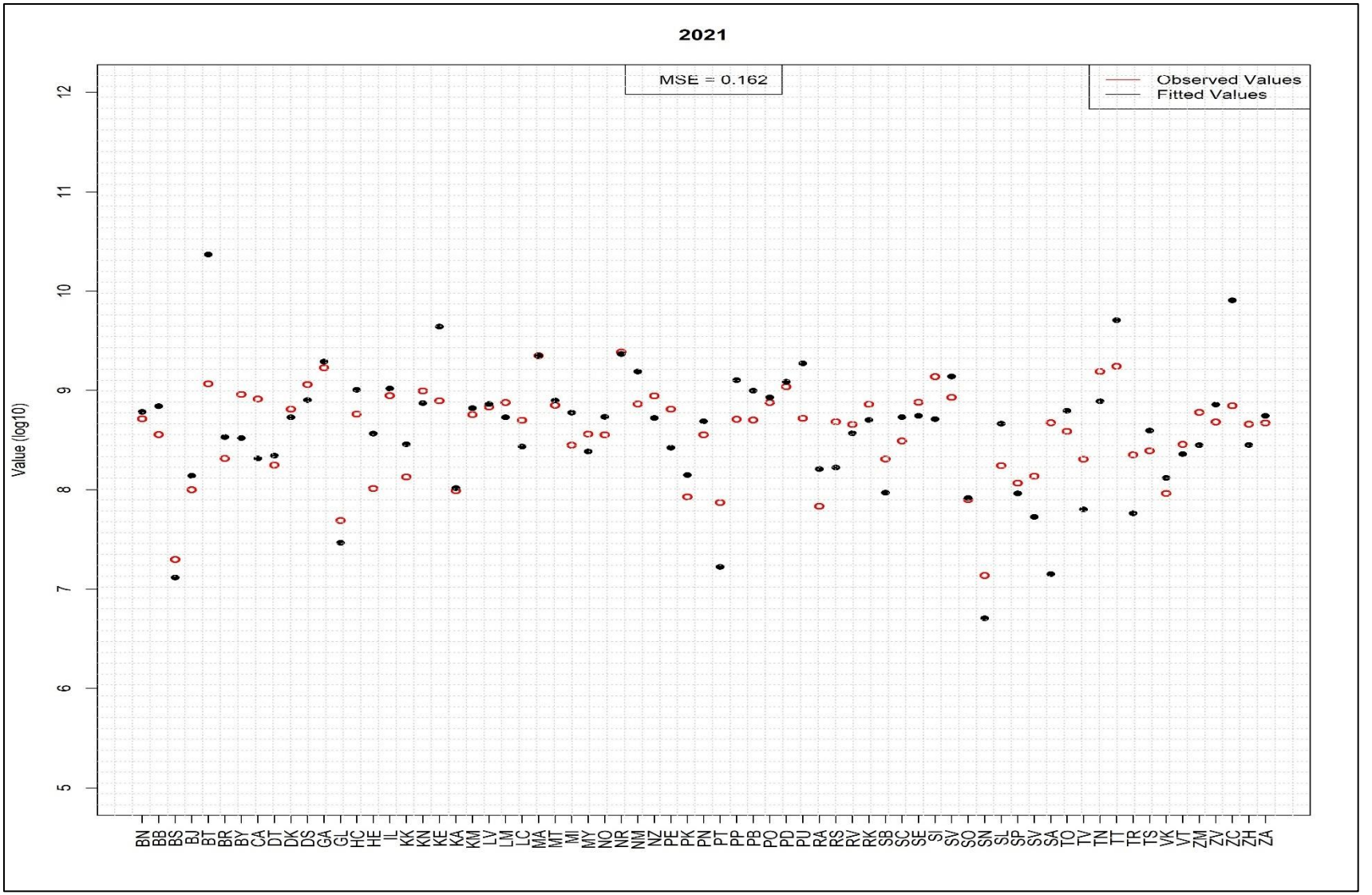

Figure 41. The fitted values (black dots) alongside the observed values (red dots) of the response for districts (2021), including the mean squared error (MSE)
Source: Author's construction

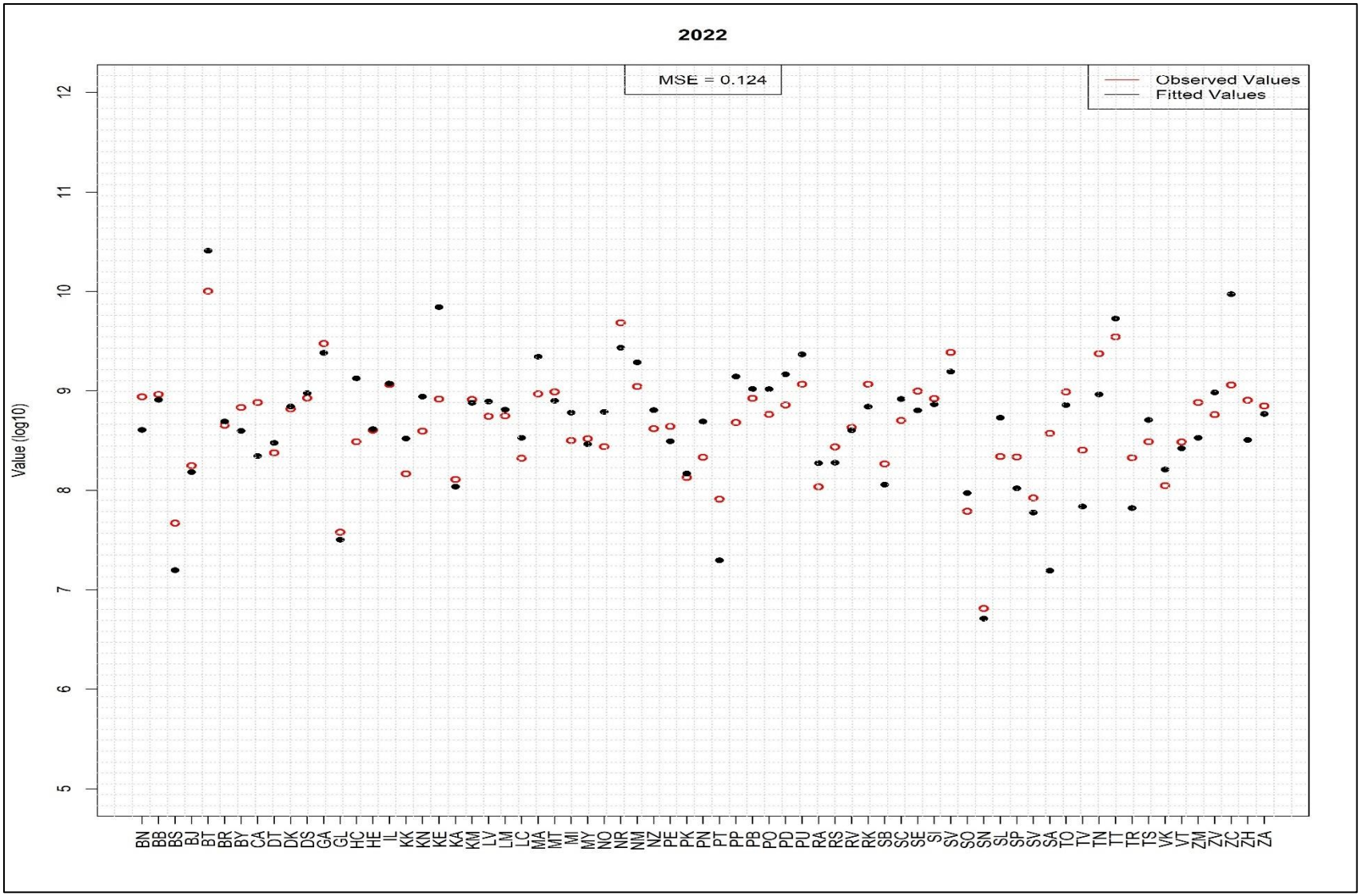

Figure 42. The fitted values (black dots) alongside the observed values (red dots) of the response for districts (2022), including the mean squared error (MSE)
Source: Author's construction

Figures 38-42 show that the difference between the model's estimated values and the observed values of the response is relatively small for most of districts. However, there are districts with relatively large differences across 2018-2022, e.g. Žarnovica, Šaľa, Komárno and Bratislava. Closer looks at the observed data reveals that these districts have unusually large counterintuitive differences between the intensity of the freight and the industrial production, e.g. both Bratislava Šaľa have relatively low intensity of the freight compared to their industrial production.

Moreover, we compare the functional spatial error model (SEM) with its functional counterpart, where the error term is not spatially correlated, refer to eq. (3.8). The performance is assessed by comparing the fitted values of the estimated model with the observed values of the response, including the mean squared error, between these two versions of the model. Figures 43-47, in Appendix, show the fitted values (black dots) alongside the observed values (red dots) of the response for districts, including the mean squared error (MSE), of the model in eq. (3.8); refer to Section 3.2.2 for details of the estimation. Similarly, the model is estimated for 2018-2022, respectively. The results show that neither of the models is superior, i.e. the performance of both models is very similar.

We can conclude that the proposed econometric model shows that there is a strong relationship between the industrial production and the intensity of the road transport, i.e. the industrial production is heavily reliant on the road transport of goods and material. The dynamics of the intensity of the freight, expressed through a mathematical function, serves a good predictor of the annual industrial production.

In general, the achievement of the empirical research is that we demonstrate that a new data source can provide an early and useful indication of the economic performance. The study shows that *high-frequency* electronic records from the digital toll system in Slovakia can be used, in the future, for production of new experimental and official statistics.

5 Discussion

In this section, we discuss the use of the new data source for experimental and official statistics, the proposed econometric model and the main contributions of the dissertation thesis. A brief summary of further research is provided.

The research of new data sources is a high agenda point in National Statistical Institutes (NSIs) since it leads to huge changes in dissemination of new experimental and official statistics. If NSIs do not dedicate part of their resources to innovations and research, the other private institutions could take a role of the leader in producing new relevant and reliable statistics. To overcome the challenges, which are related to the application of complex statistical and econometric models, NSIs need to actively collaborate with research institutions and academia.

The general goal of the empirical research of the dissertation thesis is to analyse a new data source, which contains electronic records that are automatically generated by the digital toll system in Slovakia. We infer that *high-frequency* electronic records, which represent time stamps of passages of vehicles through monitored sections of the road network, can be used to explain the dynamics of the industrial production. We propose an econometric model, which uses the aggregated electronic records that reflect the intensity of the freight, for analysis of the dynamics of the industrial production of districts in Slovakia. It follows that the specific goal of the dissertation thesis is the estimation of the proposed econometric model. The econometric model is defined within the theoretical framework of the functional spatial error model since the observations are associated with geographical locations, i.e. districts, and the covariate matrix, i.e. the intensity of the freight, is observed in functional space. To the best of our knowledge, the dissertation thesis represents a first study of such econometric model.

One of the main contributions of the dissertation thesis is the proposal and the estimation of the discussed econometric model. The results of the empirical research show that the intensity of the freight explains the dynamics of the industrial production of districts in Slovakia. The other important contribution is the enhancement of the iterative procedure, which is used for estimation of the model's parameters. Specifically, the proposed iterative procedure leads to a robust solution of the spatial autoregressive parameter, which is empirically validated by the estimation of the econometric model for five consecutive years.

Based on the results in the dissertation thesis, we believe that we achieved all stated objectives of the research.

Therefore, we believe that the results of the research in the dissertation thesis lead to a valuable study of the methodological concepts and the application of the functional spatial error model. It opens up possibilities for further research in the field of functional spatial regressions, which can be used for *real-world* applications.

Conclusion

This section makes concluding remarks that concern the theoretical concepts of statistical models used in empirical analyses, methods of the estimation of models and the interpretation of the estimated parameters. The achievements of the dissertation thesis are highlighted.

The statistical methodology of functional data analysis is a *well-researched* field, particularly, its application in parametric regression models. In the past two decades, researchers and academics have been studying the extension of functional regression models to data points that are associated with geographical locations. The extension of the methodology proves to be challenging since the problems of having too many parameters in the model's theoretical framework bring analytical and computational complexities.

In the initial part of the dissertation thesis, we show the technique of transforming the *real-valued* observations to mathematical objects, i.e. curves, which is usually the first step in functional data analysis. We follow with a description of the theoretical concepts of fundamental analysis of functional data. A good tool for exploration of functional data, i.e. observed curves, which makes an analysis of the inherent variability between and within the observed curves is the functional analysis of variance. The computation strategy for estimation of the functional analysis of variance is similar to its classical counterpart, except that the usual summation in the estimator function is replaced by integration over a continuous domain. The main difference is that the unknown parameters are estimated as mathematical functions. The advantage of the functional analysis of variance is that the functional parameters uncover details of variations on the entire domain of the observed interval, which can then be compared to the overall mean function.

In the empirical analysis, we explore the intensity of the freight, which is observed as mathematical functions, i.e. curves, for districts in Slovakia. The curves for districts are categorised to eight regions. The functional analysis of variance reveals interesting features of the intensity of the freight between and within regions in Slovakia. The significant differences between the regional profiles and the overall mean function show the concentration of the freight on the road network in Slovakia.

The focus of the dissertation thesis is to study the theoretical concepts and the application of functional regression models, where the observations are associated with geographical locations. Particularly, we study the theoretical framework and the estimation strategies of the functional spatial error model (SEM). SEM belongs to a family of spatial autoregressive models, for which we assume a spatial autocorrelation in the error term structure. We discuss the extension of SEM to functional space and outline its estimation procedure. The functional SEM consists of the functional covariate and the *real-valued* response, where the unknown regression parameter is functional and the spatial autoregressive parameter is *real-valued*. The estimation of the functional SEM is carried out by the generalised method of least squares, which leads to a system of two equations. The iterative procedure is used to find an optimal solution of the system of equations for functional regression parameter and spatial autoregressive parameter, respectively, by minimising the difference between the previous and the recent estimate of the spatial autoregressive parameter. One of the most important methodological achievements of the dissertation thesis is that we propose the enhancement of the iterative procedure. The main difference is that we propose to minimise the generalised *cross-validation* score of estimated models within the iterative procedure. Based on the results from the empirical analysis, we can conclude that the iterative procedure proposed in the dissertation thesis leads to a more robust solution for spatial autoregressive parameter when compared to its previous version.

Moreover, in the empirical analysis, we illustrate and discuss the estimated parameters of the proposed econometric model. The proposed econometric model analyses the dynamics of the annual industrial production and the intensity of the freight, which is measured as a mathematical function, for districts in Slovakia. The model is estimated for five consecutive years 2018-2022, respectively. The estimated spatial autoregressive parameters show that there is a significant degree of the spatial autocorrelation in residuals. The estimated functional regression parameter reveals a complex dynamics of the intensity of the freight across the observed time interval and its effect on industrial productions in districts. Based on the results, we can conclude that the estimated functional regression parameter is statistically significant for all years. Additionally, we assess the model's performance by comparing the fitted values and the observed values. We can conclude that the model's performance is satisfactory, i.e. the mean squared error of the estimated models for 2018-2022, respectively, is relatively small.

We believe that the main achievements of the dissertation thesis make a notable contribution to the current research. In the empirical analysis, we propose an econometric model, which utilises a new data source to explain the dynamics of the industrial production of districts in Slovakia. The model conceptualises a theoretical framework of the functional spatial error model. In the methodological part, we propose to enhance the iterative procedure for estimation of the model's parameters. The proposed iterative procedure leads to a more robust solution for spatial autoregressive parameter. To the best of our knowledge, there is no previous study of both empirical and methodological contributions of the dissertation thesis.

Based on the results in the dissertation thesis, we believe that we achieved all stated objectives of the research.

Resumé

Dizertačná práca sa zaoberá výskumom dvoch hlavných problematík, ktoré sa týkajú aplikácie štatistickej vedy v ekonometrii. Prvá problematika sa zaoberá metodologickým aspektom, ktorého počiatočným cieľom je štúdium teoretického rámca priestorových regresných modelov a ich výpočtových mechanizmov rozšírených do funkcionálnych priestorov. V tejto časti je hlavným prínosom dizertačnej práce zlepšenie iteratívnej procedúry, ktorá sa používa na odhad parametrov funkcionálneho priestorového regresného modelu, v rámci zovšeobecnej metódy najmenších štvorcov. Druhá problematika sa zaoberá empirickým aspektom, ktorej počiatočnou úlohou je štúdium nového zdroja údajov. Predpokladáme, že v budúcnosti budú tieto nové zdroje údajov využívané v experimentálnych a oficiálnych štatistikách. Súbor údajov pozostáva s denných elektronických záznamov o vozidlách, ktoré prechádzajú cez monitorované úseky cestnej siete na Slovensku. Tieto elektronické záznamy sú automaticky zberané digitálnym mýtnym systémom. Tento nový zdroj údajov má potenciál poskytnúť analytické informácie o ekonomických ukazovateľoch, ktoré môžu byť použité na formovanie rozhodovaní pre tvorcov politík. Konkrétne hlavným prínosom dizertačnej práce je návrh ekonometrického modelu, ktorého cieľom je použiť intenzitu nákladnej dopravy na cestách na vysvetlenie dynamiky priemyselnej produkcie v okresoch na Slovensku.

Výskum dizertačnej práce sa v prvej časti zaoberá štúdiom metodológie a metodiky funkcionálnych a priestorových regresných modelov v danom poradí. Hlavným cieľom tejto časti je rozšírenie priestorových regresných modelov do funkcionálneho priestoru. Konkrétne sa zameriame na priestorový autoregresný model, ktorý zahŕňa priestorovú závislosť v poruchovom člene modelu. Priestorová závislosť sa týka pozorovaní, ktoré sú skúmané pre geografické lokality, v našom prípade sú to okresy na Slovensku. Závislosť medzi jednotlivými geografickými lokalitami je vyjadrená pomocou priestorovej matice váh, ktorej prvky vyjadrujú závislosť medzi jednotlivými susednými lokalitami. Priestorová matica váh je zakomponovaná do autoregresného modelu pre poruchový člen. Tento typ priestorového regresného modelu sa tiež nazýva priestorový chybový model (*spatial error model*, SEM). V našom prípade je model transformovaný do funkcionálneho priestoru. Konkrétne model zahŕňa funkcionálnu nezávislú premennú a závislú premennú, ktorej skúmania sú reálne čísla. Hlavným rozdielom tohto modelu je, že odhadovaný regresný parameter je funkcionálny. Pre odhad parametrov funkcionálneho

SEM použijeme modifikovanú zovšeobecnenú metódu najmenších štvorcov. Riešenie tejto odhadovej metódy vedie k sústave dvoch rovníc pre výpočet funkcionálneho regresného parametru a priestorového autoregresného parametru v danom poradí. Výpočet tejto sústavy rovníc je možný iteratívnou procedúrou. V súčasnej literatúre je riešenie iteratívnej procedúry navrhnuté tak, aby sa minimalizoval rozdiel medzi priestorovými autoregresnými parametrami, ktoré boli vypočítané v poslednom a predchádzajúcom kroku. Toto riešenie môže byť v niektorých prípadoch neefektívne, t. j. nemusí vždy konvergovať k definovanému rozdielu alebo môže viesť k relatívne veľkým odhadnutým priestorovým autoregresným parametrom. Jedným z najdôležitejších metodologických úspechov dizertačnej práce je návrh na vylepšenie iteratívnej procedúry. V rámci iteratívnej procedúry navrhujeme minimalizovať zovšeobecnené skóre *krížovej-validácie* odhadnutých modelov, pre ktoré dostaneme zodpovedajúci priestorový autoregresný parameter. Na základe výsledkov v empirickej časti dizertačnej práce môžeme konštatovať, že iteratívna procedúra navrhovaná v dizertačnej práci vedie k efektívnemu a robustnému riešeniu pre priestorový autoregresný parameter. Pokiaľ je nám známe, neexistuje žiadna predchádzajúca štúdia, ktorá by navrhovala takéto riešenie danej iteratívnej procedúry.

Hlavným cieľom empirickej časti dizertačnej práce je návrh ekonometrického modelu, ktorý využíva teoretický rámec funkcionálneho priestorového chybového modelu (SEM) a výpočtový mechanizmu, ktorý sme navrhli v metodologickej časti. Úlohou ekonometrického modelu je vysvetliť dynamiku ročnej priemyselnej produkcie v okresov na Slovensku pomocou agregovaných elektronických záznamoch o vozidlách, ktoré prechádzajú cez monitorované úseky cestnej siete na Slovensku, z mýtneho systému. Agregovaný súbor údajov vyjadruje týždennú intenzitu nákladnej dopravy v okresoch na Slovensku. Tento agregovaný súbor je transformovaný do funkcionálneho priestoru, t. j. každá jednotka pozorovania je transformovaná do matematickej funkcie. Ekonometrický model je odhadovaný jednotlivo pre obdobia 2018-2022. Jedným z najdôležitejších empirických úspechov dizertačnej práce je odhad navrhovaného ekonometrického modelu. Výsledky odhadu modelu ukazujú, že nový zdroj údajov má vysoký potenciál využitia v odhade ekonomických ukazovateľov, t. j. intenzita nákladnej dopravy vysvetľuje priemyselnú produkciu v okresoch na Slovensku. Ďalším dôležitým úspechom empirickej časti je demonštrácia efektívnosti a robustnosti navrhovaného riešenia pre iteratívnu procedúru pre odhad priestorového autoregresného parametru. Pokiaľ je nám známe, neexistuje

žiadna predchádzajúca štúdia, ktorá by navrhovala inovatívne empirické a metodologické prínosy v tejto dizertačnej práci.

Veríme, že výsledky výskumu v dizertačnej práci vedú k hodnotnému štúdiu metodologických konceptov a aplikácie funkcionálneho priestorového chybového modelu. Výsledky výskumu otvárajú možnosti pre ďalší výskum v oblasti funkcionálnych priestorových regresných modelov, ktoré je možné využiť pre aplikácie v aplikovanej ekonomickej problematike.

Na základe výsledkov v dizertačnej práci sa domnievame, že sme splnili všetky stanovené ciele výskumu.

References

Anderson, T. W. (2003). An Introduction to Multivariate Statistical Analysis. Third edition. New York: Wiley.

Aneiros, G., Cao, R., Fraiman, R., Genest, C., Vieu, P. (2019). Recent advances in functional data analysis and high-dimensional statistics, Journal of Multivariate Analysis, Volume 170, p. 3-9, ISSN 0047-259X. https://doi.org/10.1016/j.jmva.2018.11.007.

Anselin, L. (1988). Spatial econometrics: methods and models. Dordrecht: Kluwer Publisher. ISBN 978-94-015-7799-1.

Anselin, L., Rey, S. J. (2014). Modern Spatial Econometrics in Practice: A Guide to GeoDa, GeoDaSpace and PySAL. GeoDa Press LLC, Chicago, IL. ISBN: 0986342106.

Anselin, L., Morrison G. (2019). Distance-Based Spatial Weights. R Notes. Available in:

https://spatialanalysis.github.io/lab_tutorials/Distance_Based_Spatial_Weights.html. [accessed on 20 December 2024].

Arnone, E., Azzimonti, L., Nobile, F., Sangalli, L. M. (2019). Modeling spatially dependent functional data via regression with differential regularization, Journal of Multivariate Analysis, Volume 170, p. 275-295, ISSN 0047-259X. https://doi.org/10.1016/j.jmva.2018.09.006.

Cliff, A., Ord, J. K. (1973). Spatial autocorrelation. London: Pion.

Cliff. A., Ord, J. K. (1981). Spatial processes: Models and applications, London: Pion.

Cox, M., Triebel, J., Linz, S., Fries, C., Flores, L. F., Lorenz, A., Ollech, D., Dietrich, A., LeCrone, J., Webel, K. (2020). Daily truck toll mileage index based on digital process data from toll collection system. Federal Statistical Office (Statistisches Bundesamt). https://www.destatis.de/EN/Methods/WISTAScientificJournal/Downloads/truck-toll-mileage-index-042020.pdf?__blob=publicationFile.

de Boor, C. (2001). A Practical Guide to Splines, Revised Edition. New York: Springer.

Ferraty, F., Vieu, P. (2006). Nonparametric Functional Data Analysis. Springer New York.

Ferraty, F., Vieu, P. (2006b). S-Plus/R code accompanying Ferraty, F. and Vieu, P. 2006. http://www.mat.univ-toulouse.fr/staph/npfda/.

Furková, A. (2013). Modely priestorovej ekonometrie - model SAR a model SEM. Využitie kvantitatívnych metód vo vedecko-výskumnej činnosti a v praxi X: zborník príspevkov zo seminára : 29. – 31. máj 2013, Čingov – Slovenský raj. Bratislava: Ekonóm, 2013. ISBN 978-80-225-3646-2, p. 1–6.

Furková, A. (2013). Spatial dependence econometrics models. In Ekonomika a informatika. ISSN 1336-3514, 2013, Roč. 11, č. 2, s. 41–52.

Furková, A. (2017). Modely a metódy priestorovej ekonometrie. Habilitačná práca, Ekonomická univerzita v Bratislave, Fakulta hospodárskej informatiky, Evidenčné číslo: 103003/H/2017/004365094.

Furková, A., Chocholatá, M. (2016). Softvérové nástroje na priestorovú analýzu dát. Nové trendy v ekonometrii a operačním výzkumu: zborník príspevkov: mezinárodní vědecký seminář Praha. Bratislava: Ekonóm, 2016, ISBN 978-80-225-4319-4, s. 44–49.

Furková, A., Knížat, P. (2024). Beyond Parametric Bounds: Exploring Regional Unemployment Patterns Using Semiparametric Spatial Autoregression. Business Systems Research, Vol. 15, no. 2, p. 48-66. https://doi.org/10.2478/bsrj-2024-0017.

Giraldo, R., Caballero, W., Camacho-Tamayo, J. (2018). Mantel test for spatial functional data. AStA Adv Stat Anal 102, p. 21–39. https://doi.org/10.1007/s10182-016-0280-1.

Greene, W.H. (2003). Econometric Analysis, 5th edition. Prentice Hall, 2003. ISBN-10: 0130661899.

Haining, R. (1993). Spatial Data Analysis in the Social and Environmental Sciences. Cambridge University Press.

Horváth, L., Kokoszka, P. (2012). Inference for Functional Data with Applications, Springer Science & Business Media, New, York.

Kelejian, H., Prucha, I. (1997). Estimation of spatial regression models with autoregressive errors by two-stage least squares procedure: a serious problem. Int. Reg. Sci. Rev. 20, p. 103-111.

Kelejian, H., Prucha, I. R., Yuzefovich, Y. (2004). Instrumental variable estimation of a spatial autoregressive model with autoregressive disturbances: Large and small samples results. In

LeSage, J. P., Pace, R. K. (2004). Advances in Econometrics: 139 Spatial and Spatiotemporal Econometrics. Oxford: Elsevier Science Ltd., p. 163–198. ISBN 978-0-7623-1148-4.

Knížat, P. (2023a). Web scraped data in consumer price indices. Statistical Journal of the IAOS, Vol. 39, no. 1, p. 203–212. https://doi.org/10.3233/SJI-220115.

Knížat, P. (2023b). Hedonic consumer price index: diagnostics and analysis of variance. Slovak Statistics and Demography, Vol. 33, no. 3, p. 21–38.

Knížat, P. (2024). Regional Intensity of the Freight: Functional Analysis of Variance. In Mathematical Methods in Economics 2024. The 42nd International Conference on Mathematical Methods in Economics, Proceedings, Praha. The Czech Society for Operations Research, 2024. ISBN 978-80-7561-482-7. ISSN 2788-3965.

Knížat, P., Ceľuchová Bošanská, D., Janík, M., Nguyen, F. (2024). New data sources in statistics: the effect of freight intensity on macroeconomic indicators (in Slovak). Slovak statistics and demography, Vol. 34, no. 1, p. 7–26.

Knížat, P., Furková, A., Glaser-Opitzová, H., Peťko, P. (2025). Nowcasting industrial production index with high-frequency toll data. Statistical Journal of the IAOS, Vol. 41, no. 1, p. 102–111. https://doi.org/10.1177/18747655241307540.

Knížat, P., Glaser-Opitzová, H. (2023). Consumer price index from web-scraped data: analysis of specific product category. Slovak Statistics and Demography, Vol. 33, no. 1, p. 37–49 (in Slovak).

Knížat, P., Glaser-Opitzová, H., Furková, A., Vojtková, M. (2024). Multilateral indices in official price statistics and a new additive splicing method. Quality and Quantity Journal, Vol. 58, no. 5, p. 4207–4222. https://doi.org/10.1007/s11135-024-01848-3.

Krugman, P. (1991). Increasing Returns and Economic Geography. Journal of Political Economy, Vol. 99, no. 3, p. 483–499.

Lee, D. T., Schacter, B. J. (1980). Two algorithms for constructing a Delaunay triangulation, International Journal of Computer and Information Sciences, Vol. 9, no. 3, pp. 219– 242. https://doi.org/10.1007/BF00977785.

Lee, L. F. (2003). Best spatial two-stage least squares estimators for a spatial autoregressive model with autoregressive disturbances. Econometric Reviews, 2003, Vol. 22, no. 4, p. 307–335.

LeSage, J., Pace, K. R. (2009). Introduction to spatial econometric. Taylor and Francis Group LLC, CRC Press.

Liu, C., Ray, S., Hooker, G. (2017). Functional principal components analysis of spatially correlated data, Statistical Computing, Vol. 27, p. 1639-1654. https://doi.org/10.1007/s11222-016-9708-4.

Menafoglio, A., Secchi, P. (2017). Statistical analysis of complex and spatially dependent data: a review of object oriented spatial statistics. European Journal of Operational Research, Vol. 258, Issue 2, p. 401-410, ISSN 0377-2217. https://doi.org/10.1016/j.ejor.2016.09.061.

Paelinck, J., Klaassen, L. (1979). Spatial econometrics. Farnborough: Saxon House.

Pineda-Ríos, W., Giraldo, R., Porcu, E. (2019). Functional SAR models: With application to spatial econometrics, Spatial Statistics, Vol. 29, p. 145-159, ISSN 2211-6753, https://doi.org/10.1016/j.spasta.2018.12.002.

Ramsay, J. O., Silverman, B. W. (2002). Applied Functional Data Analysis Methods and Case Studies. Springer Series in Statistics, Springer.

Ramsay, J. O., Silverman, B. W. (2005). Functional Data Analysis. Second Edition. Springer Series in Statistics, Springer.

Ramsay, J. O., Hooker G., Graves, S. (2009). Functional Data Analysis with R and MATLAB. Use R!, Springer.

Římalová, V., Fisherová, E., Menafoglio, A., Pini, A. (2022). Inference for spatial regression models with functional response using a permutation approach, Journal of Multivariate Analysis, Vol. 189, 104893, Elsevier.

Romano, E., Cozza, V., Verde, R. (2011). Clustering Spatial Functional Data: A Method Based on a Nonparametric Variogram Estimation. In: Proceedings of the 7th Conference of the Classification and Data Analysis Group of the Italian Statistical Society, Catania, September 9 - 11, 2009. https://doi.org/10.1007/978-3-642-11363-5.

Shi, J. Q., Choi, T. (2011). Gaussian Process Regression Analysis for Functional Data. CRC Press, Taylor & Francis Group.

Schabenberger, O., Gotway, C. A. (2005). Statistical Methods for Spatial Data Analysis. CRC Press, Taylor & Francis Group.

Stock, J. H., Watson, M. W. (2016). Factor Models and Structural Vector Autoregressions in Macroeconomics, In: Handbook of Macroeconomics, Elsevier, ISSN: 1574-0048.

Waller, L., Gotway, C. (2004). Applied Spatial Statistics for Public Health Data. John Wiley and Sons.

Whittle, P. (1954). On stationary processes in the plane. Biometrika, vol. 41, p. 434–49.

Appendix

	2018					2019					...	2022					2023				
DISTRICT	**w 1**	**w 2**	**...**	**w 51**	**w 52**	**w 1**	**w 2**	**...**	**w 51**	**w 52**	**...**	**w 1**	**w 2**	**...**	**w 52**	**w 53**	**w 1**	**w 2**	**...**	**w 51**	**w 52**
Bánovce nad Bebravou	36578	65532	...	76579	17615	28123	60762	...	78782	18895	...	25671	42578	...	32558	10867	20150	39462	...	32928	10760
Banská Bystrica	92462	184997	...	170878	33874	48157	136230	...	192121	31338	...	73519	135246	...	106625	43392	61058	124303	...	103775	34703
Banská Štiavnica	2659	5728	...	6127	1704	2458	5189	...	6695	1461	...	2521	4457	...	3917	1524	2476	4170	...	4086	1440
Bardejov	12322	24613	...	24435	5592	7709	18732	...	30212	5252	...	12684	24585	...	21395	11405	11251	25109	...	26549	8064
Bratislava	109544	196726	...	199449	32394	79651	200588	...	199752	30705	...	134698	213408	...	230249	80127	165599	289294	...	270998	73296
Brezno	35484	85960	...	81384	15617	26842	68846	...	92399	12993	...	35051	60987	...	51997	18873	27058	59603	...	59161	14582
Bytča	44936	85733	...	85799	12539	27033	65039	...	97431	12183	...	36942	62159	...	47620	16477	33896	62106	...	56555	13331
Čadca	124814	215367	...	220492	34006	91733	209599	...	241203	35282	...	132709	204164	...	156115	56480	117342	201215	...	194183	47460
Detva	27886	49870	...	54130	11791	22593	49372	...	65755	11457	...	36060	58639	...	40557	15283	23343	50284	...	45226	12557
Dolný Kubín	134948	250073	...	258482	45753	100779	225513	...	279511	46875	...	93424	181189	...	135757	54240	85419	168994	...	170525	43998
Dunajská Streda	116748	217328	...	211372	39653	87670	199843	...	223830	32984	...	86261	148700	...	117413	41225	78947	144162	...	120696	29751
Galanta	202169	386245	...	377649	77181	155985	344304	...	442965	75504	...	218196	360406	...	277343	92741	177976	328182	...	309009	86661
Gelnica	227	592	...	704	121	175	490	...	648	112	...	37	79	...	62	32	43	72	...	64	31
Hlohovec	43984	77146	...	83626	23109	35237	74420	...	83403	20444	...	47211	75131	...	62781	24954	38932	72251	...	76235	22080
Humenné	10747	20499	...	22054	4543	8333	16729	...	21621	4578	...	9394	18193	...	15406	6954	8501	17016	...	19439	6528
Ilava	71295	127232	...	137520	31678	53323	114332	...	143908	30991	...	83472	131620	...	112692	47006	70671	127384	...	127652	37784
Kežmarok	21259	39478	...	38956	8493	13648	31041	...	41234	8788	...	18233	30358	...	26385	11768	15204	28968	...	31128	9874
Komárno	83971	172267	...	157894	29865	67546	147445	...	118211	22122	...	102268	172269	...	132091	45025	89291	173626	...	169637	39825
Košice	205266	368841	...	396778	96017	173198	349031	...	432620	90296	...	213101	329049	...	280611	132711	185505	326648	...	361329	120230
Krupina	92986	165371	...	168685	36858	67704	147221	...	202589	40778	...	119498	173198	...	134357	66146	101024	165102	...	155650	55507
Kysucké Nové Mesto	199284	341963	...	354934	54033	153802	344274	...	381727	55459	...	135272	209106	...	160752	57329	119451	204831	...	199750	48725
Levice	133811	236396	...	245375	48563	98917	215827	...	285816	54635	...	144143	227085	...	166077	67163	115269	211272	...	181637	52810
Levoča	31935	56230	...	61712	15657	24789	53405	...	67807	15515	...	32926	50838	...	44877	19695	26561	48702	...	51854	17114
Liptovský Mikuláš	57504	115077	...	113495	24355	40899	98645	...	133725	23295	...	62952	110640	...	88180	36128	53673	106372	...	109047	32843
Lučenec	169281	305252	...	312644	65834	134459	290748	...	347933	66061	...	117493	180485	...	92809	30386	74064	151490	...	138255	34371
Malacky	97250	171387	...	178399	26478	73696	189544	...	183723	26544	...	121520	187528	...	136802	42642	106953	180553	...	164288	40311
Martin	198250	360759	...	398317	74064	162845	355745	...	418868	76563	...	160250	282916	...	210389	82724	136557	266971	...	259627	70286
Medzilaborce	105	204	...	222	49	67	249	...	257	36	...	64	117	...	131	44	78	129	...	133	40
Michalovce	62405	119947	...	129398	30002	48327	97291	...	130948	30322	...	43679	83806	...	77229	37238	42744	82830	...	99499	49370
Myjava	1627	3226	...	3236	485	1057	2281	...	3515	502	...	864	1539	...	1112	426	682	1478	...	1223	281
Námestovo	12097	24512	...	24478	4182	6436	18032	...	27784	4471	...	12795	21382	...	17647	7387	11724	21321	...	20123	5443
Nitra	159379	307651	...	284809	55158	108757	254869	...	322110	65385	...	153985	286867	...	226953	75610	138000	268543	...	251175	62335

Nové Mesto nad Váhom	63538	114279	...	125273	32000	50123	105898	...	133050	32077	...	76378	123042	...	110297	46126	64058	119142	...	122642	38703
Nové Zámky	63793	125670	...	121903	30329	55219	109957	...	126827	25926	...	48295	82188	...	61490	21326	38036	75773	...	70549	19930
Partizánske	18946	39293	...	38393	7304	13421	30630	...	37081	6211	...	17378	32742	...	25214	8934	15053	31137	...	28307	8192
Pezinok	1103	2841	...	3544	541	938	1985	...	3429	491	...	225	419	...	380	109	161	446	...	473	108
Piešťany	22316	39611	...	43335	12088	17429	36842	...	43238	12118	...	22429	35809	...	31350	12698	18960	34910	...	36146	10894
Poltár	4566	8149	...	8356	1780	3533	7854	...	9923	1871	...	5676	8952	...	5993	2635	4288	8133	...	7823	2205
Poprad	57214	106846	...	111295	25166	41898	94328	...	124547	24548	...	66587	112268	...	94308	41062	57968	107943	...	108881	35355
Považská Bystrica	54867	98545	...	103191	18047	41708	92029	...	108581	18696	...	63748	105663	...	84698	31147	56290	101851	...	98960	26046
Prešov	179619	314045	...	348765	73097	146132	315777	...	391391	80205	...	198154	293952	...	253803	121426	183001	293604	...	323049	107546
Prievidza	140738	266914	...	272769	49148	95684	219213	...	284977	57735	...	127421	219124	...	156401	54588	103842	200700	...	173509	48402
Púchov	61144	107153	...	109241	18119	49776	104214	...	108456	15992	...	72094	115307	...	82035	27209	59812	106194	...	93624	23387
Revúca	6991	12025	...	13393	3068	5673	11784	...	15124	3616	...	8048	13096	...	8563	4749	6420	11634	...	12247	4090
Rimavská Sobota	99807	176599	...	178970	36166	74256	171604	...	201982	37167	...	96138	150964	...	98078	41638	73678	136264	...	134234	35304
Rožňava	107032	193113	...	207765	50225	92845	181561	...	240217	54890	...	72077	113823	...	90326	42898	54524	100433	...	98853	34503
Ružomberok	190931	357326	...	392244	77932	134561	332801	...	414261	73317	...	98734	175932	...	135675	55579	82368	163201	...	160498	48641
Sabinov	13708	25017	...	28045	6390	9774	20638	...	29823	6887	...	14138	23971	...	20278	8620	11748	22921	...	23204	7058
Šaľa	35039	68731	...	57674	18320	30469	56505	...	59193	12179	...	28167	44565	...	32462	10586	20734	39049	...	37924	11511
Senec	89034	173107	...	195496	36298	67818	160473	...	193710	33782	...	117201	210998	...	196025	71963	107650	216221	...	236346	63349
Senica	149126	253084	...	229426	32970	108535	245270	...	240334	32725	...	135456	198102	...	130158	41293	114493	189183	...	158729	39092
Skalica	66678	121340	...	97231	13681	42535	94694	...	88192	11679	...	34451	57946	...	44236	11241	33003	62487	...	47806	8568
Snina	4966	8893	...	9072	1997	3335	6235	...	8887	1822	...	3700	7162	...	6204	3149	3552	6799	...	9127	2760
Sobrance	17009	32019	...	35612	9625	12316	22982	...	34744	9073	...	11346	20979	...	24190	12590	12740	22708	...	27842	24565
Spišská Nová Ves	1595	3132	...	3215	828	1254	2868	...	3384	733	...	264	608	...	520	212	277	525	...	520	131
Stará Ľubovňa	20687	40044	...	40354	9128	12825	28993	...	47136	9840	...	17345	29991	...	24000	10687	13967	28869	...	25851	7515
Stropkov	30310	56164	...	59801	10566	21842	48229	...	61257	12332	...	26950	45109	...	34763	15667	22921	42821	...	19554	7429
Svidník	329675	549745	...	598593	110227	267218	586958	...	669565	128742	...	351748	493941	...	412267	204965	336709	500950	...	601302	207900
Topoľčany	23915	48247	...	51074	12774	18014	41601	...	52968	9833	...	22839	40150	...	36368	15041	22587	44908	...	41817	10299
Trebišov	75859	137969	...	153719	29941	57576	125937	...	157304	33432	...	66832	113635	...	108094	51007	69392	124151	...	118215	42995
Trenčín	193545	340185	...	384892	88911	148281	322722	...	398398	87679	...	176323	281780	...	235554	89710	144798	265529	...	248546	74442
Trnava	237668	434251	...	427105	74310	177811	392747	...	455180	77811	...	255638	420310	...	322617	105554	208561	393040	...	359805	97181
Turčianske Teplice	50321	88247	...	106642	19630	48824	102413	...	128695	23681	...	79340	119700	...	89725	39192	67169	118137	...	111944	33013
Tvrdošín	38924	73747	...	69915	10791	24916	59670	...	76995	12024	...	33862	67117	...	49352	20135	31648	62641	...	62108	15788
Veľký Krtíš	9686	19405	...	19290	5459	7861	16319	...	21527	5216	...	8267	14233	...	16132	5721	7774	16489	...	14803	3925
Vranov nad Topľou	99037	180497	...	206180	48289	80255	165736	...	212646	50488	...	73911	132131	...	108544	49574	62444	121673	...	139445	46425
Žarnovica	35334	66570	...	74008	13570	26655	66730	...	80051	13130	...	46950	80020	...	64382	22935	37812	77147	...	67953	18507
Žiar nad Hronom	83916	152153	...	169661	31921	72453	158169	...	195468	36164	...	124462	194983	...	150082	61509	103311	189704	...	173620	50258

Žilina	285362	510108	...	551644	98832	215973	478638	...	589515	99629	...	213720	348853	...	272267	106417	185671	332926	...	322232	86907
Zlaté Moravce	30714	63854	...	66212	11044	22511	59438	...	69801	11182	...	33254	61065	...	49161	17804	28424	59109	...	55771	14668
Zvolen	215506	385335	...	402755	85770	159944	359812	...	463059	87799	...	245587	387255	...	304270	128101	201636	368291	...	339503	107054

Table 3. Toll data: counts of passages through sections of monitored roads in Slovakia (2018-2023; w = week); abridged.
Source: Author's construction.

	2018					2019						2022					2023				
DISTRICT	w 1	w 2	. . .	w 51	w 52	w 1	w 2	. . .	w 51	w 52	. . .	w 1	w 2	. . .	w 52	w 53	w 1	w 2	. . .	w 51	w 52
Bánovce nad Bebravou	4.563219956	4.816453423		4.884109691	4.245882648	4.449061647	4.78363206		4.896427002	4.276346896		4.409442787	4.629185258		4.512657719	4.036109667	4.30427505	4.596179092		4.517565353	4.031812271
Banská Bystrica	4.965963283	5.267164686		5.232686152	4.529866483	4.682659424	5.134272757		5.283574838	4.496071276		4.866399591	5.131124429		5.027859044	4.637409668	4.785742575	5.094480161		5.016092742	4.54036702
Banská Štiavnica	3.424718337	3.758003009		3.787247788	3.231469959	3.390581879	3.715083671		3.825750581	3.164650216		3.401572846	3.649042634		3.592953572	3.182984967	3.393750064	3.620136055		3.611298362	3.158362492
Bardejov	4.090681204	4.391164552		4.388012343	3.747567163	3.886998046	4.272584149		4.480179476	3.720324717		4.103256233	4.390670213		4.330312291	4.057095529	4.051191125	4.399829417		4.424048168	3.906550519
Bratislava	5.039588595	5.293861762		5.299831863	4.510464578	4.901191233	5.302304948		5.300491136	4.487209102		5.129361147	5.329210696		5.362197753	4.903778883	5.219057771	5.461339426		5.432966086	4.865080274
Brezno	4.550032571	4.934296407		4.910539031	4.193597761	4.428814872	4.837878713		4.965667271	4.113709438		4.544700413	4.785237271		4.715978287	4.27584094	4.432295692	4.77526812		4.772035506	4.163817094
Bytča	4.652594411	4.933148021		4.933482226	4.098262902	4.431894244	4.813173855		4.98869716	4.085754244		4.567520404	4.793504019		4.677789391	4.216878142	4.530148451	4.793133559		4.752471007	4.124862728
Čadca	5.096263302	5.333179158		5.343392837	4.531555555	4.962525597	5.321389206		5.382382705	4.547553196		5.122900377	5.309979166		5.193444633	4.751894688	5.069453486	5.303660353		5.288211206	4.676327734
Detva	4.445386223	4.697839368		4.733438027	4.071550639	4.353973902	4.69348072		4.817928782	4.059070913		4.557025722	4.768186555		4.608065823	4.184208613	4.36815667	4.701429818		4.655388178	4.098885894
Dolný Kubín	5.130166452	5.398066804		5.412430305	4.660419576	5.003370045	5.353171582		5.446398904	4.670941281		4.970458458	5.258131828		5.132762232	4.734319681	4.931554483	5.227871286		5.231788058	4.643432935
Dunajská Streda	5.067249449	5.337115683		5.325047457	4.598276605	4.942851007	5.300688941		5.349918295	4.518303322		4.935814488	5.172310969		5.069716185	4.615160664	4.897335631	5.158850799		5.081692877	4.473501568
Galanta	5.305714563	5.586862287		5.577088339	4.887510401	5.193082837	5.536942068		5.646369413	4.877969996		5.338846785	5.556792012		5.443017209	4.967271774	5.250361442	5.516114757		5.489971129	4.937823696
Gelnica	2.356025857	2.772321707		2.847572659	2.082785537	2.243038049	2.690196608		2.811575006	2.049218023		1.568201724	1.897627091		1.792391689	1.505149978	1.633468456	1.857332496		1.806179974	1.491361694
Hlohovec	4.643294722	4.887313413		4.922341324	4.363781153	4.546998926	4.871689666		4.921181672	4.310565872		4.6740432	4.875819169		4.797828229	4.397140171	4.590306715	4.858843862		4.882154405	4.343999069
Humenné	4.031287249	4.311732675		4.34348737	3.657342737	3.920801382	4.223469981		4.334875777	3.660675788		3.972850556	4.25990432		4.187689894	3.842234686	3.929470016	4.230857477		4.28867392	3.814780146
Ilava	4.853059073	5.104596354		5.138365864	4.500757754	4.726914575	5.058167801		5.158084938	4.49123559		4.921540819	5.119321886		5.051893087	4.672153296	4.849241237	5.105114882		5.106027624	4.577307933
Kežmarok	4.327542832	4.596355143		4.590574357	3.929061124	4.135069014	4.491935704		4.615255467	3.943890048		4.260858132	4.482273157		4.421357099	4.07070266	4.181957861	4.461918512		4.493151218	3.994493123
Komárno	4.924129325	5.236202091		5.198365627	4.475162519	4.829599636	5.168635005		5.072657891	4.344824388		5.009739763	5.236207133		5.120873228	4.653453721	4.950807687	5.239614476		5.229520583	4.600155784
Košice	5.312317019	5.566839191		5.598547584	4.982348133	5.238542873	5.542864002		5.636106593	4.955668512		5.328585488	5.517260575		5.448104691	5.122906922	5.268355562	5.514080003		5.557902819	5.080012847
Krupina	4.968417566	5.218459353		5.227076465	4.566531766	4.830614328	5.167969763		5.306615861	4.610425922		5.077360637	5.238542873		5.128260298	4.820503587	5.004424456	5.217752334		5.192149125	4.744347756

Kysucké Nové Mesto	5.299472432	5.533979118		5.550147603	4.732659081	5.186961983	5.536904225		5.581752879	4.743972034		5.131207911	5.320366494		5.206156385	4.758374366	5.07718979	5.311395685		5.300486788	4.687751848
Levice	5.126491816	5.373640124		5.389830313	4.686305508	4.995270936	5.334105774		5.456086537	4.737470947		5.158793557	5.356188448		5.220309491	4.827130087	5.06171 2526	5.324841944		5.25920 432	4.722716167
Levoča	4.504266 92	4.749968084		4.790369622	4.194708552	4.394259007	4.727581919		4.831 27 453	4.190 75 178		4.517538974	4.706188457		4.652023817	4.294355985	4.424244422	4.687546796		4.714782263	4.233351528
Liptovský Mikuláš	4.759698055	5.060988532		5.054976729	4.386588134	4.611712689	4.994075077		5.126212607	4.367262715		4.799009532	5.043912167		4.945370094	4.557843 92	4.729755 87	5.026827325		5.037613722	4.51644 282
Lučenec	5.228608216	5.484658518		5.495050098	4.818450243	5.128589877	5.463516736		5.541495622	4.819945144		5.070011993	5.256441114		4.967590093	4.482673534	4.869607163	5.180383966		5.140680847	4.536192168
Malacky	4.98788961	5.233977877		5.251392416	4.422885178	4.867443916	5.277710041		5.264163528	4.423966369		5.084647761	5.273066122		5.136092447	4.629837566	5.029192971	5.256604709		5.215605843	4.605423572
Martin	5.297213196	5.557217174		5.600228842	4.869607163	5.211774429	5.551138805		5.622077183	4.884018942		5.204798038	5.451657509		5.323023029	4.917631526	5.135313967	5.426464088		5.414349855	4.846868828
Medzilaborce	2.021189299	2.309630167		2.346352974	1.690 19608	1.826074803	2.396199347		2.409933123	1.556302501		1.806179974	2.068185862		2.117271296	1.643452676	1.892094603	2.110 58971		2.123851641	1.602059991
Michalovce	4.795219388	5.078 98939		5.111927564	4.477150207	4.684189836	4.988072667		5.117 09887	4.481757843		4.640272687	4.923275113		4.887780411	4.570986348	4.630875161	4.918187661		4.997818716	4.693463127
Myjava	3.211387553	3.508664363		3.510008513	2.685741739	3.024074987	3.358125285		3.545925329	2.700703717		2.936513742	3.187238 62		3.046104787	2.629409599	2.833784375	3.169674434		3.087426457	2.448706 32
Námestovo	4.082677681	4.389378748		4.388 77593	3.621384028	3.808616035	4.256043899		4.443 79477	3.650 40467		4.107 04029	4.330048325		4.246670886	3.868468099	4.069 07581	4.328 80757		4.303692727	3.735838334
Nitra	5.202431098	5.488058331		5.454553709	4.741608511	5.036457219	5.406317015		5.508004208	4.815478128		5.187478417	5.457680592		5.355935928	4.878579238	5.139879086	5.429013836		5.399976411	4.794731964
Nové Mesto nad Váhom	4.803033 54	5.057966431		5.097857478	4.505149978	4.700037057	5.024887758		5.124014879	4.506193744		4.882968282	5.090053382		5.042563 7	4.663945795	4.806573375	5.076064886		5.088639224	4.587744 63
Nové Zámky	4.804773026	5.099231615		5.086014394	4.481858091	4.742088537	5.041222882		5.103 21172	4.413735517		4.683 90217	4.914808412		4.788804493	4.328909405	4.580194839	4.879514482		4.848490862	4.299507299
Partizánske	4.277517533	4.594315188		4.584252049	3.863560765	4.127784876	4.486146997		4.569151438	3.793161529		4.239999793	4.515105204		4.401641748	3.951045948	4.177623062	4.493276767		4.451893845	3.913389944
Pezinok	3.042575512	3.453471234		3.549493713	2.733197265	2.972202838	3.297760511		3.535167485	2.691081492		2.352182518	2.622214023		2.579783597	2.037426498	2.206825876	2.649334859		2.674861141	2.033423755
Piešťany	4.348616353	4.597815807		4.636838801	4.082354451	4.241 27247	4.566343198		4.635865597	4.083430948		4.350809911	4.553992193		4.496237545	4.103735323	4.277838333	4.542949849		4.558060244	4.037187371
Poltár	3.659535907	3.911104318		3.921998431	3.250420002	3.548143637	3.895090897		3.996642991	3.272073788		3.754042387	3.951920074		3.777644278	3.420 78062	3.632254777	3.910250772		3.893 37333	3.343408594
Poprad	4.757502312	5.028758268		5.046475654	4.400814192	4.622193292	4.974640626		5.095333271	4.390016115		4.823389449	5.050255986		4.974548535	4.613440099	4.763188317	5.033194484		5.036952101	4.548450842
Považská Bystrica	4.739311215	4.993634594		5.013641821	4.256405018	4.620219365	4.963924703		5.035753837	4.271748699		4.804466564	5.023922937		4.927873155	4.493416223	4.750431249	5.007965297		4.995459687	4.415741036
Prešov	5.254352274	5.496991883		5.542532895	4.863899553	5.164745328	5.499380494		5.592610835	4.904201443		5.297002844	5.468276419		5.404496751	5.084311689	5.262453463	5.467761968		5.509268401	5.031594262
Prievidza	5.148411375	5.426371354		5.435795012	4.691 50585	4.980839322	5.340866306		5.454 80981	4.761 43917		5.105241009	5.340689947		5.194239526	4.737097183	5.016373044	5.302547372		5.239322007	4.684863307
Púchov	4.786353847	5.030004335		5.038385667	4.258134225	4.697019994	5.017926066		5.035253583	4.203902781		4.857899122	5.061855673		4.913999182	4.434712581	4.776788325	5.026 09998		4.971387192	4.368974516

Revúca	3.844539302	4.080085085		4.126877869	3.486855355	3.753812784	4.071292734		4.179666669	3.558228422		3.905687968	4.117138666		3.932625944	3.676602217	3.807535028	4.065729059		4.088029718	3.611723308
Rimavská Sobota	4.999161002	5.246988824		5.252780238	4.558300478	4.870731551	5.234527407		5.305312668	4.570157508		4.982895083	5.178873395		4.991571601	4.619489861	4.867337828	5.134381133		5.127862532	4.547823914
Rožňava	5.029513641	5.285811511		5.317572388	4.700919945	4.967758521	5.259022566		5.380603739	4.739493231		4.857796702	5.056230028		4.955812778	4.632437045	4.736587709	5.001876436		4.994989854	4.537856858
Ružomberok	5.280876447	5.553064618		5.593556309	4.891715822	5.128919206	5.522184623		5.617274049	4.865204686		4.994466732	5.245344484		5.132499983	4.744910729	4.915758521	5.212722816		5.205469625	4.687002495
Sabinov	4.136974096	4.398235229		4.447855444	3.805500858	3.990072335	4.314667608		4.474551329	3.838030083		4.150387977	4.379686152		4.307025119	3.935507266	4.069963938	4.360233561		4.365562857	3.848681654
Šaľa	4.544551703	4.837152663		4.760980073	4.262925469	4.483858201	4.752086879		4.772270351	4.085610163		4.449740594	4.648993911		4.511375273	4.024731189	4.316683094	4.591609917		4.578914138	4.061113054
Senec	4.949555885	5.238314630		5.291137876	4.559882696	4.831344978	5.205401972		5.287152041	4.528685358		5.068931317	5.324278339		5.292311463	4.857109260	5.032014034	5.334897872		5.373548257	4.801739764
Senica	5.173553369	5.403264690		5.360642633	4.518118947	5.035569811	5.389644431		5.380815215	4.514879655		5.131798247	5.296888860		5.114470867	4.615876436	5.058778935	5.276882108		5.200656280	4.592087890
Skalica	4.823982565	5.084003991		4.987804752	4.136117843	4.628746437	4.976322462		4.945429192	4.067405658		4.537201833	4.763023462		4.645775849	4.050804948	4.518553419	4.795789675		4.679482407	3.932879458
Snina	3.696006715	3.949048292		3.957703042	3.300378065	3.523095838	3.794836458		3.948755180	3.260548373		3.568201724	3.855034317		3.792671789	3.498172661	3.550472957	3.832445041		3.960328051	3.440909082
Sobrance	4.230678781	4.505407764		4.551596365	3.983400738	4.090469680	4.361387820		4.540879816	3.957750911		4.054842779	4.321784783		4.383635868	4.100025730	4.105169428	4.356178886		4.444700429	4.390316768
Spišská Nová Ves	3.202760687	3.495821753		3.507180977	2.918030337	3.098297536	3.457579147		3.529430354	2.865103975		2.421603927	2.783903579		2.716003344	2.326335861	2.442479769	2.720159303		2.716003344	2.117271296
Stará Ľubovňa	4.315697514	4.602537453		4.605880659	3.960375631	4.108057374	4.462293156		4.673352725	3.992995098		4.239174304	4.476990947		4.380211242	4.028855809	4.145103133	4.460431741		4.412477348	3.875928985
Stropkov	4.481585936	4.749458031		4.776708446	4.023910606	4.339292403	4.683308257		4.787155724	4.091033516		4.430558770	4.654263200		4.541117248	4.194985843	4.360233561	4.631656805		4.291235611	3.870930358
Svidník	5.518086015	5.740161288		5.777131634	5.042287988	5.426865709	5.768607026		5.825792744	5.109720252		5.546231637	5.693675077		5.615178573	5.311679707	5.527254725	5.699794381		5.779092648	5.317854489
Topoľčany	4.378670385	4.683470314		4.708199872	4.106326912	4.255610158	4.619103770		4.724013575	3.992686039		4.358677085	4.603685550		4.560719418	4.177276711	4.353858552	4.652323714		4.621352873	4.012795058
Trebišov	4.880007113	5.139781517		5.186727551	4.476266301	4.760241490	5.100153344		5.196739766	4.524162358		4.824984458	5.055512116		5.033801588	4.707629781	4.841309405	5.093950222		5.072672587	4.633417953
Trenčín	5.286781956	5.531715160		5.585338884	4.948955495	5.171085506	5.508828572		5.600317150	4.942895588		5.246308966	5.449910165		5.372090483	4.952840857	5.160762563	5.424111960		5.395406778	4.871818033
Trnava	5.375970712	5.637740827		5.630534656	4.871047261	5.249958624	5.594112876		5.658183171	4.891040997		5.407625411	5.623569723		5.508687248	5.023474696	5.319233101	5.594436751		5.556067194	4.987581364
Turčianske Teplice	4.701749263	4.945699950		5.027928281	4.292920300	4.688633357	5.010355088		5.109561674	4.374400038		4.899492196	5.078090415		4.952913467	4.593197426	4.827168883	5.072385938		5.049000821	4.518684992
Tvrdošín	4.590217464	4.867744358		4.844570362	4.033061693	4.396478322	4.775756038		4.886462523	4.080048968		4.529712605	4.826832536		4.693304757	4.303951634	4.500340627	4.796858682		4.793147544	4.198327118
Veľký Krtíš	3.986144465	4.287913647		4.285332228	3.737113094	3.895477796	4.212693542		4.332983511	3.717337583		3.917347938	4.153296449		4.207688213	3.757471948	3.890644536	4.217194318		4.170349739	3.593839661
Vranov nad Topľou	4.995797476	5.256469988		5.314246535	4.683848212	4.904472099	5.219416853		5.327657218	4.703188167		4.868709078	5.121004722		5.035605822	4.695253962	4.795490715	5.085194216		5.144402947	4.666751912
Žarnovica	4.548192805	4.823278557		4.869278668	4.132579848	4.425778687	4.824321125		4.903366762	4.118264726		4.671635597	4.903198547		4.808764464	4.360498744	4.577629649	4.887319042		4.832208635	4.267336025

Žiar nad Hronom	4.923844774	5.18228052		5.229582022	4.504076488	4.860056373	5.199121369		5.291075669	4.558276461		5.095036775	5.289996748		5.176328609	4.788938666	5.014146565	5.278076488		5.239599752	4.701205202
Žilina	5.45539614	5.707662135		5.741658899	4.994897584	5.334399461	5.680007175		5.77049486	4.998385771		5.329845166	5.542642462		5.434995006	5.027011012	5.268744076	5.522347713		5.508168667	4.939054758
Zlaté Moravce	4.48733638	4.805188108		4.820936706	4.043126398	4.352394788	4.774064188		4.843861645	4.048519488		4.521843892	4.785792361		4.691620708	4.250517586	4.453685194	4.771653612		4.746408432	4.166370901
Zvolen	5.333459366	5.585838458		5.605040941	4.93333541	5.203967953	5.556075643		5.66563633	4.943489569		5.390205374	5.587997034		5.483259134	5.10755252	5.304568073	5.566191106		5.530843616	5.029602899

Table 4. Toll data: counts of passages (log10) through sections of monitored roads in Slovakia (2018-2023; w = week); abridged.
Source: Author's construction.

No.	Region	District	2018	2019	2020	2021	2022
	Bratislava		**28 381 270 358**	**28 418 710 673**	**26 255 541 584**	**28 674 050 371**	**38 129 203 026**
1		Bratislava I	2 324 888 777	2 710 764 586	2 559 652 932	2 831 897 293	3 508 261 564
2		Bratislava II	9 906 999 036	9 277 502 227	8 476 799 944	10 549 454 276	18 059 478 411
3		Bratislava III	833 091 393	884 341 682	769 354 780	796 023 347	1 309 279 951
4		Bratislava IV	11 921 219 439	12 082 883 054	11 305 370 799	11 282 822 694	11 522 077 735
5		Bratislava V	279 041 795	257 390 248	217 397 165	222 813 898	274 978 934
6		Malacky	2 427 844 418	2 388 313 464	2 231 545 560	2 204 827 055	2 542 868 913
7		Pezinok	322 327 766	349 698 285	140 970 668	147 904 398	169 746 716
8		Senec	365 857 734	467 817 127	554 449 736	638 307 410	742 510 802
	Trnava		**11 304 213 615**	**11 592 239 689**	**11 263 932 715**	**12 814 885 705**	**14 805 700 391**
9		Dunajská Streda	706 606 665	764 253 953	801 783 688	948 929 489	1 091 081 852
10		Galanta	2 612 639 163	2 279 283 924	1 950 506 314	2 413 365 360	2 671 003 742
11		Hlohovec	895 338 899	1 049 085 615	1 015 455 245	1 339 597 989	1 771 883 159
12		Piešťany	471 309 528	471 547 389	490 105 705	493 241 659	468 366 358
13		Senica	602 270 311	554 294 762	516 324 036	735 020 471	943 594 702
14		Skalica	1 317 507 510	1 358 184 955	1 386 291 287	1 561 231 933	1 751 208 046
15		Trnava	4 698 541 539	5 115 589 091	5 103 466 440	5 323 498 804	6 108 562 532
	Trenčín		**9 596 299 870**	**9 578 506 025**	**8 570 203 040**	**9 900 204 964**	**11 985 926 751**
16		Bánovce nad Bebravou	660 804 353	669 260 633	607 729 171	406 500 876	466 165 678
17		Ilava	1 026 968 678	1 095 550 615	1 046 958 719	1 188 667 796	1 581 339 680
18		Myjava	284 664 536	262 117 548	242 763 990	290 956 524	326 208 832
19		Nové Mesto nad Váhom	1 797 946 943	1 780 546 823	1 549 131 085	1 943 217 854	2 226 994 982
20		Partizánske	317 065 802	315 319 100	266 433 694	311 541 346	389 570 624
21		Považská Bystrica	1 173 532 465	1 118 849 002	995 540 664	1 048 105 322	1 214 010 935
22		Prievidza	1 347 154 920	1 335 799 900	1 217 472 031	1 467 783 136	1 754 750 164
23		Púchov	2 203 572 985	2 210 415 659	1 868 121 193	2 321 169 193	2 901 220 789
24		Trenčín	784 589 188	790 646 745	776 052 493	922 262 917	1 125 665 067
	Nitra		**6 530 472 981**	**6 398 279 444**	**5 774 101 621**	**6 883 647 933**	**8 764 675 058**
25		Komárno	791 054 933	826 601 177	746 202 062	874 075 605	1 035 739 193

26		Levice	729 739 196	717 299 259	729 329 477	780 138 045	975 642 255
27		Nitra	2 646 843 828	2 628 513 318	2 323 003 253	2 722 727 155	3 406 907 550
28		Nové Zámky	543 788 027	525 484 504	527 769 207	639 357 946	766 200 781
29		Šaľa	681 648 120	646 089 611	540 411 053	825 470 517	1 323 430 636
30		Topoľčany	779 190 414	712 247 251	624 159 463	720 770 374	900 443 738
31		Zlaté Moravce	358 208 463	342 044 324	283 227 106	321 108 291	356 310 905
	Žilina		**14 440 463 320**	**14 989 418 847**	**12 641 656 406**	**14 802 041 456**	**18 252 313 536**
32		Bytča	364 394 146	382 152 300	331 260 839	396 259 615	421 508 984
33		Čadca	228 213 489	241 192 153	207 038 745	221 152 675	286 709 188
34		Dolný Kubín	632 950 974	619 468 447	535 712 648	698 661 720	676 376 461
35		Kysucké Nové Mesto	843 599 256	797 885 779	661 593 173	759 859 931	903 931 552
36		Liptovský Mikuláš	591 475 702	574 464 524	535 917 794	649 143 911	768 879 807
37		Martin	900 997 230	955 187 709	788 981 036	794 592 329	849 545 284
38		Námestovo	594 079 209	644 895 952	541 804 493	613 685 916	778 748 096
39		Ružomberok	607 051 159	577 999 094	503 866 293	693 556 477	973 339 992
40		Turčianske Teplice	63 886 837	62 167 636	58 021 859	66 479 398	80 703 315
41		Tvrdošín	321 103 484	373 906 480	393 500 530	510 192 107	693 450 089
42		Žilina	9 292 711 834	9 760 098 773	8 083 958 996	9 398 457 377	11 819 120 768
	Banská Bystrica		**4 045 512 765**	**3 961 532 965**	**3 678 997 303**	**4 516 343 877**	**5 580 453 081**
43		Banská Bystrica	701 559 550	712 376 796	694 880 862	811 599 798	893 106 730
44		Banská Štiavnica	9 498 430	11 577 607	13 114 750	15 698 925	16 935 208
45		Brezno	426 400 004	391 043 201	339 715 661	491 591 826	689 453 338
46		Detva	164 019 620	185 427 900	219 928 379	298 688 735	442 404 227
47		Krupina	97 924 172	103 384 798	103 881 231	109 300 933	118 535 430
48		Lučenec	311 773 786	288 237 806	271 794 964	338 490 442	361 488 280
49		Poltár	18 348 132	18 314 714	16 715 157	19 717 099	27 577 193
50		Revúca	194 363 941	195 423 135	161 432 374	187 236 808	226 494 243
51		Rimavská Sobota	150 233 798	159 533 370	167 577 564	189 238 570	253 580 667
52		Veľký Krtíš	120 320 231	141 087 793	131 383 435	161 861 848	206 961 983
53		Zvolen	704 244 590	643 629 874	554 087 207	586 084 866	690 550 147

54		Žarnovica	261 807 312	278 905 669	281 091 624	338 587 166	390 253 630
55		Žiar nad Hronom	885 019 199	832 590 302	723 394 095	968 246 861	1 263 112 005
	Prešov		**3 664 271 810**	**3 723 304 271**	**3 494 162 632**	**3 994 288 700**	**4 760 202 260**
56		Bardejov	128 945 647	117 529 837	138 201 271	152 917 615	174 895 212
57		Humenné	426 020 846	406 502 608	368 945 535	411 868 601	508 681 383
58		Kežmarok	307 837 991	292 401 831	287 373 168	331 560 926	402 316 382
59		Levoča *)					
60		Medzilaborce *)					
61		Poprad	1 194 459 318	1 249 897 682	1 272 736 010	1 388 934 429	1 612 747 269
62		Prešov	950 458 484	996 487 468	852 061 687	1 042 379 588	1 214 517 645
63		Sabinov	107 978 977	108 489 366	93 202 361	114 241 352	133 881 767
64		Snina	99 491 374	95 296 556	82 679 827	93 742 471	110 748 638
65		Stará Ľubovňa	84 381 882	86 940 785	91 932 382	104 614 817	152 160 469
66		Stropkov	57 673 327	64 245 839	53 238 454	59 932 241	69 857 489
67		Svidník	23 230 053	21 071 643	14 174 195	15 510 774	24 599 027
68		Vranov nad Topľou	274 068 536	274 199 211	229 389 671	264 164 399	338 665 467
	Košice		**7 420 958 636**	**6 677 895 739**	**5 920 686 373**	**8 605 856 086**	**9 834 613 434**
69		Gelnica	35 160 033	39 387 679	29 295 954	31 906 559	44 734 746
70		Košice I	1 040 375 736	1 070 365 890	982 994 048	1 011 751 365	1 414 843 938
71		Košice II	3 358 161 976	2 626 912 278	2 114 302 005	4 485 898 577	4 839 297 110
72		Košice III	13 617 079	12 112 058	13 400 319	13 430 133	17 167 180
73		Košice IV	615 411 356	627 830 934	587 319 019	649 413 882	627 276 594
74		Košice-okolie	773 112 107	721 455 623	696 699 745	796 912 948	1 065 104 728
75		Michalovce	647 037 849	637 801 326	596 247 563	606 164 397	701 575 766
76		Rožňava	387 768 932	381 537 322	370 846 439	401 510 832	454 424 450
77		Sobrance	4 776 340	4 582 898	5 072 527	5 144 750	5 907 238
78		Spišská Nová Ves	479 036 118	493 453 102	461 104 517	534 912 978	578 141 300
79		Trebišov	66 501 110	62 456 629	63 404 237	68 809 665	86 140 384

Table 5. Industrial production output by district and region in Slovakia (2008-2022).
Source: Statistical Office of the Slovak Republic. Author's construction.

No.	Region	District	Latitude	Longitude	Population	Size (km²)	Population density (pop./km²)	Total no. of cities and villages	No. of cities (out of total)
	Banská Bystrica		**48.5**	**19.5**					
1		Okres Banská Bystrica	48.73333	19.15	107 199	809.4	132	42	1
2		Okres Banská Štiavnica	48.44961	18.9082	15 407	292.3	53	15	1
3		Okres Brezno	48.8	19.75	58 297	1 265.2	46	30	1
4		Okres Detva	48.55985	19.42044	30 562	449.2	68	15	2
5		Okres Krupina	48.35736	19.06334	21 271	584.9	36	36	2
6		Okres Lučenec	48.33333	19.66667	69 018	825.6	84	57	2
7		Okres Poltár	48.42895	19.79488	20 247	476.2	43	22	1
8		Okres Revúca	48.683	20.11568	37 816	730.3	52	42	3
9		Okres Rimavská Sobota	48.38333	20.03333	79 655	1 471.1	54	107	3
10		Okres Veľký Krtíš	48.2	19.35	40 900	848.2	48	71	2
11		Okres Zvolen	48.58333	19.13333	65 578	759	86	26	2
12		Okres Žarnovica	48.48438	18.72076	24 668	425.3	58	18	2
13		Okres Žiar nad Hronom	48.58333	18.86667	43 738	517.7	84	35	2
	Bratislava		**48.33333**	**17.16667**					
14		Okres Bratislava I	48.14653	17.10584	47 375	9.6	4 940		
15		Okres Bratislava II	48.1528	17.17764	125 980	92.5	1 362		
16		Okres Bratislava III	48.18543	17.1379	77 594	74.7	1 039		
17		Okres Bratislava IV	48.22787	16.99722	105 043	96.7	1 086		
18		Okres Bratislava V	48.11122	17.09444	122 048	94.2	1 296		
19		Okres Malacky	48.43458	17.02166	79 689	949.5	84	26	2
20		Okres Pezinok	48.28785	17.26799	69 953	375.5	186	17	3
21		Okres Senec	48.22187	17.40328	105 075	359.9	292	29	1
	Košice		**48.66667**	**21.33333**					
22		Okres Gelnica	48.85207	20.93385	31 736	584.4	54	20	1
23		Košice I	48.72914	21.25004	62 965	85.4	737		
24		Košice II	48.69753	21.22273	78 385	73.9	1 061		
25		Košice III	48.73242	21.29047	27 551	16.9	1 634		
26		Košice IV	48.68691	21.2657	56 143	60.9	922		
27		Okres Košice-okolie	48.77897	21.41373	131 955	1 541.3	86	114	2
28		Okres Michalovce	48.75	21.93333	108 132	1 019.3	106	78	3
29		Okres Rožňava	48.66667	20.53333	58 422	1 173.3	50	62	2

30		Okres Sobrance	48.74247	22.18197	22 226	538.2	41	47	1
31		Okres Spišská Nová Ves	48.95	20.56667	98 556	587.4	168	36	3
32		Okres Trebišov	48.63333	21.71667	103 002	1 073.5	96	82	4
	Nitra		**48.16667**	**18.33333**					
33		Okres Komárno	47.76667	18.13333	99 273	1 100.1	90	41	4
34		Okres Levice	48.21667	18.6	108 479	1 551.1	70	89	4
35		Okres Nitra	48.31667	18.08333	164 820	870.7	189	62	2
36		Okres Nové Zámky	47.98333	18.16667	135 003	1 347.1	100	62	3
37		Okres Šaľa	48.1477	17.87306	50 384	355.9	142	13	1
38		Okres Topoľčany	48.56667	18.18333	69 577	597.6	116	54	1
39		Okres Zlaté Moravce	48.38294	18.39842	40 765	521.2	78	33	1
	Prešov		**49.16667**	**21.25**					
40		Okres Bardejov	49.28333	21.28333	75 547	936.2	81	86	1
41		Okres Humenné	48.93333	21.91667	58 688	754.2	78	62	1
42		Okres Kežmarok	49.13593	20.42929	75 188	839.3	90	41	3
43		Okres Levoča	49.01986	20.57688	33 192	357.3	93	33	2
44		Okres Medzilaborce	49.27062	21.902	10 665	427.3	25	23	1
45		Okres Poprad	49.05	20.3	102 038	1 105.4	92	29	3
46		Okres Prešov	49	21.25	174 741	933.7	187	91	2
47		Okres Sabinov	49.10115	21.09844	61 542	483.5	127	43	2
48		Okres Snina	48.98579	22.15059	34 071	804.7	42	34	1
49		Okres Stará Ľubovňa	49.3	20.7	53 107	624	85	44	2
50		Okres Stropkov	49.20249	21.65025	19 553	389	50	43	1
51		Okres Svidník	49.3	21.56667	31 183	549.8	57	68	2
52		Okres Vranov nad Topľou	48.9	21.68333	79 295	769.5	103	68	2
	Trenčín		**48.83333**	**18.25**					
53		Okres Bánovce nad Bebravou	48.71908	18.25773	35 186	461.9	76	43	1
54		Okres Ilava	48.99641	18.23374	56 920	358.5	159	21	3
55		Okres Myjava	48.75	17.58333	25 013	327.4	76	17	2
56		Okres Nové Mesto nad Váhom	48.75466	17.83506	61 349	580	106	34	2
57		Okres Partizánske	48.62613	18.37957	43 509	301	145	23	1
58		Okres Považská Bystrica	49.11667	18.45	60 577	463.2	131	28	1
59		Okres Prievidza	48.76667	18.63333	128 613	959.8	134	52	4
60		Okres Púchov	49.12406	18.32477	43 710	375.3	116	21	1

61		Okres Trenčín	48.8952	18.04264	113 225	674.8	168	37	3
	Trnava		**48.41667**	**17.66667**					
62		Okres Dunajská Streda	47.99635	17.60937	127 025	1 074.6	118	67	4
63		Okres Galanta	48.2	17.71667	95 457	641.7	149	36	3
64		Okres Hlohovec	48.41667	17.75	43 197	267.2	162	24	2
65		Okres Piešťany	48.59064	17.82679	62 015	381.1	163	27	2
66		Okres Senica	48.68333	17.36667	58 980	683.5	86	31	2
67		Okres Skalica	48.75	17.16667	46 916	357.1	131	21	3
68		Okres Trnava	48.36667	17.6	132 524	741.3	179	45	1
	Žilina		**49.16667**	**19.16667**					
69		Okres Bytča	49.22267	18.55844	31 308	281.6	111	12	1
70		Okres Čadca	49.43333	18.78333	87 135	760.6	115	23	3
71		Okres Dolný Kubín	49.2	19.3	38 734	491.9	79	24	1
72		Okres Kysucké Nové Mesto	49.30206	18.78603	32 642	173.7	188	14	1
73		Okres Liptovský Mikuláš	49.08333	19.61667	71 405	1 341.1	53	56	2
74		Okres Martin	49.06667	18.93333	93 041	735.7	126	43	3
75		Okres Námestovo	49.40651	19.4767	64 385	690.5	93	24	1
76		Okres Ružomberok	49.07494	19.30083	56 488	646.8	87	25	1
77		Okres Turčianske Teplice	48.86283	18.85759	15 753	392.8	40	26	1
78		Okres Tvrdošín	49.33421	19.55461	35 745	478.9	75	15	2
79		Okres Žilina	49.22303	18.74044	160 538	815.1	197	53	3

Table 6. Geographical information on 79 districts and 8 regions of Slovakia.
Source: Author's construction.

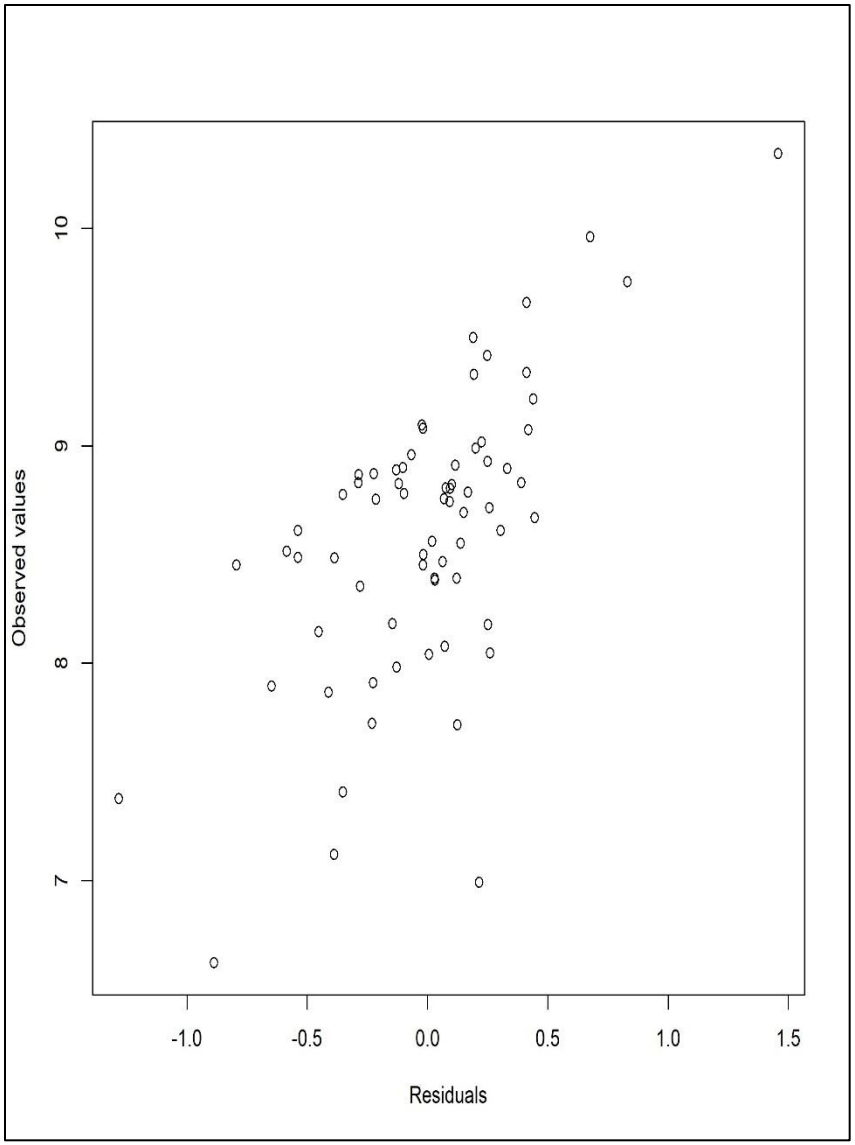

Figure 28. The observed values against residuals $\hat{\rho} = \pm 1.795$ (2018).
Source: Author's construction.

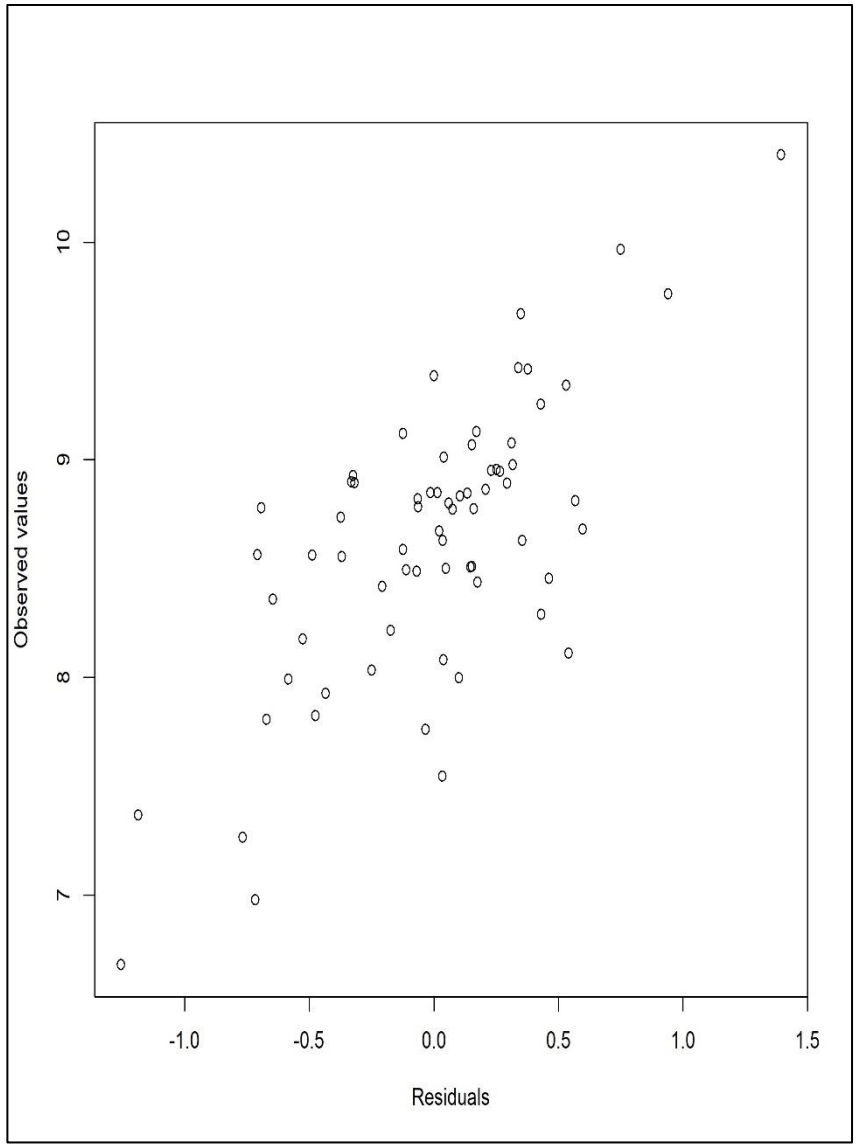

Figure 29. The observed values against residuals $\hat{\rho} = \pm 1.455$ (2019).
Source: Author's construction.

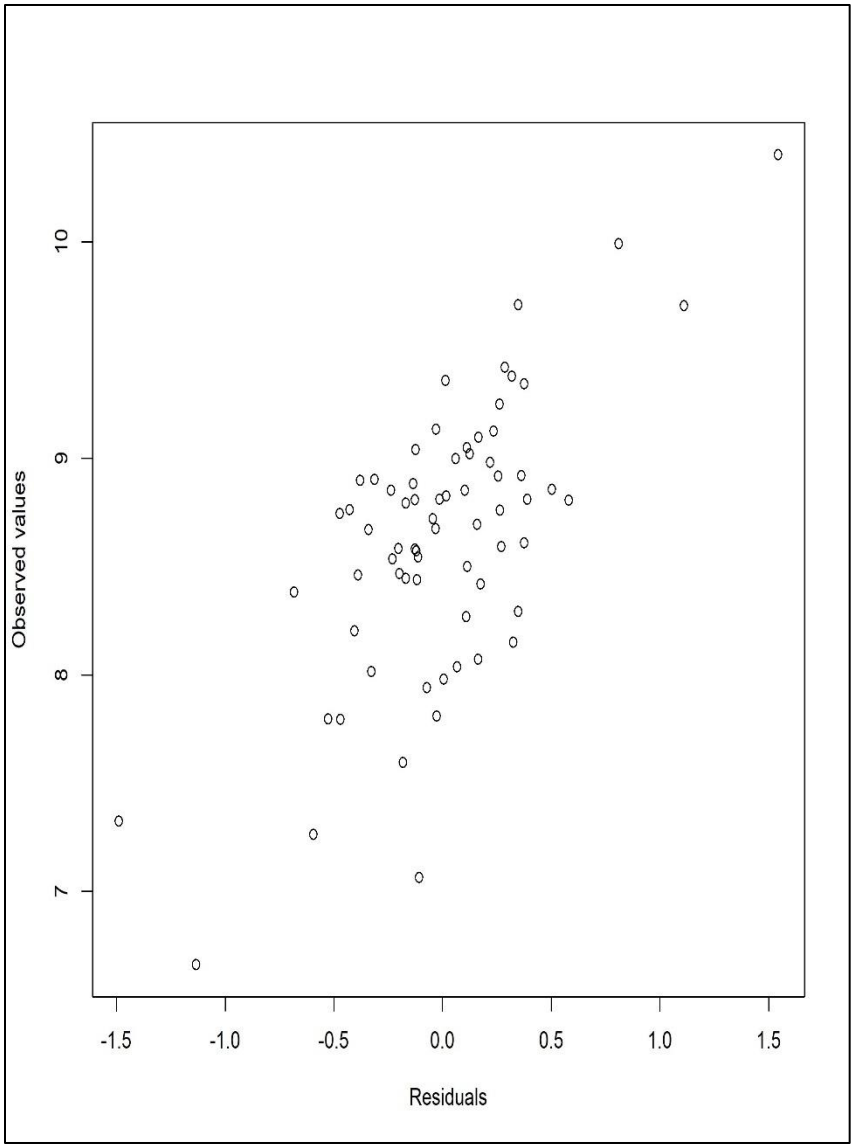

Figure 30. The observed values against residuals $\hat{\rho} = \pm 1.565$ (2020).
Source: Author's construction.

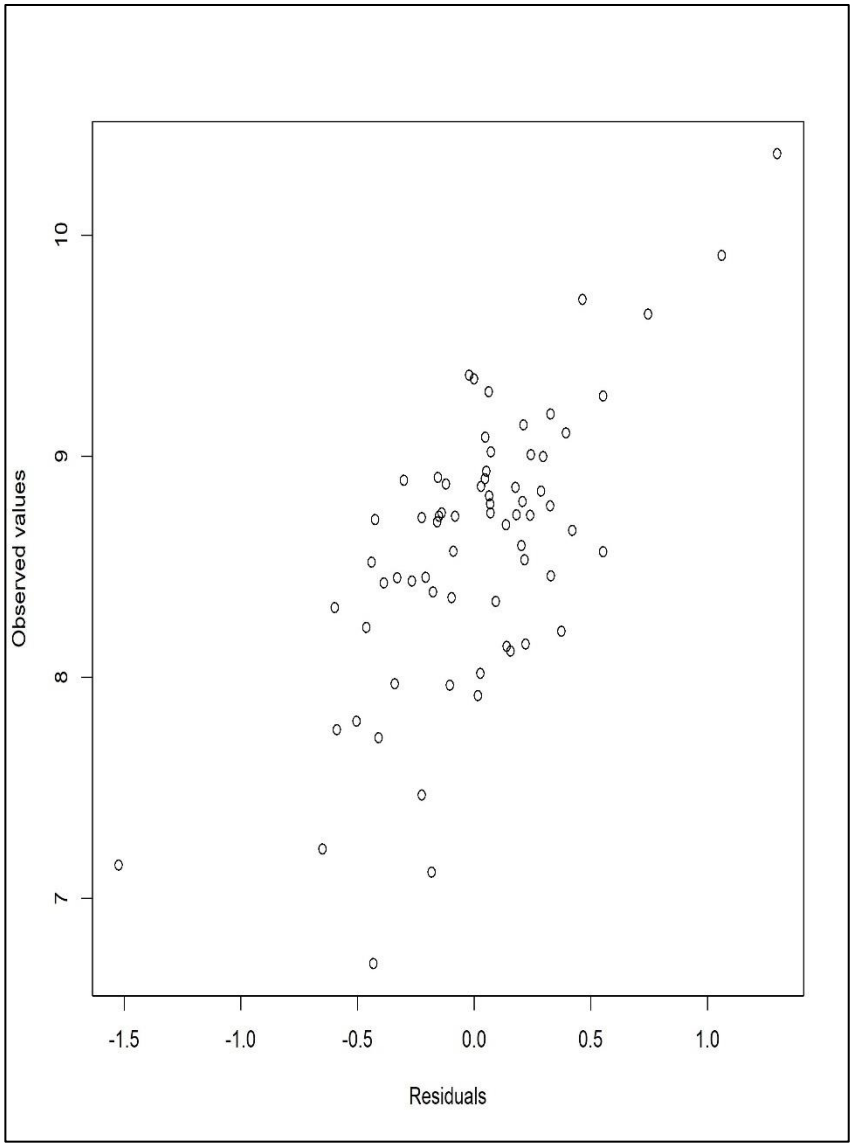

Figure 31. The observed values against residuals $\hat{\rho} = \pm 1.405$ (2021).
Source: Author's construction.

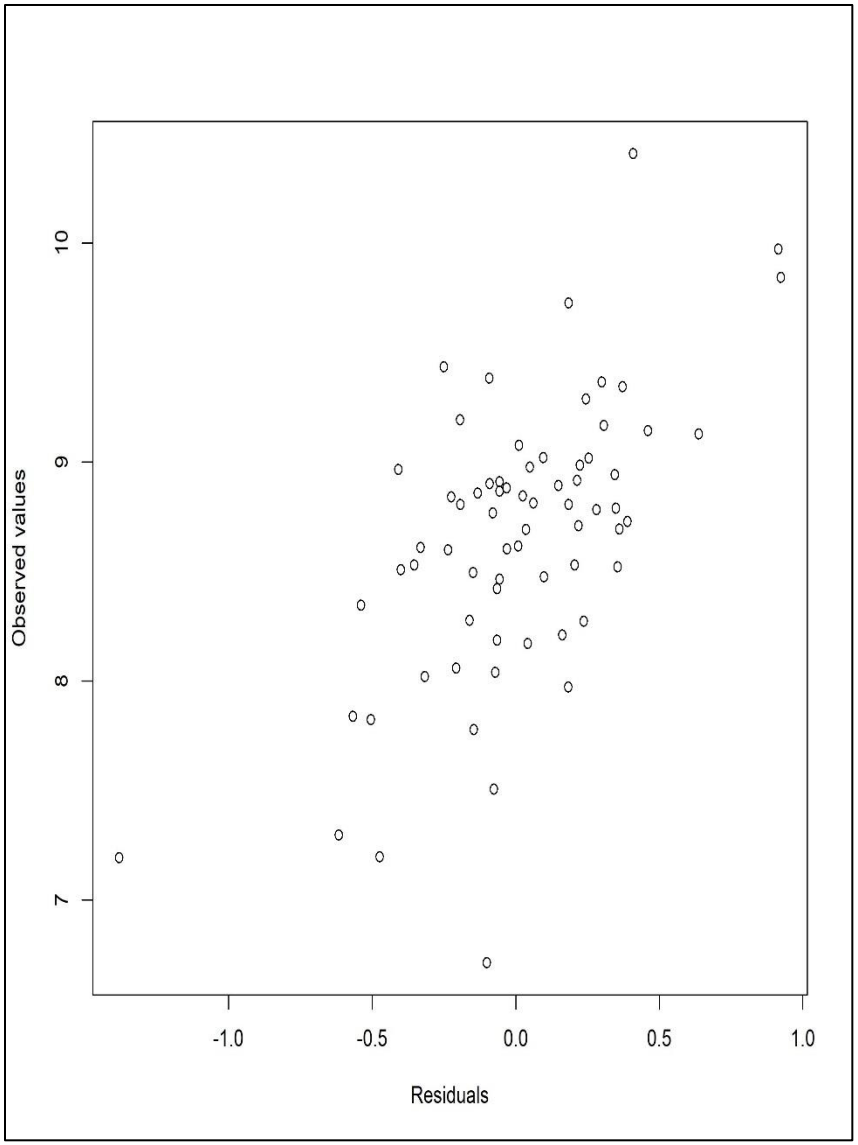

Figure 32. The observed values against residuals $\hat{\rho} = \pm 1.545$ (2022).
Source: Author's construction.

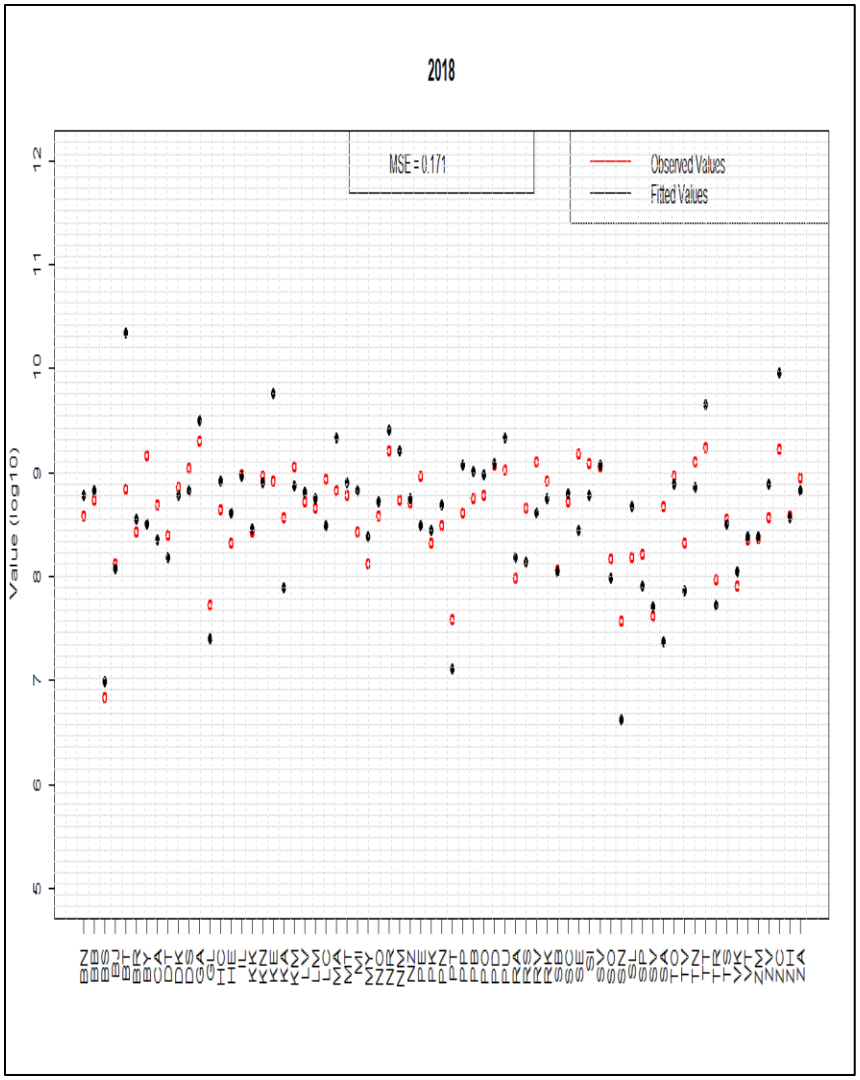

Figure 43. The predicted versus the observed responses (2018), including Mean Squared Error (MSE): Functional linear model.
Source: Author's construction.

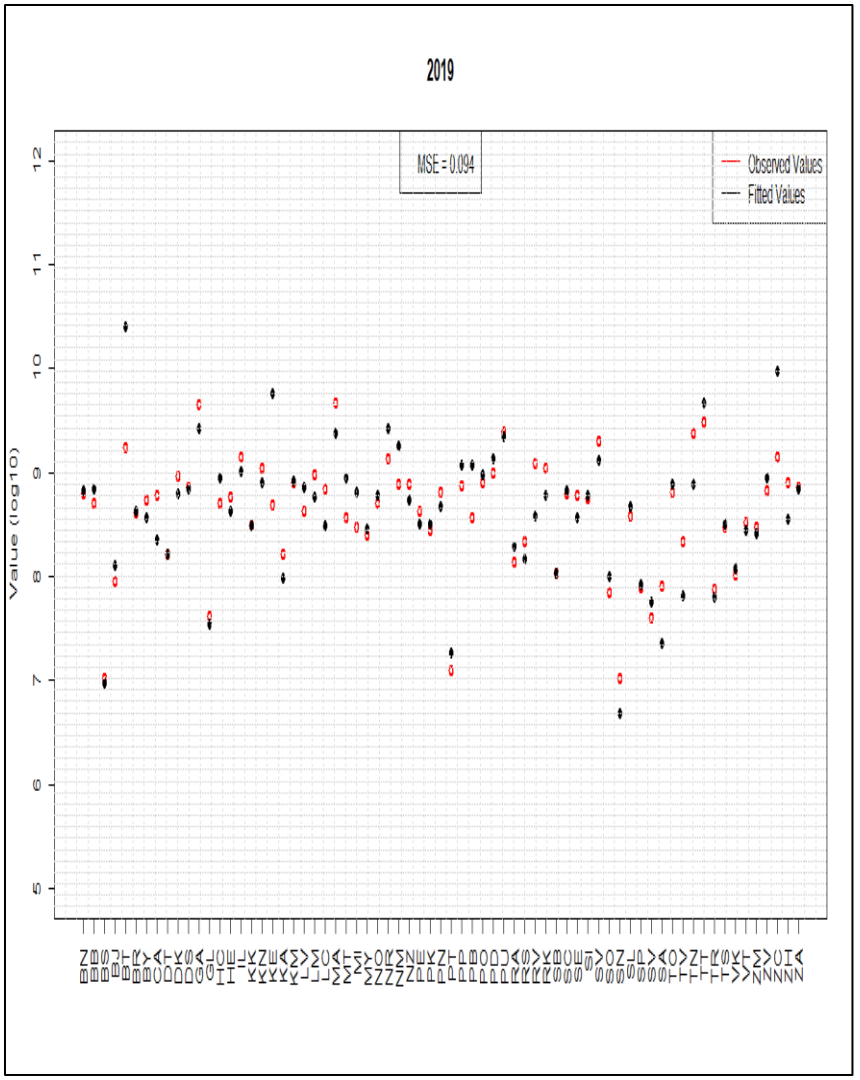

Figure 44. The predicted versus the observed responses (2019), including Mean Squared Error (MSE): Functional linear model.
Source: Author's construction.

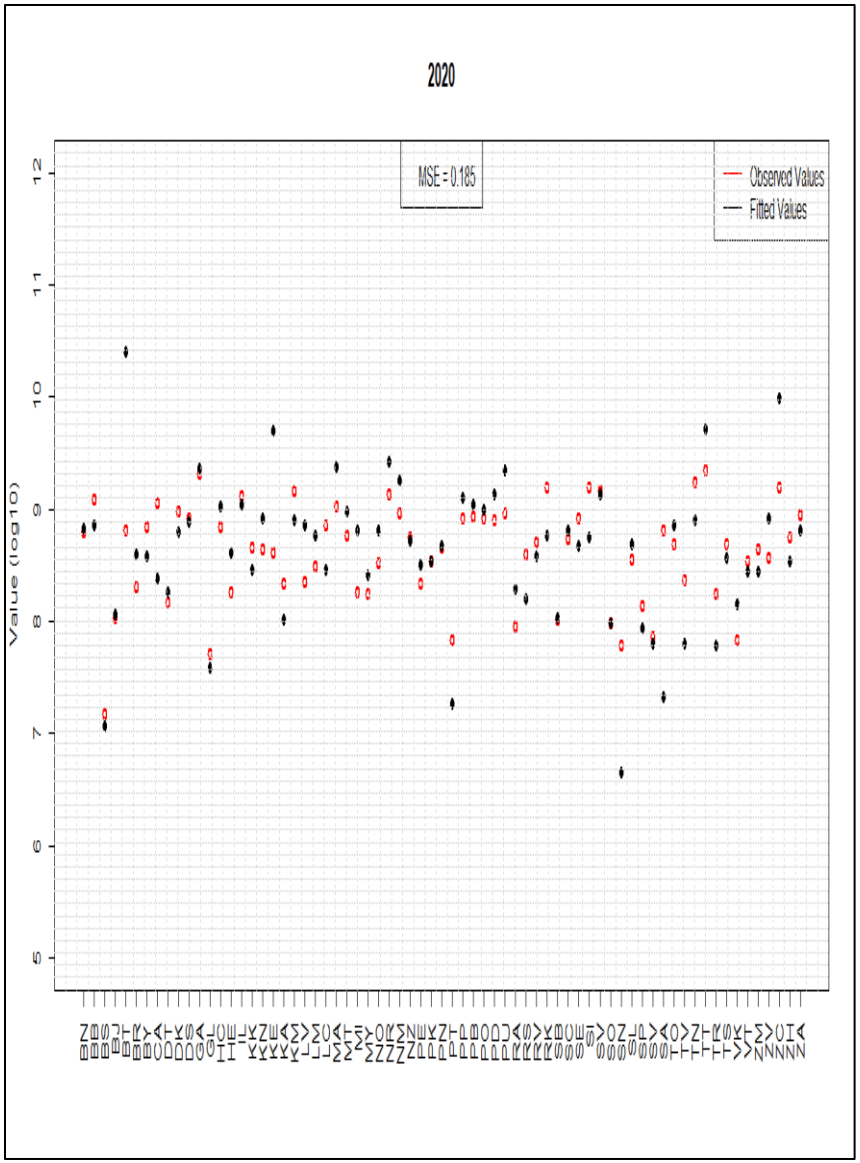

Figure 45. The predicted versus the observed responses (2020), including Mean Squared Error (MSE): Functional linear model.
Source: Author's construction.

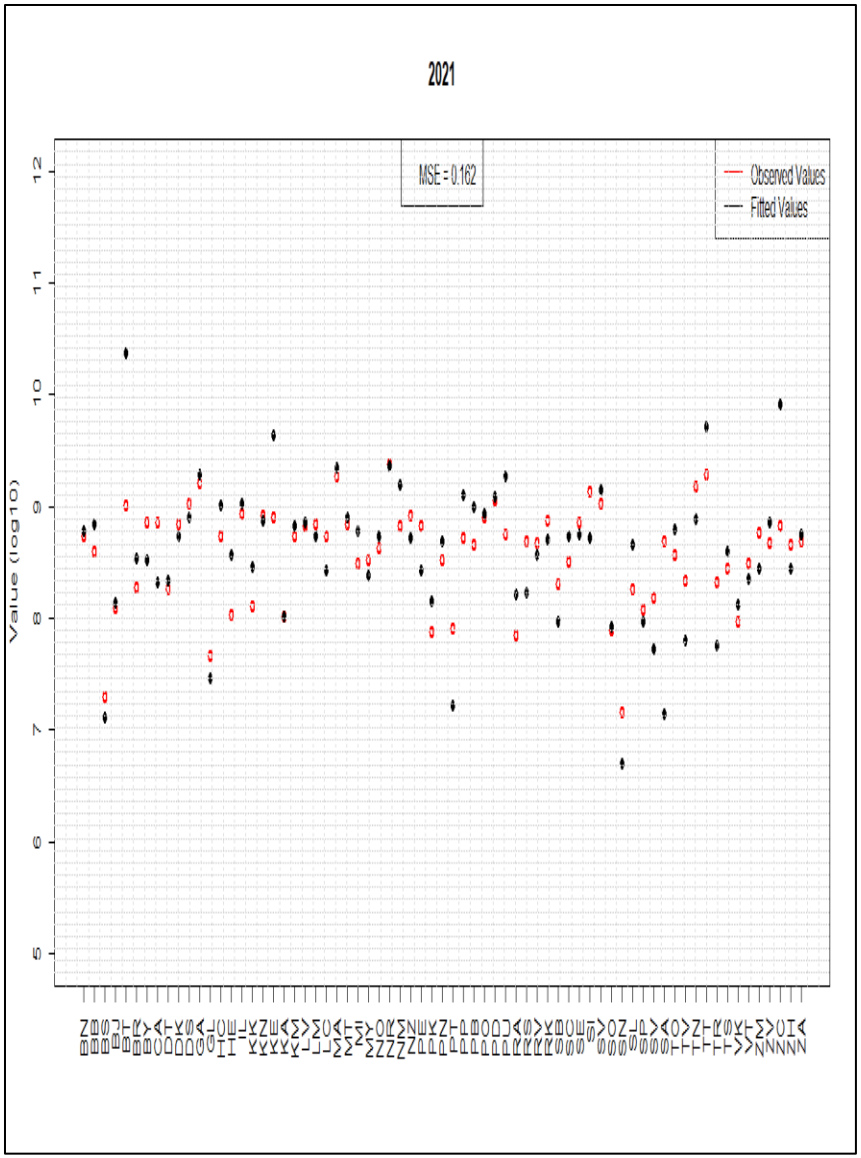

Figure 46. The predicted versus the observed responses (2021), including Mean Squared Error (MSE): Functional linear model.
Source: Author's construction.

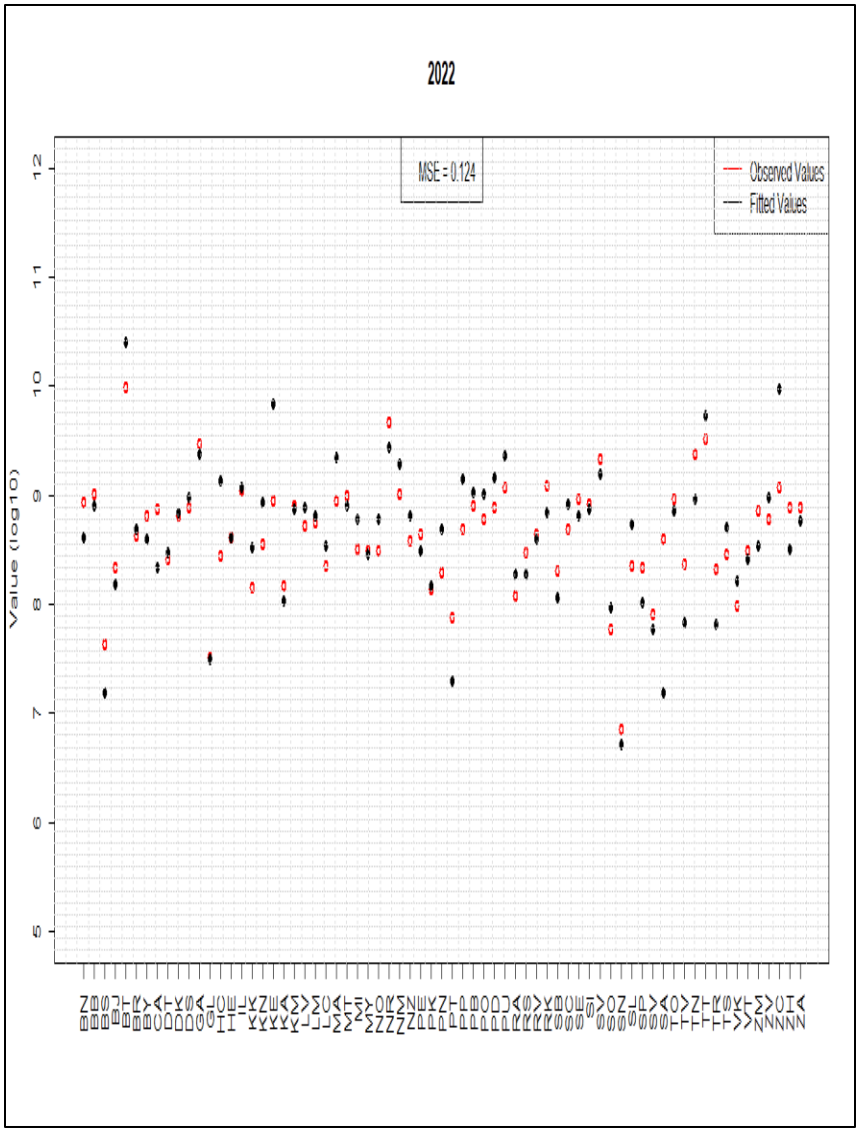

Figure 47. The predicted versus the observed responses (2022), including Mean Squared Error (MSE): Functional linear model.
Source: Author's construction.