Stochastic cellular automaton model of culture formation

Klausen, Frederik Ravn; Lauritsen, Asbjørn Bækgaard

Published in: Physical Review E

DOI: 10.1103/PhysRevE.108.054307

Publication date: 2023

Document version: Early version, also known as pre-print

Stochastic cellular automaton model of culture formation

Frederik Ravn Klausen
QMATH, University of Copenhagen, Universitetsparken 5, Copenhagen Ø, Denmark

Asbjørn Bækgaard Lauritsen†
Institute of Science and Technology Austria, Am Campus 1, 3400 Klosterneuburg, Austria
(Dated: November 9, 2023)

We introduce a stochastic cellular automaton as a model for culture and border formation. The model can be conceptualized as a game where the expansion rate of cultures is quantified in terms of their area and perimeter in such a way that approximately geometrically round cultures get a competitive advantage. We first analyse the model with periodic boundary conditions, where we study how the model can end up in a fixed state, i.e. freezes. Then we implement the model on the European geography with mountains and rivers. We see how the model reproduces some qualitative features of European culture formation, namely that rivers and mountains are more frequently borders between cultures, mountainous regions tend to have higher cultural diversity and the central European plain has less clear cultural borders.

I. INTRODUCTION

The topic of border formation between nations or cultures is complex and it has been subject of interdisciplinary discussion for centuries [1, Chapter 2]. The role of natural boundaries, such as rivers and mountains, in the formation of borders was the basis of many early thoughts on borders [2, p.22-23], but has since been under criticism in the academic literature [3].

Simultaneously, the statistical physics of social dynamics aims at making simple models of complex social phenomena to capture some, but not all aspects of the phenomena. This point of view has with some success been applied to areas such as traffic, networks, economics [4, Chapter 1]. See also the recent collection [5] and the review [6] for more applications of physics to the study of social phenomena.

In particular, models of culture and language inspired by statistical physics have been intensively studied in recent decades. Some of the most studied examples are the voter model [7], the Axelrod model [8] along with many others (see for example the review [4]).

In this paper, we apply the methods of statistical physics to another complex topic of the social sciences: The problem of border formation. We show that a simple model taking only the locations of seas, rivers and mountains as input can reproduce significant features of actual border locations.

The model is a stochastic cellular automaton with coarsening dynamics constructed such that approximately round cultures spread faster. It is inspired by, but yet substantially different from, the agent-based model from [9] which took spreading of information as a starting point. Instead, the inspiration for our model stems from ideas in popular culture about the role of geography and military power in border formation [10]. For instance, the popular strategic games Risk [11] and Civilization [12] which involve using armies to conquer territory.

Thus, the model becomes a concrete mechanisation of 19th century naturalistic thoughts on borders where power and natural geography played a central role (although a major difference is that we present a probabilistic rather than deterministic point of view).

The failure of naturalistic models (such as the one considered in this paper) in describing more than just some overall probabilistic correlations may well be used as yet another argument against the 19th century naturalistic point of view that was implicitly build to fuel contemporary imperialistic agendas [1, Chapter 3]. With these objections in mind we nevertheless construct a simple model for borders that take only natural boundaries and power into account. In particular, we ignore the myriad of other factors such as climate, diseases, individuals, cities, trade routes, taxation, technology, natural disasters, religion, crops, ideologies etc. Adding model features to account for these factors would come at the expense of the simplicity of the model and we believe they may be better dealt with using other approaches.

Over the last 25 years, stochastic cellular automata have been used to model similar spacial forms and reproduce dynamic spatial behaviour [13]. Examples of the approach include urban growth [14, 15] and forest fires [16].

We study our model both with periodic boundary conditions (i.e. on a torus) with no geographical features and on a map of Europe with the geographical features of rivers and mountains. On the European map we compare our model to historical data of border locations in the years 1200-1790 from [17–19]. Although we compare outcomes of our model to historical data, we emphasise that the purpose of the model is not to predict the actual borders of Europe, but rather to demonstrate a (computable, random) process capable of forming reasonable cultural borders. Since the current European borders to some extent reflect the cultural/linguistic boundaries our model is in turn a model of border formation. This is also the reason we focus on the European map as it could be
argued that it is the region where cultural identities and political borders are the most interrelated [20].

II. MODEL: THE GAME OF EUROPE

We consider a $200 \times 200$ pixelated map of either a torus (meaning with periodic boundary conditions) or of Europe, which consists of 40000 cells. A country is a set of cells and we assign each (land) cell to a country, see for example Figure 1. The countries should not necessarily be thought of as real countries, but could just as well be interpreted as cultures, tribes, or even ideas.

One can envision our model as a game of Europe where countries with varying power compete against each other using armies. We emphasise again that we only use this terminology to reflect our inspiration and for clarity of presentation.

Figure 1. Snapshot of the dynamics of the model on the map of Europe at times 200 and 400. Every country is coloured in a random colour. Note the presence of countries of different sizes, that some rivers form clear borders and that some rivers are fully contained in one country.

At each time step, every cell is updated based on interactions between neighbouring countries. To define the interactions we first need to define the powers of the different countries. A naive approach is to define the power of a country to be its area divided by its perimeter. This is motivated by the following:

1. The size of the army a country has is roughly proportional to its population, which again is roughly proportional to its area.
2. This army should be positioned in the border regions - having size the perimeter, as this is where the interaction with neighbouring countries takes place.

The validity of point 1. above can well be criticised. See for instance the reference [17], where the size of an army is argued to be given by the wealth of the corresponding country, or [21], where fertility is considered. The model presented here, though more naive, is much simpler, since we do not need any a priori information of wealth of different areas or some other descriptor of population, the size of an army, etc.

If we define the power of a country to be its area divided by its perimeter the power has dimensions of length, and thus large countries are stronger (more powerful). This leads to an uninteresting dynamics, where one country quickly dominates. For a more interesting dynamics we therefore choose the power of a country to be some dimensionless quantity. Denoting the area of country $C$ by $A_C$ and its perimeter by $\pi_C$ we define its power as

$$\Pi_C := \frac{\sqrt{\alpha_C}}{\pi_C}, \quad (1)$$

There are many ways of constructing such a dimensionless power, e.g. $\alpha_C/\pi_C^2$ is another possible choice. Equation (1) is a natural choice and leads to a more local dynamics as will be explained below.

More precisely a cell is a pair of integers $(i, j)$ and the area of a country is simply defined as the number of cells the country consists of. The perimeter is defined as the number of cells $(i, j)$ owned by the country such that at least one of the cells in the $3 \times 3$ square\footnote{For the cell $(i, j)$ the $3 \times 3$ square around it means the set $\{i-1, i, i+1\} \times \{j-1, j, j+1\}$, i.e. the 9 cells $(i-1, j-1), (i-1, j), (i-1, j+1), (i, j-1), (i, j), (i, j+1), (i+1, j-1), (i+1, j), (i+1, j+1)$.} around the cell $(i, j)$ belongs to a different country.

For the “battle” of a cell $(i, j)$ we need to define the neighbours of $(i, j)$. All cells $(i', j')$ such that the centre of $(i', j')$ is closer to $(i, j)$ than some number $R$ are considered the neighbours of $(i, j)$. See Figure 2 for an illustration. More formally, the relation is that $(i', j')$ is a neighbour of $(i, j)$ if $(i' - i)^2 + (j' - j)^2 \leq R^2$. Notice also that a cell is a neighbour of itself.

We call $R \in [1, \infty)$ the radius of influence. Countries occupying more of the neighbouring cells are stronger, as they have more of their army in the neighbourhood. Denote by $N_C(i, j)$ the number of neighbours of cell $(i, j)$ of country $C$. The local power of the country $C$ is its (global) power times its number of neighbourhood cells.

Figure 2. The radius of influence determines which cells a cell influences. The blue (darker) areas are the neighbourhoods corresponding to radius of influence $R = 1$ and $R = 4$. 

$\Pi_C := \frac{\sqrt{\alpha_C}}{\pi_C}$.  

(1)
Finally, we add some randomness to the dynamics. This is controlled by the fluctuation $p \in [0,1]$. Each country’s power at the cell $(i, j)$ is multiplied by independent fluctuations uniformly in $(1 - p, 1 + p)$. We denote the (random) fluctuation of country $C$ in cell $(i, j)$ by $\Phi_C(i,j)$. (All $\Phi_C(i,j)$ are independent.) In total, the local power of country $C$ in the battle of cell $(i, j)$ is given by

$$
\Pi_{loc}^C(i,j) := \Pi_C \times N_C(i,j) \times \Phi_C(i,j).
$$

The country with the highest local power at cell $(i, j)$ wins the battle and conquers the cell $(i, j)$ (or, defends, if the country already occupied cell $(i, j)$).  

The dynamics is computed at all cells simultaneously. After each time step the map is updated accordingly, and the process repeats.

We can now describe how the choice of the power in Equation (1) gives a more local dynamics. The other relevant choice of the power $\Pi_C$ is the square of the choice in Equation (1). If a cell has $N_A$ neighbours of country $A$ and $N_B$ of country $B$ the local powers (without the fluctuation) are $\Pi_A N_A$ and $\Pi_B N_B$, whereas for the other choice of the power it would be $\Pi^2_A N_A$ and $\Pi^2_B N_B$. For the second choice the global powers $\Pi_A$ and $\Pi_B$ matter more for the battle of some cell, while for the first (Equation (1)) more weight is put on the local effect of how many neighbours $N_A$ and $N_B$ the cell has of the different competing countries.

On the map of Europe, the above-described dynamics is modified by geographic parameters. Further, we have to describe boundary conditions. These geographic parameters and the boundary conditions are explained in Section IV. On the torus (meaning the $200 \times 200$ cell grid has periodic boundary conditions) there are neither geographic features nor boundary conditions and it is thereby a model with no geography.

### III. FREEZING TRANSITION ON THE TORUS

The model presented above is a Markov chain on the set of configurations. As such, any realisation of the dynamics will for large enough time end up in some irreducible recurrence class [22, Theorem 1.40].

Clearly, any configuration of only one country occupying all cells is an irreducible recurrence class (of just one configuration) and one can ask the question if these are the only ones.

To investigate this we define a configuration to be freezing if it is a recurrence class (of just one configuration, hence in particular irreducible). That is, the probability of staying in a freezing configuration is 1 and as such the dynamics “freezes” if it reaches such a configuration. In particular, the one-country configurations are freezing.

For small fluctuations $p$, there may be other freezing configurations, but for $p$ large there are none. Even, for large enough $p$ ($p = 1$ is trivially large enough) any other configuration is transient, meaning that, with probability 1, the dynamics eventually ends up in a one-country configuration. Thus there exists a smallest $p$, the critical $p_c$, such that for any $p > p_c$ the only irreducible recurrence classes are the one-country configurations.

A necessary condition for the only irreducible recurrence classes being the one-country configurations is that no other configuration is freezing. We calculate here the minimal value $p_{file}$ for a tiling configuration (meaning all countries are $k \times k$ squares for a $k > 2R$, see Figure 3) to be non-freezing. Trivially then $p_c \geq p_{file}$.

**Figure 3. Schematic of the battle of a corner cell (●) in a tiling configuration with the defending country (top right) in light blue and a strongest attacking country (top left) in light red. The dark red shaded region (left darker region) is the attacking (red) cells in the neighbourhood. Note that the neighbourhood (see also Figure 2) can be decomposed into the origin (meaning the corner cell) and 4 regions congruent to the dark red shaded region (left darker region). Hence $N_R = 4N_A + 1$, with $N_R$ denoting the number of cells in a neighbourhood of radius $R$ and $N_A$ denoting the number of attacking cells of a strongest attacking country. The defending country has exactly $R + 1$ cells more than a strongest attacking one (say in the central column (in dark blue)) so $N_D = N_A + R + 1 = \frac{N_R}{4} + R + \frac{3}{4}$.**

It is a simple calculation to see that for a tiling configuration the cells that are easiest to conquer are the corners. All countries have the same power (they all have the same shape) and thus we need

$$
(1 - p)N_D < (1 + p)N_A
$$

for such a tiling configuration to be non-freezing. Here $N_D$ ($N_A$) denotes the number of neighbours of the corner of the defending (respectively strongest attacking) country and we suppress in the notation the dependence on


\[ 2 \] One may argue that technically in order for the model to be a cellular automaton the dynamics has to be local and since the countries might be arbitrarily large the computation of power and perimeter may depend on arbitrarily many cells which contradicts locality. Since we know that our countries stay local we still call the model a cellular automaton.

\[ 3 \] We recall that a set of configurations being a recurrence class means that the dynamics will stay in the class. The set is irreducible if any configuration in the class can be reached from any other configuration, see [22, Chapter 1].
the cell \((i, j)\). A counting argument (see Figure 3) finds that \(N_R = \frac{N_R}{2} + R + \frac{3}{4}\) and \(N_A = \frac{N_A}{2} - \frac{1}{4}\) for integer \(R\), where \(N_R\) denotes the number of cells in the neighbourhood corresponding to radius of influence \(R\) (see Figure 2). Thus, Equation (2) gives
\[
p_{\text{tile}} = \frac{2(R + 1)}{N_R + 2R + 1}.
\]
For large \(R\) we have \(N_R \approx \pi R^2\) and so \(p_{\text{tile}} \approx 2/\pi R\). For small \(R\) the values of \(p_{\text{tile}}\) are given in Table I.

Table I. Minimal value \(p_{\text{tile}}\) for a tiling configuration to be non-freezing. We see that \(p_{\text{tile}}\) decreases in \(R\) meaning that for larger \(R\) less fluctuation is needed for the tiling configuration to be non-freezing.

<table>
<thead>
<tr>
<th>(R)</th>
<th>(p_{\text{tile}})</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>0.500</td>
</tr>
<tr>
<td>2</td>
<td>0.333</td>
</tr>
<tr>
<td>3</td>
<td>0.222</td>
</tr>
<tr>
<td>4</td>
<td>0.172</td>
</tr>
<tr>
<td>5</td>
<td>0.130</td>
</tr>
<tr>
<td>6</td>
<td>0.111</td>
</tr>
<tr>
<td>7</td>
<td>0.0976</td>
</tr>
<tr>
<td>8</td>
<td>0.084</td>
</tr>
</tbody>
</table>

Finally, an empirical observation is that for large \(R\) \((R = 6, 7, 8)\) the shapes of the countries in the simulations turn out to be approximately hexagonal after many time steps. A configuration of hexagonally shaped countries is however non-freezing for \(p\) smaller than \(p_{\text{tile}}\), but it does signify a metastable configuration. For the value \(R = 4\) used in most of the simulations however, we do not see such hexagonal shapes. Moreover, countries of different sizes coexist, (as we also see on the European map in Figure 1).

IV. GEOGRAPHICAL FEATURES

We now describe the geographical features of Europe taken into account in the model. This differs from the work [9], where only the coastline is incorporated. The parameters of the model are summarised in Table II.

Coasts and boundary conditions: We implement the sea as cells that cannot be conquered. Further, if a cell is on the perimeter of a country only because it borders the sea (meaning its 3 \(\times\) 3 square (see Footnote 1 on page 2) contains only sea and its own country) we add to the perimeter the sea-border parameter \(s\in [0,1]\) instead of 1.

The parameter \(s\) models how much easier it is to defend a country which has the sea instead of neighbours. In all of the following we set \(s = 0.5\) we therefore do not include it as a parameter. Similarly, for the boundaries of the map we choose the same parameter \(s\) - our main goal with penalising the boundaries (which are beyond the Ural, Kaukasus and Sahara) is to make sure that the boundary conditions do not influence the centre of the map.

Rivers are like the sea implemented as cells that cannot be conquered and give the same effect to the perimeter as the sea, through the parameter \(s\) (that we set to 0.5). Additionally, if a country occupies both sides of a river, it gets a bonus to its area through the parameter \(A_r\ge 0\) and the river no longer counts as a border. Technically this is implemented as follows. Suppose cell \((i, j)\) is part of the river. If all cells in the 3 \(\times\) 3 square around cell \((i, j)\) (see Footnote 1 on page 2) are either part of the river, or belong to the same country, then that country gets a bonus to its area of \(A_r\).

The parameter \(A_r\) is supposed to capture the effect that rivers work as trade routes as well as giving benefits to fertility and infrastructure, as many civilisations arose around rivers [23, p.1]. The location of the implemented rivers is seen in Figure 4.

Mountains are implemented by a variable \(m(i, j)\in [0,1]\) at each cell \((i, j)\) indicating how mountainous the cell is. Mountains have two effects. The first effect is that when calculating the local powers at cell \((i, j)\) the country already owning the cell gets a bonus parametrised by the parameter \(D_m > 1\), i.e. for the defending country \(D\) its local power is
\[
\tilde{\Pi}^{\text{loc}}(i, j) = [(1 - m(i, j)) + D_m m(i, j)] \times \Pi^{\text{loc}}(i, j).
\]
Thus a “fully mountainous” region with \(m(i, j) = 1\) gets the defensive bonus \(D_m\).

The second effect is that the perimeter contribution of cell \((i, j)\) is weighted by
\[
\pi_w(i, j) = (1 - m(i, j)) + P_m m(i, j)
\]
where \(P_m \in [0,1]\) is a parameter. Here a “fully mountainous” cell with \(m(i, j) = 1\) contributes only \(P_m\) to the perimeter and a “completely flat” cell with \(m(i, j) = 0\) contributes 1 to the perimeter. For a country \(C\) a cell \((i, j)\) in \(C\) is in the perimeter of \(C\) if at least one of the cells in the 3 \(\times\) 3 square around the cell \((i, j)\) belongs to a different country (see Footnote 1 on page 2). The total perimeter \(\pi_C\) is then calculated as the sum of \(\pi_w(i, j)\) for all cells in the perimeter of \(C\).

The parameters \(D_m, P_m\) model how much easier it is to defend mountainous territory. The parameter \(D_m\) models a local effect - it is easier to defend, while \(P_m\) models a global effect - the army of the country may be stationed elsewhere. The values of \(m\) are plotted in Figure 4.

V. METHODS

In this section, we briefly describe the sources of data and how we prepare the data. The mountainous
Simulations on the torus: To gauge the effect of the parameters $p, R$ on the torus we perform simulations for various values of $p$ and $R$. As a starting point for the investigations of the parameters we choose $(p, R) = (0.2, 4)$, which were found to give reasonable dynamics on the European map. Interestingly, the choice $(p, R) = (0.2, 4)$ is just above the parameters of freezing for the torus (see Section III).

Fluctuations: In Figures 5(a) and 5(c) we vary the fluctuations and plot the average country size as a function of time respectively the fluctuation. In Figure 5(c) we see that at time step number 1000 the average country size has a maximum (in $p$) around $p = 0.2$. In particular, the model is non-monotone in the fluctuation $p$ and as such, $p$ cannot be interpreted as an effective temperature of the model.

Radius: For completeness, we also study the effect of the radius. These findings are shown in Figure 10 in the Appendix. We see that the average size of countries is monotone increasing in the radius of influence $R$.

Simulations on the map of Europe: In this section, we vary the parameters to determine their effects in the European geography. As a starting point, we again choose $(p, R) = (0.2, 4)$.

Fixing $R = 4$ the value $p = 0.2$ is close to the value of fastest growth on the torus. On the European map, however, the choice $(p, R) = (0.2, 4)$ is very slow evolving and barely above frozen (see Figure 5(b)). As for the geographical parameters $(A_r, D_m, P_m)$ we investigate their effects in the Appendix and find that reasonable choices are $(A_r, D_m, P_m) = (8, 2, 0.5)$. This leads us to the standard choice of values $(p, R, A_r, D_m, P_m) = (0.2, 4, 8, 2, 0.5)$ reported in Table II that we use for most of our investigations.

Fluctuations: In Figures 5(b) and 5(c) we vary the fluctuation size and plot the average country area as a function of time respectively the fluctuation. As for the torus we see a non-monotone dependence of the average country size in the fluctuation $p$. For small fluctuation $p$ (i.e. below $\simeq 0.8$) larger fluctuation leads to a larger average size at time step 1000, but for large fluctuation sizes the dynamics reverses.

Interestingly, the fluctuation $p$ giving the maximal growth is quite different for the European map compared to that of the torus. This may be understood as follows. On the map of Europe, some small island and peninsula states exist forcing the average country size to be small, even if mainland Europe is split between few large

---

4 The simulations were conducted on a 2010 laptop with a dual-core Intel Core i5-5257U processor and 8 GB RAM for about one year.
countries. This effect explains why (for large times) the average country size on the map of Europe is much smaller than that on the torus (see Figure 5(c)). This additionally explains the difference in which p's give the largest average sizes. Namely, on the European map, the average country size is essentially given by the reciprocal of the number of such island and peninsula states. The dependence of the number of such states on the fluctuation p is a completely different dynamics than that of the number of countries on the torus.

Geographical parameters: In the Appendix, we discuss the effects of the parameters $A_r$, $P_m$ and $D_m$ and show how the parameters $P_m$ and $D_m$ contribute in two very different ways.

Noticeably the parameter $D_m$ is the most influential. The effect of changing the parameter $P_m$ in comparison is much smaller. Finally, the average country size is monotone increasing in the parameter $A_r$.

![Figure 5.](image)

Figure 5. ((a) and (b)) Averaged areas of countries over time for different fluctuations p for simulations on the torus (a) and the map of Europe (b). For the simulations on the torus the parameters are chosen to be $R = 4$ and for the map of Europe $A_r = 8, D_m = 2, P_m = 0.5, R = 4$. (c) Plot of the average areas as a function of the fluctuation parameter p for both simulations on the torus and on the map of Europe at time step 20,100 and 1000. All plotted data are averages over 20 simulations and the shaded regions are between the 5% and 95% quantiles. Notice the very different behaviour of the dynamics on the torus and the map of Europe (for large times) with the fastest evolution taking place for very different values of p. In (c) the fastest evolution corresponds to maximum of the curves shown. In the inserts of (a) and (b) we see that for very short times, the evolutions on the torus and map of Europe are very similar.

VI.1. Findings

We finally present the main findings. Namely, that the model, with appropriate parameters, reproduces some of the qualitative features of the cultural borders in Europe. In particular, we find that mountainous regions have a higher frequency of borders, which can be interpreted as higher cultural diversity in mountains. This effect is also present in historical data.

Locations of borders: In the following, we say that a land cell corresponding to one country is a border if at least one cell in its $3 \times 3$ square (see Footnote 1 on page 2) belongs to a different country. In Figure 6 the frequency for each cell to be a border is shown and in Figure 7 we plot the correlation of the border frequency and the mountainous parameter m.

We see that mountain regions are more frequently borders and that larger areas that are flatter tend to have a very low density of borders. In this way the model reproduces the idea of mountains acting as natural borders, which we can also confirm in historical data in Figure 7 (see also the discussion below). Further, this reflects the higher cultural diversity that is often seen in mountainous areas [32].

Inspecting Figure 6 one sees that the frequency of borders close to rivers are significantly increased (compare also with the snapshot Figure 1), an effect observed in current subnational borders in [33].

Historical comparison: We next compare to the historical data from [17–19]. In Figure 7 we plot the correlation between border frequencies and the mountainous parameter m also for the historical data from [17–19]. The historical borders are of course correlated since the different data points of [17–19] are only separated by 5 year intervals. These correlations, in the form of static borders, are the cause of the border frequency being 1 for many cells with low mountain parameters. However, there is no inherent reason to use one year over another. We thus plot in Figure 7 the average border frequency of the historical data and compare the simulated data to these. As with the simulated data, we see a correlation between border frequencies and how mountainous an area is: More mountainous regions are more often borders. The verifies historically the idea of mountains acting as natural borders as we also see for our simulated data.

In Figure 8 we plot the border frequencies of the historical data from [17–19]. We see that the area that was then the Holy Roman Empire has a very high density of borders. This is because the Holy Roman Empire was not considered a country in the source of the data [17–19]. Instead, all the smaller individual states, usually German, that were part of the Holy Roman Empire were considered (a phenomenon sometimes known as Kleinstaaterei [34]). This digression illustrates the point that our model is not a model for borders between territorial
Figure 6. Heatmap of the frequencies of when cells are borders (colour gradient) at the time steps $t = 100$ (left) and $t = 1000$ (right). The parameters are chosen to be $p = 0.2, A_r = 8, D_m = 2, P_m = 0.5, R = 4$ and the frequencies are computed over an average of 20 iterations. We see that the border frequencies around rivers are significantly increased. Simultaneously, the figure shows the coarsening effect of the model from timestep 100 to timestep 1000. Furthermore, comparing to Figure 4 we see that mountain areas tend to have higher border frequency.

**Contested areas:** Finally, we consider which cells are most contested. For each square we find how often it was conquered. This is then averaged over 20 simulations. The results are shown in Figure 9. We see that the average number of times a cell changed countries is higher on the central European plain than elsewhere, in particular in the mountainous regions in the Alps and Caucasus. This might be connected to the, from one point of view, less clear cultural borders on the central European plain (as can be seen on old historical maps [35, 36], although one should be wary of context of such maps). Together with the above discussion on border location we reach the conclusion that not only are mountain regions more often borders, they are also much more stable borders.
VII. DISCUSSION AND OUTLOOK

We have seen that for suitable choices of the parameters, the model efficiently reproduces many of the features of historical political and cultural borders:

There is a high density of borders in mountain regions and along rivers and clear/stable borders in mountain regions and unclear/unstable borders on the large central European plain. We underline that this is another way (popularised in [10]) of approaching the problem of culture spreading than what was done in [9]. Although the model is designed in such a way borders in mountainous regions and along rivers should occur more often, the results show a (somewhat realistic) probabilistic model capturing the naturalistic approach to border formation exists.

The model provides a simple framework for thinking about how the geography plays a role in border formation, which to some extent (on a statistical level) captures properties of the interplay between geography and borders. Thus, the model could provide some qualitative insight into the influence of geography on the political map of Europe.

As discussed one could imagine adding additional geographical features, but this would come at the expense of simplicity of the model.

Another point of discussion in regards to the model is that we have left out the possibility of new “countries/cultures” to form. With such a possibility one might be able to get a continual dynamics instead of our coarsening dynamics and then study the steady state as was done in [9].

One potential way to incorporate the possibility of new countries forming could entail a probability for fracturing of large countries every time step. However, this would introduce more parameters and thus also come at the expense of simplicity of the model. Additionally, many models of our inspiration in statistical physics (see the review [4]) do not include the possibility of new countries, parties etc. forming. Thus our model could still function as a starting point for such investigations.

ACKNOWLEDGMENTS

Thanks to Kim Sneppen, Svend Krojer, Peter Wildemann, Peter Rasmussen and Kent Bækgaard Lauritsen for discussions and suggestions. FRK acknowledges support from the Villum Foundation for support through the QMATH center of Excellence (Grant No. 10059) and the Villum Young Investigator (Grant No. 25452) programs.

[31] corteva/geocube: 0.3.3 Release (2022).
Appendix: Additional plots

Dependence of country size on radius $R$

In Figure 10 we plot the average size of countries for simulations on the torus for various values of the radius of influence $R$. As one could expect, the average size of countries increases in as $R$ increases. Effectively countries much smaller than $R$ don’t have enough local power to defend themselves and get conquered. This leads to a larger average country size for large $R$.

Figure 10. Averaged areas of countries over time on the torus (i.e. with periodic boundary conditions) for different values of the parameter $R$ with the fluctuation fixed at $p = 0.2$. The results were averaged over 20 iterations and the shaded regions are between the 5% and 95% quantiles. We see clearly how the rate of growth of countries increases with $R$ increases.

Influence of the parameter $A_r$

To study the effect of the parameter $A_r$ we plot in Figure 11 the average country size against time for different values of the parameter $A_r$. We see that countries tend to get larger (and larger quicker) when the parameter $A_r$ is large. This is the behaviour one would expect, as increased $A_r$ gives countries around rivers more military strength to expand and get larger.

Figure 11. Plot of the averaged country size over time for different value of the river area bonus $A_r$. Here the parameters were $p = 0.2$, $R = 4$, $D_m = 2$, $P_m = 0.6$ instead of the normal choice. The results were averaged over 20 simulations and the shaded regions are between the 5% and 95% quantiles. The rate of growth of countries increases when the parameter $A_r$ is increased (due to countries forming around rivers).

Comparison of the two mountain parameters

We investigate whether the two mountain parameters have different effects. To do this we plot the frequency of borders compared to the mountainous parameter $m$. In Figure 12(a) we make a plot as in Figure 7. In Figure 12(b) we bin all cells into 4 bins depending on the value $m$. Judging from Figure 12(a) it looks like only the mountain defence parameter has a significant effect on the model, but in Figure 12(b) we see that the two mountain parameters have different effects.
Figure 12. Distribution of border frequencies for different choices of the parameters $D_m$ and $P_m$. The parameters are chosen to be $p = 0.2, R = 4, A_r = 0$ and averaged over 20 iterations. (a) Plot as in Figure 7. The shaded areas denote the 25 and 75% quantiles. Note that these error bars indicate the variation of the points with a fixed (binned) mountain parameter and not the variations over iterations. (b) Violin plot. The category “flat” corresponds to the 50% land cells with the lowest mountain parameter $m$. The remaining 50% are split evenly into the three remaining groups. Note that the two parameters $D_m$ and $P_m$ are clearly different and that $D_m$ is the more important of the two.