Proceedings of the 35th European Safety and Reliability & the 33rd Society for Risk Analysis Europe Conference
Edited by Eirik Bjorheim Abrahamsen, Terje Aven, Frederic Bouder, Roger Flage, Marja Ylönen

doi: 10.3850/978-981-94-3281-3_ESREL-SRA-E2025-P7195-cd

Spatiotemporal Crime Analysis for Risk Management Using the Non-Stationary Moving Average Method

Karolina Dlouhá

Department of Security Services, Faculty of Safety Engineering, VSB-TUO, Ostrava, Czech Republic.
E-mail: karolina.dlouha@vsb.cz

Lukáš Pospíšil

Department of Mathematics, Faculty of Civil Engineering, VSB-TUO, Ostrava, Czech Republic.
E-mail: lukas.pospisil@vsb.cz

Radomír Ščurek

Department of Security Services, Faculty of Safety Engineering, VSB-TUO, Ostrava, Czech Republic.
E-mail: radomir.scurek@vsb.cz

This paper presents a novel approach to spatiotemporal crime analysis, tailored for risk management and safety applications, by introducing the Non-Stationary Moving Average (NSMA) method. The NSMA method extends the standard Moving Average technique by incorporating non-stationarity, addressing the dynamic nature of crime patterns. By combining temporal smoothing with spatial clustering through the K-means algorithm, this approach enables the identification of distinct crime clusters and provides insights into temporal trends. The proposed methodology is formulated as a multicriteria optimization problem, balancing the objectives of spatial clustering and temporal regularization through a regularization parameter. The problem is solved using a subspace algorithm, similar to the approach used in K-means, which alternates between optimizing cluster centers and cluster moving averages. Applied to real-world crime data from the Czech Republic, this method demonstrates its potential to improve resource allocation and decision-making in crime prevention. The NSMA method contributes to advancing the fields of spatiotemporal analysis and risk evaluation, offering a versatile tool for addressing complex urban safety challenges.

Keywords: Moving average, modeling, clustering, crime analysis, regression.

1. Introduction

Security is one of the most important pillars of a functioning modern society. In an environment where criminal activities are constantly evolving, it is crucial to have tools that not only help understand current threats but also enable their prevention. Spatiotemporal analysis, which combines spatial and temporal data, represents a modern approach to identifying high-risk areas and understanding the dynamics of crime. This analytical framework contributes to effective decision-making in deploying security forces, preventing crime, and developing long-term public safety strategies.

The dataset used for this analysis contains information about criminal incidents in the Czech Republic for the year 2024. It includes precise timestamps of reports made by witnesses or the police, exact geographic coordinates, and the types of criminal activities. These data offer a detailed view of the spatial and temporal characteristics of crime, enabling the identification of patterns and the dynamics of its occurrence. Additionally, the dataset provides information on the investigative stages of each incident, offering a comprehensive overview from the moment of reporting to the current status.

The inherently multifaceted nature of crime requires the integration of traditional methods with innovative technologies to effectively address emerging threats. By combining advanced analytical techniques with spatial and temporal data, it becomes possible not only to allocate resources more effectively but also to strengthen society's overall capacity to respond to security challenges.

The inherently multifaceted nature of crime requires the integration of traditional methods with innovative technologies to effectively address emerging threats. By combining advanced analytical techniques with spatial and temporal data, it becomes possible not only to allocate resources more effectively but also to strengthen society's overall capacity to respond to security challenges.

Spatiotemporal crime analysis is crucial for understanding and predicting criminal activities across different areas and times. By employing clustering methods, this approach can help identify areas with varying levels of criminal activity, which is essential for crime prevention and the efficient allocation of resources in law enforcement. Recognizing locations with higher concentrations of crime allows security forces to be strategically deployed, enabling faster and more effective responses to emerging threats.

Additionally, clustering analysis offers the potential to tailor preventive measures to the unique needs of specific regions. Strategies such as enhanced surveillance, the installation of advanced security systems, or community-focused crime prevention programs can be developed based on these insights. Beyond immediate tactical advantages, this methodology provides a robust foundation for long-term planning and improving public safety. By integrating spatial and temporal patterns, spatiotemporal crime analysis contributes to creating safer environments while equipping society with the tools needed to address evolving security challenges.

As a standard method for spatio-temporal analysis, STARMA (Spatio-Temporal Autoregressive, Pfeifer and Deutsch (1980)) models can be mentioned. These models analyze the autocorrelation (i.e., the linear dependency) between the temporal component (consecutive time steps) and, simultaneously, the spatial component (data in neighboring regions).

However, this approach has certain limitations, including the necessity of time discretization, which can be performed, for example, on a daily basis, as well as fixed spatial discretization, such as box discretization with a predefined box size parameter.

In the novel method presented in this paper, we address spatial discretization by introducing adaptive clustering methods based on a more general centroid-based approach. The assumption of linear dependency between consecutive time steps is removed and replaced with an assumption based on the smoothness of the data. This is achieved by generalizing the moving average method.

Traditional methods, such as the Moving Average (MA), provide a simple tool for analyzing data but lack the flexibility needed for dynamic environments where crime patterns can change rapidly. The extension of this method presented in this paper, called the Non-Stationary Moving Average (NSMA) method, offers an advanced cluster-based approach that dynamically separates records into groups and applies the moving average to individual clusters. This model is better at describing the data and provides deeper insights into underlying patterns.

2. Mathematical aspect of modelling process

Our approach builds on a standard, well-established regression technique where we fit a parametric model to the given data. To enhance this, we integrate spatial clustering using the K-means method with a modified version of the widely used moving average technique, which smooths data over time. This modification incorporates an indicator function, allowing us to introduce piecewise non-stationarity.

By combining these two methods through multi-objective minimization, we develop a spatiotemporal approach. In the final formulation, the temporal clustering effectively acts as a regularization mechanism for the spatial clustering.

This paper examines a dataset consisting of ordered pairs:

$$\{[\tau_i, X_i], i \in \mathcal{N}\}, \quad \tau_i \in \mathbb{R},\ X_i \in \mathbb{R}^2, \qquad (1)$$

where $\mathcal{N} := \{1, \ldots, N\} \subset \mathbb{N}$ is the index set, and $N \in \mathbb{N}$ represents the total number of data points. Here, τ_i denotes the time associated with the event of interest, while X_i corresponds to its geographic location.

In this work, we focus on crime incidents reported in the official statistics provided by The Police of the Czech Republic (2024). Specifically, we analyze offenses categorized as *general criminal activity* for the year 2024. Each crime record includes details of the date and time of occurrence, along with the corresponding location described by latitude and longitude coordinates.

2.1. *Spatial clustering: K-means*

In this section, we focus on finding groups of data with similar spatial characteristics. We adopt the standard K-means method (see, e.g., MacQueen (1967)), which cluster the given spatial part of data X_i (1) into $K \in \mathbb{N}$ clusters by solving the optimization problem

$$[C^*, \Gamma^*] := \arg \min_{\Gamma \in \Omega_\Gamma} \underbrace{\sum_{i=1}^{N} \sum_{k=1}^{K} \Gamma_{k,i} \cdot \|X_{:,i} - C_{:,k}\|_2^2}_{=:f(C,\Gamma)}, \tag{2}$$

where $C \in \mathbb{R}^{2,K}$ is a matrix of so-called *centroids* and $\Gamma \in \mathbb{R}^{K,N}, \Gamma_{k,i} = \gamma_k(i), \gamma_k : \{1, \ldots, N\} \to \{0, 1\}$ are *cluster indicator functions* defined by

$$\gamma_k(i) = \begin{cases} 1 & \text{if } X_i \text{ belongs to } k\text{-th cluster,} \\ 0 & \text{elsewhere.} \end{cases} \tag{3}$$

We suppose that every datapoint X_i belongs to exactly one cluster. This condition defines the feasible set of problem (2) as

$$\Omega_\Gamma := \left\{ \Gamma \in \{0, 1\}^{K,N} \mid \forall i : \sum_{k=1}^{K} \Gamma_{k,i} = 1 \right\}. \tag{4}$$

The problem (2) is solved using *subspace algorithm*, see Alg. 1. The problem is addressed by alternately fixing one variable and minimizing the objective function with respect to the other. Beginning with an initial guess, the method iteratively updates each variable until convergence. This approach simplifies the optimization by breaking it into a series of single-variable subproblems. The algorithm ensures a sequence of objective function values that do not increase.

Both inner problems have analytical solutions. In the case of the C-problem (when the affiliation to clusters Γ is fixed), the new mean value is computed as the average of points affiliated with the given clusters. For the Γ-problem (when cluster mean values C are fixed), the points are assigned to the clusters with the smallest distance. It can be easily proven that the sequence converges to a locally optimal point depending on the initial approximation. In practice, a technique called *annealing* is applied: the algorithm is run from several random initial guesses, and the best solution, i.e., the one with the smallest objective function value, is selected.

Alg. 1: Subspace algorithm: K-means

> Choose initial approximation $\Gamma^{\langle 0 \rangle} \in \Omega_\Gamma$
> Set initial $f_\alpha^{\langle 0 \rangle} = \infty$
> Set iteration counter $\text{it} = 0$
> **repeat**
> $\quad C^{\langle \text{it}+1 \rangle} = \arg \min_{C \in \mathbb{R}^{2,K}} f(C, \Gamma^{\langle it \rangle})$
> $\quad \Gamma^{\langle \text{it}+1 \rangle} = \arg \min_{\Gamma \in \Omega_\Gamma} f(C^{\langle \text{it}+1 \rangle}, \Gamma)$
> $\quad f^{\langle \text{it}+1 \rangle} = f(C^{\langle \text{it}+1 \rangle}, \Gamma^{\langle \text{it}+1 \rangle})$
> $\quad \text{it} = \text{it} + 1$
> **until** $|f^{\langle \text{it} \rangle} - f^{\langle \text{it}-1 \rangle}| < \varepsilon$;

Horenko (2010) introduced a relaxed version of the K-means algorithm with regularization. The key idea is to relax the binarity of the affiliation functions Γ by interpreting them as probabilities of a point's affiliation to clusters. Additionally, to handle time-series data and incorporate the assumption of cluster affiliation persistence, the author introduced, among other ideas, the H^1 regularization, which involves minimizing the Euclidean norm of the discrete derivative. This approach is further analyzed and extended in Pospíšil et al. (2018).

In the following section of the paper, we adapt this idea and introduce regularization using the Moving Average (MA) method.

2.2. *Temporal clustering: Nonstationary moving average method*

In this part, we focus on temporal part of the data τ_i (1). The first step of our analysis involves time discretization. We calculate the daily number of crime records and represent it as $\varphi(t)$, where

$t = 1, \ldots, 365$. This time series displays fluctuations, making it suitable for smoothing using the standard MA method, as demonstrated in Fig. 1.

The MA method (see, e.g., Frost (2024)) is a statistical technique that smooths short-term fluctuations in a dataset, emphasizing longer-term trends by calculating the average of a specified number of consecutive data points. Let us call the number of consecutive data points by $P \in \mathbb{N}$ and define a new function MA by averaging previous P function values, i.e.,

$$\hat{\varphi}(t) := \frac{1}{P} \sum_{p=1}^{P} \varphi(t-p). \tag{5}$$

See Fig. 1 for an example for various values of parameter P.

The Non-Stationary Moving Average (NSMA) extends the standard MA method by accounting for non-stationarity, which assumes the existence of regimes between which the analyzed process alternates. This extension is achieved by introducing the affiliation of individual records to specific groups, referred to as clusters or regimes. Let $K \in \mathbb{N}$ denote the number of clusters, and define the regime indicator functions Γ as specified in (3).

The number of events in time t in cluster k can be easily computed as a sum of values of corresponding indicator functions of records in given day and we can define the moving average function of the cluster (a simple modification of (5))

$$\varphi_k(t) := \sum_{i \in \mathcal{N}} \gamma_k(i), \quad \hat{\varphi}_k(t) := \frac{1}{P} \sum_{p=1}^{P} \varphi_k(t-p). \tag{6}$$

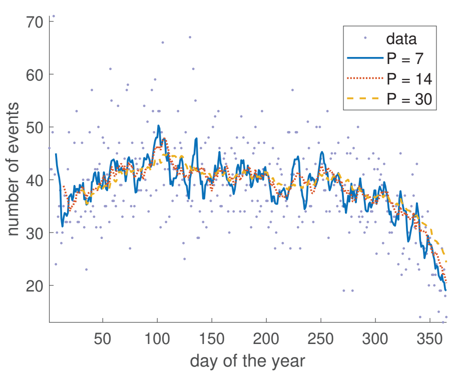

Fig. 1. Moving average: number of events labeled as *General Criminal Activity* in 2024. The data is smoothed using the MA method with various window sizes P.

The optimal values of the indicator functions can be obtained by solving a regression problem. In this case, we aim to minimize the error of the NSMA by finding the best partition of events into clusters, such that the smoothing error is minimized. By using a squared Euclidean distance function, the optimization problem can be formulated as follows:

$$\Gamma^* = \arg\min_{\Gamma \in \Omega_\Gamma} \underbrace{\sum_t (\varphi(t) - \hat{\varphi}(t))^2}_{=:g(\Gamma)}, \tag{7}$$

where we substitute the property

$$\hat{\varphi}(t) = \sum_{k=1}^{K} \hat{\varphi}_k(t). \tag{8}$$

This method allows for the identification of groups of events with similar smooth characteristics in time.

It can be shown that the optimization problem (7) is a convex quadratic optimization problem. However, due to the binary nature of the feasible set, the problem becomes practically unsolvable when dealing with large datasets. To overcome this challenge, we relax the problem by replacing the binary cluster affiliation with probabilistic affiliation. We relax the feasible set (4) to the form

$$\hat{\Omega}_\Gamma := \left\{ \Gamma \in [0,1]^{K,N} \mid \forall i : \sum_{k=1}^{K} \Gamma_{k,i} = 1 \right\}. \tag{9}$$

This modification allows the problem (7) to be solved using methods like the Spectral Projected Gradient (SPG) method introduced by Birgin et al. (2000). For further details regarding its application to quadratic programming (QP), see Pospíšil et al. (2018).

2.3. *Spatio-temporal clustering: multicriteria optimization*

The objective of the proposed methodology is to simultaneously cluster events both in time and space. To achieve this, we combine the minimization problems (7) and (2) into a single multicriteria optimization problem

$$[C^*, \Gamma^*] := \arg \min_{\Gamma \in \hat{\Omega}_\Gamma} \underbrace{\frac{1}{N} f(C,\Gamma) + \frac{\varepsilon}{T} g(\Gamma)}_{:=L_\varepsilon(C,\Gamma)}, \tag{10}$$

where $\varepsilon \geq 0$ is the aggregation parameter that balances the importance of time regularization via NSMA and spatial clustering via K-means, enabling a unified spatiotemporal solution approach.

We address the numerical solution of (10) by combining the numerical methods for both K-means and NSMA. The subspace algorithm simplifies the problem into a sequence of C-problems and Γ-problems. The C-problems yield the same solution as the K-means algorithm since the additional term remains constant when Γ is fixed and does not affect the minimizer. The Γ-problems, on the other hand, are still QP problems, as the first term introduces only a new linear component, which can be solved efficiently using the SPG method.

3. Results and Discussion

The proposed methodology was implemented in a MATLAB environment and applied to the official crime statistics for the year 2024. A moving average window size of $P = 7$, corresponding to a weekly average, was selected. The number of clusters was set to $K = 30$.

To determine the optimal regularization parameter ε, the L-curve approach was employed, as described in Hansen and O'Leary (1993) and illustrated in Fig. 2. Based on the analysis, we selected $\varepsilon = 10^{0.5}$, a value that achieves a balance between sufficiently low temporal clustering error and an acceptable level of spatial clustering error.

This analysis led to the emergence of several empty clusters, which were subsequently removed. Due to the limitations of the Euclidean distance measure used in K-means for spatial clustering - particularly its ineffectiveness in grouping events that are not close to general centroids - we reapplied K-means clustering. This second step was designed to group clusters with similar moving averages. Through this process, the number of clusters was reduced to $K = 5$, as shown in Fig. 3. The final results effectively identify events with similar moving averages and highlight regions characterized by comparable occurrences of the crime activities of interest.

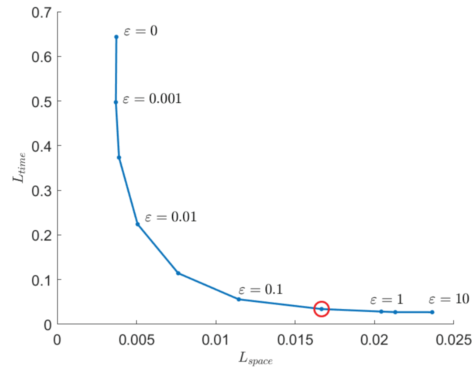

Fig. 2. L-curve: The highlighted value represents the selected Pareto-optimal aggregation parameter $\varepsilon = 10^{0.5}$. The value L_{space} corresponds to the spatial clustering term in the cost function, $f(C^*, \Gamma^*)/N$, while L_{time} represents the error from the non-stationary moving average, $g(\Gamma^*)/T$, which is the second part of the objective function in (10).

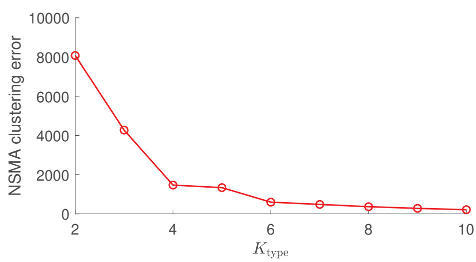

Fig. 3. Reclustering: The error of clustering NSMA of clusters; we aim to identify clusters of the NSMA to define types of crime activity and simplify the categorization of clusters identified in the previous analysis. Based on this curve, the optimal number of cluster classes was determined to be $K_{\text{type}} = 6$.

We present our final classification and typization of crime activities in the Czech Republic, based on the official data for the year 2024. The optimal spatial clustering ensures that events within each identified type are geographically close to each other, while the temporal clustering guarantees that the identified types exhibit similar moving averages over time. This dual clustering approach allows us to simultaneously smooth short-term fluctuations in the dataset and reduce the dispersion of event locations.

By leveraging spatial clustering, we grouped events into localized clusters, reflecting the assumption that criminal activities tend to form spatially cohesive patterns. Similarly, temporal clustering captures the inherent dynamics of these events, revealing distinct types of crime activities characterized by comparable temporal trends. This methodology assumes the existence of underlying types of criminal activities and effectively combines spatial and temporal dimensions to enhance the interpretability of the data.

Overall, this approach not only smooths temporal noise but also enforces spatial consistency, providing a comprehensive framework for understanding and categorizing crime activity patterns in a way that aligns with real-world observations of clustering and typization in criminal behavior.

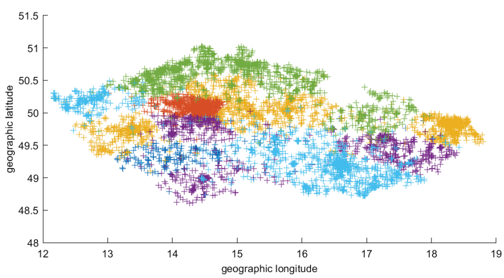

Fig. 4. Map of crime activities: events have been grouped into $K_{\text{type}} = 6$ types based on the results of NSMA analysis with optimal parameters.

The spatial classification of crime activities is presented on the map in Fig. 4, where the identified clusters are visualized to highlight their geographical distribution. The corresponding moving averages for each crime activity type, reflecting their temporal trends, are shown in Fig. 5. Together, these figures provide a comprehensive view of the spatial and temporal characteristics of the classified crime activity types, illustrating the effectiveness of the proposed clustering approach.

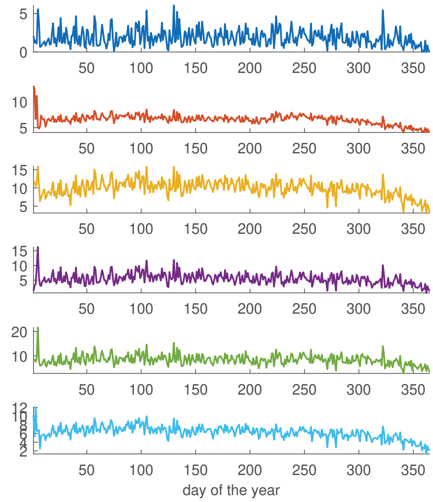

Fig. 5. Moving averages of types: the value of moving averages for individual identified types of criminal activities.

4. Conclusion

In this paper, we introduced a non-stationary extension of the popular Moving Average method, combined with K-means clustering, to achieve a unified spatiotemporal approach for event clustering. This methodology integrates temporal smoothing and spatial clustering into a single multicriteria optimization problem, balanced by a regularization parameter. By relaxing binary cluster affiliations into probabilistic ones, we addressed computational challenges..

The proposed approach was implemented in MATLAB and applied to real-world data on general criminal offenses. Using the L-curve method, we identified an optimal regularization parameter and demonstrated the method's effectiveness in capturing both spatial and temporal patterns while maintaining computational feasibility.

Future work will focus on a more in-depth analysis of the methodology, its application to larger datasets, and evaluating its performance using standard tools such as cross-validation. Additionally, we plan to extend the application of this method to other areas of security, further exploring its potential in data-driven solutions for complex systems.

Acknowledgement

This contribution was made possible thanks to the funding provided by Department of Mathematics at Faculty of Civil Engineering, VSB-TUO.

References

Birgin, E. G., J. M. Martínez, and M. M. Raydan (2000). Nonmonotone spectral projected gradient methods on convex sets. *SIAM Journal on Optimization 10*, 1196–1211.

Frost, J. (2024). Moving averages and smoothing methods for time series analysis. Online. Online. Available from: https://statisticsbyjim.com/time-series/moving-averages-smoothing/ [ref. 2024-04-16].

Hansen, P. C. and D. P. O'Leary (1993). The use of the l-curve in the regularization of discrete ill-posed problems. *SIAM Journal on Scientific Computing 14*, 1487–1503.

Horenko, I. (2010). Finite element approach to clustering of multidimensional time series. *SIAM J. Sci. Comput 32*(1), 62–83.

MacQueen, J. B. (1967). Some methods for classification and analysis of multivariate observations. In L. M. L. Cam and J. Neyman (Eds.), *Proc. of the fifth Berkeley Symposium on Mathematical Statistics and Probability*, Volume 1, pp. 281–297. University of California Press.

Pfeifer, P. and S. J. Deutsch (1980). A three-stage iterative procedure for space-time modeling. *Technometrics 22*(1), 35–47.

Pospíšil, L., P. Gagliardini, W. Sawyer, and I. Horenko (2018). On a scalable nonparametric denoising of time series signals. *Communications in Applied Mathematics and Computational Science 13*, 107–138.

The Police of the Czech Republic (2024). Crime statistics. Online. Online. Available from: https://kriminalita.policie.cz [ref. 2025-01-07].