Spatial Autocorrelation of Global Stock Exchanges Using Functional Areal Spatial Principal Component Analysis

Tzung Hsuen Khoo, Dharini Pathmanathan, Sophie Dabo-Niang

▶ **To cite this version:**

Tzung Hsuen Khoo, Dharini Pathmanathan, Sophie Dabo-Niang. Spatial Autocorrelation of Global Stock Exchanges Using Functional Areal Spatial Principal Component Analysis. Mathematics , 2023, 11 (3), pp.674. 10.3390/math11030674 . hal-04392798

HAL Id: hal-04392798
https://hal.science/hal-04392798

Submitted on 5 Apr 2024

HAL is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L'archive ouverte pluridisciplinaire **HAL**, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.

Distributed under a Creative Commons Attribution 4.0 International License

mathematics

MDPI

Article

Spatial Autocorrelation of Global Stock Exchanges Using Functional Areal Spatial Principal Component Analysis

Tzung Hsuen Khoo [1], Dharini Pathmanathan [1,*] and Sophie Dabo-Niang [2]

[1] Institute of Mathematical Sciences, Faculty of Science, Universiti Malaya, Kuala Lumpur 50603, Malaysia
[2] Université Lille, CNRS, UMR 8524-Laboratoire Paul Painlevé, INRIA-MODAL, F-59000 Lille, France
* Correspondence: dharini@um.edu.my

Abstract: This work focuses on functional data presenting spatial dependence. The spatial autocorrelation of stock exchange returns for 71 stock exchanges from 69 countries was investigated using the functional Moran's I statistic, classical principal component analysis (PCA) and functional areal spatial principal component analysis (FASPCA). This work focuses on the period where the 2015–2016 global market sell-off occurred and proved the existence of spatial autocorrelation among the stock exchanges studied. The stock exchange return data were converted into functional data before performing the classical PCA and FASPCA. Results from the Monte Carlo test of the functional Moran's I statistics show that the 2015–2016 global market sell-off had a great impact on the spatial autocorrelation of stock exchanges. Principal components from FASPCA show positive spatial autocorrelation in the stock exchanges. Regional clusters were formed before, after and during the 2015–2016 global market sell-off period. This work explored the existence of positive spatial autocorrelation in global stock exchanges and showed that FASPCA is a useful tool in exploring spatial dependency in complex spatial data.

Keywords: functional data analysis; spatial autocorrelation; principal component analysis; stock market

MSC: 62R10; 62H11; 91B72

check for updates

Citation: Khoo, T.H.; Pathmanathan, D.; Dabo-Niang, S. Spatial Autocorrelation of Global Stock Exchanges Using Functional Areal Spatial Principal Component Analysis. *Mathematics* **2023**, *11*, 674. https://doi.org/10.3390/math11030674

Academic Editors: Mustapha Rachdi and Jin-Ting Zhang

Received: 27 December 2022
Revised: 21 January 2023
Accepted: 25 January 2023
Published: 28 January 2023

Copyright: © 2023 by the authors. Licensee MDPI, Basel, Switzerland. This article is an open access article distributed under the terms and conditions of the Creative Commons Attribution (CC BY) license (https://creativecommons.org/licenses/by/4.0/).

1. Introduction

Functional data analysis (FDA) is widely applied in many disciplines of science, economics and so on. FDA expresses discrete observations in the form of functions which creates functional data. The entire measured function is represented as a single observation. Statistical concepts from multivariate analysis are then applied to model and perform analysis from a collection of functional data. A comprehensive overview of the fundamental principles and applications of FDA is available in [1]. FDA has found many recent applications due to its ability to simplify analyses especially in multivariate, spatial and time series analyses. Some recent works involving FDA include [2,3]. A systematic review of the applications of FDA is available in [4].

As complex and high-dimensional spatial data have been more readily available over the past two decades, it prompted many new investigations on spatial dependence of subjects in various fields. The authors of [5] analyzed the spatial dependence of bankruptcy in Spain using Moran's I index and the local association index. A relatively new branch of statistics known as the spatial functional statistics (SFS) was developed to analyze this type of data. SFS incorporates the spatial structure in the FDA framework [6]. One of the earliest developments of SFS can be found in [7], which provides different approaches to integrating spatial data into FDA based on the type of spatial data, e.g., geostatistical data, point patterns and areal data. The authors of [8] proposed a spatial FDA approach where spatial clustering using local spatial autocorrelation was combined with distances

calculated directly from functional data. The proposed procedure was then applied to areal diversity profiles. Furthermore, reference [9] devised a dimension reduction technique appropriate for functional data which is indexed by spatial locations on a grid.

One of the most essential tools in FDA is functional principal component analysis (FPCA). This work aims to combine FPCA and spatial statistics to analyze spatial autocorrelations among global stock exchanges. The idea of using PCA for spatial data was first proposed by [10] by developing multivariate spatial principal component analysis (sPCA) to find subtle spatial patterns in genetic data. By reducing the dimensionality of the georeferenced genetic data, the principal components were used to create maps of projected PC scores which provided more insights to the spatial patterns. FASPCA is a combination of sPCA developed by [10] and FDA. Similarly, FASPCA reduces the dimensionality of data with a spatial component which is an important aspect in investigating spatial autocorrelations among areal data points. Spatial autocorrelation characterizes the spatial relationship in data. A set of data where data points have similar values as their neighboring data points will have, by definition, a higher spatial autocorrelation. One of the widely used measures of spatial autocorrelation is the Moran's I statistic. The implementation of the Moran's I statistic in the FDA framework is implemented in this work.

This study seeks to analyze the spatial autocorrelations of global stock exchanges using the daily return data. The spatial dependencies of global and regional stock exchanges represent important information for policy makers and investors to build an optimally diversified portfolio. There has been research on the spatial dependencies of stock exchanges which used spatial econometric methods to investigate the impact of a financial crisis on the spatial autocorrelations of stock exchanges. Most of these studies in the literature focused on how financial crises such as the United States (USA) subprime crisis (2008) and the European sovereign debt crisis (2011) affected the spatial autocorrelations of stock exchanges [11,12]. The spatial dependence of between risks of stock markets was analyzed in [13] by studying the impact of systemic risk on spatial dependence related to some of the most significant financial crises.

Particularly, the impacts from these two financial crises were observed to increase the spatial autocorrelations of global stock exchanges. Most studies relating to the spatial analysis of stock exchanges focused on the 2008–2013 period. The Asian region, specifically East and Southeast Asia, witnessed a financial crisis in the late 1990s. The foreign exchange market data based on the synchrony level of the Canadian Dollars/United States Dollars and Singapore Dollars/United States Dollars foreign exchange rate time series were analyzed by [14] before and after the 1999 Asian financial crisis using cross-sample entropy. It is uncertain how other financial crises such as the 2015–2016 market sell-off impacted the spatial autocorrelations of global and regional stock exchanges. The 2015–2016 global market sell-off lasted from approximately June 2015 to June 2016. It was triggered by the crash of the Shanghai stock exchange in China which was followed by other international events, for example the Greek debt default in June 2015, the fall in petroleum prices and the Brexit vote announcement in February 2016. This work aims to use FASPCA to analyze spatial autocorrelations in global stock exchanges and reveal spatial patterns in the data. This work also investigates the impact of the 2015–2016 global market sell-off on the spatial dependencies of global stock exchanges.

2. Data Description

The daily closing prices from 71 stock exchanges of 69 countries worldwide were obtained from Investing.com and Yahoo Finance for the period from 1 June 2014 to 1 June 2017. As mentioned earlier, one of the recent major global stock market crashes/sell-offs occurred from June 2015 to June 2016. The sample period was divided into period 1 (one year before the market-crash), period 2 (during the market crash) and period 3 (one year after the market crash). This allowed the spatial autocorrelation in the sample period to be analyzed individually and comparatively. The days where the stock data were not available due to holidays in any countries were imputed with values of the trading day

before. Secondly, for stock exchanges which operated from Sunday to Thursday, e.g., the Saudi stock exchange, the closing price for Friday will be imputed with Sunday's closing price. This allows each closing price from the 71 stock indices to be analyzed based on standardized trading days. This allows each closing price from the 71 stock indices to have the same sampling date.

3. Methodology

3.1. Data Smoothing for Functional Data Analysis

Consider n spatial locations i, one observed discrete measurement $Y_{i,x,t}$ taken at time t of location $i \in I \subset \mathbb{Z}^2$, I a lattice region V, for a given $x \in X = [T_1, T_2]$. In this study, $Y_{i,x,t}$ is the log return observed on x where x is a date indexed by $T_1 = 0$ and $T_2 = 259$, in year $t \in D = \{(1\text{ June }14 - 1\text{ June }15), (1\text{ June }15 - 1\text{ June }16), (1\text{ June }16 - 1\text{ June }17)\}$ for each stock exchange at location i. Assume that, for a given t these measurement points $Y_{i,x,t}$ are noisy observations of a smooth areal stochastic functional process $\{S_{i,t}\}_{i\in I,\ t\in X}$:

$$Y_{i,x,t} = \mu_t(x) + S_{i,t}(x) + \epsilon_{i,x,t} \tag{1}$$

where μ_t is the mean function at time t (year). The n functions $S_{i,t}(.)$ are the centered spatial squared integral functional random variables on the space–time domain $I \times D$, namely $S_{i,t}(.)$ is valued in the Hilbert space $L^2(X)$ endowed with the inner product $\langle f, g\rangle = \int_X f(x)g(x)dx$, for f, g in $L^2(X)$. The unobserved variables $\{\epsilon_{i,x,t},\ i = 1, \ldots, n\}$ are i.i.d with zero mean Gaussian measurement errors with variance σ^2.

A Karhunen–Loève expansion for the n functions $S_{i,t}(x)$ [15] is postulated as follows:

$$S_{i,t}(x) = \sum_{k=1}^{\infty} \beta_{k,i,t}\phi_k(x) \tag{2}$$

where ϕ_k's are the orthonormal eigenfunctions and $\beta_{k,i,t}$ are the autocorrelated scores (functional principal components, FPC, which will serve as a foundation in constructing FASPCA, are discussed in Section 3.3). In practice, the sum is a truncated finite integer, K which is to be chosen.

To compute the FPCs, the sample data $(S_{i,t})_{i=1,\ldots,n}$ can be expressed by means of a truncated basis expansion:

$$S_{i,t}(x) = \sum_{m=1}^{\infty} c_{i,m,}B_m(x) \approx \sum_{m=1}^{p} c_{i,m}B_m(x),\ x \in X \tag{3}$$

where $B_m(.)$ is some collection of basis functions, $c_{i,m} = \langle S_{i,t}, B_m\rangle$ have zero mean.

Two main basis systems for building functions were presented by [1]. The Fourier basis system is commonly used for period data. The B-spline basis system is preferable for nonperiodic data. For log return data of stock exchanges, the B-spline basis system provides more flexibility. Once the basis system is selected, the technique of roughness penalized least square is adopted to achieve data fitting and smoothing. The criterion of roughness penalized fitting can be expressed in general as:

$$F(\beta) = \sum_j [Y_{i,j,t} - S_{i,t}(x_j)]^2 + \lambda \int [L(S_{i,t}(x))]^2 dx \tag{4}$$

where $\lambda \geq 0$ is the smoothing parameter, and $L(.)$ is the differential operator. The second term is often called as the total curvature of the range of x. Therefore, λ controls the "roughness" of the function relative to the least square specified in the first term. It is customary in the roughness penalty approach to choose a number of basic functions comparable to the number of data observed. The value of λ is determined through general cross validation to ensure that the functions are not over-smoothed or overfitted. The

fitting and smoothing of the curves with the B-spline basis system was performed using the fda [16] package in the R software.

3.2. Functional Moran's I Statistic and Its Implementation on Spatial Weight Matrices

The classical univariate Moran's I statistic [17] of a n row vector X_m of components $\{c_{i,m}\}_{i=1,\ldots,n}$ is expressed as:

$$\widetilde{I}(X_m) = \frac{X_m^T W X_m}{X_m^T X_m} \tag{5}$$

where $W = (W_{ij})$ is a spatial weight matrix and W_{ij} represents the neighboring relation between two locations, i and j. Moran's I statistic is a measure of spatial autocorrelation which quantifies the spatial dependency among observations in a geographic space [10]. The Moran's I statistic was extended to the functional context [18] resulting in the functional Moran's I statistic. The functional Moran's I statistic of an n row vector $\mathbf{S}(x)$ with components $\{S_i(x)\}_{i=1,\ldots,n}$ can be described as:

$$I(\mathbf{S}(x)) = \frac{C_n(\mathbf{S}(x))}{\sigma_n(\mathbf{S}(x))} \tag{6}$$

where

$$C_n(\mathbf{S}(x)) \approx \frac{1}{n}\mathbf{B}(x)^T\mathbf{X}^T W\mathbf{X}\mathbf{B}(x) \tag{7}$$

$$\sigma_n(\mathbf{S}(x)) \approx \frac{1}{n}\mathbf{B}(x)^T\mathbf{X}^T\mathbf{X}\mathbf{B}(x) \tag{8}$$

where $\mathbf{X}$ is a $n \times\ p$ matrix that consists of the scores $(c_{i,m})_{i=1,\ldots,n;m=1,\ldots,p}$ of $\mathbf{S}(x)$; $\mathbf{B}(x)$ is a $p \times 1$ vector of components $B_m(x)$, $m = 1, \ldots, p$. This study seeks to examine the existence of spatial autocorrelation in the log returns for 71 global stock exchanges. The functional Moran's I statistic (6) is used here.

Choices of Weight Matrices

Spatial weight matrices can be classified into two groups by how the weights are constructed, i.e., weights based on boundaries (more commonly known as contiguity-based weights) and weights based on distances. Contiguity-based weights can be used when neighboring locations share a common boundary and/or vertices. In this study, contiguity-based weights are unsuitable as there exist countries which have no common boundaries with any countries in their proximity.

Weights based on distance can generally be constructed using criteria such as KNN, radial distance weights, power distance decay weights [19] and proximity graphs [10,20,21]. However, weights such as radial distance weights require a specification of a critical distance which is inappropriate in this study. Therefore, this study proceeded with KNN and the graph-based Gabriel neighbor which is one of the proximity graphs.

The great-circle distance was chosen to be the measure of distance between countries. The KNN matrix was constructed by setting k = 4 which selected the four nearest neighbors for each location i. This was because the average number of links for each country in the Gabriel graph is more than 3 and less than 4. k = 4 was selected to match it and ensured the comparison between two weight matrices were fair at this stage. The entries of the KNN matrix are given as:

$$w_{ij} = \begin{cases} N_k(i)^{-1}, \ if\ j \in N_k(i) \\ 0, \ otherwise \end{cases}$$

where $N_k(i)$ is the set of k closest neighbours of location i. The KNN matrix is standardized such that the sum of each column and row equals to 1. The graph-based Gabriel neighbor matrix was constructed using the following inequality:

$$d(i,j) \leq \min((d(i,q)^2 + d(j,q)^2)^{\frac{1}{2}} \mid q \in Q) \tag{9}$$

where Q is the set of all locations. The entries of the graph-based Gabriel neighbor matrix are given as:

$$w_{ij} = \begin{cases} N_k(i)^{-1}, \ if \ (i,j) \ satisfies \ (9) \\ 0, \ otherwise \end{cases}$$

where $N_k(i)$ is the set of neighbors of location i. The graph-based Gabriel neighbor matrix is also standardized such that the sum of each column and row equals to 1.

In this study, a Shapiro–Wilk test is used to test the normality of the log return data. If the log return data are shown to violate the normality assumption, a Monte Carlo test is performed on the Moran's I statistic. The spdep [22] package in the R software was used to generate spatial weight matrices as well as perform the Monte Carlo test on the Moran's I statistics.

3.3. FASPCA

3.3.1. Spatial Principal Component Analysis (sPCA)

The classical principal component analysis (PCA) finds scaled vectors $\mathbf{u_k}$, $k = 1, \ldots, p$ ($\| \mathbf{u_k} \| = 1$), such that the variances of the scores $\varphi_k = \mathbf{X}\mathbf{u_k}$ ($\mathbf{X}$ is a $n \times p$ data matrix) are maximized [10]. This can be expressed as:

$$\| \mathbf{X}\mathbf{u_k} \|^2_{1/n} = \frac{1}{n}\mathbf{u_k}^T\mathbf{X}^T\mathbf{X}\mathbf{u_k} \tag{10}$$

where $\| \mathbf{X}\mathbf{u_k} \|^2_{1/n} = var(\varphi_k)$.

The classical PCA is designed to summarize variabilities in the data with the scores φ_k. However, it is not designed to reveal spatial patterns as it lacks spatial components in the analysis.

sPCA was designed to produce scores which summarize variability and reveal spatial structures in the data simultaneously [10]. sPCA finds scaled vectors $\mathbf{v_k}$, $k = 1, \ldots, p$ ($\| \mathbf{v}_k \| = 1$), such that the scores $\chi_k = \mathbf{X}\mathbf{v_k}$ are both scattered and spatially autocorrelated. Instead of maximizing (10), it identifies the extreme values of the following criterion $C(\mathbf{v_k})$:

$$C(\mathbf{v_k}) = var(\mathbf{X}\mathbf{v_k})I(\mathbf{X}\mathbf{v_k}) = \frac{1}{n}\mathbf{v_k}^T\mathbf{X}^T W\mathbf{X}\mathbf{v_k} \tag{11}$$

where $I(.)$ is the Moran's I statistic. The components of the classical PCA are calculated based on the positive definite covariance matrix which led to positive eigenvalues. The components of $\mathbf{v_k}$ are calculated based on Equation (11) and is related to $\frac{1}{n}\mathbf{X}^T W\mathbf{X}$ which is not positive definite because of W (see [10] for more details). This causes some of the principal component scores to be associated with negative eigenvalues.

$C(\mathbf{v_k})$ is highly positive when $\mathbf{X}\mathbf{v_k}$ has a large variance and exhibits a global spatial structure and is highly negative when $\mathbf{X}\mathbf{v_k}$ has a large variance but with a local spatial structure.

3.3.2. Implementation of FASPCA

A study where the sPCA is extended to functional framework to consider spatial autocorrelation on the variable of interest in the sampling locations is of interest. FASPCA was developed as a functional version of sPCA which pertains to areal spatial data [18] being implemented. In FASPCA, discrete sample data $(S_{i,t})_{i=1,\ldots,n}$ are first fitted into functions according to the procedure presented in Section 3.1. The functions are expressed as:

$$S_{i,t}(x) = \sum_{m=1}^{\infty} c_{i,m,}B_m(x) \approx \sum_{m=1}^{p} c_{i,m}B_m(x)$$
$$\mathbf{S}(x) \approx \mathbf{C}\mathbf{B}(x)$$

where $\mathbf{S}(x)$ is the $n \times 1$ vector of functions $S_{i,t}(x)$, $\mathbf{B}(x)$ is the $p \times 1$ vector of components $B_m(x)$, $m = 1, \ldots, p$, and $\mathbf{C} = (c_{i,m})_{i=1,\ldots,n;m=1,\ldots,p}$ is a $n \times p$ coefficient matrix.

$\mathbf{C}$ is then used along with Equation (11) to obtain $\mathbf{v_k}$, $k = 1, \ldots, p$ and subsequently the scores $\chi_k = \mathbf{C}\mathbf{v_k}$.

$$\mathbf{C} \approx \hat{\mathbf{C}} = \sum_{k=1}^{p} \chi_k \mathbf{v}_\mathrm{k}^\mathrm{T} \tag{12}$$

$$\mathbf{S}(x) \approx \hat{\mathbf{C}}\mathbf{B}(x) = \left(\sum_{\boldsymbol{k}=\mathbf{1}}^{\boldsymbol{p}} \chi_k \mathbf{v}_\mathrm{k}^\mathrm{T} \right) \mathbf{B}(x) \tag{13}$$

where $\hat{\mathbf{C}}$ is the approximation of $\mathbf{C}$ using orthonormal vectors $\mathbf{v_k}$ and the scores χ_k.

The FASPCA decomposition is then achieved by setting $\hat{\phi}_k(x) = \mathbf{v}_k^T \mathbf{B}(x)$ to be the estimated eigenfunctions and the n row vector of functional scores $\hat{\boldsymbol{\beta}}_\mathbf{k} = \mathbf{S}(.), \hat{\phi}_k(.)$. For each entry $S_{i,t}(x)$ of $\mathbf{S}(x)$, this can be summarized as:

$$S_{i,t}(x) \approx \hat{\mu}_t(x) + \sum_{k=1}^{p} \hat{\beta}_{k,i,t} \hat{\phi}_k(x) \tag{14}$$

where $\hat{\mu}_t(x) = \frac{1}{n} \sum_{i=1}^{n} S_{i,t}(x)$, is the empirical mean with $\hat{\beta}_{k,i,t} = \int_0^T S_{i,t}(x) \hat{\phi}_k(x) \mathrm{dx}$.

3.4. Summary

In summary, the first part of the study consists of fitting and smoothing the log return data of the stock exchanges using a suitable basis system. The next step involves the calculation of the functional Moran's I statistic to check if there exists spatial autocorrelation in log return data. Monte Carlo tests are performed on the Moran's I statistic to verify if the log return data violate the normality assumption. For the second part of this study, the classical PCA is used on the log return data. However, the classical PCA cannot provide information about the spatial dependency in the data. Therefore, FASPCA is used to generate PCs which contain information about global and local structures in the log return data [18]. A global structure corresponds to a positive eigenvalue resulting from a strong variance and positive spatial autocorrelations. Thus, it is a pattern where each data point has a PC score similar to its neighbors. A local structure corresponds to a negative eigenvalue resulting from a strong variance but a negative spatial autocorrelation. It is a pattern where each data point has a PC score dissimilar to its neighbor (Jombart et al. 2008). FASPCA was implemented using the fda [16], adegenet [23], adespatial [24] and ade4 [25–28] packages from the R software.

4. Results and Discussions

4.1. Stationarity in Time and Space

From both spatial and temporal perspectives, stationarity indicates whether the data have constant mean and variance in each location or time. It is important for the space and time variables to be stabilized before applying PCA. Daily closing prices of each stock were converted into daily logarithmic returns using $\ln p_t - \ln p_{t-1}$, where p_t is the closing price at time t. This transformation is essential to ensure stationarity in time. The stationarity in time of daily logarithmic returns for each stock exchange was inspected using three tests, namely, the Augmented Dickey–Fuller test [29], the KPSS test for level and trend stationarity. The test results of all the data indicated that all daily logarithmic returns of the stock indices were stationary in time.

Spatial stationarity of the functional data was examined using a trace variogram [30] with the aid of the geofd package [31]. These findings showed that an exponential trace variogram fits the functional data best. Then the functional spatial data are not far from satisfying the spatial stationarity requirement from a Gaussian field with exponential trace variogram. This is applicable to all three time periods investigated.

4.2. Functional Moran's I Statistic

The log return data were transformed into functions using Equation (3) where the basis system had been chosen to be B-spline. The presence of spatial autocorrelation in the data was verified for the three periods separately. The Shapiro–Wilk test [32] was performed to assess the normality of the log return data. It was found that the log return data violated the normality assumption. Therefore, Monte Carlo permutation tests were performed using 999 random permutations of the log return data. The Monte Carlo permutation tests were carried out for both KNN and graph-based Gabriel connection networks. Table 1 shows the existence of positive spatial correlations in the log return data with both connection networks. In Table 1, a classical Moran's I statistic was calculated by viewing the log return data matrices as panel datasets. Moreover, the classical Moran's I statistic indicates that the spatial autocorrelation increased in period 2 which subsequently decreased in period 3. The same behavior is observed for both connection networks.

Table 1. Classical Moran's I statistic for log return data of stock exchanges in 71 countries based on KNN and Gabriel connection network.

Weight Matrix	Period 1	Period 2	Period 3
KNN	0.242 **	0.375 **	0.283 **
Gabriel	0.258 **	0.399 **	0.306 **

Note: ** p-value < 0.001.

Functional Moran's I Statistics—Results and Discussions

Figure 1 illustrates the functional Moran's I statistic for both the KNN and graph-based Gabriel connection networks. It can be seen that the functional Moran's I statistic of both connection networks show comparable movements. In period 1 (black curves) which is a year before the start of the global market sell-off, the spatial autocorrelation was fluctuating at a relatively low level until October 2014 where the spatial autocorrelation increased substantially. This could be explained by the end of the quantitative easing program in the USA in October 2014. According to [33], the quantitative easing program from 2008–2014 led to positive spillover effects on emerging market countries such as Brazil, India and Indonesia, as well as other smaller emerging markets. The spatial autocorrelation dropped quickly to almost zero and then steadily increased until June 2015 (end of the black curves).

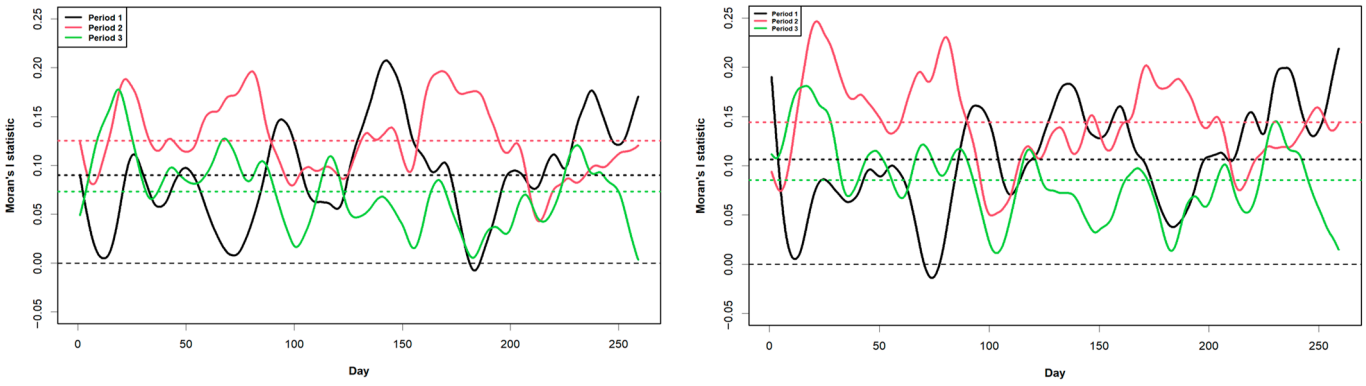


Figure 1. Functional Moran's I statistic (**a**) calculated using KNN connection network and (**b**) calculated using Gabriel graph-based connection network.

In period 2 (red curves), the spatial autocorrelation was fluctuating at a relatively high level with period 1. The Chinese stock market crash in June 2015 corresponded to the increase in spatial autocorrelations. Results from [34] suggested that financial spillovers on regional stock markets have been steadily growing through increasing trade linkages with other countries prior to the stock market crash. Furthermore, the announcement of the

Brexit referendum was found to increase the spatial autocorrelations noticeably in February 2016. According to [35], the announcement increased all the market co-volatilities in the eurozone where the co-movements between Germany, France and Italy (Western Europe) increased the most.

In period 3 (green curves), the spatial autocorrelation peaked at the start of the period (June 2016) which slowly decreased and fluctuated at a similar level as period 1. The peak matched the release of the Brexit referendum results. The analysis presented by [36] indicated that the neighboring equity markets in the eurozone, specifically the PIIGs group (Portugal, Ireland, Italy, Greece and Spain), were most affected. On the other hand, the BRICs countries (Brazil, Russia, India, China and South Africa) experienced positive returns during this period. A similar trend in spatial autocorrelations was observed during the financial crisis in 2008, where negative shocks emanating from the USA stock markets led to the increase in spatial autocorrelations among the global equity markets. This eventually decreased to the pre-crisis level after the financial crisis subsided [11].

4.3. FASPCA's Results

Three cases were performed for comparative purpose. The first case uses functional PCA (FPCA) to generate functional PCs. The second case uses the FASPCA with KNN connection network to generate positive and negative functional PCs. Similarly, the third case uses FASPCA with the graph-based Gabriel connection network to generate positive and negative functional PCs.

4.3.1. FPCA

In the first case, four PCs were calculated using the classical FPCA. From Table 2, the first PC accounts for approximately 20–32% of the total variabilities of the data for three separate periods. Figure A1 shows the magnitudes of the eigenvalues generated from FPCA. Figure A2 shows the maps of projected first PC scores onto a world map where the positive and negative scores of the PCs are represented by black and white squares of different sizes, respectively. The size of the squares is proportional to the absolute values of the PC scores. The maps of projected first PC scores in the three periods do not unravel any spatial patterns in the data due to the lack of spatial factors in the model.

The classical FPCA was performed for comparative purposes to identify significant autocorrelations in all principal components as shown in Table 2. Insignificance in some of the spatial autocorrelations when viewed from a classical perspective proves the importance of adding the spatial aspect in FPCA. Therefore, this is not an approximation of the original process. As a matter of fact, it paves the way to using FASPCA as the right dimensionality reduction approach used when dealing with functional spatial data.

4.3.2. FASPCA with KNN Connection Network

In the second case, two positive and two negative PCs were generated from FASPCA using the KNN connection network. The magnitudes of the generated eigenvalues are displayed in Figure 2. From Table 2, significant positive spatial autocorrelations exist for the first two global structures. The four functional PCs account for approximately 55% of the total variability in the data. Similar to the first case, the two positive PC scores in the three periods were mapped to identify any spatial patterns in the data.

Figure 3a illustrates the map of the first PC scores in period 1. Approximately six major clusters were formed. These include Central and South America, the Middle East and South Asia, East Asia, Southeast Asia (SEA) and Europe. In Europe, Southern European countries including Hungary, Romania, Bulgaria, Turkey, Greece, Serbia, Bosnia and Herzegovina formed a cluster while the remaining European countries formed another cluster. In period 2, the two clusters in East Asia and SEA became indistinguishable as the PC scores of several SEA countries changed sign. The two major European clusters were observed to merge together and formed a single cluster. Other clusters remained similar compared to period 1 (Figure 3b). Hence, there are three major clusters remaining in period 2. In

period 3, many neighboring clusters became more integrated compared to periods 1 and 2 (Figure 3c). The variability explained by the first PC decreased from approximately 31% in period 1 to 27% in period 2 and decreased further to approximately 22% in period 3.

The maps of projected second positive PC scores (Figure A3) indicate that the second positive PC were dominated by the two sub-clusters formed in Europe. These two sub-clusters are the Western and Eastern Europe clusters. This can be observed throughout the three periods. The spatial autocorrelations of the second positive PCs are found to be positive and significant throughout the three periods. The variability explained by the second PC increased from 11% in period 1 to 15% in period 2 and dropped to approximately 13% in period 3 (Table 2).

Table 2. Moran's test on principal components using FPCA and FASPCA, using KNN and Gabriel spatial weight matrices for log return data.

	Period 1 (Pre-Crisis Period)		**Period 2 (Crisis Period)**		**Period 3 (Post-Crisis Period)**	
	Moran's I	**Variability (%)**	**Moran's I**	**Variability (%)**	**Moran's I**	**Variability (%)**
Classical FPCA						
1st PC	0.061	25.46	−0.004	31.82	0.164 *	20.41
2nd PC	0.375 ***	11.92	0.119 *	12.77	0.045	12.78
3rd PC	0.079	10.06	0.416 **	9.45	−0.027	8.12
4th PC	0.053 ***	8.62	0.146	8.58	0.276 **	6.88
Total		56.06		62.62		48.19
FASPCA (KNN (2,2))						
1st positive PC	0.561 ***	31.43	0.509 ***	27.55	0.476 ***	21.99
2nd positive PC	0.279 ***	10.75	0.210 **	15.29	0.359 ***	13.3
1st negative PC	−0.162 **	7.32	−0.187 **	7.63	−0.214 **	8.54
2nd negative PC	−0.203 ***	5.51	−0.164 *	3.98	−0.216 ***	5.27
Total		55.01		54.45		49.1
FASPCA (Gabriel (2,2))						
1st positive PC	0.483 ***	22.37	0.434 **	23.9	0.450 ***	17.93
2nd positive PC	0.320 ***	13.68	0.376 ***	14.36	0.445 ***	12.09
1st negative PC	−0.297 **	9.56	−0.369 ***	12.21	−0.329 ***	11.14
2nd negative PC	−0.434 ***	6.39	−0.308 **	5.19	−0.415 ***	8.75
Total		52		55.66		49.91

Note: * $p < 0.05$, ** $p < 0.01$; *** $p < 0.001$.

4.3.3. FASPCA with Gabriel Graph-Based Connection Network

In the third case, two positive and two negative PCs were obtained from FASPCA using the graph-based Gabriel connection network. The magnitudes of the eigenvalues are displayed in Figure A4. From Table 2, there exist significant spatial autocorrelations for the first two global structures, and the four functional PCs account for more than 50% of the total variability in the data throughout the three periods. The purpose of examining the third case is to compare the effects of different choices of connection network on the PCs and see whether they yield similar spatial patterns. From the map of the projected first positive PC scores in period 1 (Figure A5), several major clusters were formed during this period. These include Central and South America, East Asia, SEA and Europe. Similarly, many Southern European countries form a cluster while the remaining European countries form a separate cluster. In period 2, the demarcation of two clusters in East Asia and

SEA became unclear as the PC scores of most of the SEA countries changed signs. The two major European clusters were observed to merge and formed a single cluster. Other clusters remained similar compared to period 1 (Figure A5b). In period 3, neighboring clusters became more integrated compared to periods 1 and 2 (Figure A5c). The variability explained by the first PC decreased slightly from approximately 22% in period 1 to 24% in period 2 and decreased further to approximately 18% in period 3.

The projected maps of the second positive PC scores in the third case (Figure A6) appear to indicate that the second positive PCs were dominated by the two sub-clusters formed in Europe. The two sub-clusters are the Western and Eastern Europe clusters. The spatial autocorrelations of second positive PC are found to be significant throughout the three periods. The variability explained by the second positive PC increased from 13.5% in period 1 to 14% in period 2 and decreased to approximately 12% in period 3 (Table 2).

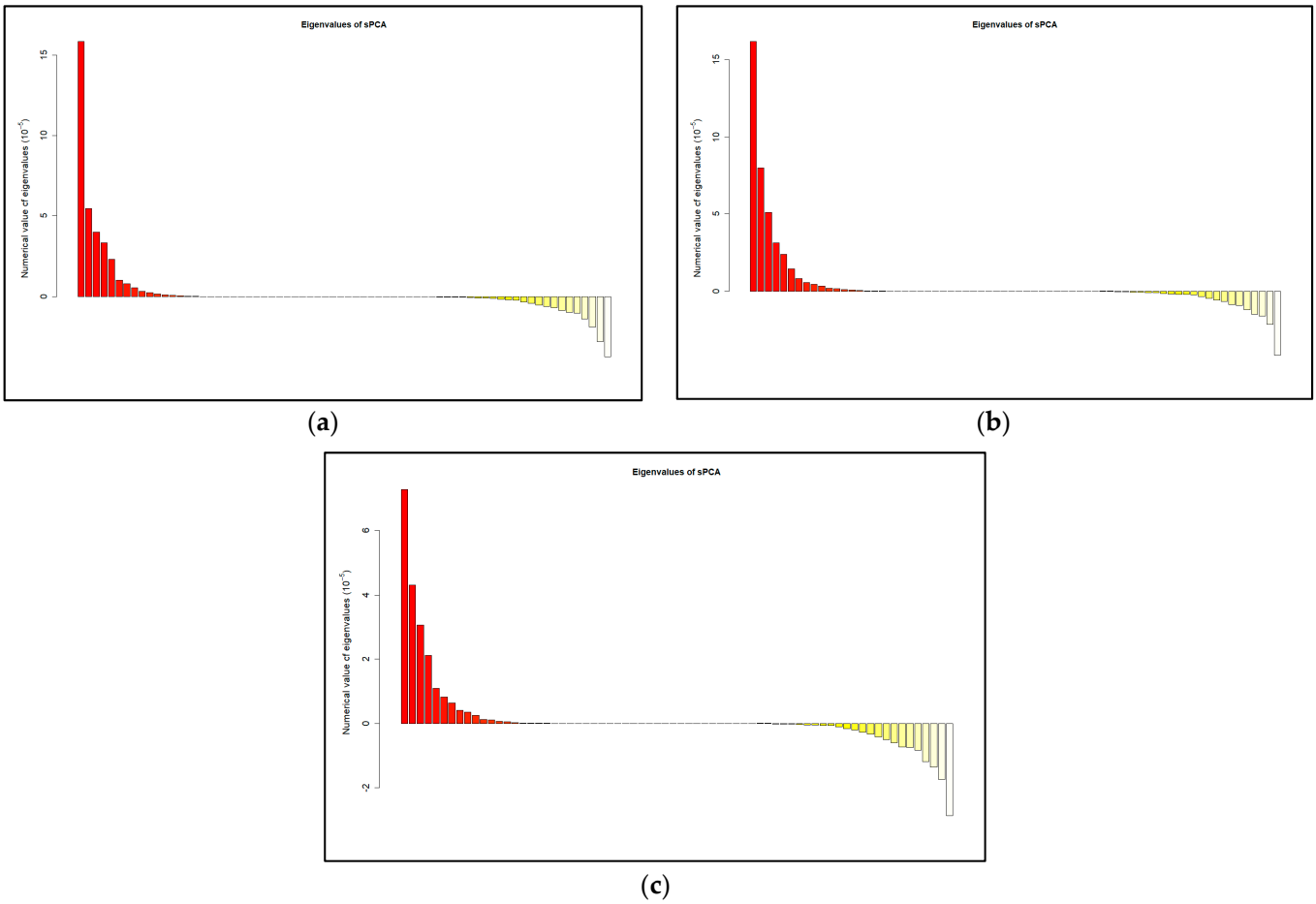
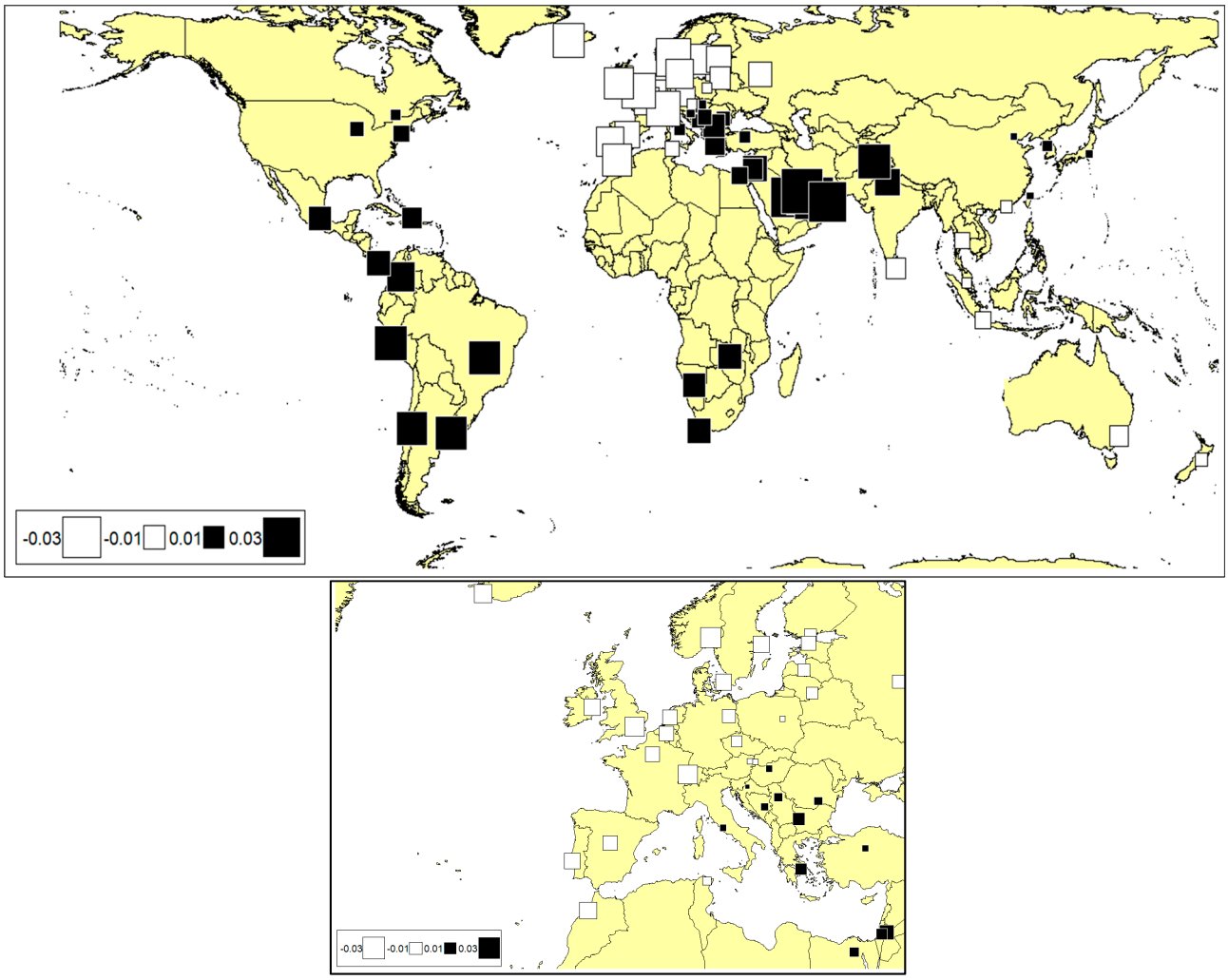


Figure 2. Distribution of eigenvalues based on FASPCA and KNN connection network: (**a**) period 1; (**b**) period 2; (**c**) period 3. Eigenvalues in red and yellow correspond to positive and negative eigenvalues respectively.



(**a**)

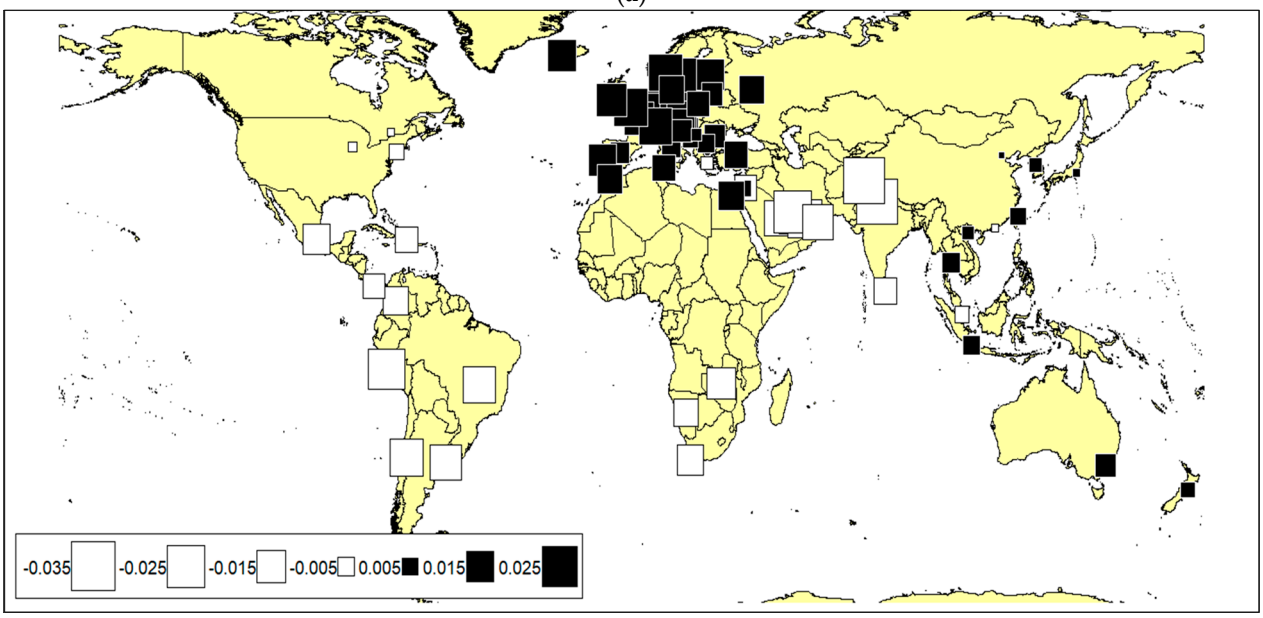


Figure 3. *Cont.*

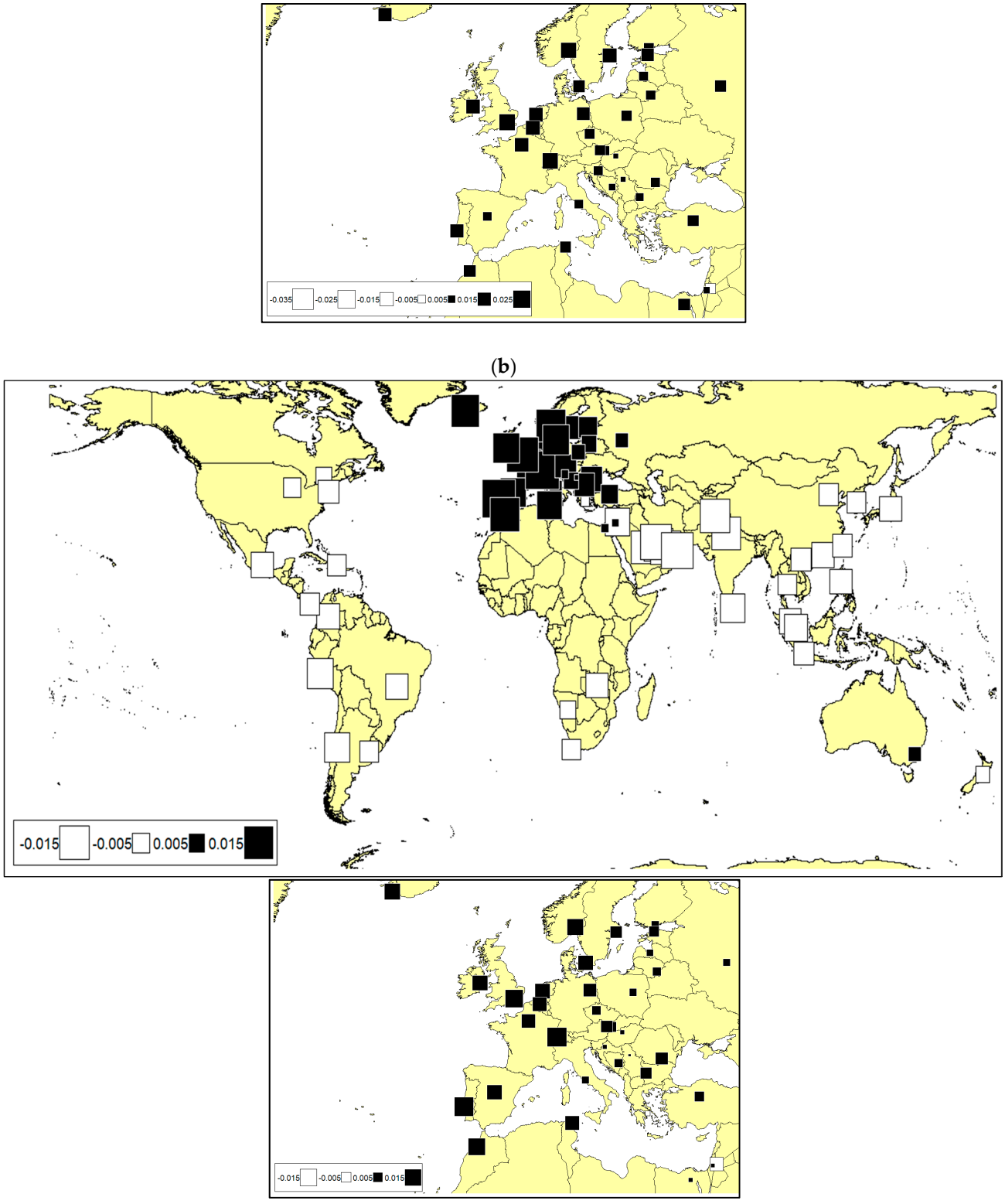


Figure 3. Map of projected first positive PC scores from FASPCA and KNN Connection Network; the smaller map below each bigger map is the zoom-in of Europe: (**a**) period 1; (**b**) period 2; (**c**) period 3.

4.3.4. FASPCA—Discussions

Several conclusions can be made by comparing the second and third case of this work. The neighbor-selecting mechanisms of the KNN and graph-based Gabriel connection networks are fundamentally different. KNN (K = 4) chooses four nearest neighbors for each data point while the graph-based Gabriel connection network selects the neighbor based on nine (9). Therefore, it was expected for the second and third case not to yield exactly the same spatial patterns. However, it was also expected that both cases would produce similar spatial patterns if significant spatial autocorrelations existed among the data. The two positive PCs in both cases were tested to be significant during the three periods (Table 2). The maps of projected positive PC scores in both cases also produced similar spatial clustering. Comparing the projected maps of the first positive PC scores in both cases, similar clusters formed in the major regions throughout the three periods.

In Europe, the northern and southern regions formed clusters respectively in period 1. These two clusters did not last as the signs of the first PC scores (see Figure 3a,b) suggested that they became more integrated in periods 2 and 3. Results from [37] which used minimum spanning trees with partial correlations on 57 global equity markets (2005–2014) observed that the European equity markets could be viewed as two separate clusters. During period 2, the magnitudes and signs of the first PC scores of the Western and Northern European equity markets changed considerably. This could be attributed to the fear of Greece's eurozone exit due to the Greek debt crisis as well as the Chinese equity market crash [38]. The Western European countries were more affected by the Chinese equity market crash compared to the southern European countries due to higher bilateral trades [34].

Maps of the second positive PC scores in both cases (see Figures A3 and A6) were dominated by the formation of the Western and Eastern Europe clusters throughout the three periods. This implies that the spatial autocorrelations among the Western and Eastern European equity markets were only slightly affected by the events that happened during the period from June 2015 to June 2016.

The East Asian and SEA countries formed two separate clusters during period 1. Results from the analysis by [39] showed that, from 1991 to 2014, the equity market movement in the SEA region was predominantly affected by the economic situations in the SEA region. Notwithstanding, the market integration of the SEA equity markets with other East Asian equity markets could not be neglected. Particularly, China showed stronger market integration with the SEA countries compared to the USA.

In period 2, the East Asian and SEA clusters weakened and became more integrated in period 3. China's influence on regional equity markets was observed to be comparable to that of Japan's [34]. Moreover, the integration between the Chinese and the SEA equity markets was foreseen to strengthen due to increasing trades between China and the SEA countries [40]. Hence, the Chinese equity market crash in June 2015 may have produced negative spillovers to its neighboring countries as well as the SEA countries and caused adverse effects on their returns in the equity markets.

5. Concluding Remarks

The functional Moran's I statistic, classical PCA and spatial functional PCA were used to analyze the spatial autocorrelation of the log returns of 71 stock exchanges in 69 countries. The functional Moran's I statistic showed that the spatial autocorrelation of stock exchanges was exacerbated by the occurrences of a financial crisis/bear market such as the 2015–2016 global market sell-off. Moran's tests of the spatial functional PCs indicated that positive spatial autocorrelation existed in the data throughout the three separate periods.

The projected maps of the first and second positive spatial functional PC scores showed that spatial clusters were formed in all three periods. Moreover, it was found that similar spatial cluster patterns were formed in three periods even when different spatial weight matrices were considered. The findings of this study show that the inferences drawn based on the functional Moran's I statistic and FASPCA correspond well with the events that

happened in all three periods. This verifies the effectiveness of these tools in measuring continuous spatial autocorrelations of global equity markets and identifying spatial patterns in complex spatial data, respectively.

This study emphasized the viability of FASPCA as an exploratory technique on complex spatial data by reducing the dimensionality of data with spatial information. It can further be enhanced to be applied in the spatio-temporal framework where the implementation of FDA reduces the space–time dimensions simultaneously, involving traditional time series forecasting and spatial prediction. This study involves a two-stage approach where discrete data are first converted into functional data before performing FASPCA. A simultaneous approach similar to [41] can be applied in future works to further enhance the study.

Author Contributions: Conceptualization, D.P., S.D.-N. and T.H.K.; methodology, T.H.K., D.P. and S.D.-N.; software, T.H.K.; validation, T.H.K., D.P. and S.D.-N.; formal analysis, T.H.K., D.P. and S.D.-N.; investigation, T.H.K., D.P. and S.D.-N.; resources, T.H.K., D.P. and S.D.-N.; data curation, T.H.K.; writing—original draft preparation, T.H.K.; writing—review and editing, T.H.K., D.P. and S.D.-N.; visualization, T.H.K.; supervision, D.P. and S.D.-N.; project administration, D.P.; funding acquisition, D.P. All authors have read and agreed to the published version of the manuscript.

Funding: This work was supported by the Universiti Malaya, Faculty Research Grant [GPF088A-2020].

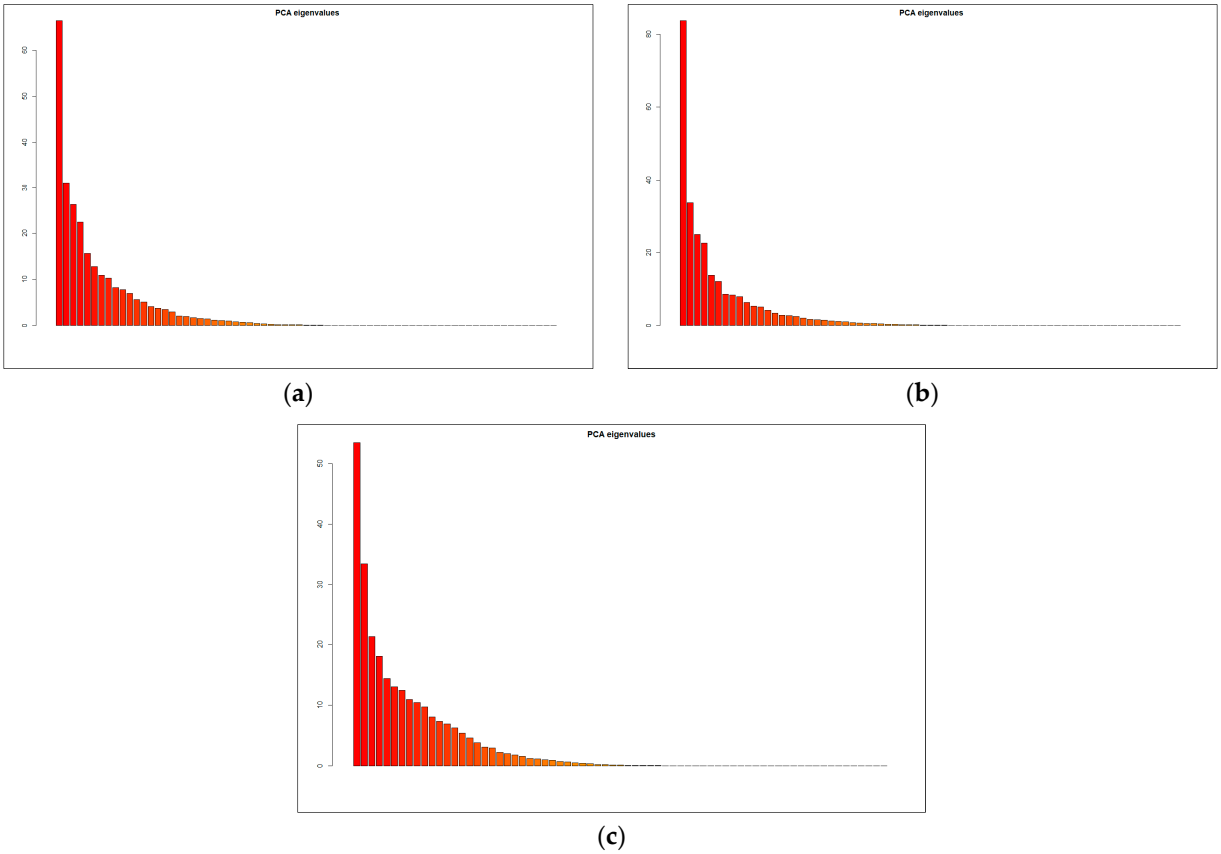
Institutional Review Board Statement: Not applicable.

Informed Consent Statement: Not applicable.

Data Availability Statement: Not applicable.

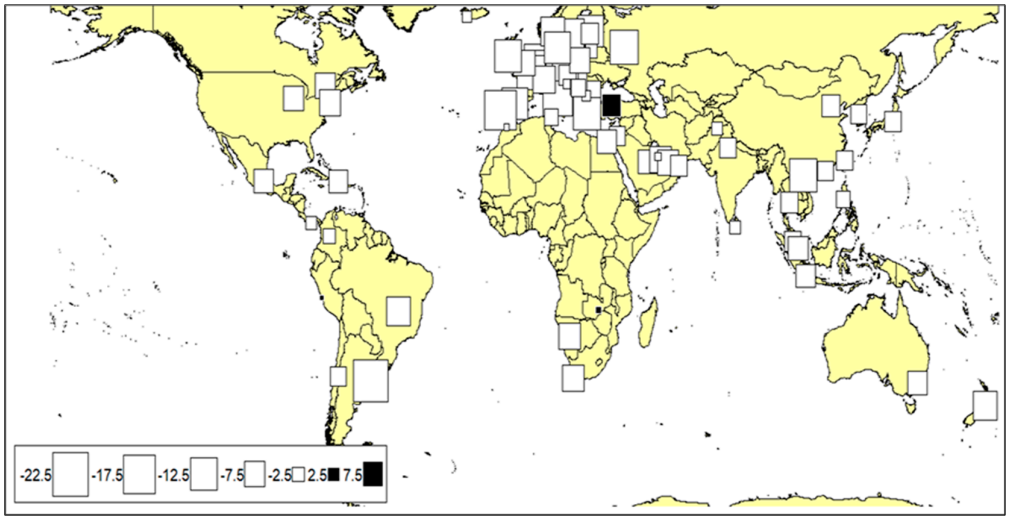
Conflicts of Interest: The authors declare no conflict of interest.

Appendix A

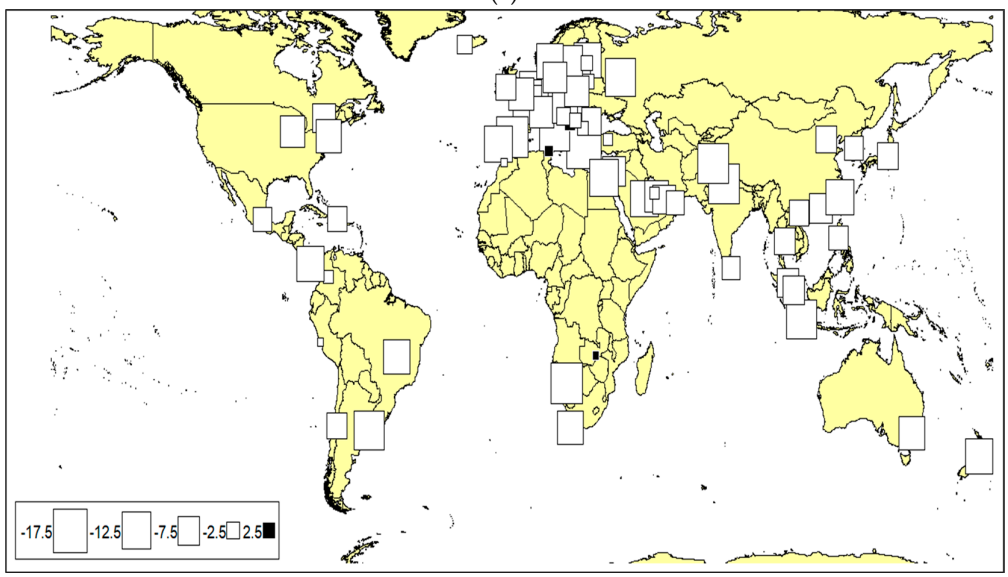


(**a**) (**b**) (**c**)

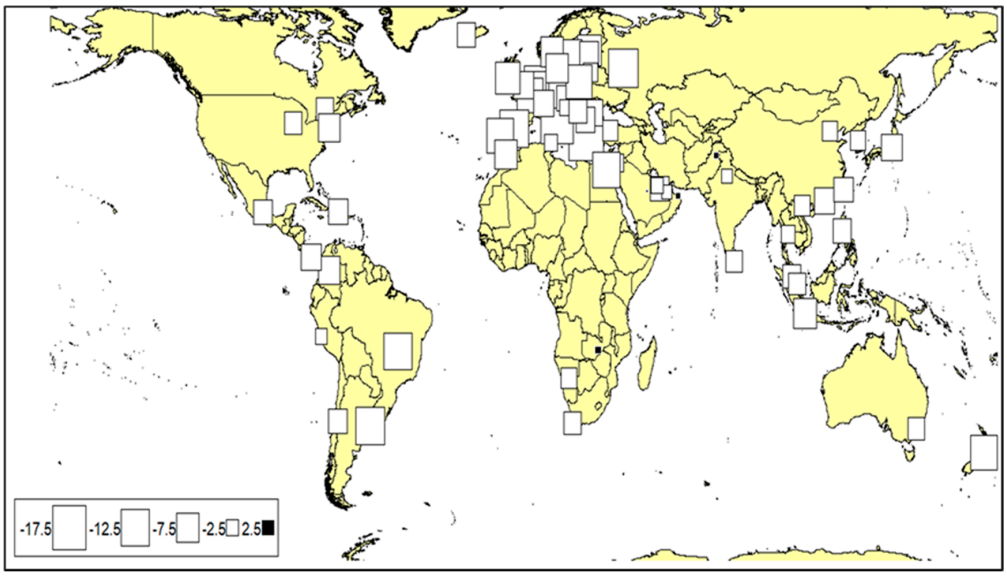
Figure A1. Distribution of eigenvalues based on FPCA: (**a**) period 1; (**b**) period 2; (**c**) period 3.



(**a**)

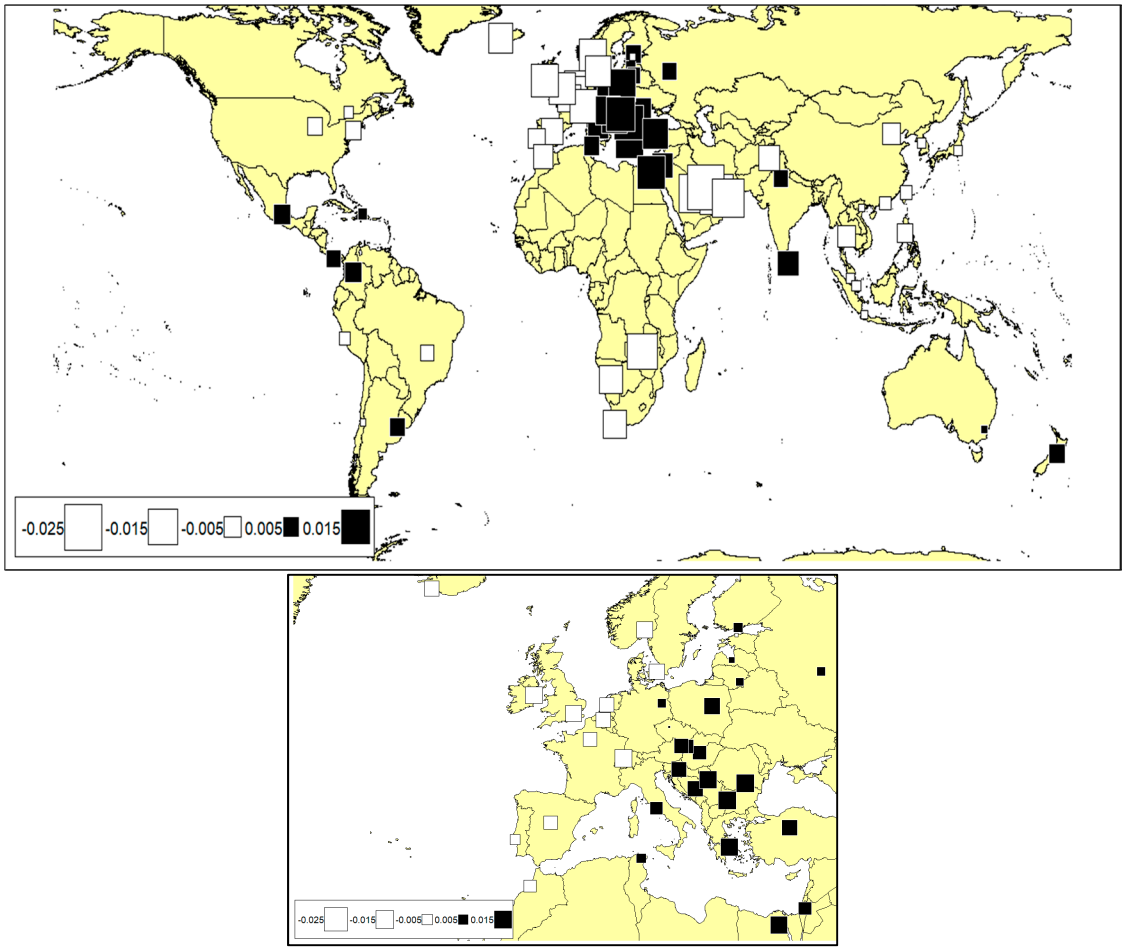


(**b**)



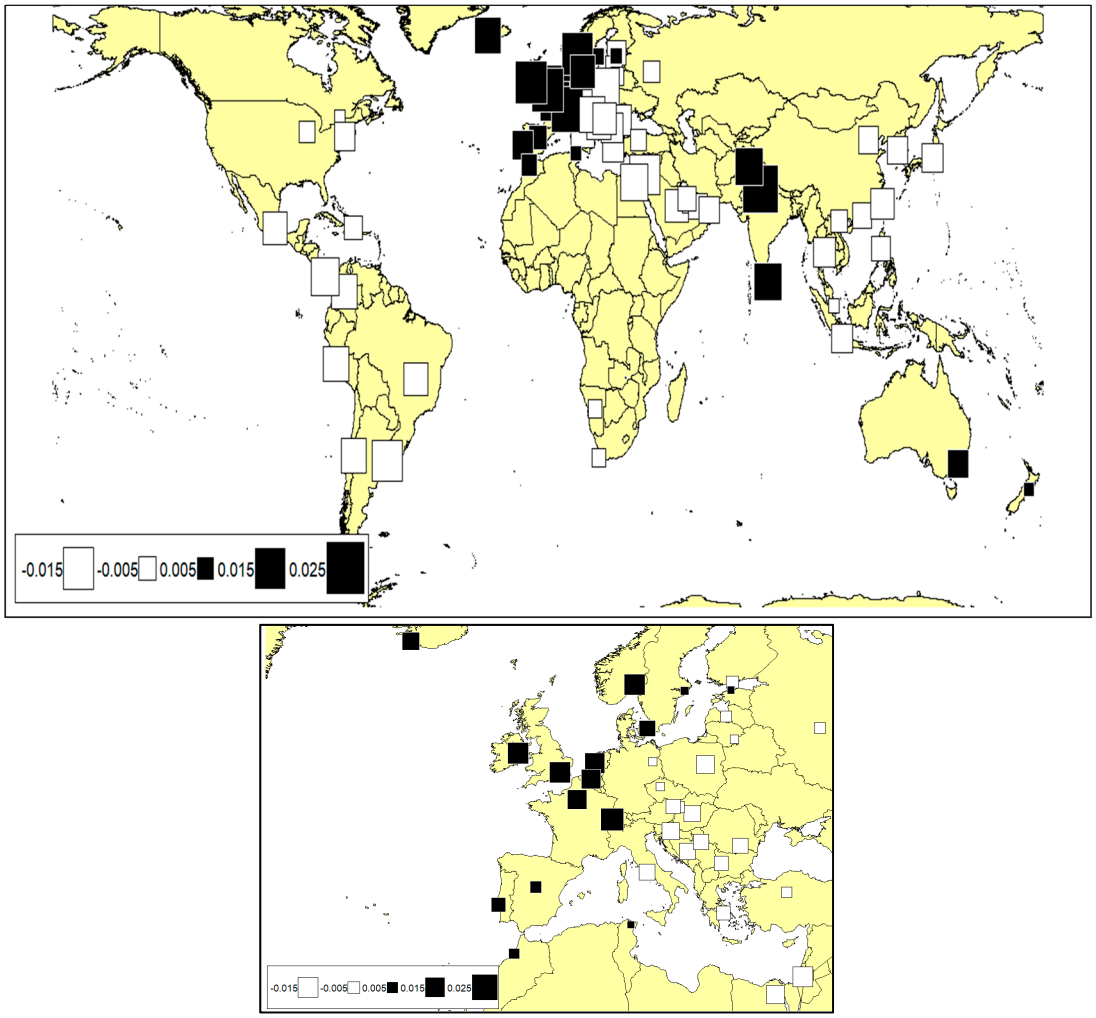
(**c**)

Figure A2. Map of projected first PC scores from FPCA: (**a**) period 1; (**b**) period 2; (**c**) period 3.



(**a**)

Figure A3. *Cont.*



(**b**)

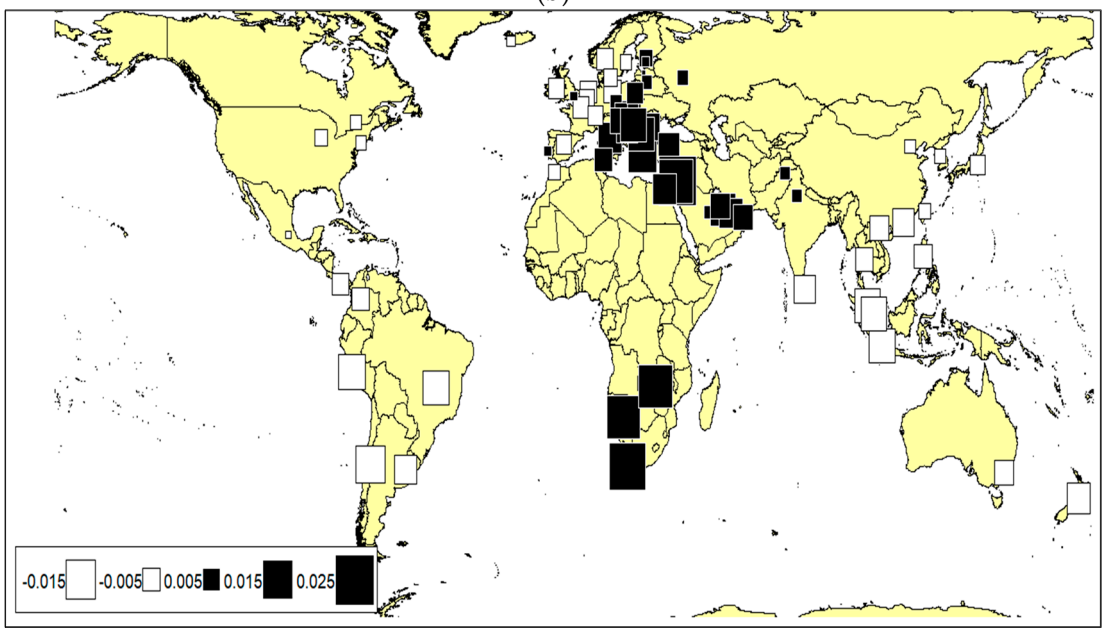
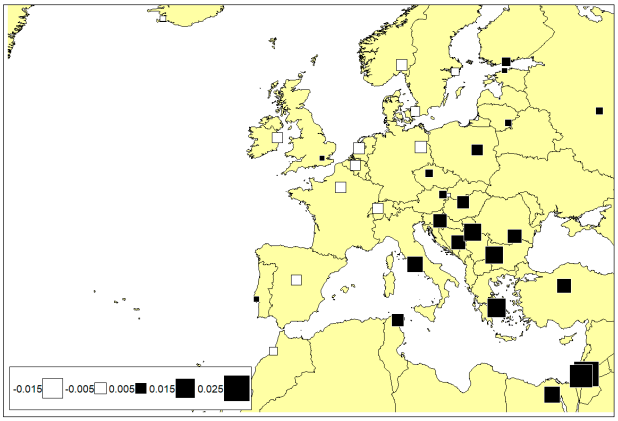
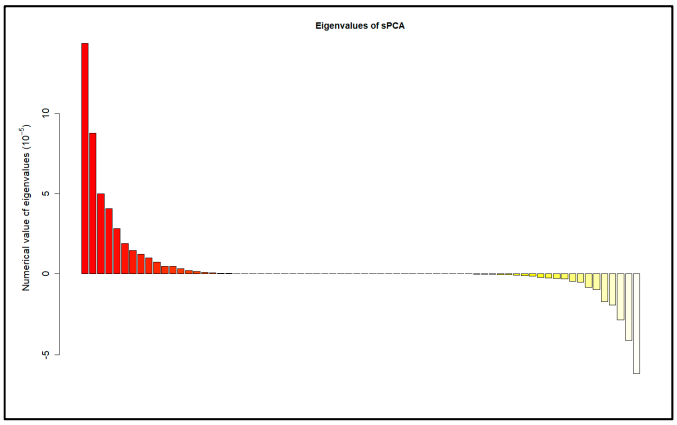


Figure A3. *Cont.*

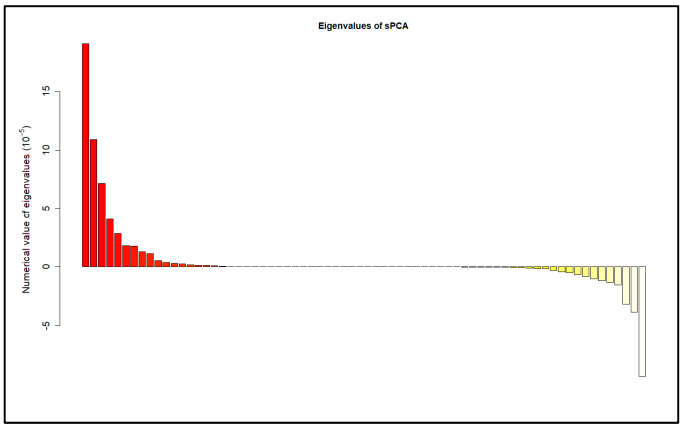


(**c**)

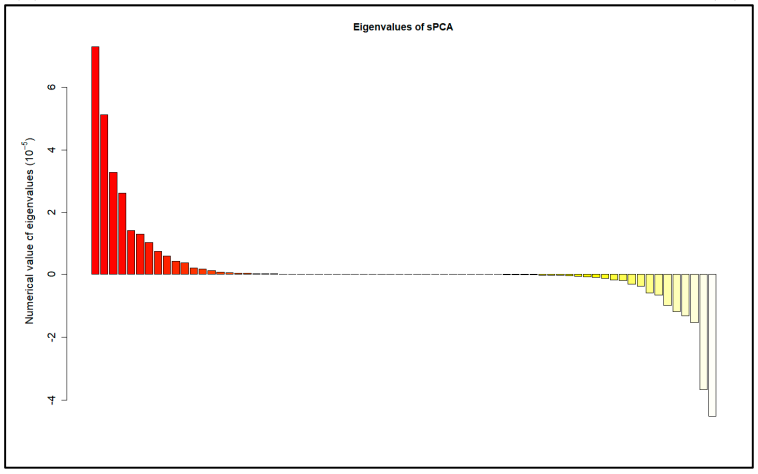
Figure A3. Map of projected second positive PC scores from FASPCA and KNN connection network; the smaller map below each bigger map is the zoom-in of Europe: (**a**) period 1; (**b**) period 2; (**c**) period 3.



(**a**)

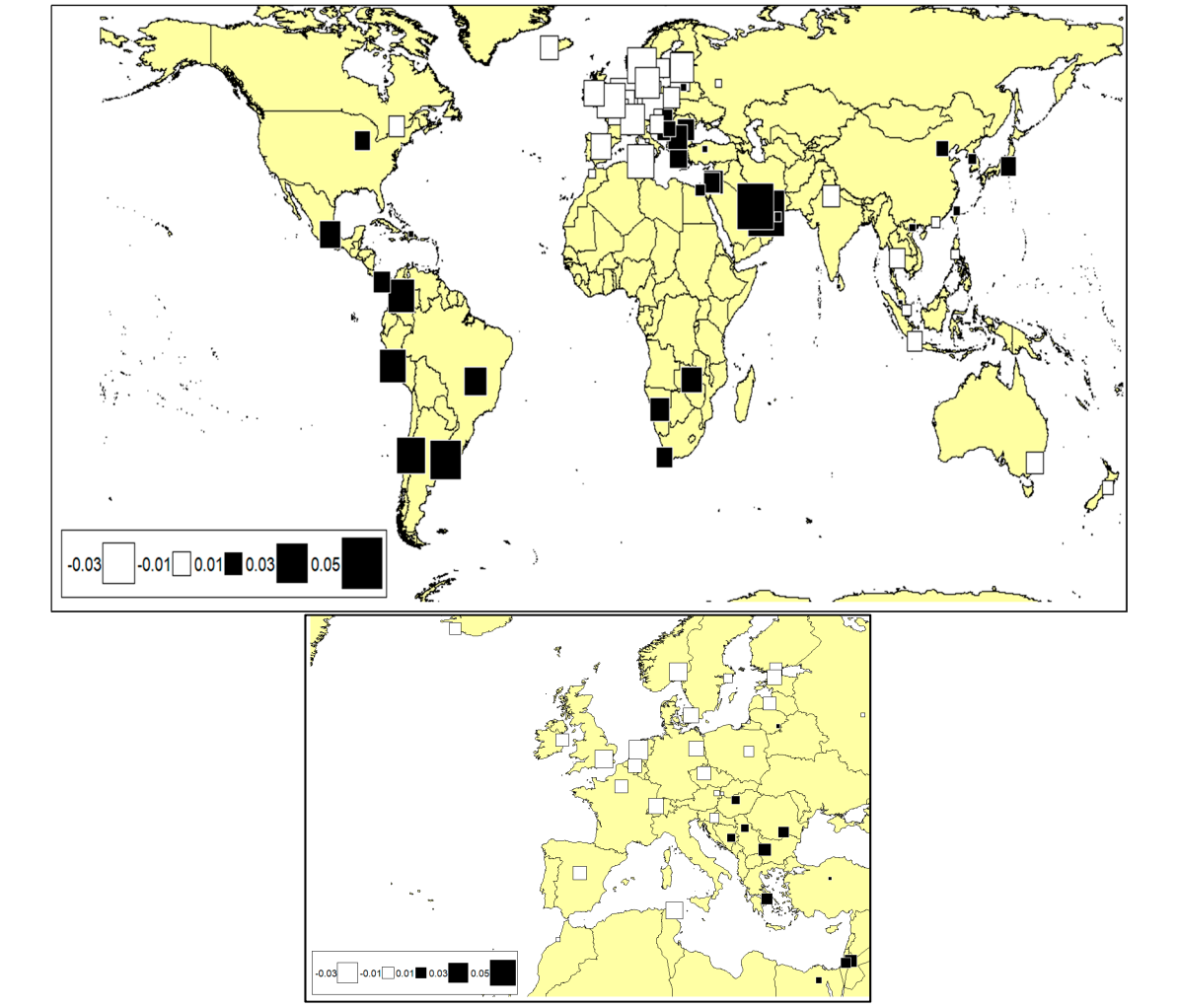


(**b**)



(**c**)

Figure A4. Distribution of eigenvalues based on FASPCA and Gabriel graph-based connection network: (**a**) period 1; (**b**) period 2; (**c**) period 3. Eigenvalues in red and yellow correspond to positive and negative eigenvalues respectively.



(**a**)

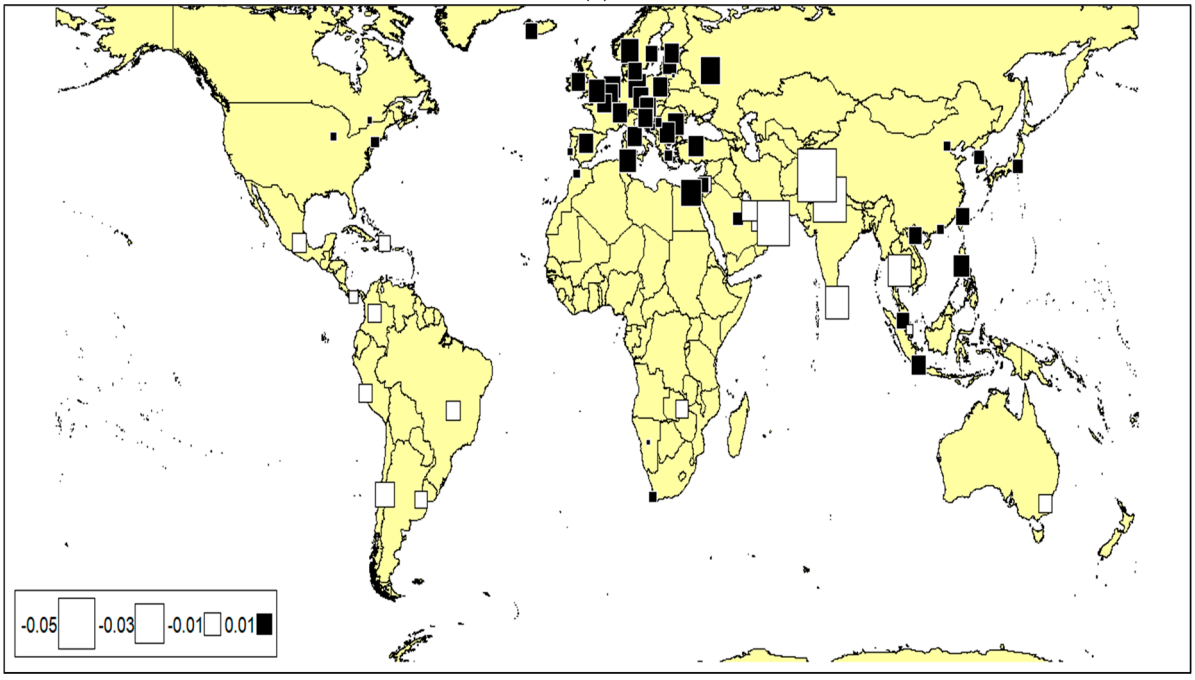
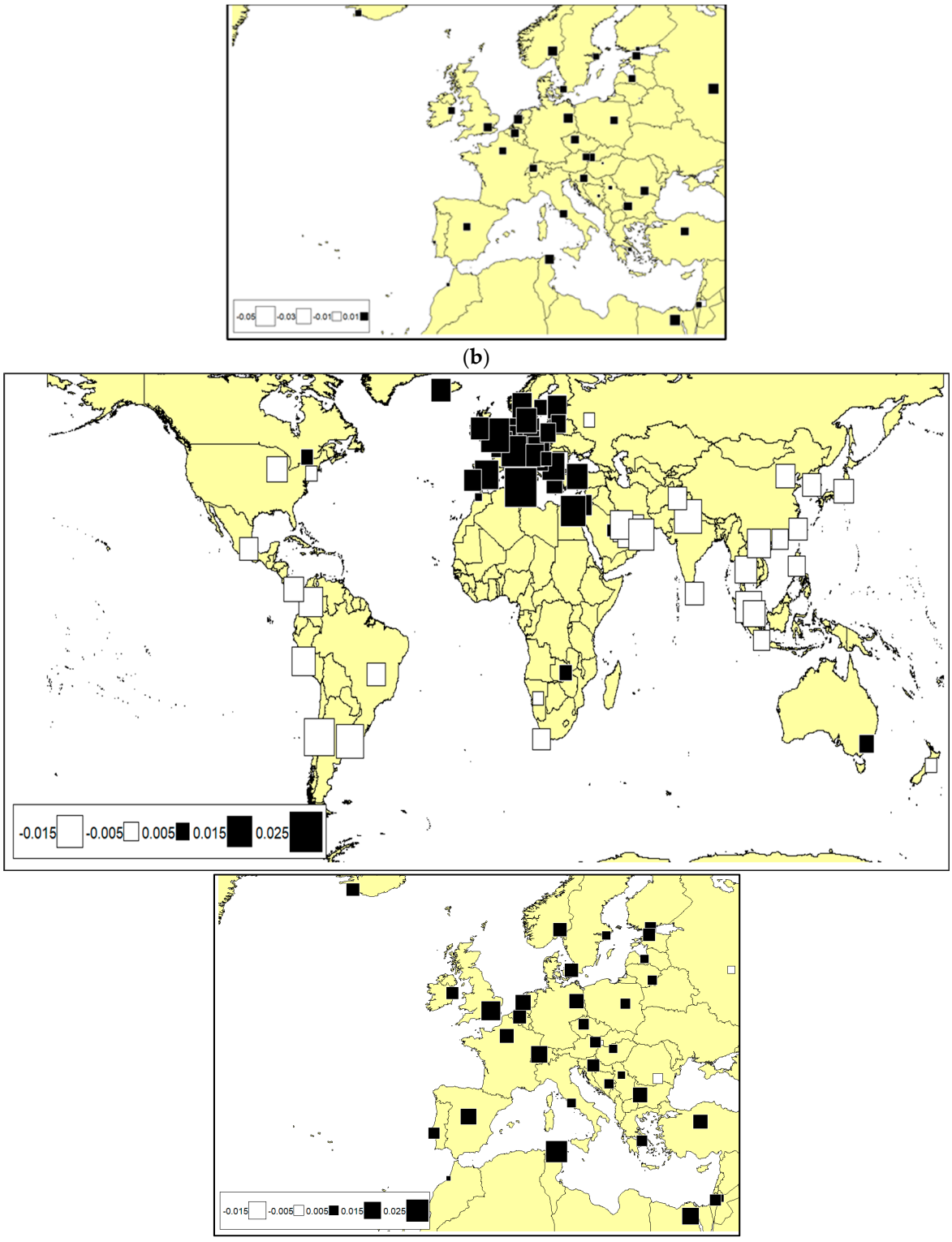


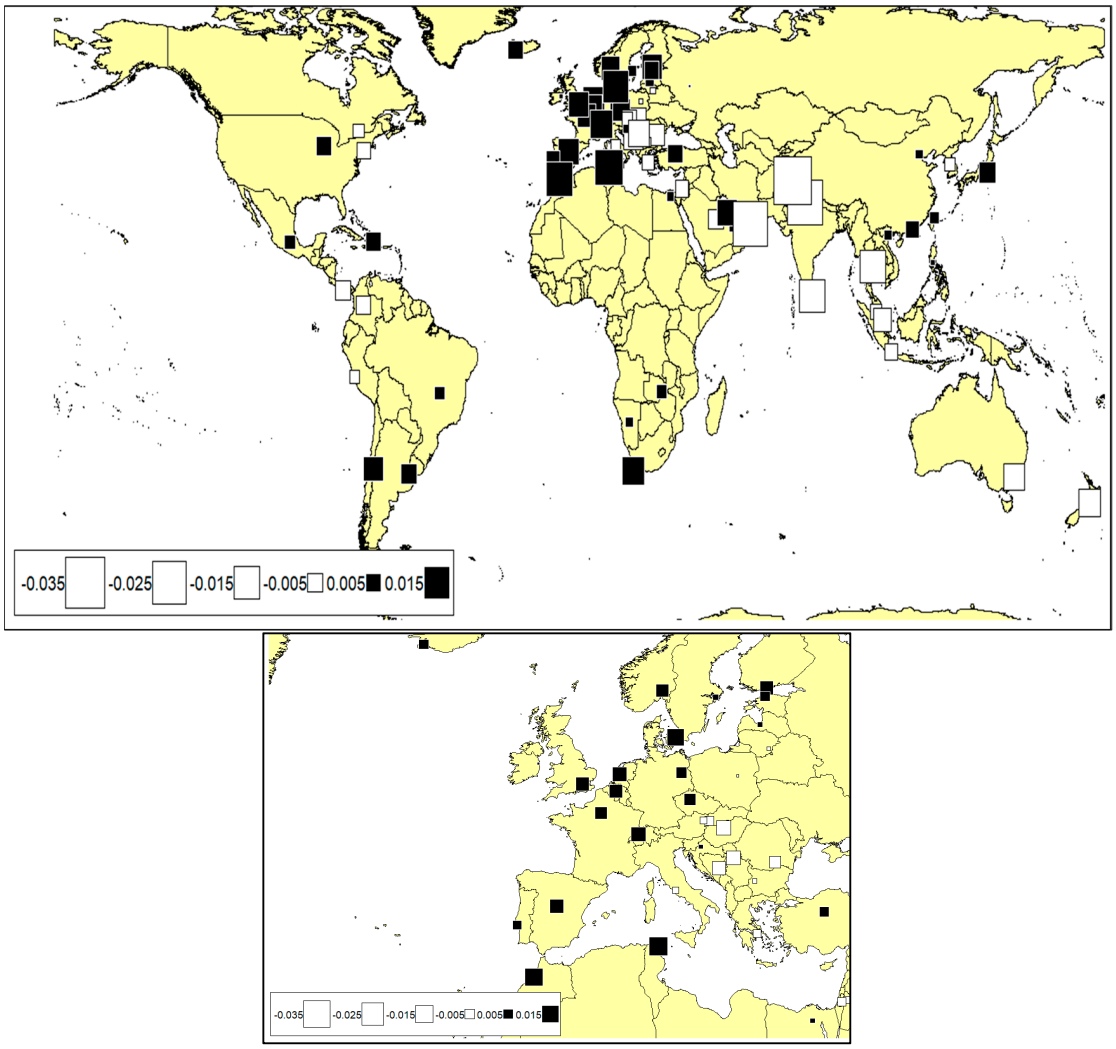
Figure A5. *Cont.*



(**b**)

(**c**)

Figure A5. Map of projected first positive PC scores from FASPCA and Gabriel graph-based connection network; The smaller map below each bigger map is the zoom-in of Europe: (**a**) period 1; (**b**) period 2; (**c**) period 3.



(**a**)

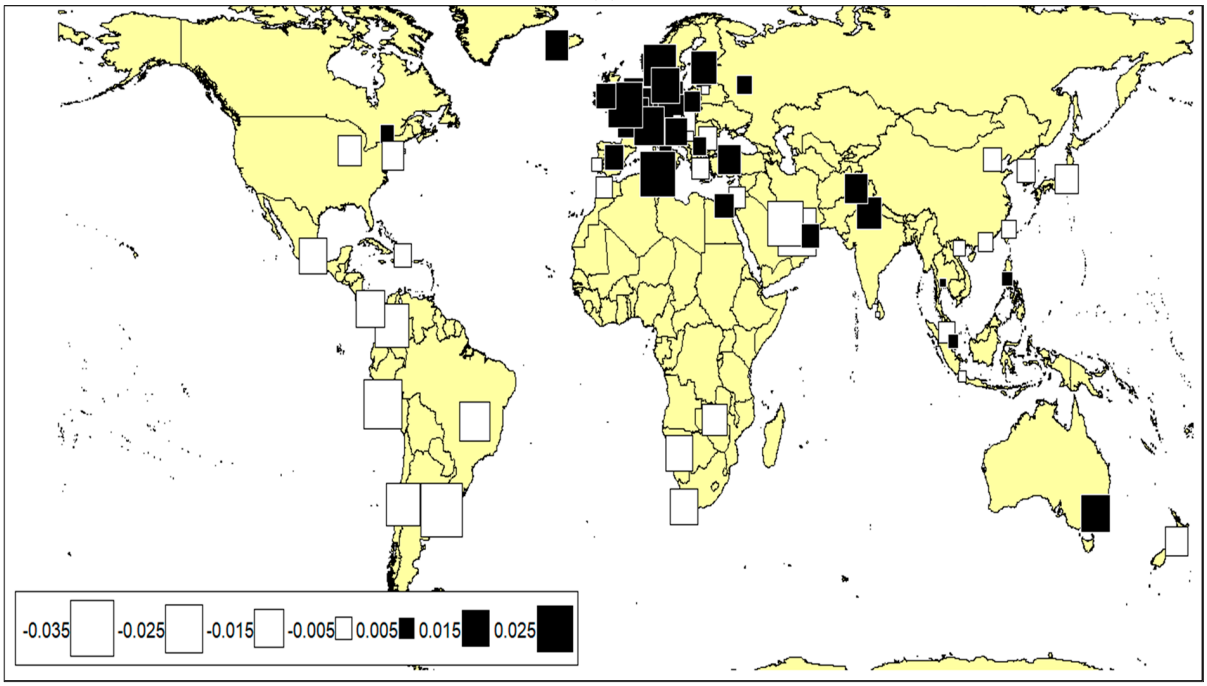


Figure A6. *Cont.*

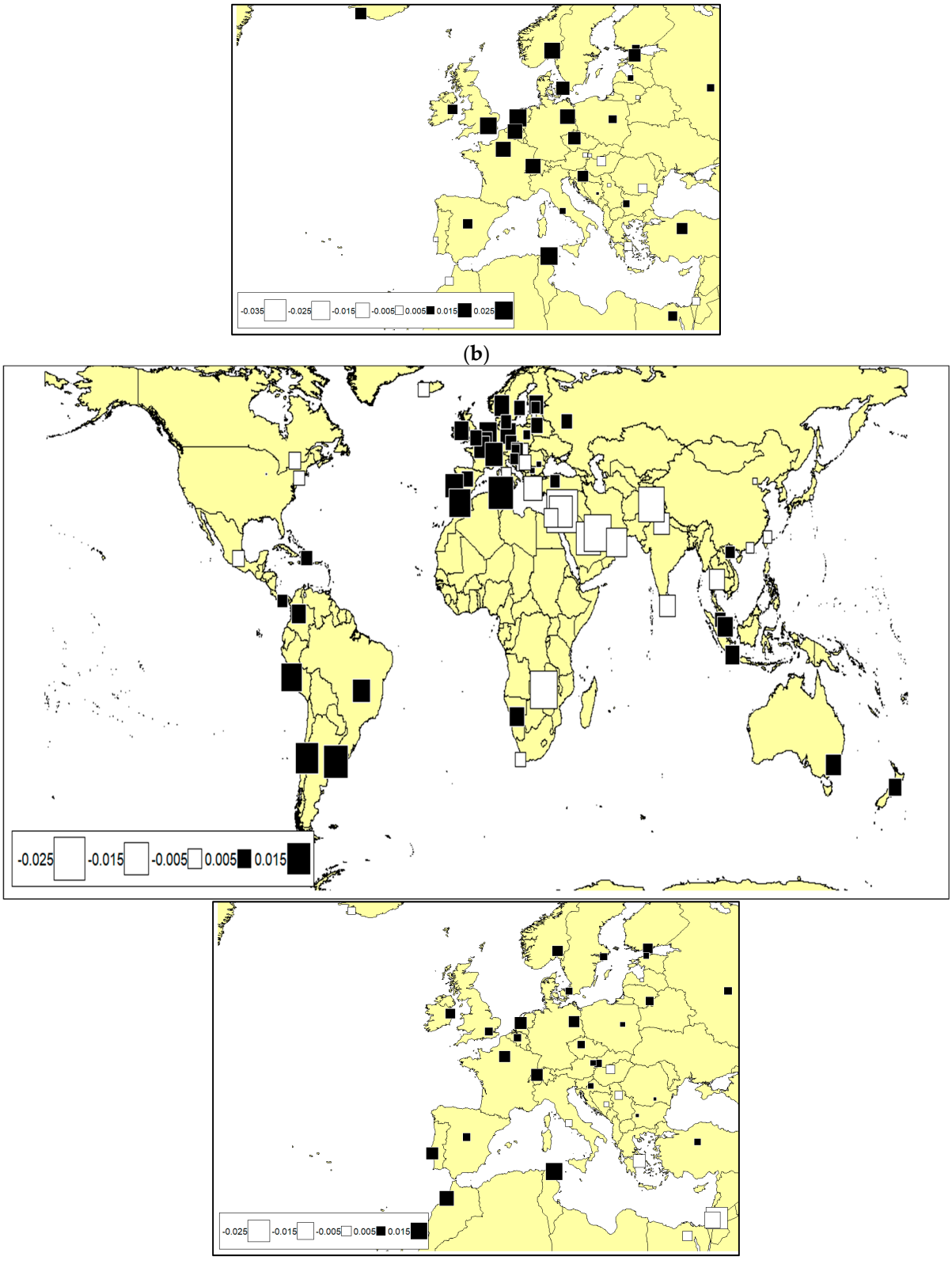


Figure A6. Map of projected second positive PC scores from FASPCA and Gabriel graph-based connection network; the smaller map below each bigger map is the zoom-in of Europe: (**a**) period 1; (**b**) period 2; (**c**) period 3.

References

1. Ramsay, J.O.; Silverman, B.W. *Functional Data Analysis*, 2nd ed.; Springer: Berlin/Heidelberg, Germany, 2005.
2. Alshahrani, F.; Almanjahie, I.M.; Elmezouar, Z.C.; Kaid, Z.; Laksaci, A.; Rachdi, M. Functional ergodic time series analysis using expectile regression. *Mathematics* **2022**, *10*, 3919. [CrossRef]
3. Shah, I.; Muhammad, I.; Ali, S.; Ahmed, S.; Almazah, M.M.A.; Al-Rezami, A.Y. Forecasting day-ahead traffic flow using functional time series approach. *Mathematics* **2022**, *10*, 4279. [CrossRef]
4. Ullah, S.; Finch, C.F. Applications of functional data analysis: A systematic review. *BMC Med. Res. Methodol.* **2013**, *13*, 43. [CrossRef]
5. Rico, M.; Cantarero, S.; Puig, F. Regional disparities and spatial dependence of bankruptcy in Spain. *Mathematics* **2021**, *9*, 960. [CrossRef]
6. Mateu, J.; Romano, E. Advances in spatial functional statistics. *Stoch. Environ. Res. Risk Assess.* **2017**, *31*, 1–16. [CrossRef]
7. Delicado, P.; Giraldo, R.; Comas, C.; Mateu, J. Statistics for spatial functional data: Some recent contributions. *Environmetrics* **2010**, *21*, 224–239. [CrossRef]
8. Fortuna, F.; Di Battista, T. Functional unsupervised classification of spatial biodiversity. *Ecol. Indic.* **2020**, *111*, 106027. [CrossRef]
9. Kuenzer, T.; Hörmann, S.; Kokoszka, P. Principal component analysis of spatially indexed functions. *J. Am. Stat. Assoc.* **2021**, *116*, 1444–1456. [CrossRef]
10. Jombart, T.; Devillard, S.; Dufour, A.B.; Pontier, D. Revealing cryptic spatial patterns in genetic variability by a new multivariate method. *Heredity* **2008**, *101*, 92–103. [CrossRef]
11. Weng, Y.; Gong, P. Modeling spatial and temporal dependencies among global stock markets. *Expert Syst. Appl.* **2016**, *43*, 175–185. [CrossRef]
12. Zhang, W.; Zhuang, X.; Li, Y. Spatial spillover around G20 stock markets and impact on the return: A spatial econometrics approach. *Appl. Econ. Lett.* **2019**, *26*, 1811–1817. [CrossRef]
13. Bolancé, C.; Acuña, C.A.; Torra, S. Non-normal market losses and spatial dependence using uncertainty indices. *Mathematics* **2022**, *10*, 1317. [CrossRef]
14. Ramírez-Parietti, I.; Contreras-Reyes, J.E.; Idrovo-Aguirre, B.J. Cross-sample entropy estimation for time series analysis: A nonparametric approach. *Nonlinear Dyn.* **2021**, *105*, 2485–2508. [CrossRef]
15. Ash, R.; Gardner, M.F. *Topics in Stochastic Processes*; Academic Press: New York, NY, USA, 1975.
16. Ramsay, J.O.; Graves, S.; Hooker, G. Fda: Functional Data Analysis [Computer Software Manual]. Available online: https://cran.r-project.org/web/packages/fda/fda.pdf (accessed on 1 June 2022).
17. Eckardt, M.; Mateu, J. Partial and semi-partial statistics of spatial associations for multivariate areal data. *Geogr. Anal.* **2021**, *53*, 818–835. [CrossRef]
18. Hassan, A.A. Exploring spatial patterns of mortality in Europe using functional spatial principal components for areal data. In *Spatial Data Analysis: Applications to Population Health*; HAL Open Science: Lyon, France, 2021; pp. 15–30.
19. Hassan, A.A. Spatial data analysis. In *Spatial Data Analysis: Applications to Population Health*; HAL Open Science: Lyon, France, 2021; pp. 9–10.
20. Matula, D.W.; Sokal, R.R. Properties of gabriel graphs relevant to geographic variation research and the clustering of points in the plane. *Geogr. Anal.* **1980**, *12*, 205–222. [CrossRef]
21. Sánchez, J.S.; Pla, F.; Ferri, F.J. Prototype selection for the nearest neighbour rule through proximity graphs. *Pattern Recognit. Lett.* **1997**, *18*, 507–513. [CrossRef]
22. Bivand, R.S.; Wong, D.W.S. Comparing implementations of global and local indicators of spatial association. *Test* **2018**, *27*, 716–748. [CrossRef]
23. Jombart, T. Adegenet: A R package for the multivariate analysis of genetic markers. *Bioinformatics* **2008**, *24*, 1403–1405. [CrossRef]
24. Dray, S.; Bauman, D.; Blanchet, G.; Borcard, D.; Clappe, S.; Guenard, G.; Jombart, T.; Larocque, G.; Legendre, P.; Madi, N.; et al. Adespatial: Multivariate multiscale spatial analysis [Computer Software Manual]. Available online: https://cran.r-project.org/web/packages/adespatial/adespatial.pdf (accessed on 1 June 2022).
25. Bougeard, S.; Dray, S. Supervised multiblock analysis in R with the Ade4 package. *J. Stat. Softw.* **2018**, *86*, 1–17. [CrossRef]
26. Chessel, D.; Dufour, A.B.; Thioulouse, J. The Ade4 package—I: One-table methods. *R News* **2004**, *4*, 5–10.
27. Dray, S.; Dufour, A.B. The Ade4 package: Implementing the duality diagram for ecologists. *J. Stat. Softw.* **2007**, *22*, 1–20. [CrossRef]
28. Dray, S.; Dufour, A.B.; Chessel, D. The Ade4 Package—II: Two-table and K-table methods. *R News* **2007**, *7*, 47–52.
29. Dickey, D.A.; Fuller, W.A. Distribution of the estimators for autoregressive time series with a unit root. *J. Am. Stat. Assoc.* **1979**, *74*, 427–431. [CrossRef]
30. Giraldo, R.; Delicado, P.; Mateu, J. Ordinary kriging for function-valued spatial data. *Environ. Ecol. Stat.* **2011**, *18*, 411–426. [CrossRef]
31. Giraldo, R.; Delicado, P. Package 'Geofd': Spatial Prediction for Function Value Data [Computer Software Manual]. Available online: https://cran.r-project.org/web/packages/geofd/geofd.pdf (accessed on 1 June 2022).
32. Shapiro, S.S.; Wilk, M.B. An analysis of variance test for normality (complete samples). *Biometrika* **1965**, *52*, 591–611. [CrossRef]
33. Bhattarai, S.; Chatterjee, A.; Park, W.Y. *Effects of US Quantitative Easing on Emerging Market Economies*; Social Science Research Network: Rochester, NY, USA, 2018; ISBN 8133593557.

34. Arslanalp, S.; Liao, W.; Piao, S.; Seneviratne, D. *China's Growing Influence on Asian Financial Markets*; IMF Working Papers Series; International Monetary Fund: Washington, DC, USA, 2016; 37p. [CrossRef]
35. Li, H. Volatility spillovers across European stock markets under the uncertainty of Brexit. *Econ. Model.* **2020**, *84*, 1–12. [CrossRef]
36. Burdekin, R.C.K.; Hughson, E.; Gu, J. A First look at Brexit and global equity markets. *Appl. Econ. Lett.* **2018**, *25*, 136–140. [CrossRef]
37. Wang, G.J.; Xie, C.; Stanley, H.E. Correlation structure and evolution of world stock markets: Evidence from Pearson and partial correlation-based networks. *Comput. Econ.* **2018**, *51*, 607–635. [CrossRef]
38. European Stock Markets Drop on Greece Exit Fears. Available online: https://web.archive.org/web/20181015114506/https://www.businessinsider.com/afp-european-stock-markets-drop-on-greece-exit-fears-2015-6 (accessed on 1 December 2022).
39. Lee, G.; Jeong, J. An investigation of global and regional integration of ASEAN economic community stock market: Dynamic risk decomposition approach. *Emerg. Mark. Financ. Trade* **2016**, *52*, 2069–2086. [CrossRef]
40. Marukawa, T. Dependence and competition: Trade relationship between Asian countries and China. *J. Contemp. East Asia Stud.* **2021**, *10*, 246–261. [CrossRef]
41. Halleck Vega, S.; Elhorst, J.P. A regional unemployment model simultaneously accounting for serial dynamics, spatial dependence and common factors. *Reg. Sci. Urban Econ.* **2016**, *60*, 85–95. [CrossRef]

Disclaimer/Publisher's Note: The statements, opinions and data contained in all publications are solely those of the individual author(s) and contributor(s) and not of MDPI and/or the editor(s). MDPI and/or the editor(s) disclaim responsibility for any injury to people or property resulting from any ideas, methods, instructions or products referred to in the content.