

How Dykes Effect River Network Topology - Insights From Numerical Modelling And Graph Theory

Research Paper For 'ESPM'

Malte Springer

September 25, 2025

springer1@uni-potsdam.de

Abstract

River networks organize mainly according to the underlying topography and then start to restructure the surface through erosion and deposition. This interaction is strongly affected by rapid changes in surface lithology when geological anomalies such as vertical dykes are present. Landscape evolution models (LEMs) can be used to understand the effect of dykes on river network system organization and model a wide range of different conditions such as dyke strength, position, shape and thickness. In this paper I present evidence that vertical dykes stronger than the surrounding material generate topographic highs which in turn cause drainage divide migration and the restructuring of the entire river network system. I also show that this river network reorganization can be expressed by representing rivers as graph objects and computing graph metrics such as the number of nodes, longest path, eccentricity and Shreve stream order. Finally it could be determined that the likelihood of rivers eroding trough and crossing a dyke is positively related to dyke material erodibility and it's nearness to the river sink while being negatively related to the dyke thickness. The results of this paper can also be considered a successful test of the LEM FastScape's ability to accurately model the effect of dyke structures on river network system development.

Keywords: Landscape Evolution Modelling, Graph Theory, River Network Topology, FastScape, Dykes

1 Introduction

Water from precipitation always follows the steepest gradient in the terrain. This tendency leads to the formation of rivers where a sufficient amount of water arrives downhill. Thus, river systems first follow the topographic structure provided by the surface terrain. But rivers also exert erosive forces on the surface material, transport loosened sediment and deposit it further downstream. Through this mechanism, river systems actively restructure the underlying surface and possibly their own spatial organization pattern, for example when drainage divides migrate and streams are captured by other rivers. Multiple factors such as the surface uplift rate, climatic conditions, surface lithology and singular geological events such as earthquakes, damming by volcanic lavas, glacial movement or landslides can moderate or directly impact the way in which rivers and the earth surface interact.

One aspect particularly worth investigating with landscape evolution models (LEMs) is the influence of vertical dykes on river network structures. Dykes are geological phenomena which describe structural anomalies which occur when new material fills vertical cracks in pre-existing horizontally organized layers. These cracks can either be filled by rising magmas or deposited sediments. Consequently dykes can be considered as parts of the surface which are mainly characterized by an abrupt change in material properties. The central metric for landscape evolution modelling is the erodibility (or inverse hardness) of the material. From a modelling perspective, dykes also share some characteristics with other linear features such as dams made of lava, ice or sediment. Damming events such as these were modelled with LEMs in the past using for example the LAPSUS software (Van Gorp et al., 2015; Van Gorp et al., 2014). Modelling of dykes can take inspiration from this research when it comes to computationally representing abrupt linear changes in surface properties. However to my knowledge, no peer-reviewed paper modelling specifically dykes and their influence on river networks has been published. This leads to the guiding research question: *How do vertical dykes influence the topological characteristics of river networks?*

One way to approach river networks is to represent them as graph objects. Graphs are mathematical entities which are comprised of nodes connected by edges. River network graphs use edges for river segments with linear flow patterns and nodes for positions where the flow direction changes or where multiple rivers unite or one river branches into different streams. Thus graphs highlight primarily the topological characteristics of a river network - the connectivity of river segments - rather than a concrete representation of the true spatial shape of the river even though the latter is partially approximated by the graph. Importantly, river networks are a case of directed graphs where each edge has an associated direction from one node to the other. This direction represents the flow direction of the water going from node to node.

The characteristics of river network graphs can be expressed by various graph theoretical metrics. These can be focussed either on the level of the entire graph or on the level of individual nodes. On the graph level, metrics such as the number of nodes, longest path length, diameter and radius describe the scale and shape of the entire graph. On the node level, metrics such as eccentricity and closeness/farness centrality express the centrality (importance) of individual nodes while node degree, closeness and Katz centrality and pagerank metrics capture the connectivity of nodes (Lee et al., 2022). This paper focusses on the number of nodes, longest path, eccentricity and also the Shreve stream order which is a measure for river branching behaviour to detect changes in river network topology.

Modelling of landscape evolution depends crucially on capable software solutions. Each of these has different strengths and weaknesses. One particularly efficient LEM is the software package FastScape which is especially well suited for modelling surface and river network development by means of the stream power law. Thus, FastScape was selected to be the computational backbone for this modelling study.

To summarize, systematic research using LEMs to assess the impact of dykes on the arrangement of river network topology is lacking. To fill this research gap, I will demonstrate how the LEM FastScape can be used to model surface development in areas with linear dykes over the course of 10^7 years. I will also analyse graph representations of the resulting river networks and describe systematic differences related to dyke presence and properties using graph metrics such as node number, eccentricity, Shreve stream order and longest path. Finally the impact of dyke characteristics such as erodibility, thickness and position on the likelihood of river crossings through the dyke is estimated to evaluate what dykes are most likely to generate complete separations of different river network areas.

2 Methods

To answer the research question, I will utilize the landscape evolution model (LEM) FastScape which is available as a module in the Python programming language (Bovy & Lange, 2023). FastScape in Python works within the xarray-simlab environment

which allows for efficient modelling of processes described by (partial) differential equations by storing data in array formats from the xarray library (Bovy et al., 2021).

FastScale includes a range of process-based models to represent different earth surface dynamics, such as erosion, sediment transport, and deposition. In this study, I used the *sediment model*, which accounts for both channel incision and sediment redistribution. Channel erosion is modelled by an extended version of the stream power law (SPL). The basic version of the SPL can mathematically be written as:

$$\frac{\delta h}{\delta t} = U - K_f A^m S^n = 0 \quad (1)$$

where h is the surface elevation, U is the tectonic uplift rate, A is the upstream drainage area (a proxy for water discharge), and S is the local channel slope. The coefficients m and n are positive exponents that describe the relative sensitivity of erosion to drainage area and slope, respectively. Typical values of m range between 0.3 and 0.7, while n is often close to 1. Finally, K_f is the erodibility coefficient, which integrates the influence of lithology, climate, and hydrology on erosion efficiency.

The extended SPL used in FastScale incorporates additional terms to account for sediment transport and deposition, allowing for a more realistic simulation of landscape evolution. By coupling uplift, fluvial incision, and sediment dynamics, the model provides a powerful tool to investigate the interactions between tectonics, climate, and surface processes over geological timescales.

Table 1: Model and simulation parameters

Parameter	Values
Grid Sizes (nx, ny, xl, yl)	200, 100, 40e3, 20e3
Maximum Time	1e8
Time Steps	400
Uplift Rate	1e-4
Flow Slope Exponent	1.0
SPL Area Exponent	0.5
SPL Slope Exponent	1.0
Erodibility Soil	1e-4
Erodibility Bedrock (Background)	1e-5
Erodibility (Dyke)	1e-6 to 1e-4 (depending on simulation run)
Diffusivity Bedrock	0.0
Diffusivity Soil	0.1
Dyke Positions	7500, 10000, 20000, 30000 (depending on simulation run)
Dyke Thicknesses	100, 1000, 2500, 5000 (depending on simulation run)

The model setup follows the parameters as listed in table 1. The model is run until a time of 1e8 years with 400 time steps.

The simulated area is twice as long in the x direction (40000 units) as it is in the y direction (20000 units). The bottom and top boundaries receive 'looped' structure, while the left and right boundaries are 'fixed value'. This setup allows to model a linear dyke feature orthogonal to the x direction and measure the impact of the dyke on river networks who's final sink is on the left or right side of the dyke.

The simulation structure works by assuming different values for dyke position, dyke thickness and dyke erodibility (K_{f-dyke}). For each combination of these variables 100 complete model runs are executed with 100 randomly generated initial topography maps (random values between 0 and 1). These initial noises are stored and used for each of the variable combinations (each first simulation run for every parameter combination is initialized with the same random noise pattern).

The dyke structure is represented by a change in bedrock erodibility where the dyke is situated. The area without dyke has a constant background erodibility of 1e-5 while K_{f-dyke} is varied 10 times in logarithmic space between 1e-6 and 1e-4. Thus the simulation tests for the effect of dykes which are less erodible than the background ($K_{f-dyke} < K_{f-background} = 1e-5$), equal to the background, and more erodible than the background bedrock but less than the soil ($K_{f-background} = 1e-5 < K_{f-dyke} < K_{f-soil} = 1e-4$). The dyke then receives four different positions and thicknesses. The dyke centre is positioned at x-indices 7500, 10000 ($\frac{1}{4}$ of the x axis), 20000 (half of the x axis) and 30000 ($\frac{3}{4}$ of the x axis). The different thicknesses are 100, 1000, 2500 and 5000.

The impact of the dyke on the the river network topology is assessed by first extracting the river network from the final topography of the model simulation and transforming it into a graph. To achieve this, the Python implementation of TopoToolbox is used, which can extract stream networks from array data structures and NetworkX which can build directed graphs from the coordinates

of the stream network provided by TopoToolbox. Here a threshold of 25 upstream array cells was used to limit the detail of the stream network for Topotoolbox. This keeps the detail and complexity of the river network at a reasonable level which also limits the computational intensity.

Afterwards, the NetworkX DiGraph representation of the river network of each simulation run can be used to compute graph metrics which describe the final topography. Here I used the number of nodes and longest path on the graph level and eccentricity and Shreve stream order on the node level. Node eccentricity is the maximum distance of the node to any other node in the graph. Eccentricity is normally defined for undirected graphs. Here it could be computed only along the direction of the river but then it would less well represent the overall shape of the river network. Thus the river is assumed to be undirected only for the computation of eccentricity. The longest path is the maximum distance between nodes only along the direction of all edges. The Shreve stream order is constructed by giving each source node the number 1. Whenever two streams converge their current numbers are added. This way the Shreve stream order measures all contributing sources uphill to a specific river segment. The number of nodes in a graph is a basic measure for graph size/extension. For each of the metrics and simulation setups, histograms of all graphs/nodes over all runs with different noises but same parameters are created to approximate the theoretical distribution of the metric. These histograms are grouped for all rivers that have their sink left/west and right/east side of the dyke. Since these metrics mainly give insight about the extension of a river network system, they can be used to see the influence of dykes on the length and size of the river networks. If the dyke has a significant influence of the structure of the river network systems, then I expect that the distribution of graph metrics left and right of the dyke are different. To detect systematic differences between the histograms beyond just visual inspection, KS distance, Wasserstein-1-metric and Energy Distance are reported for each west/east comparison. KS distance measures the vertical distance between two distributions and gives a value between 0 and 1 which captures if the distributions are the same or not. W1 gives the area difference between the two distributions in the units of the variable. Thus larger values indicate larger distribution differences. Energy Distance is based on pairwise absolute differences between the two distributions and is 0 exactly and only when the distributions are the same.

Finally, a last metric called *Dyke Crossing Index* (DCI) is computed which notes if a dyke was crossed by one of the rivers or if it is never breached. For each simulated river graph a binary indicator Y_{ij} is recorded that equals 1 if the river *crosses* the dyke (its outlet lies west/east of the dyke while at least one source lies on the opposite side), and 0 otherwise. Index i denotes the simulation run and j a graph within that run. The probability of a crossing is related to dyke properties using a population-averaged logistic model fitted with Generalized Estimating Equations (GEE). GEE accounts for correlation among graphs from the same run by using an exchangeable working correlation and robust (sandwich) standard errors. Predictors are z -scored, so effects are per 1-SD.

$$\text{logit}\{\Pr(Y_{ij} = 1)\} = \beta_0 + \beta_1 \log K_{f\text{-dyke}}^{(z)} + \beta_2 \text{thickness}_{ij}^{(z)} + \beta_3 \text{distCentre}_{ij}^{(z)}.$$

Here $\text{logit}(p) = \log\{p/(1-p)\}$. Coefficients are interpreted as marginal (population-averaged) effects: a one-SD increase in predictor k multiplies the odds of a crossing by $\exp(\beta_k)$, i.e. changes the odds by $(\exp(\beta_k) - 1) \times 100\%$.

The Python code used for the entire modelling workflow can be accessed on my github together with additional material: (https://github.com/maltespringer/Landscape_Evolution_Modelling_FastScape_Graph_Theory.git)

3 Results

The simulation and result analysis was performed in a Jupyter Notebook environment and parallelization of tasks was done whenever applicable. The time for $\text{Centres} * \text{Thicknesses} * K_{f\text{-dyke}} * \text{Noises} = 4 * 4 * 10 * 100 = 160 * 100 = 16000$ landscape simulations with 17 parallel workers was ≈ 5.5 hours on an intel core. The time for the computation of the resulting graph structure and metric distribution took about 45 minutes for each metric over all simulation results with the same number of parallel workers.

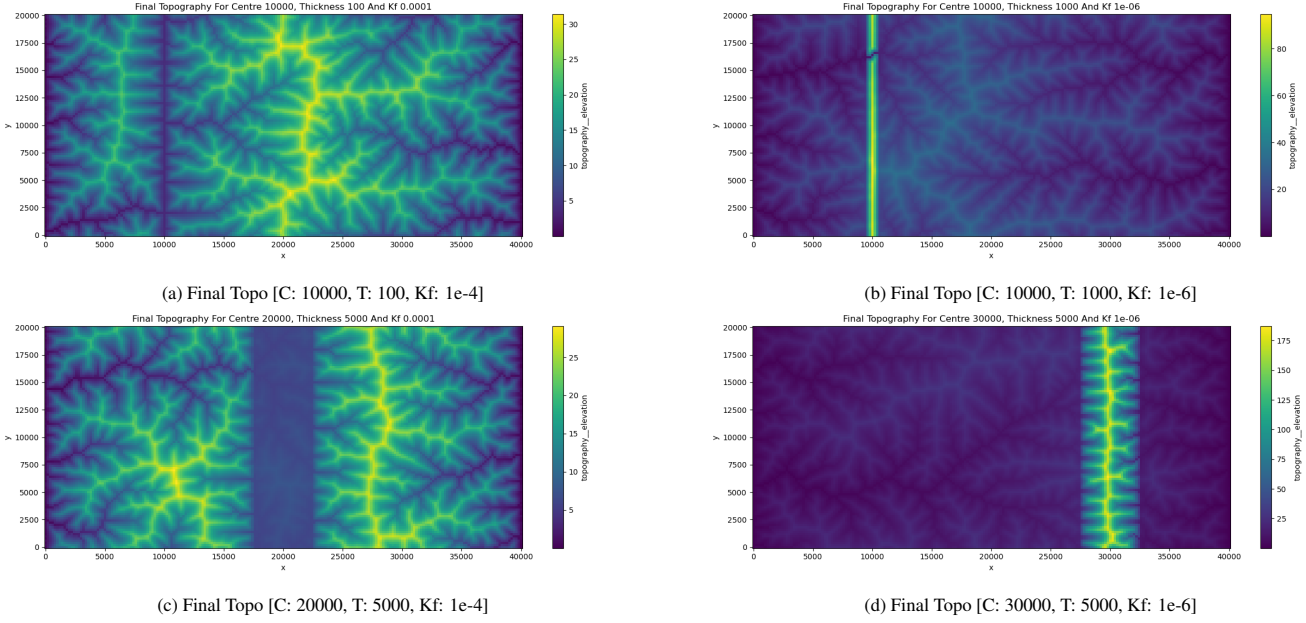


Figure 1: Final Topography for selected simulation runs [C = Dyke centre position, T = Dyke thickness, Kf = Dyke erodibility] over the first random noise field

3.1 Final Topography

The final status of topographical development for three selected cases (figure 1) immediately shows that the resulting topography is vastly different for the various dyke parameter combinations. These parameter combinations are a) a dyke centre of 10000 (more to the left), a thickness of 100 units and K_{f-dyke} of 1e-4, b) a dyke centre of 10000, thickness of 1000 units and K_{f-dyke} of 1e-6, c) a dyke centre directly in the middle of the field at 20000, a thickness of 5000 units and a K_{f-dyke} of 1e-4 and d) a dyke centre more to the right at 30000, a thickness of 5000 units and a K_{f-dyke} of 1e-6.

The essential observation to be made here is that a lower dyke erodibility K_{f-dyke} than the background erodibility causes areas with topographic elevations as in visible in cases b) and d). Assuming equal erodibility, a higher dyke thickness further contributes to higher elevations. Case d) has a dyke thickness five times as high as case b) and reaches an elevation more than 100% higher. Dykes which have a higher erodibility than the background on the other hand generate topographic depressions which are close to 0 meters in elevations for cases a) and c).

3.2 Graph Metrics

The observations made for final simulation topography are reflected in the distribution of river network graph metrics. These distributions computed for the graph metrics from all simulation runs are reported for selected parameter combinations in figures 2 and 3 as histograms in log space for the y-axis since way more graphs and nodes with low metric values were observed. The complete distributions for each of the model runs can be found in the supplementary material.

The main takeaway from the graph metric results is that the distributions of all selected graph metrics are significantly different on the two sides of the simulated fields whenever a dyke stronger than the background material is created that is not directly centred in the middle of the simulated area. The number of nodes is most differently distributed for case d) where a very thick and hard dyke is situated at the right side of the area. Here the graphs with most nodes on the larger western side have up to almost 1200 nodes while the graphs on the right side have only a little more than 200 at maximum. Another strong difference is visible for case b) where a less thick but equally strong dyke is located closer to the west. Here the graphs on the western side have significantly less nodes than the graphs on the eastern side. In case a), where the erodibility of the dyke is higher than the background erodibility, the rivers have more nodes when they run towards the west. Finally, in case c) where the dyke is exactly in the centre of the area, the distributions look basically the same. These visual observations are confirmed especially by the W1 and energy distance which have significantly higher value for cases b) and d), slightly lower ones for case a) and the lowest for case c) where the dyke was in the centre.

The distributions for node eccentricity show a similar pattern. The largest difference in node eccentricity distributions can be

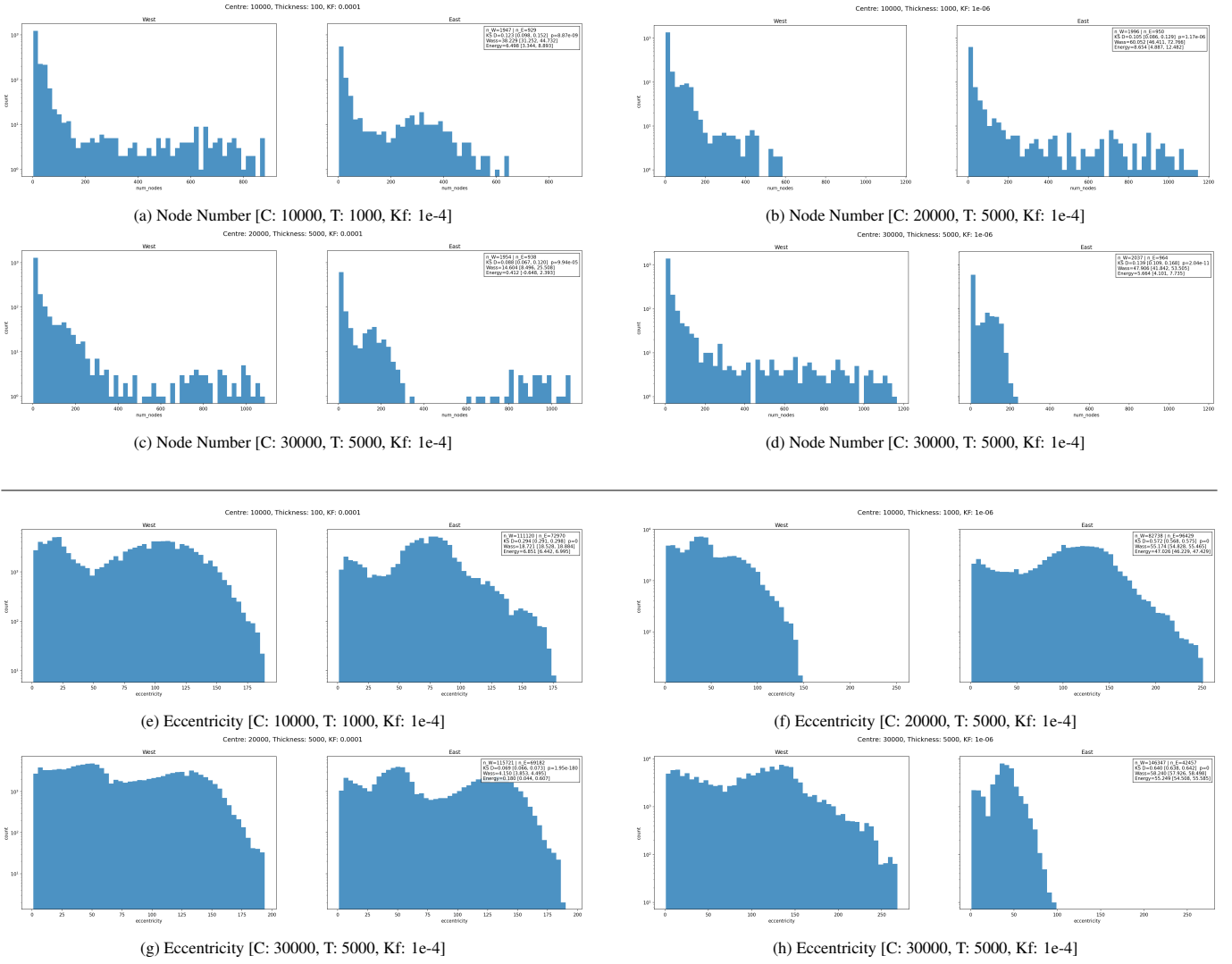


Figure 2: Distributions of selected simulation results [C = Dyke Centre, T = Dyke Thickness, Kf = Dyke Erodibility] with distribution comparison metrics (1)

found in cases b) and d) where a stronger dyke than the background is situated more on the western or eastern side. In both cases W1 and energy distance indicate strongly different distributions and visual inspection confirms that the eccentricity distribution is pushed towards lower values for the side where the dyke is situated implying that the rivers eventually running to this side have shorter maximum path lengths between the nodes which corresponds well with the previous results that in these cases less nodes are present in the first place. Case a) where the dyke is weaker than the background has lower W1 and energy distance scores but still slightly more nodes with higher eccentricities for the side without dyke. Also the western side (with dyke) has a particularly low number of nodes with eccentricity values around 50. Case c) where the dyke is in the middle of the region has very similar distributions on both sides and low scores in the comparison metrics.

The longest path metric has large differences for cases b) and d). Here the maximum path length is about 112 and 70 for the sides with dyke for cases b) and d) respectively while the opposite side has rivers with path lengths of up to 175 nodes. The A weaker dyke on the western side results in longer maximum paths in the west while again average values are less well represented. Case c) with a centralized dyke again results in very similar distributions to both sides. For the longest path metric though all comparison metrics are not very sensitive. Case c) still is considered the most similar on both sides but the metric values for the other three scenarios are very similar indicating that longest path is less effected by dyke characteristics than the previous graph metrics.

The Shreve stream order finally corresponds to the previous metrics. maximum source counts at the river sink do not exceed 25 and 60 sources for cases d) and b) on the side with dyke while the opposite side can accumulate up to 130 sources until the sink is reached. The weaker dyke in case a) is associated with slightly higher source counts at sink position for the side with dyke than for the other even though here the distributions look more similar than for the previous metrics. Case c) is very similar again even

Table 2: Dyke Crossing Index (DCI): GEE logit with exchangeable working correlation and cluster-robust SEs (clusters = runs). Predictors are z -scored, so effects are per 1 SD.

Predictor	Coef	SE	95% CI	OR	95% OR CI	p
Intercept	-2.9856	0.006	[-2.997, -2.974]	0.0505	[0.0499, 0.0511]	< 0.001
$\log K_f - \text{dyke} - z$	0.3101	0.005	[0.300, 0.320]	1.3636	[1.3503, 1.3770]	< 0.001
Thickness $_z$	-0.1857	0.006	[-0.197, -0.175]	0.8306	[0.8214, 0.8398]	< 0.001
Dist. Centre $_z$	0.2832	0.006	[0.271, 0.295]	1.3274	[1.3112, 1.3438]	< 0.001
<i>Model:</i> GEE (Binomial, logit link), exchangeable working correlation, robust covariance.						
<i>Obs.:</i> 475,427 <i>Clusters:</i> 16,000 <i>Mean cluster size:</i> 29.7						

4 Discussion

The results have shown that dykes which are not centred at the middle position of the simulation area change the distribution of graph metrics reflecting the river network characteristics. The explanation for the observed pattern is primarily the impact that a lower erodibility K_f has on the behaviour of the stream power law modelling. At steady state, the condition is that uplift is equal to erosion $U = K_f A^m S^n$. Because uplift U is assumed constant across the landscape, any spatial reduction in K_f requires a compensating increase in $A^m S^n$ to maintain equilibrium. In practice, drainage area A is constrained by the network geometry and cannot grow arbitrarily in the dyke zone. Therefore, the adjustment is primarily achieved through steeper slopes S . This steepening propagates into the surrounding landscape, leading to higher local relief. Over time, this generates elevated topography over and around the dykes compared to more erodible regions, where channels can more effectively lower the landscape.

This increased topography in the dyke zones with reduced erodibility effectively works as an obstacle to rivers which most of the time leads to the establishment of a drainage divide where the dyke is located. Thus, most of the rivers having their final sink on the side where the dyke is located will have their source also on this side which leads to shorter and smaller river networks. The opposite happens on the other side of the dyke where more space is available to form long, extensive river networks. Often times a small topographic high established in the centre of the entire region during the calibration phase of the model will reorganize and migrate towards the dyke zone to finally melt with it (see video in the supplementary material).

Only sometimes very large rivers can establish enough erosive power due to a high number of upstream contributing cells to erode through the dyke and keep their sink on the side where the dyke is at even though some sources are on the opposite site. This is well illustrated by the GEE-Model estimating the influence of dyke properties on the likelihood of rivers crossing the dykes. It's results pointed towards positive effects for dyke erodibility and distance to the grid centre while thickness had a negative effect. These observations are in line with the theoretical expectations. Thicker dykes are harder to erode completely at one position. So the thicker a dyke, the more likely it works as a complete barrier between two regions who's river systems are entirely separated. Erodibility on the other hand by definition makes breaches more likely since the material is less resistant. The distance from the simulation grid centre indirectly effects the likelihood of crossings because the larger this distance, the more space between the dyke and the grid centre where an initial topographic high develops in the calibration phase of the model.

The simulation runs where the dyke was established in the centre of the area simply did not lead to any variation of the distributions between the two sides since the effects of the dyke only further enforced present tendencies in the landscape without causing any biased difference to any of the two sides. It should be noted though that the topography is influenced by dykes in the centre. Thus, variables such as river channel steepness or flow direction will be influenced by dykes in the centre. Such changes are simply not reflected by my study design.

Dykes weaker than the background material cause the opposite effect to strong dykes. Here zones with very low topography are found which cause the main drainage divide in the centre of the area to migrate away from the dyke zone. This in turn leads to longer and larger river systems on the side with the weak dyke and shorter and smaller dykes on the opposite side.

The results in this work add to the established literature. First I could demonstrate that the FastScape software package is very well capable of modelling dykes and their influence on river systems. Modern ways of running multiple simulations in parallel also make it possible to generate distributions of parameters for multiple random starting conditions. This is an advance from approaches such as that of Van Gorp et al. (2014) who were able to run multiple scenarios but each of them only once which enforces biases induced by the chosen starting conditions.

The results presented here should be taken with a grain of salt due to methodological limitations and partially unrealistic modelling assumptions. First, the modelling setup was chosen to be rather static with constant climate, uplift rate and area shape. This allowed me to isolate the effect of erodibility changes but cannot be assumed for realistic settings. Secondly, dykes are likely

to change with depth such that different dyke layers will be exposed over time to interact with the river. Here I also assumed a uniform dyke so deep that it will not be completely excavated over millions of years, which is also rather unrealistic. Finally, the regression estimation of dyke properties effects on the occurrence of crossings suffers from a low number of explored dyke property values (only four thicknesses and positions). A statistically sound analysis should consider much more different cases perhaps with a lower number of repeated measurements per parameter combination to reduce computation time.

5 Conclusion

In this paper I have followed the question how river networks react to the presence of vertical dykes. By using the software packages FastScape, Topotoolbox and NetworkX, I was able to show that strong dykes cause zones of topographic elevation which lead to drainage divide migration and force entire river networks to change their shape. Dykes weaker than the surrounding material also change river networks but cause drainage divide migration not towards the dyke zone but away from it. Finally it was confirmed that the likelihood of rivers eroding through a dyke and establishing a permanent crossing is positively impacted by dyke erodibility and its distance to the simulation grid centre while being negatively impacted by dyke thickness.

Future studies should try to model more diverse landscapes which also consider differently shaped study regions and non-linear dyke features. Also the response of river systems could be assessed with different graph metrics or analyses not coming from graph theory at all. Finally, A dynamical analysis of river network response to dykes is needed. This should include the precise tracking of positional changes made by river network nodes to their eventual steady-state organization and secondly changes in dyke characteristics through time since dykes are continually exposed and eroded. A dyke feature completely unchanged along the entire material column which is uncovered over millions of years is not likely found in reality and should be adapted by future research projects.

References

- Bovy, B., & Lange, R. (2023, September). *Fastscapelib/fastscapelib: Release v0.1.0* (Version 0.1.0). Zenodo. <https://doi.org/10.5281/zenodo.8375653>
- Bovy, B., McBain, G. D., Gailleton, B., & Lange, R. (2021, January). *Benbovy/xarray-simlab: 0.5.0* (Version 0.5.0). Zenodo. <https://doi.org/10.5281/zenodo.4469813>
- Lee, F., Simon, K. S., & Perry, G. L. (2022). River networks: An analysis of simulating algorithms and graph metrics used to quantify topology. *Methods in Ecology and Evolution*, 13(7), 1374–1387.
- Van Gorp, W., Temme, A., Veldkamp, A., & Schoorl, J. (2015). Modelling long-term (300 ka) upland catchment response to multiple lava damming events. *Earth Surface Processes and Landforms*, 40(7), 888–900.
- Van Gorp, W., Temme, A. J., Baartman, J. E., & Schoorl, J. M. (2014). Landscape evolution modelling of naturally dammed rivers. *Earth Surface Processes and Landforms*, 39(12), 1587–1600.

6 Appendix

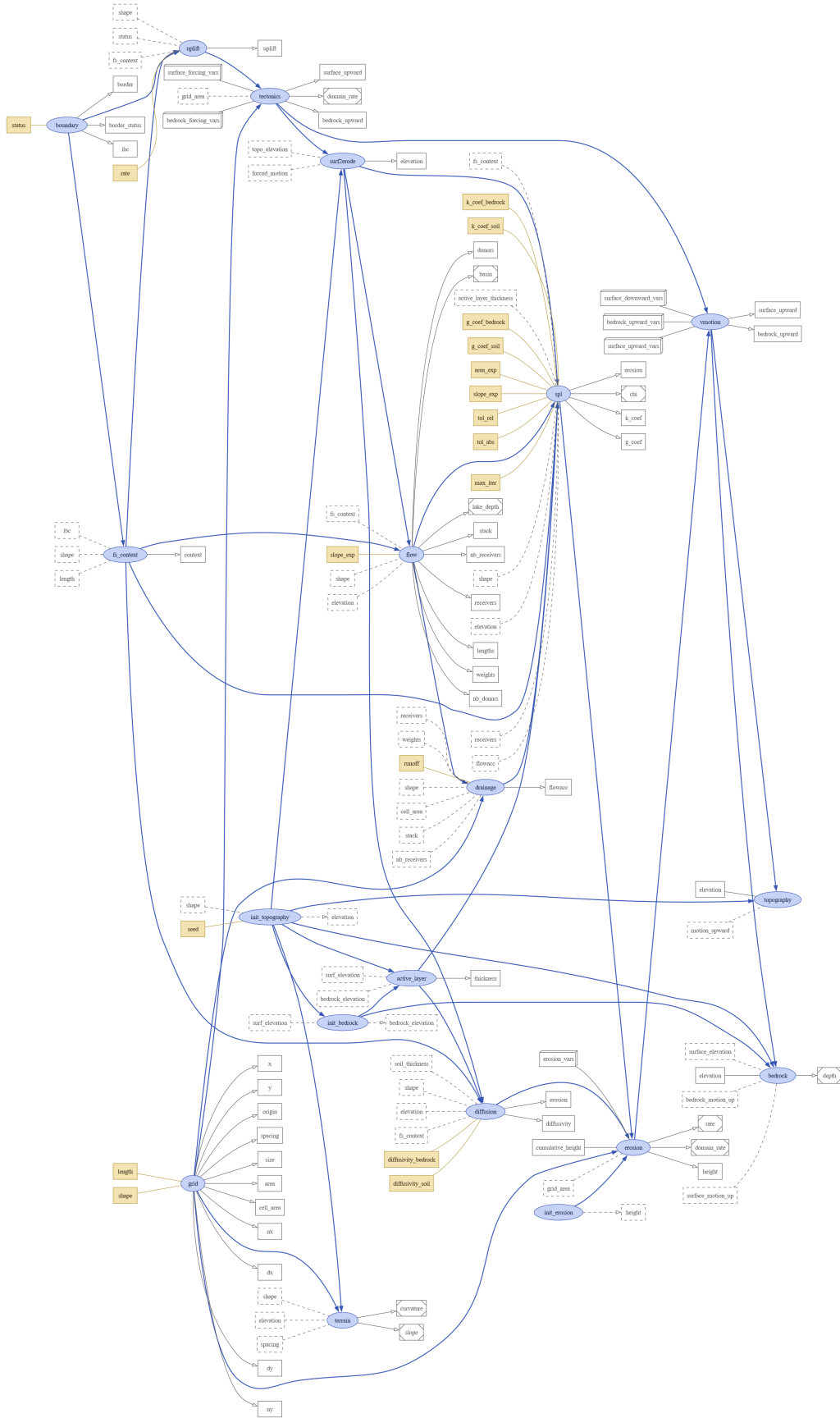


Figure 4: Visualized graph of the entire modified sediment model

7 Supplementary Material

For a notebook containing the simulation workflow code, any plots generated which were not shown in this paper or the video file displaying an example of the dynamical development of the simulated topography please consider the github repository for this project under this link: https://github.com/maltespringer/Landscape_Evolution_Modelling_FastScape_Graph_Theory.git