Computational Social Systems

Computational Modelling of Social Systems
Graz University of Technology

A Statistical Model of Criminal Behaviour

Project Report

SS 2022

Group 1

Graz, July 10, 2022

Contents

1 Introduction

The observation that *crime*, or what is being socially defined and identified as it, follows certain patterns, regularities and therefore seems to be the expression of an underlying *order* has been a guiding theme of sociological inquiry since the early days of the discipline (see for example Durkheim 1992 or Merton 1938) that later inspired and influenced the field of Criminology. With a multitude of theoretical constructs trying to explain the empirical distribution of certain forms crimes, the overreaching empirical consensus is that criminal occurrences aggregate both spatially and temporally, leading to *hotspots* with typical lifespans and variations due to seasonal, economic or geographical factors. Enabled by comprehensive ways of data collection, processing, and analysis, hotspots can be visualized to represent these dynamics.

By using criminal statistics, investigating the aforementioned regularities and focusing especially on the occurrences of burglaries, Short et al. (2008) propose a model that aims to explain the spatio-temporal patterns found in the formation of hotspots and therefore predict their future dynamics, expansions or disappearances. By taking into account how crimes committed in the past effect the likeliness of a certain target to be victimized repeatedly, they build on the work of Wilson and Kelling (1982) and the so-called *Broken Windows Effect* and it's central assumption that, if a broken window pane is not repaired quickly, all the panes in the house will soon be broken. A single broken window might therefore lower the threshold for others to break another one, starting a dynamic that gradually normalizes deviant behaviour by contributing to the perception of missing social integration and control (that is suggested by the presence of broken windows). Put into policy, their theory suggests that criminal behaviour has to be blighted immediately, legitimizing the rigorous crackdown on even minor offenses in order to stop them from "spreading". Despite its flaws and widespread criticism (see for example Kunz and Singelnstein 2016), the *Broken Windows Theory* was highly influential throughout the 1980s and 90s and a driving force behind the infamous "stop-and-frisk" and "zero-tolerance" practices introduced by police forces in some North American cities (Fagan and Davies, 2001).

Building on these ideas, Short et al. (2008) formalized the idea in their model that preceding occurrences of burglaries encourage criminals to burglarize a property again. Moreover, they take into account that crime affects not just a single property, but an entire neighborhood, making it more likely for neighboring houses to be burglarized when crimes were committed somewhere around them, suggesting a "crime tolerant" area. These assumptions about a raised attractiveness of a house to be burglarized after successful burglaries had happened and the *neighborhood effect* of such an event derive from the empirical observation that burglars tend to return to places that they have already burglarized in the past due to thorough knowledge of the site, familiarity with the inhabitants schedules and information on goods to be stolen. By this, the authors try to create a "quantitative mathematical model that captures the essential dynamics of hotspot formation" (Short et al., 2008, p. 1250), considering the theoretical implications of the "Broken Windows Theory". The focus on burglaries roots in the fact that this form of crime is considered to be "simple", modeling two kinds of agents and their behavior:

- Offenders burglarize houses with a certain probability and in regards to their *attractiveness*. They rest (become inactive) after a successful burglary but return after a *cooldown* phase to look for new targets. If they decide not to burglarize, they move towards a more *attractive* site.

- A set of immobile house agents with a certain *attractiveness* to be burglarized, consisting of a static (initial) and a dynamic part that increases due to events of victimization or near-victimization in the neighborhood and that decreases again over time by a certain rate.

The following report describes our replica of the model suggested by Short et al. (2008) in an attempt to review their findings and contribute genuine ideas to the further development of models that explain the criminal dynamics described. We are following the original lines of investigation to ask how and by what parameters different distributions and aggregations of crime appear and stabilize. Section 2 of this paper will describe the original model specifications and our adoptions. The subsequent chapter will discuss the results obtained by our simulation, compare them to the visualizations of the original paper and ground them in further analytical elaborations. Finally, we will discuss our findings and experiences with the proposed model and talk about its underlying limitations, theoretical flaws and possible adaptions.[1]

2 Model description

In this section, the composition and behaviour of our replica of the Statistical Model of Criminal Behavior by Short et al. (2008) will be explained in further detail. As already mentioned, the model consists of two types of agents: (immobile) *houses* and *burglars* that walk towards attractive targets. The model is instantiated with several model-level parameters: *initial number of active burglars*, *grid size*, *spacing*, *intrinsic attractiveness* of the houses, *attractiveness increase* per burglary event, *neighborhood-effect*, *attractiveness-decay-rate*, as well as *burglary-regeneration-rate*.

[1] All resources to our replica can be found here: `https://github.com/ohdearaugustin/CMSS`

$$B_s(t+\delta t) = \left[B_s(t) + \frac{\eta \ell^2}{z} \Delta B_s(t) \right] (1 - \omega \delta t) + \theta E_s(t),$$

Δ is the discrete spatial Laplacian operator

$$\Delta B_s(t) = \left(\sum_{s' \sim s} B_{s'}(t) - z B_s(t) \right) \Big/ \ell^2.$$

Figure 1: Calculation of dynamic attractiveness (Short et al., 2008, p. 1254)

Our model is located on a 50x50 multi grid with a spacing of 1 in all simulations, meaning that each grid cell accommodates one house. Due to the multi grid, more than one burglar can stand on one grid cell at a time. As in the original model, toroidal wrapping has been disabled, and *Van Neumann* neighborhood applied. In our model, *simultaneous activation* has been applied. Regarding the number of active burglars, the authors stated that "burglars are also generated at *each lattice site* at a rate Γ (p. 1252). To keep the model stable and to avoid highs and lows in burglary behaviour, we decided to only instantiate a particular number of burglars at the beginning, whereas the other burglars are instantiated as *inactive* and given a randomly assigned *regeneration* value (to appear as active after some time steps). Each house has a particular attractiveness, which can be translated into the probability that a burglar chooses to break into the house. The house's attractiveness is composed of an initial intrinsic attractiveness (equal for all houses) and a dynamic component that is adjusted for each house at every time step (see Figure 1).

Following the *broken-windows-effect*, each burglary event increases the attractiveness of a site by θ. The parameter *neighborhood-effects*, ranging from 0 to 1, indicates how strong a burglary effects the neighborhood. The higher this parameter, the higher the influence of the *Broken Windows Effect* on the neighbors of the victimized property (p. 1254). This is being calculated by multiplying the discrete spatial Laplacian operator with the neighborhood-effect η and the grid spacing , divided by the number of neighbors. Subsequently, this term is added to the previous dynamic attractiveness. Furthermore, at each time step, this dynamic term is multiplied by $1-\omega$ to account for a monotonous decay in attractiveness, meaning that if no burglary event happens in this neighborhood for a certain time, the attractiveness returns to its intrinsic level. The formula is depicted in Figure 1 and the parameter descriptions can be seen in Figure 2.

After the attractiveness of all houses is calculated, the burglars choose whether to rob the property on the grid cell they are currently standing on or move towards a more attractive site. The higher the attractiveness, the higher the probability for the burglars to commit a crime. If they decide to break in, they vanish from the grid for a certain amount of time tied to the burglar's regeneration rate Γ. Once their regeneration status is back to 1, they return to the grid cell where they have committed the crime and continue to search for further targets. This simulates them fleeing and then returning

Parameter name	Meaning
ℓ	Grid spacing
δt	Time step
ω	Dynamic attractiveness decay rate
η	Measures neighborhood effects (ranging from 0 to 1)
θ	Increase in attractiveness due to one burglary event
A_s^0	Intrinsic attractiveness of site s
Γ	Rate of burglar generation at each site

Figure 2: Parameters of the criminal model (Short et al., 2008, p. 1255)

$$\overline{B} = \frac{\theta\Gamma}{\omega}, \qquad \overline{n} = \frac{\Gamma\delta t}{1 - e^{-\overline{A}\delta t}}.$$

Figure 3: Formulas of homogeneous equilibrium of attractiveness and burglars (Short et al., 2008, p. 1255)

to active status. Short et al. (2008, p. 1253) explain this mechanism with the tendency of burglars to commit crimes in their surroundings, as journey-to-crime distributions show a decrease in criminal activity with increasing distance traveled to commit the crime. If a burglar chooses not to rob the house on the cell they are on, they evaluate the neighborhood for the most attractive house and move towards it. If the highest attractiveness is shared by more than one house, the new target is chosen at random. This random walk process concludes the burglar activity of a particular time step.

Both the initial dynamic attractiveness and the initial number of active burglars are set to their homogeneous equilibrium attractiveness, which is the rate at which the attractiveness remains fixed. The formula is depicted in Figure 3. The homogeneous equilibrium for burglars is understood as the number of burglars that are instantiated *at each lattice site* (e.g. 0.01097). This value is multiplied by the number of grid cells and the *spacing* parameter, leading to the total number of burglars that are initially placed on the grid.

3 Model analysis

This section examines the visualizations and results of the simulations, for which we re-implemented the model described in the last chapter and recreated the scenarios mentioned in the original paper. We used the suggested parameters on our model and could observe the three distinct pattern described, namely *spatial homogeneity*, *dynamic hotspots*, and *stationary hotspots*. The original paper observed that the difference between the patterns *"lies essentially in the relative amount of stochasticity present for the parameters chosen"* (Short et al., 2008, p. 1257). For instance, a small neighbouring effect (η) directly (negatively) influences the size of the hotspots. The rate of burglar generation at each site (Γ) directly has a positive effect on the number of burglars. The difference between simulations A versus C, as well as B versus D is the decreased neighborhood-effect in the latter combinations (0.2 to 0.03).

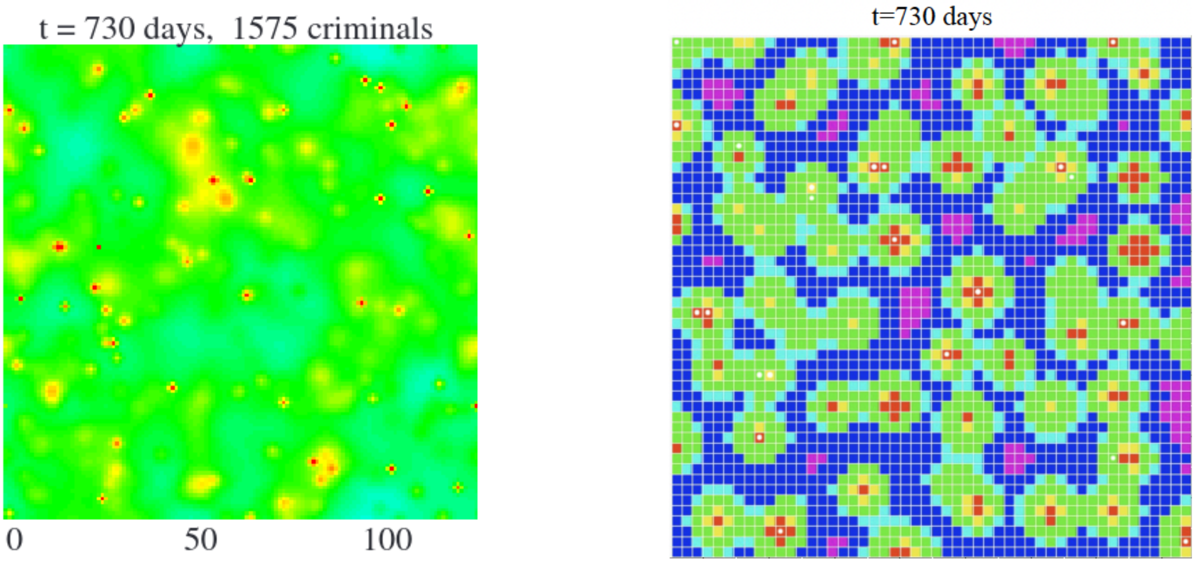
Those effects are observed best by animating the simulation and watching the model iterate over the steps. A more detailed figure can be found in the appendix, where we reproduced all four variations of the model. The simulation was conduced with 730 steps, which is equivalent to 2 years. The patterns will be explained briefly in the sections below. For a better understanding, Table 1 illustrates the color scheme for the model visualizations. Similar to the original, the color scheme has been approximated from the equilibrium attractiveness times 2.

color	attractiveness
red	a $>$ 0.5
yellow	0.3 $<$ a $<$ 0.5
green	0.1 $<$ a $<$ 0.3
light blue	0.075 $<$ a $<$ 0.1
dark blue	0.04 $<$ a $<$ 0.075
violet	0.04 $>$ a

Table 1: Color scheme for the model

3.1 Spatial homogeneity

Spatial homogeneity is characterized by a constant level of attractiveness of houses throughout the simulation. In this simulation, a large number of criminals is present on the grid (1575 in the original and 271 in the replica). Local increases due to burglaries disappear quickly and no distinct hotspots can be observed. Our replica exhibits no extensive hotspot formation. However, whereas in the beginning of the simulations the original and the replica look alike (see appendix), as they both start from their neutral equilibrium attractiveness, at step 730 blue and purple areas are visible in the replica. This can be explained by the *attractiveness-decay-rate* that is affecting all houses at every step and the fact that some low-activity areas tend to fall below this threshold.



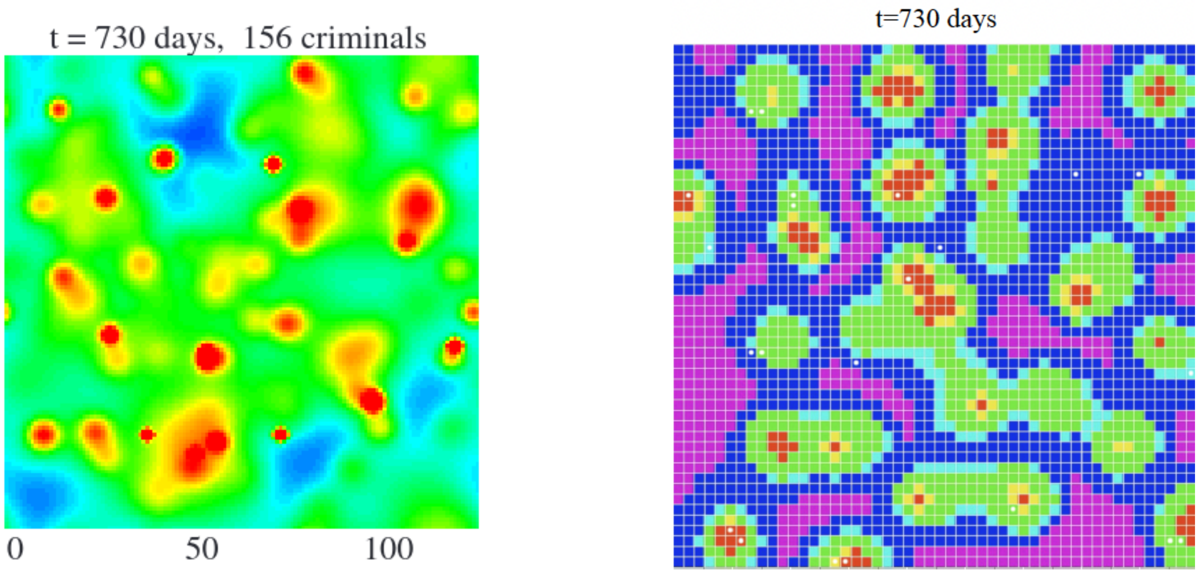
(a) original

(b) recreated

Figure 4: Simulation A: Spatial homogeneity ($\eta = 0.2, \theta = 0.56, \Gamma = 0.019$)

3.2 Dynamic hotspots

In this scenario, localized hotspots form and remain for varying lengths of time. In this simulation, less criminals are present on the grid compared to simulation A (156 in the original and 27 in the replica). The location of the hotspots is constantly changing (see appendix).



(a) original

(b) recreated

Figure 5: Simulation B: dynamic hotspots ($\eta = 0.2, \theta = 5.6, \Gamma = 0.002$)

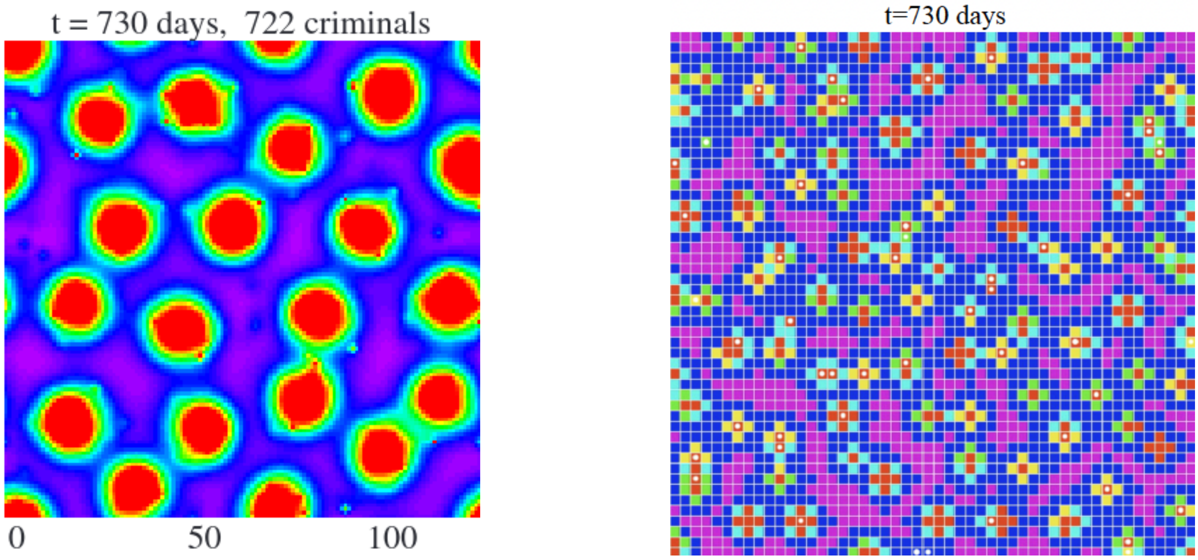
3.3 Stationary Hotspots

As the term *stationary* implies, hotspots in this scenario remain constant over time. Similar to simulation A, the number of burglars on the grid is high. Hotspots are often surrounded by areas of low attractiveness. Their size depends on the chosen parameters, especially the neighborhood effects. In simulations C and D, (η) has been decreased from 0.2 to 0.03. From the visualizations it is noticeable that the hotspot size is significantly smaller compared to scenarios A and B.

According to Short et al. (2008), the number of criminals on the grid has the most influence on whether hotspots are dynamic or stationary. Based on the original model, B and D are dynamic, and A and C stationary, while A is also spatially homogeneous. So far, we mainly discussed the model visualizations. The subsequent section therefore aims to quantify these assumptions.

3.4 Descriptive statistics

To quantify our model behavior, we started off by comparing the simulation's means, standard deviations and medians of burglaries per house normalized by the amount of burglars on the grid as depicted in Table 2. The statistics are based on a batch run



(a) original

(b) recreated

Figure 6: Simulation C: stationary hotspots ($\eta = 0.03, \theta = 0.56, \Gamma = 0.019$)

with 5 iterations and 730 steps. Regardless of the simulation, more than half of all houses have never been robbed. Thus, the means are slightly above 0, with the means in simulations B and D being higher than those in A and C. The medians in all simulations are equal to 0. Due to the long tails in simulations A and C, the respective standard deviations of approximately 0.19-0.2 are noticeably higher than those from simulations B and D. This is also depicted in the probability density function in Figure7.

simulation	burglars	mean	std	median
A	271	0.04930	0.18620	0
B	27	0.05362	0.08957	0
C	271	04950	0.20143	0
D	27	0.05395	0.13282	0

Table 2: Mean, std, and median of burglaries per house (730 steps) normalized by number of burglars

We continued by investigating the number of burglaries that occurred at each house. We thereby noticed that in simulations A and C, where a large number of burglars was present and stationary hotspots were expected, long tails were noticeable. In 730 steps, certain houses were burgled exceptionally often, some of them more than 500 times. In Figure 8 it is noticeable that, despite the amount of burglars in simulations A and C being tenfold, the number of houses being burgled 0-7 times is quite similar. The box plot in Figure 9 depicts the number of burglaries per house in the different simulations normalized by the respective number of burglars present on the grid. Despite

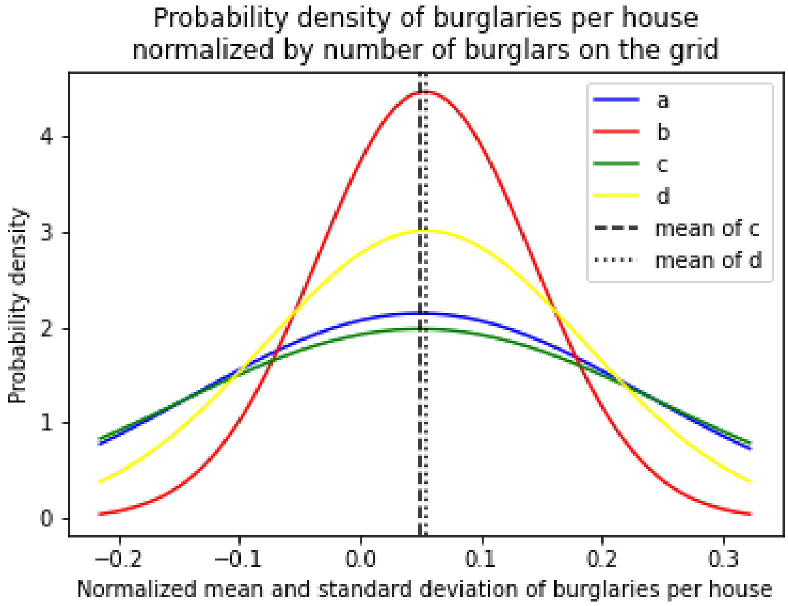


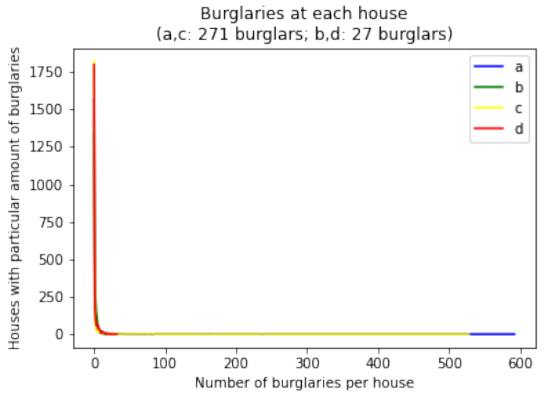
Figure 7: Probability density function of burglaries per house

the normalization, in simulations A and C, where 271 burglars were present on the grid, certain houses have been targeted more than twice as often as in simulations B and D, where only 27 burglars were present. This indicates that more burglars lead to stationary hotspots, whereas less burglars lead to more dispersion of targeted houses, and thus more spatial dynamics in the occurrence of crime events.

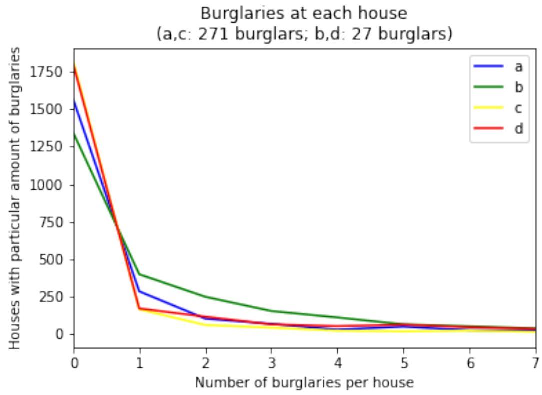
3.5 Hotspot Analysis

For our analysis, a hotspot was defined as a house with an attractiveness over 0.5. This is derived from the equilibrium attractiveness times 2, similar to the original (p. 1257). For reasons of simplicity, we did not consider the size of distinct clusters with a shared dynamic attractiveness above this threshold.

Figure 10 shows the mean total number of hotspots after 730 steps and 5 iterations. Subfigure (a) indicates that many hotspots are created in scenario A (mean 33403.2) and C (mean 33524.8, highest of all simulations). This is an interesting finding, as C depicts stationary hotspots in its visual analysis. This supports the assumption of a spatially homogeneous distribution of hotspots, as a high amount of hotspots are created that then disappear rather than consolidate. Nonetheless, it has to be highlighted that in simulations A and C, 271 as opposed to 27 burglars in B and D have been on the grid. Furthermore, we can see the difference between scenarios B and D towards scenario A



(a) Burglaries unscaled



(b) Burglaries scaled

Figure 8: Number of burglaries per house

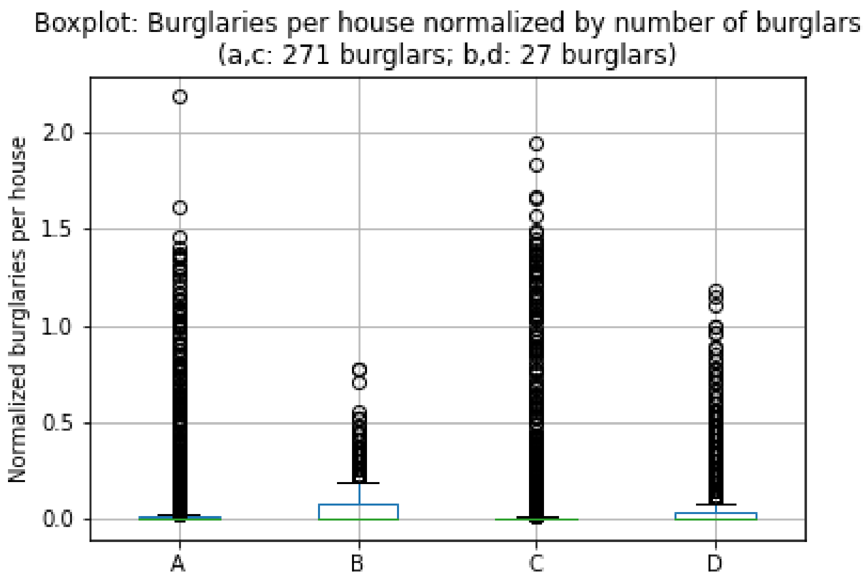
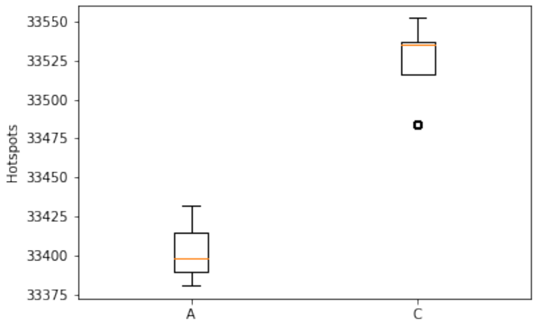
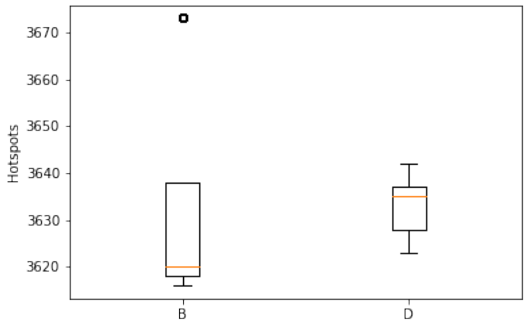


Figure 9: Box plot of normalized burglaries per house



(a) Model A vs C



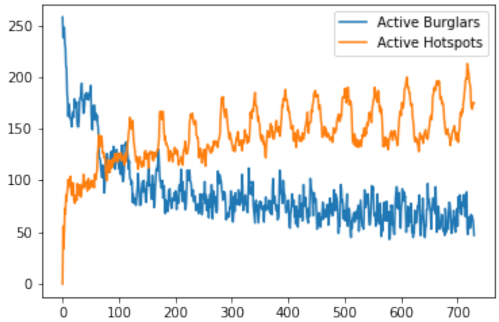
(b) Model B vs D

Figure 10: Total number of hotspots

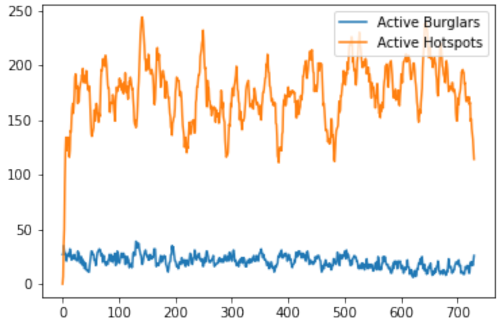
and C in subfigure (b). B and D show dynamic hotspots and have a similar mean (B: 3621.4, D: 3633.0). In an attempt to explain the difference between scenarios B and D towards A and C it could be argued that hotspots in their dynamic appearance often are smaller than compared to the stationary hotspots with a larger spatial expansion and spatial homogeneity.

Figure 11 compares active hotspots and active burglars over time. It is notable that scenarios A and C start with a high initial number of active burglars. Furthermore, hotspot formation is slower in scenario A compared to all other simulations, with a maximum of active hotspots at a time at around 200. This supports the assumption of spatial homogeneity. Visually, we can see that a spike in active burglars is followed by a spike in active hotspots.

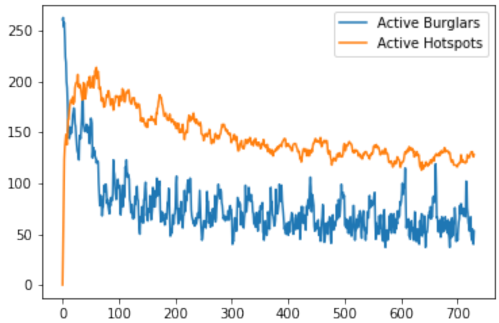
For B and D we can see that they start with a lower initial number of active burglars. B shows the highest peak of active hotspots at around 233. This could also be explained by the dynamic effect of the hotspots. Advancing to model C, we see a fast creation of many hotspots in the beginning of the simulation and a decline over time. This can be typical for stationary hotspots, as these do not wander over the grid and have a tendency to aggregate on certain fields without *spreading* towards other locations. Scenario C peaks at an active hotspot rate of 210. The last scenario D has the lowest maximal active hotspots, with a peak at around 170. Moreover, it is noticeable that by comparing simulations A and B, as well as B and D the number of active hotspots fluctuates greater where the parameter *neighborhood-effect* is larger (A and B), compared to those where it is smaller (B and C).



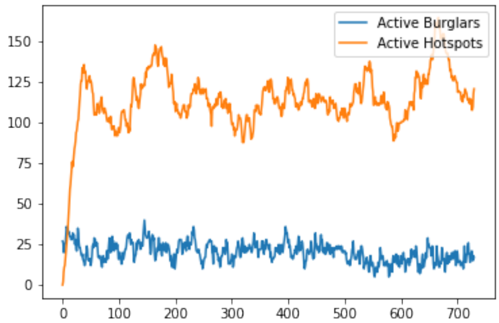
(a) Scenario A: Burglars vs Hotspots



(b) Scenario B: Burglars vs Hotspots



(c) Scenario C: Burglars vs Hotspots



(d) Scenario D: Burglars vs Hotspots

Figure 11: Active burglars and hotspots over time

4 Conclusions

The model presented here aims to explain the regularities of burglaries within urban areas and the formation of crime-hotspots. As with every formalization of models, simplifications are necessary and not all parameters of reality can be represented properly. Before addressing our findings, we would therefore like to discuss some central limitations and theoretical restrictions to the model replicated here.

- The premises of the *Broken Windows Theory* presented are flawed insofar as they focus on symptoms rather than causes of crime and neglect the systematic discrimination and social disintegration of inhabitants of certain districts, the socially embedded recognition of what classifies as a crime as well as people's willingness to cooperate with police forces and even report burglaries. Crime statistics therefore often fail to correctly represent actual occurrences of delinquent behavior since they only depict such cases that were recognized by police forces.

- Furthermore, while the theory's original explanation of lowering the threshold for smashing windows once has been broken, intuitively appears to be reproducible for some crimes, we want to question if this applies to all categories of delinquency, especially the case of burglaries described here. Since burglarized houses can not simply be identified as such and have no detectable or measurable *attractiveness* for bystanders, it is questionable whether and how they would contribute to the public perception of a "crime tolerant area" that is assumed here.

- The uniform distribution of house's attractiveness in the model does not reflect reality and neglects the important aspect of *barriers* that occur when burglarizing a property: Surveillance and home security systems, the structure of the neighborhood (density of houses) or presence of police forces are important considerations that might prohibit burglaries. Especially an increase in police presence might be a direct result of previous burglaries and therefore hinder a property increasing in *attractiveness*. The role of law enforcement has been discussed by some of the original authors as an addition to the model in a later paper (see Jones et al. 2010).

- The assumption that knowledge about how to burglar a certain house might increase its risk to be targeted again neither takes into account the inhabitant's reaction to a previous victimization nor finite resources (goods to be stolen) within houses to be burglarized. With the results of the simulation above, it seems to be unlikely that a single house will be burglarized over 500 times, simply because there will be nothing left to steal. An upper limit of possible burglaries per house would seem as a reasonable consideration to further develop the model.

- We might also ask if a single type of criminal that repeatedly burglars represents the wide array of reasons to commit such a crime, from organized and professional groups to individuals making use of a specific opportunity.

One fundamental difference between the visualization of our replicas as opposed to the original is that decreased neighborhood effects result in smaller hotspot sizes. Unfortunately, as the measure was too complex to implement in the scope of this project, we were not able to quantify the effect of neighborhood-effect on the hotspot size. It was, however, noticeable from the burglars versus hotspots simulations that, ceteris paribus, in scenarios A compared to C, as well as B compared to D, the number of active hotspots fluctuates greater where the neighborhood-effect is larger (A and B), compared to those where it is smaller (B and C). Since the original paper fundamentally builds on the idea that crime is likely to *spread* and that the strength of this assumption is represented by the *neighborhood-effect* in the model, we would suggest further research towards the quantification and analysis towards the characteristics of this parameter.

Regarding the dynamics of the hotspot creation, we noticed that in simulations A and C, where a large amount of burglars was present, the number of burglaries per house exhibited long tails. Certain houses were robbed particularly often. This was prevalent even after the normalization process, indicating that burglaries are rather concentrated on certain houses and thus more stationary than those in simulations B and D that exhibited no long tails. This indicates more dispersion and dynamism regarding the targeted houses. Yet, as already mentioned, the effect observed here raises questions about a necessary adoption of the model proposed that includes a central limit of maximum burglaries per house to better represent reality.

References

E. Durkheim. *Über soziale Arbeitsteilung : Studie über die Organisation höherer Gesellschaften*. Suhrkamp-Taschenbuch Wissenschaft 1005. 1. aufl.. edition, 1992.

J. Fagan and G. Davies. Street stops and broken windows: Terry, race and disorder in new york city. *Fordham Urban Law Journal*, 28, 02 2001.

P. A. Jones, P. J. Brantingham, and L. R. Chayes. Statistical models of criminal behaviour: The effects of law enforcement actions. *Mathematical Models and Methods in Applied Sciences*, 20(1):1397–1423, 2010.

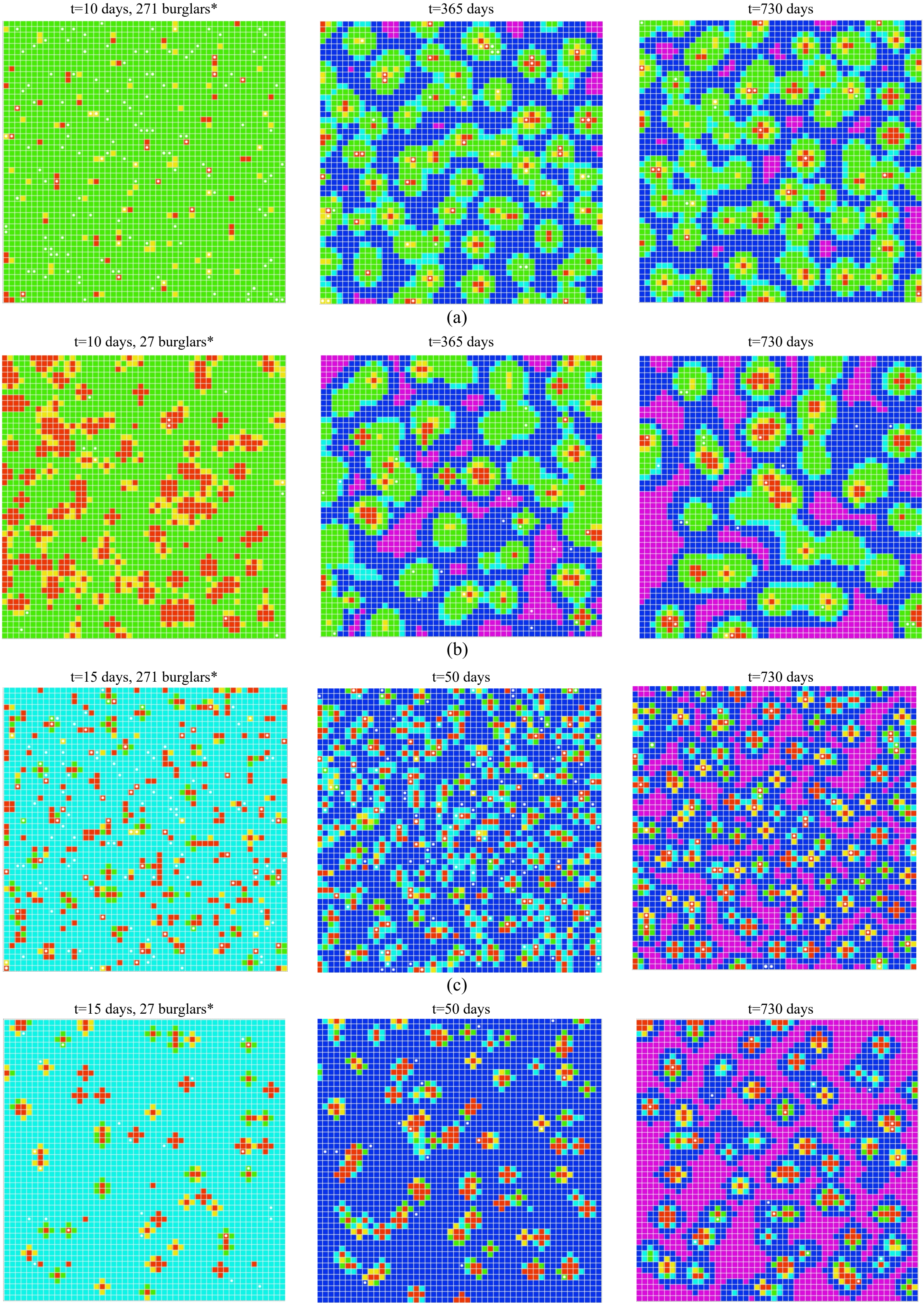
K.-L. Kunz and T. Singelnstein. *Kriminologie*. Haupt Verlag, Bern, 7., grundlegend überarbeitete auflage edition, 2016.

R. K. Merton. Social structure and anomie. *American Sociological Review*, 3(5):672–682, 1938.

M. Short, M. Dorsogna, V. Pasour, G. Tita, P. Brantingham, A. Bertozzi, and L. Chayes. A statistical model of criminal behavior. *Mathematical Models Methods in Applied Sciences - M3AS*, 18, 08 2008.

J. Wilson and G. L. Kelling. Broken windows: The police and neighbourhood safety. *The Atlantic Monthly*, 249:29–38, 1982.

Replica of the original simulations



* The number of initial burglars is lower due to the smaller 50x50 grid (as opposed to 128x128).