Spatial discretization and its impact on physical simulations

Gabriel Nilsson
Department of Computational Sciences, Minerva University
San Francisco, United States
gabriel.nilsson@uni.minerva.edu

Abstract

I explore the effect of spatial discretization choices on physical simulations by implementing a Dielectric Breakdown Model (DBM) and a Percolation Model (PM) on three regular grid types: the cube grid (CG) and two orientations of the rhombic dodecahedron grid (RDG4 & RDG6). I compare the strike location distribution, fractal dimensionality, and critical percolation probability between each grid type to find whether the grid choice has an impact on the outcomes with statistical significance. If statistically significant differences are found, it prompts further analysis of the underlying cause. It also implies that we should be more critical of our grid choices for physical simulations, and the grid choice should be properly discussed and justified when performing numerical work. If, instead, no significant results are found, or one with minimal practical significance, it prompts further exploration of what features make a simulation independent of grid choice.

1 Project description

In this project, I explore the quantitative effects on physical simulations as we change the cell shape in a regular grid. I will compare three specific grids, which I will call the Cube Grid (CG), Rhombic Dodecahedron Grid 4 (RDG4), and Rhombic Dodecahedron Grid 6 (RDG6). I will implement two separate models, the Dielectric Breakdown Model (DBM) and the Percolation Model (PM), using all three grids. I then measure the statistical and practical significance in a difference of means test in the mentioned measurement outcomes. I use the rhombic dodecahedron as my comparison with the cube grid since the shape isolates the effect of increasing the number of neighbors well; further justification of this choice is discussed in Section 5. The DBM is a common model to both visualize lightning structures for computer graphics [1] and to model real-life lightning and its impact on infrastructure [2]–[4]; the model is described in more detail in Section 3.

In the DBM, the measurement outcomes are the *strike distribution variance* and the *fractal dimensionality*, and in the PM it is the *critical percolation probability*, all described in further detail in Section 4.

This work raises and interrogates the assumption that we get the same simulation results when using a CG as when using other reasonable grids. An assumption made frequently with no or minimal discussion of its validity. This assumption is intuitively correct in some models, but erroneous in others. For example, consider estimating solutions to Laplace's equation using finite difference methods; it is then an intuitive expectation that the results from the different grids will converge to a common solution. That is, we expect to find the same equilibrium heat distribution when we use squares and hexagons. If we instead consider John Conway's Game of Life [5], it is evident that the rule set for a square grid would produce an entirely different behavior if applied on a hexagon grid. This highlights the general intuition that the assumption of grid choice invariance feels intuitive for "realistic" models, such as finding solutions to Laplace's equation, and intuitively wrong in "abstract" models like Conway's Game of Life.

I selected the DBM and PM as they represents a "realistic" and "abstract" models respectively. Additionally both models have been used to reach about real-life lightning using it. The PM model arguably falls into the abstract category. However, the PM is also important to interrogate it is a common model and used for conclusion about real life [6], [7].

While much of the previous literature on the DBM is done in 2D, I will exclusively analyze the 3D case. I do this as it is a better representation of real-life dielectric breakdown, which occurs in 3D space. Additionally, it is wise to stick to either 2D or 3D as any results from simulations done in a certain dimension would not necessarily be applicable to any other dimension.

reading mark

If statistical significance (p-value < 0.0005) and practical significance (Cohen's d > 0.5) are found in my selected measurement outcomes, I have shown that there are computational models, both abstract and realistic aspiring ones, which depend on the grid choice. In this case, we need to keep this in mind when performing computational work to make conclusions about real-life physics. If no such significance in differences is found, this prompts further exploration of exactly what makes a model or simulation independent of grid choice and whether this holds true for other models used in computational literature.

I use a very small significance level of 0.0005 because of a Bonferroni correction. I expect I will run approximately 20 difference-of-means tests, starting with

a significance level of 0.01, as I want to have a small chance of type I error, the Bonferroni correction yields a new significance level of $0.01/20 = 0.0005$.

2 Previous work

The dielectric breakdown model (DBM) has been well-studied and applied in previous literature [2]–[4], [6]–[12]. Where the earliest explicit algorithmic description was in 1984 by Niemeyer et al [9]. The DBM has been applied to analyze: the distribution of lightning strike locations given environmental features like buildings [2], [3], [8], the fractal dimensionality of DBM structures [4], [9], and analyze percolation probabilities given fraction of conductive material in composites [6], [7].

With a well-defined algorithm like the DBM and its frequent usage in previous computational work, we have an especially strong opportunity to extend the results of these studies by measuring the same or similar outcomes as they do. However, the generalizability of the previous work is limited as they only use square or cube grids. The extension is in not only using the cube grid (CG) but also the two variants of the rhombic dodecahedron grids (RDG4 & RDG6), which specifically aims to isolate the effect of increasing the number of neighbors per cell while keeping other grid features constant.

With this project, I will be able to find which results from previous work are generalizable to another grid structure, as well as be able to measure exactly how the results change as we change the grid at different resolutions. This is the primary way my project is an extension of previous work.

There is, in general, a limited amount of previous work on the comparison of using different grids; for example, in the literature review by [13], only 2 of the 64 papers in Ecological Modelling that included spatial interactions used a hexagonal grid. Of the literature that compares the effect of using different grids, very few apply it to physical simulations, and almost none perform the analysis in 3D space.

To enumerate, there is previous literature discussing whether hexagonal or square grids make for better visual rendering in raster graphics [14], [15], comparing which grids make for better mappings on a sphere [16], [17], and analyzing theoretical connectivity features of square and hexagon grids [13]. In [18], they model water flow using altitude data, finding that grid choice matters because hexagons are more self-consistent across different resolutions.

According to [19], hexagons make cellular automata more robust for changes in rotation, and according to [20] hexagons produce a smaller error than squares when using finite differences to approximate solutions to the Laplace operator, which is exactly what I do in this project.

For the spatial representation of information and memory recall, it seems to not matter whether you use a hexagon or square grid [21].

While the comparison between square grids and hexagons has been explored in a multitude of contexts, there is minimal research on the comparison of 3D grids, especially in analyzing the effect on physical simulations. This is the secondary way in which my work is an extension of previous work.

However, the gap in the academic literature on grid comparisons in 3D, which makes my work a novel avenue of exploration, is also what makes my project limited as an extension. This is because conclusions made in 3D do not necessarily give insight into the dynamics at a 2D level, which is almost exclusively where previous work has been done [3], [4], [8], [9]. Similarly, I use a specific set of models and metrics, which does not speak to what happens in other models. As noted above, we may find significance in the grid choice according to one model and metric [18], [19], but not in another [21].

Still, this builds on the previous work on these specific models, DBM and PM, giving valuable insight into whether previous findings generalize into three dimensions.

3 The Dielectric Breakdown Model

The dielectric breakdown model (DBM) models lightning, or more generally any form of dielectric breakdown across a non-conductive medium. The model divides the space into cells, usually squares or cubes, and starts in an initial charge configuration, such as Figure 1. The gray cells in the figure represent the "lightning structure" and have a fixed electric potential at $e = 0$; the black cells represent the ground and have a fixed electric potential at $e = 1$. The remaining white cells represent the insulating medium and do not have a fixed electric potential. For implementation purposes, each cell that is at the outer edges of the simulation box will also be of fixed electric potential.

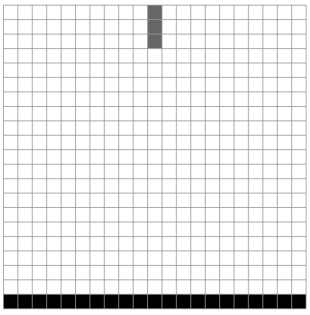
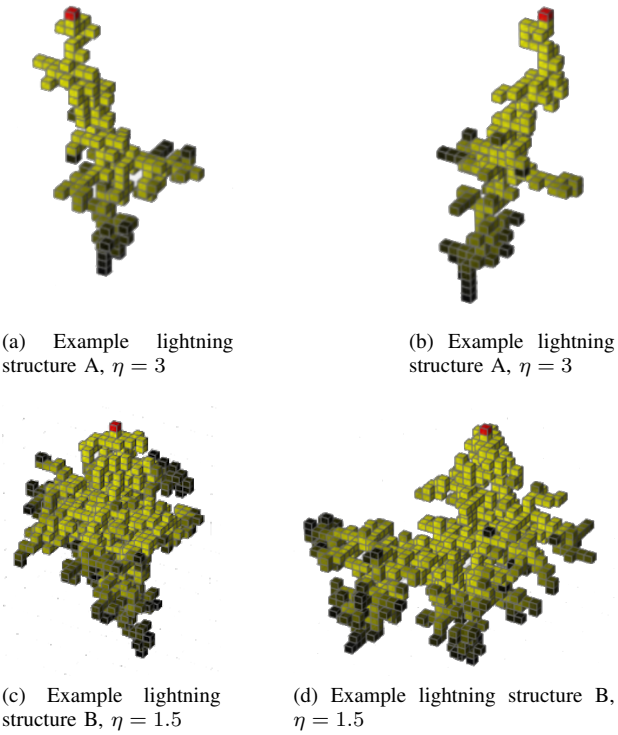


Fig. 1: Example initial DBM configuration in 2D square grid (Fig 1b in [1])

At each timestep, the lightning structure grows using a Laplacian Growth Model, the core logic of the DBM [1]–[3], [8]. It's called the Laplacian Growth Model because we calculate the electric potential in the air by solving Laplace's equation $\nabla^2 U = 0$, where U is the electric potential distribution across the simulation box, and let the probability of growth to a neighboring cell be proportional to the electric potential at that point. I find the electric potential for each cell by approximating the solution to Laplace's equation, using finite difference methods.

The DBM has two parameters, η and α. η dictates the downward tendency of the lightning structure, where if $\eta = 0$ the lightning will spread equally in all directions with no bias, and when $\eta \to \infty$ the lightning structure

will increasingly become a straight line taking the shortest distance to the ground. α is the approximation threshold and is a parameter of the iterative approximation method for solving Laplace's equation. Further details about α, and the simulation implementation can be found in Section 8.1.



(a) Example lightning structure A, $\eta = 3$

(b) Example lightning structure A, $\eta = 3$

(c) Example lightning structure B, $\eta = 1.5$

(d) Example lightning structure B, $\eta = 1.5$

Fig. 2: 3D rendition of two lightning structures in a cube grid, shown from two different angles. Visualizes the effect of changing the parameter η.

Laplace's equation is used as it is the steady-state solution to the heat equation, which also models the electrostatics in a volume of free space. This is why the model does a good job at modeling lightning structures since it approximates the electric potential in the air using the appropriate differential equations which describe electric field behavior. However, do mind that like any other model, there are idealistic assumptions made, such as relative homogeneity of the air and no external electric field. I say relative homogeneity since irregularity in the air, which makes real-life lightning structures non-deterministic, is represented in the model through the stochasticity in the selection of the growth sites. Yet we need to keep in mind that such idealized simplifications are being made and that this limits our ability to generalize results from the DBM to real life, however, this is not of concern for this work as I am looking for the effect on the model's measurement outcomes, not whether the model is an accurate representation of reality.

The fact that we're using Laplace's equation makes this simulation especially interesting to observe under a change of grid structure because the equation makes the value of every grid dependent on every other grid value. In the context of this simulation, the probability of the lightning growing to a specific new cell is dependent on the entire current lightning configuration. This is not true in, for example, simpler percolation models like the one I will explore later, there the probability of growth is a function of the local neighborhood only. The hope is that any findings from using a model with highly interconnected dependence between micro-states will be able to tell us more about the nature of physical simulations' relationship with grid choice, rather than the relationship between local interactions and the grid choice.

4 Outcome Metrics

4.1 fractal dimensionality

The simulation parameters and process are the same as for the metric "strike position" mentioned above. Instead of measuring the the strike distribution variance I instead measure the fractal dimensionality, specifically the Hausdorff dimensionality (also sometimes called box-counting dimension) used in previous analyses on the fractal dimensionality of the DBM [4], [9], [10], [12].

While this outcome metric, like the other two, is a valuable venue of exploration since it allows me to gain more diverse insight into the effect of grid choice on simulations, it also holds a unique benefit by allowing me to account for a potentially extraneous variable.

Model parameters, like simulation box size, approximation threshold α, and downward-trend parameter η all likely influence the lightning strike measurement outcome and may act as extraneous variables unless controlled for (discussed in Section 6), however since they're model parameters they are very easy to control for.

Another variable that potentially acts extraneously is the average fractal dimensionality of the lightning structure, and it might not be the case that keeping all model parameters constant means that the fractal dimensionality is kept constant. In other words, a lightning structure produced on a CG with $\eta = 6$ does not necessarily have the same fractal dimensionality as a lightning structure produced on an RDG with $\eta = 6$. It may also be the case that the variance of the lightning strike distribution depends on the fractal dimensionality. For this reason, I will analyze whether fractal dimensionality stays constant over different grid types, and if not, I will need to perform comparisons between the models where dimensionality is fixed.

4.2 Strike position variance

This is a metric of the DBM. The simulation will take in the target height of the simulation box H and produce a simulation box of dimensions $W \times W \times H$ where $W = 2H$ in physical distance. The other parameters are the downward tendency η, the approximation threshold α, and, of course, the target grid type (CG, RDG4, or RDG6). I let the lightning structure start as a single cell in the center of the top face of the box, where the opposite side of the box is ground and has an electric potential of 1 (where the lightning structure is 0). One run of the simulation lets the lightning structure grow iteratively until it reaches the opposite side (ground), and then I measure the strike location of the structure. The

measurement outcome is the variance of the distribution of striking locations over many simulation runs. This measurement outcome is not an exact replication of previous literature, however, strike distribution in 2D has been analyzed [3], [8] as well as analysis of lightning strikes on objects such as buildings and powerlines [2]–[4].

4.3 Critical percolation probability

Finally, I will implement a percolation simulation similar to [6], [7]. The parameters of the simulation are p, the probability of individual cell activation, the size of our simulation box N from which we produce a simulation box of dimensions $N \times N \times N$ where the units are cells, and the target grid type (CG, RDG4, or RDG6).

The measurement outcome is for which individual cell activation probabilities p we have critical percolation probability, which is the activation probability at which the percolation probability is independent of the size of the simulation box. Further description of this metric can be found in Section 9.1.

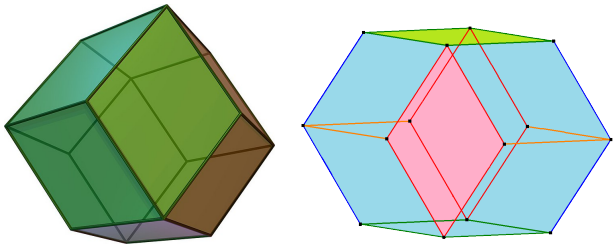
5 Why the Rhombic Dodecahedron

In my analysis, I am looking to isolate the effect of increasing the number of neighbors, with the baseline shape being the cube. To achieve this I need to ensure the new cell shape can tesselate the volume as well as be face-, edge-, and vertex-transitive.

Transitivity along a component, like edges, means that there's only one unique version of that component on the polyhedron. For example, if we have vertex-transitivity, we can color some vertex red and another blue and then translate and rotate the polyhedron so that the blue vertex is on the exact location where the red vertex was, and the set of vertex coordinates is the exact same as before. For example, a normal 2D triangle is vertex-transitive if the angles between the edges that connect to the vertex are the same as all the other vertices (e.g., an equilateral triangle). Similarly, for a given 3D polyhedron, we have vertex transitivity if all vertices have the same angles between all its edges. Notice that all vertices need to have the same number of edges connecting to them for the shape to be vertex-transitive. Similarly, for edge- and face-transitivity, I need to be able to perfectly map from any edge (face) to any other edge (face) through only translation and rotation for it to be transitive under that component.

By definition, the set of convex polyhedra, which are face-, edge-, and vertex-transitive, is the set of platonic solids, of which only the cube tessellates 3-dimensional Euclidean space. To find another shape that can produce a 3D grid, I will need to abandon one of these transitivity properties.

It happens to be the case that the polyhedral classes of rhombohedra and rhombic dodecahedra are the only space-filling polyhedra with congruent faces in 3-dimensional Euclidian space. Where if we apply the constraint of edge- and vertex-transitivity to the class of rhombohedra we get the cube, and we can't apply the constraints of edge- and vertex-transitivity to the class of rhombic dodecahedra in 3D space. However, if we apply only the constraint of edge-transitivity, rhombic dodecahedra we get the "regular" rhombic dodecahedron and the Bilinski dodecahedron. While both are space-filling, neither satisfies vertex-transitivity, and only the "regular" rhombic dodecahedron is face-transitive. Notice in Figure 3 that both polyhedra have face congruency as all faces are congruent rhombuses, but the Bilinski dodecahedron is not face transitive. We see this if we pay attention to the different colorings of faces in Figure 3b; we can not translate and rotate the Bilinski dodecahedron such that a red face ends up where the green face is without the vertices of the polyhedron now being at a new set of coordinates. Notice that we can do this for the "regular" rhombic dodecahedron (Figure 3a). Hence we opt for the "regular" rhombic dodecahedron for our selection of another grid type.



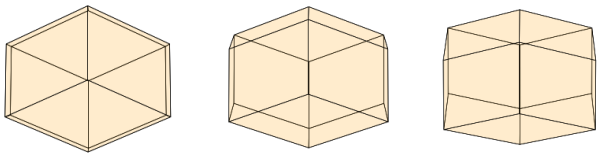
(a) "Regular" rhombic dodecahedron [22]

(b) Bilinski dodecahedron [23]

Fig. 3: The set of all space-filling and face-congruent polyhedra that are not rhombohedra

From now on when I write rhombic dodecahedron (RD), I refer to the "regular" rhombic dodecahedron. Notice that the RD is transitive over each component, except vertices. We can notice this by finding a vertex that has three connecting edges (an obtuse vertex) and another vertex that has 4 connecting edges (an acute vertex) in Figure 3a. Because of this asymmetry and the directional nature of the dielectric breakdown model (DBM), it may matter whether the vertical axis is through two obtuse vertices or two acute vertices. The potential impact of this difference in orientation is discussed in Section 6.1. Because of this, I will run the DBM on three grids I will treat as distinct, the cube grid (CG), the rhombic dodecahedron grid (RDG) oriented with acute vertices along the vertical axis (RDG4), and the rhombic dodecahedron grid with obtuse vertices along the vertical axis (RDG6). The reason for the numbers 4 and 6 in the naming of RDG4 and RDG6 is because of the number of horizontal neighbors in that orientation. In other words, the orthographic projection of a RD oriented with acute vertices vertically will be a square, while the orthographic projection of a RD oriented with obtuse vertices vertically will be a hexagon (see Figure 4).

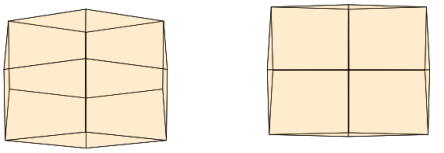
One might ask why not consider other orientations of the RD, for example, the orientation orthogonal to both Figure 4a and Figure 4e. The reason is that in all other orientations, a given horizontal layer, or slice, of a simulation box will have cells of different vertical displacements from the starting location of the lightning structure. For example, in the suggested orientation or-



(a) ($\alpha = 0°$) Obtuse perspective, hexagonal profile

(b) ($\alpha \approx 13.68°$)

(c) ($\alpha \approx 27.37°$)

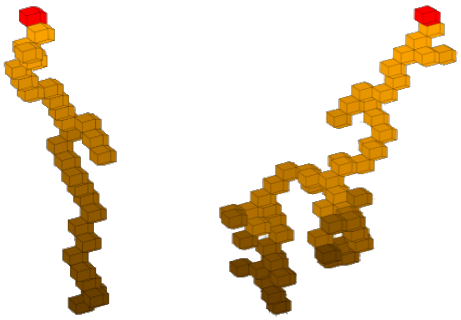


(d) ($\alpha \approx 41.05°$)

(e) ($\alpha \approx 54.74° = cos^{-1}(\frac{1}{\sqrt{3}})$) Acute perspective, square profile

Fig. 4: Nearly orthographic projection of rhombic dodecahedron from the z-axis as we rotate from obtuse vertices along z-axis to acute vertices along z-axis. it is "rolling" forward from the reader's perspective.

thogonal to both RDG4 and RDG6, every other row in a given horizontal slice is slightly higher up, meaning that the lightning structure will strike those rows with much higher frequency. As I am interested in the strike distribution, I want to omit features that bias the lightning to certain cells on the ground layer of the simulation box.



(a) $\eta = 6$ (b) $\eta = 4$

Fig. 5: 3D rendition of a lightning structures in RDG6, where simulation box height is 20, $\alpha = 0.01$, with varying values for η.

6 Theoretical Expectations & Assumptions

The broad research question is whether there is a difference in statistical and practical significance in any of the three measurement outcomes as we change grid type. Where for each outcome, I will perform a statistical test to evaluate whether there is a difference or not as I change the grid type. A benefit of gathering empirical data from simulations is that I can control the size of my sample and hence have control over the certainty of the results. I need to pay extra attention to the practical significance, as the large sample sizes may produce results that are statistically significant but not practically significant.

For the measurement outcome of strike location, I will do a t-test over the observed variance in physical distance units for each grid. For the fractal dimensionality and the percolation probability, I will perform a difference of means test between my measurements from each grid.

I hypothesize that there will be a difference in all metric outcomes with statistical and practical significance. However, the differences' practical significance will diminish as we increase the resolution, i.e., decrease the size of the cells in the simulation box.

In this analysis I am attempting to keep every property of the simulation fixed, except the cell shape. A problem arises when I try to maintain the same dimensions for the simulation box across different grid types. If I consider simulation boxes in units of cells, then it is easy to set the dimensions to be equal. However, if I consider physical distance I will be unable to produce two simulation boxes from different grids which have identical dimensions. Let us consider the distance between the center of given cell C_1 and the center of a neighbor C_2, I call this distance 1 standard distance unit (sdu) and normalize our grids such that for any two neighboring cells $||C_1 - C_2|| = 1$ sdu. We can see in Figure 6 that this results in an irrational height for RDG4 and RDG6 when we stack two cells "vertically". This means that there is no way to get the exact same height (in physical distance) between two simulation boxes of different grid types. An assumption I make in my analysis is that the inevitable discrepancy in height between various simulation boxes can be accounted for in the analysis.

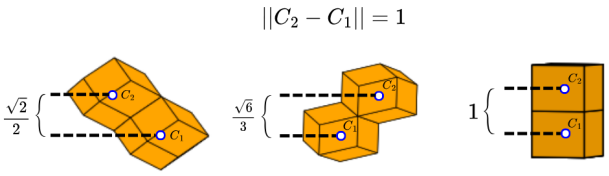


Fig. 6: Vertical distance between neighboring cells in (from left to right) RDG4, RDG6, and CG. Where C_1 and C_2 are the center-points of two cells that share a face.

Just like it is important to compare the different simulations at the same resolution (dimensions), it is also important to compare the simulations under the same parameter value η and numerical approximation thresholds. If we make any comparisons over different resolutions, different values for the parameter η, or different approximation thresholds they can act as extraneous variables, influencing the measurement outcomes we want to associate purely with neighborhood size and orientation. Note that these are still model parameters and not strictly extraneous variables, which is why I take care to say "acts as" an extraneous variable, since if not controlled for, they will be a non-independent variable that influences the dependent variable.

6.1 *Theoretical expectation for strike distribution*

In the comparison between the CG and the two RDGs (RDG4 & RDG6), I expect the increased number of neighbors to increase the strike distribution variance with statistical significance, given the same simulation box dimensions and value for the downward tendency parameter (η).

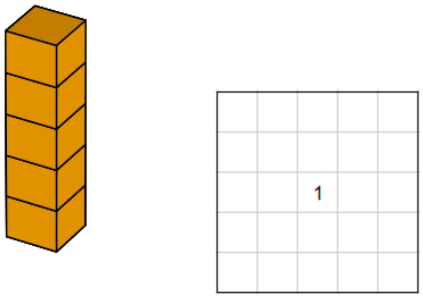
I argue for the plausibility of this expectation by showing that it is true at the extreme where $\eta \to \infty$, under some assumptions. At this extreme, the lightning structure is guaranteed to always select the neighboring cell with the largest electric potential, this will always be a cell in the next vertical (z-axis) layer of the simulation. If we assume that the electric potential at the lowest vertical layer is independent of the electric potential on all layers above, which is not strictly true in DBM, we can let all neighbors on the next layer down have an equal probability and analyze the theoretical distribution of strike location.

For each grid, imagine that the lightning structure starts n cells above the ground layer, such that it can strike the ground in exactly $n-1$ steps. I then look at which coordinates the strike can occur, and the probability distribution of these strike locations. We will visualize the specific case of $n = 5$.

6.1.1. CG

In the cube grid, each cube has exactly one neighbor that is on the next layer, leading to only one possible strike location, which is at the same horizontal coordinates as the start location.

For the case of $n = 5$, there is only one possible path, visualized in Figure 7a. We can visualize the frequency (and probability) distribution by counting the number of paths that end at each coordinate (7b)



(a) 3D rendering of only possible path

(b) Frequency & probability distribution

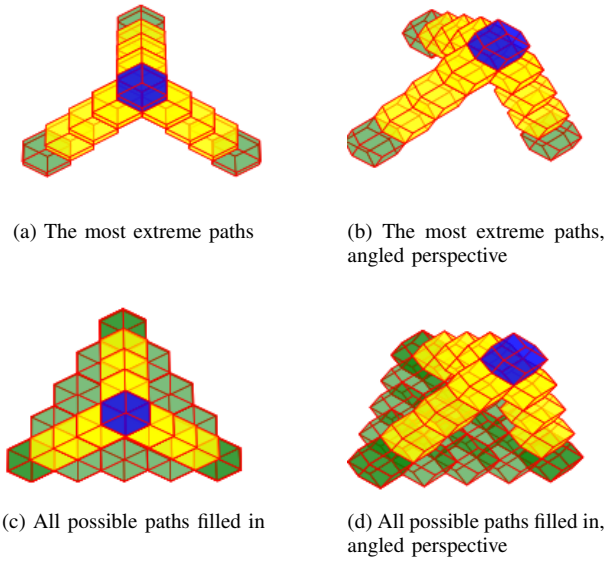
Fig. 7: The only possible lightning strike in a cube grid, when the structure starts at $z = 5$, and where the strike takes exactly $n-1$ (4) steps.

6.1.2. RDG6

In RDG-6, each individual RD has exactly three neighbors on the next layer, leading to multiple possible paths of equal probability.

At a height n there will be 3^{n-1} total possible paths to the ground since we take exactly $n-1$ steps and each step has 3 options. The number of possible striking locations will follow the triangular numbers: $\frac{n(n+1)}{2}$.

For the specific case $n = 5$:



(a) The most extreme paths

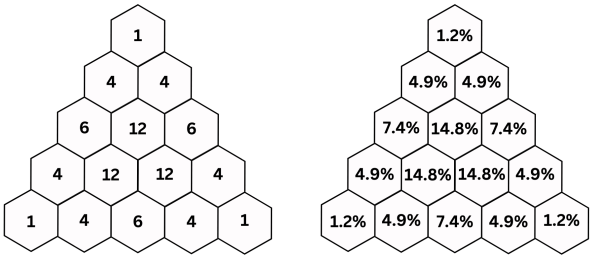
(b) The most extreme paths, angled perspective

(c) All possible paths filled in

(d) All possible paths filled in, angled perspective

Fig. 8: A visualization of the starting cell (blue) the most extreme paths (yellow), and all possible striking locations (green). For RDG6 and $n = 5$.

I count the number of paths that reach each strike location in exactly $n-1$ (4) steps, then divide by the total number of paths 3^4 (for $n = 5$) to get the probability distribution.



(a) Frequency distribution

(b) Probability distribution

Fig. 9: Frequency and probability distributions for possible strike location for RDG6 when the lightning structure starts 5 cells above the ground ($n = 5$).

I generalize the frequency distribution by noticing that the frequency count for a given cell, defined by its integer barycentric coordinates k_i, k_j, and k_z, is the trinomial coefficient $\binom{n}{k_i,k_j,k_z}$ where $n = k_i + k_j + k_z$.

By comparing Figure 7b and 9 we see that the RDG-6 grid clearly produces a much larger variance of strike locations than CG when $\eta \to \infty$.

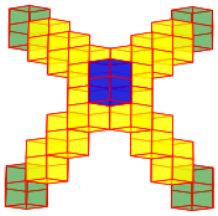
6.1.3. RDG4

In RDG4, each individual RD has exactly 4 neighbors on the next layer, leading to even more possible paths of probability, compared to RDG6.

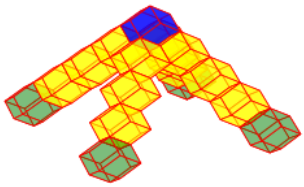
At a height n, there will be 4^{n-1} total possible paths to the ground since we take exactly $n-1$ steps, and

each step has 4 options. The number of possible striking locations will follow the square numbers: n^2.

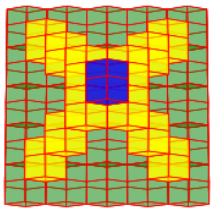
For the specific case $n = 5$:



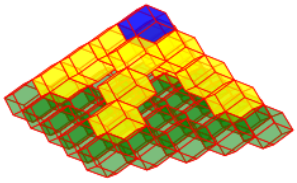
(a) The most extreme paths



(b) The most extreme paths, angled perspective



(c) All possible paths filled in



(d) All possible paths filled in, angled perspective

Fig. 10: A visualization of the starting cell (blue) the most extreme paths (yellow), and all possible striking locations (green). For RDG4 and $n = 5$

I count the number of paths that reach each strike location in exactly $n - 1$ (4) steps, then divide by the total number of paths 4^4 (for $n = 5$) to get the probability distribution.

1	4	6	4	1
4	16	24	16	4
6	24	36	24	6
4	16	24	16	4
1	4	6	4	1

(a) Frequency distribution

0.4%	1.6%	2.3%	1.6%	0.4%
1.6%	6.3%	9.4%	6.3%	1.6%
2.3%	9.4%	14.1%	9.4%	2.3%
1.6%	6.3%	9.4%	6.3%	1.6%
0.4%	1.6%	2.3%	1.6%	0.4%

(b) Probability distribution

Fig. 11: Frequency and probability distributions for possible strike location for RDG4 when the lightning structure starts 5 cells above the ground ($n = 5$).

I generalize the frequency distribution by noticing that the frequency counts for a given cell of coordinate x, y, with a height of n, will be the product of two binomial coefficients $\binom{x}{n}\binom{y}{n}$.

It is clear that in both orientations of the RDG (RDG4 & RDG6) we have a larger variance of strike location than in the CG. We can extend this argument by considering the case where η does not approach infinity, but instead $\eta = 0$. The lightning structure will then spread around it is starting point like a sphere, rather than straight lines. Even in this case, we can consider counting all the lightning strikes that reached the ground in exactly $n - 1$ steps, and we would see the exact same conclusion as above. As we continue and count the lightning strikes that reach the ground in exactly n, $n + 1$, $n + 2$, $\cdots$, $n + k$ steps, we can, at each lightning step value, imagine how the two RDG grids would have more possible striking locations and a larger variance, than the CG counterpart. Using this argument, it is plausible that the RDG grids will have a larger variance in the strike location distributions than the CG, and hence, plausible to believe that the increased size of the cell neighborhood can have a meaningful effect on the measurement outcomes.

The theoretical analysis that underlies this plausibility argument is limited in that it relies on the validity of the assumptions outlined at the beginning of Section 6.1, and it only considers the two extreme cases of the bias parameter ($\eta = 0$, $\eta = \infty$). The reason I only consider the extreme parameter values is because that is when the growth site selection is approximately independent of lightning structure shape. A similar predictive analysis of any other parameter value would require considering the ensemble of possible lightning structure shapes, which would either be intractable or require further simplifications and assumptions that would harm the strength of the argument. Still, this analysis could be improved by performing a more rigorous analysis of, for example, how probability density propagates in the different grids. However, while non-rigorous, I find this a sufficiently persuasive example as an argument for it being plausible that variance in the lightning strike distribution is larger for the RDGs compared with the CG.

7 Fractal Dimensionality

Before we start analyzing the strike position distributions, we first need to evaluate whether we need to consider fractional dimensionality as a variable that may act as an extraneous variable.

7.1 Method

I will calculate the Minkowski-Bouligand dimension, also known as the box-counting dimension, of the lightning structures produced by the DBM. Specifically, a lightning structure that is grown radially and unbiased from the origin. This lightning structure with no inherent directional bias allows for more efficient and intuitive calculation for the dimensionality, while still describing the dimensionality's dependence on the parameter η.

There are multiple valid ways to calculate the box-counting dimension, all of which are equal for well-behaved fractals. I will use the counting method of concentric spheres centered at the origin, which is appropriate for the radially grown lightning structure as it has a center.

Where r is the radius of a sphere centered at the origin, and $N(r)$ is the number of cells that are members of the lightning structure S, and inside the sphere of radius r, the fractional dimensionality is the slope of the logarithm of $N(r)$ over the logarithm of the radius.

$$\dim(S) = \frac{\log(N(r))}{\log(r)} \frac{d}{d\ \log(r)}$$

The slope will be approximated using the average rate of change between two appropriate values of r,

$$\dim(S) \approx \left(\frac{\log(N(r_2))}{\log(r_2)} - \frac{\log(N(r_1))}{\log(r_1)}\right)\frac{1}{\log(r_2) - \log(r_1)}$$

This measure relies on the notion of dimension as a power rule in scaling. If we scale all lines in a 2-dimensional object by 2, we get $2^2 = 4$ copies of the original object, and if we scale all lines in a 3-dimensional object by 2 we get $2^3 = 8$ copies of the original object. In general, with a dimension D, a linear scaling of c, and a dimension measure V (area, volume, etc), we have that

$$V \sim c^D$$

Where in this analysis, I scale r linearly and measure the increase in dimension measure V by counting the number of cells $N(r)$ included in a sphere of radius r

$$N(r) \sim r^D$$

I hence get the dimension by

$$\log(N(r)) \sim D\log(r)$$
$$D \sim \frac{\log(N(r))}{\log(r)}$$

I also note that since the lightning structures now are grown radially, and since the fractal dimensionality is evaluated with no consideration for direction, there is no difference between RDG4 and RDG6, since they are just a 90° rotation of each other.

7.2 Results

After simulating many radially grown lightning structures and evaluating their fractal dimensionality D, we get the results in Figure 12.

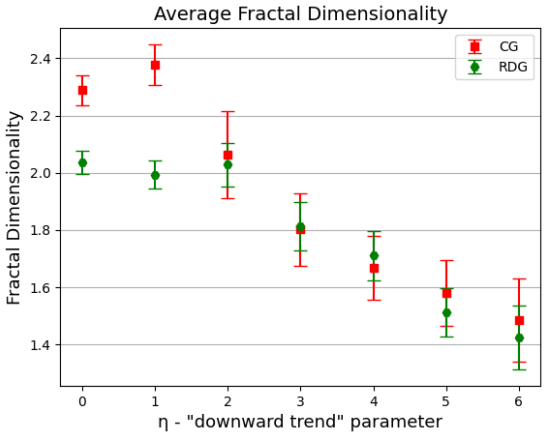


Fig. 12: Fractal dimensionality of lightning structures at different values for η. Whiskers are 95% confidence intervals

These results are noteworthy, it seems like the fractal dimensionality does depend on the grid type if $\eta = 0$ or $\eta = 1$, however, it seems like the fractal dimensionality is constant for all other integer values of η, including $\eta = 6$ which is the parameter value I used for all DBM simulation runs.

Using a significance level of 0.0005, only $\eta = 0$ and $\eta = 1$ have statistically significant differences in average fractal dimensionality, with p-values of 2.69×10^{-8} and 1.87×10^{-7} respectively. Additionally, the effect sizes are very large, with Cohen's d of 3.26 and 4.39, respectively. The difference in means is not statistically significant for any other value of η, see Table I.

This means that we do not need to control for the fractal dimensionality in the lightning strike distribution analysis since the fractal dimensionality is the same for all grid types when $\eta = 6$.

8 Strike Position Simulation & Method

8.1 Simulation description

A single run of the simulation is described in Algorithm 1. Where the most important steps are recalculating the electric potential (line 14) and the selection of new growth sites (line 16).

When calculating the new potential (Algorithm 2) I approximate the solution to the Laplace equation $\nabla^2 U = 0$ using a finite difference method. Where we let U be the electric potential air, we can approximate the solution as follows. We first make use of the approximation for the second derivative in a uni-variable function $f(x)$,

$$f_{xx} \approx \frac{f(x-h) - 2f(x) + f(x+h)}{h^2}.$$

Which we then can use in the 2D case to find an approximation for the second spatial derivative

$$\begin{aligned}\nabla^2 U(x,y) =& U_{xx} + U_{yy} \\ \approx& \frac{U(x-h,y) - 2U(x,y) + U(x+h,y)}{h^2} + \\ & \frac{U(x,y-h) - 2U(x,y) + U(x,y+h)}{h^2} \\ =& \frac{\begin{array}{r}-4U(x,y) + U(x-h,y) + U(x+h,y) \\ + U(x,y-h) + U(x,y+h)\end{array}}{h^2}.\end{aligned}$$

We can then define a stencil

$$\Delta_h = \frac{1}{h^2}\begin{bmatrix}0 & 1 & 0\\ 1 & -4 & 1\\ 0 & 1 & 0\end{bmatrix}$$

and more concisely write,

$$\nabla^2 U \approx \Delta_h U(x,y)$$

Where we note that $\Delta_h(\cdot)$ is an operation that acts on every possible local neighborhood of $U(x,y)$ simultaneously.

Crucially, the finite difference method depends on the stencil Δ_h which is derived from the specific neighborhood in the given discretization of the continuous function that you wish to approximate. Where if I wish to apply the finite difference method on my grids, I simply apply the stencil that corresponds to the graph neighborhood. In general, these stencils represent that the approximation of the solution to the Laplace's equation is when each cell equals the average of its neighbors.

Algorithm 1 Pseudocode for a lightning-strike simulation

Input : Dimensions of box (*Width, Depth, Height*), bias to ground (η), and estimation threshold (ϵ)

begin
- $W \longleftarrow$ Width
- $D \longleftarrow$ Depth
- $H \longleftarrow$ Height
- struck_ground $\longleftarrow$ **False**
- e_pot $\longleftarrow$ Matrix((W, D, H))
- fixed_potential $\longleftarrow$ Matrix((W, D, H))
- lightning_structure $\longleftarrow$ List[$(\mathrm{W}/2, \mathrm{D}/2, 0)$]

/* Fix potential for lightning structure and ground */

- e_pot$[\mathrm{W}/2, \mathrm{D}/2, 0] \longleftarrow 0$
- fixed_potential$[\mathrm{W}/2, \mathrm{D}/2, 0] \longleftarrow$ True
- e_pot$[:, :, \mathrm{H}] \longleftarrow 1$
- fixed_potential$[:, :, \mathrm{H}] \longleftarrow$ True

- **while *not*** struck_ground **do**
  - /* Update electric potential in air */
  - electric_potential $\longleftarrow$ NEW_POTENTIAL(ϵ)
  - /* Select new growth-cell */
  - **For each cell** i **adjacent (**C**) to lightning structure:**
    - weights$[i] \longleftarrow \frac{e_i^\eta}{\sum_{k \in C} e_k^\eta}$
  - new_neighbor $\longleftarrow$ random.choose(C, weights)
  - lightning_structure.add(new_neighbor)
  - e_pot[new_neighbor] $\longleftarrow 0$
  - fixed_potential[new_neighbor] $\longleftarrow$ True
  - **if** new_neighbor.z $== H$ **then**
    - struck_ground $\longleftarrow$ **True**
- strike_location $\longleftarrow$ lightning_structure.last
- **Return** strike_location

Algorithm 2 Pseudocode for NEW_POTENTIAL()

Input : Current electric potential (*e_pot*), and error threshold (ϵ)

begin
- worst_error $\longleftarrow 2\epsilon$
- old_e_pot $\longleftarrow$ copy(e_pot)
- new_e_pot $\longleftarrow$ zeros_like(e_pot)

- **while** worst_error $> \epsilon$ **do**
  - /* Set each cell to the average of neighbors */
  - new_e_pot $\longleftarrow$ convolve(old_e_pot, stencil)
  - new_e_pot $\longleftarrow$ old_e_pot $where$ fixed_pot
  - relative_errors $\longleftarrow \frac{|\text{new\_e\_pot} - \text{old\_e\_pot}|}{\text{old\_e\_pot}}$
  - worst_error $\longleftarrow$ max(relative_errors)
  - old_e_pot $\longleftarrow$ copy(new_e_pot)
  - new_e_pot $\longleftarrow$ zeros_like(new_e_pot)
- **Return** new_e_pot

$$\nabla^2 U = 0$$
$$\Delta_h U(x,y) = 0$$
$$\frac{1}{h^2}\begin{bmatrix}0 & 1 & 0\\ 1 & -4 & 1\\ 0 & 1 & 0\end{bmatrix} U(x,y) = 0$$
$$\begin{bmatrix}0 & 0 & 0\\ 0 & 4 & 0\\ 0 & 0 & 0\end{bmatrix} U(x,y) = \begin{bmatrix}0 & 1 & 0\\ 1 & 0 & 1\\ 0 & 1 & 0\end{bmatrix} U(x,y)$$
$$\begin{bmatrix}0 & 0 & 0\\ 0 & 1 & 0\\ 0 & 0 & 0\end{bmatrix} U(x,y) = \begin{bmatrix}0 & \frac{1}{4} & 0\\ \frac{1}{4} & 0 & \frac{1}{4}\\ 0 & \frac{1}{4} & 0\end{bmatrix} U(x,y)$$

Notably, this notion of cells being equal to the average of their neighborhood holds true for any graph structures (see Numerical Appendix, Section 11), for example, the 3D cube case (Figure 13) and hexagon grids (see Eq. 6 in [24]).

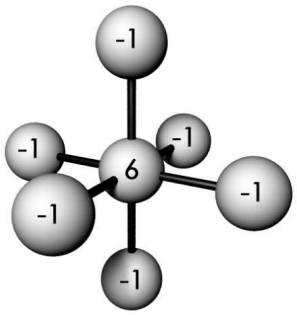


Fig. 13: 3D visualization of 3D cube stencil. Figure 1d from [25].

When I write *convolve()* in line 6, Algorithm 2, I refer to a convolution using the relevant stencil for the grid type. A single convolution will make the new value of each cell equal the average of it is neighborhood from the previous iteration. This is an iterative process that moves closer and closer to the true solution, the error decreases exponentially which means that the rate of change is proportional to the magnitude of the error between the current estimation and the true solution. This is why I can terminate the process when the worst relative rate of change reaches α, using that as a proxy for distance to the true solution.

The specific choice of η is arbitrary for my purposes since I am not aiming to make realistic lightning structures. However, I still choose the specific value of $\eta = 6$ because it produces lightning structures that have a similar fractional dimensionality to real-life lightning on a cube grid [26].

8.2 Analysis method

I will compare the lightning strike distributions (Figure 28, 29, and 30), by comparing the variances of each distribution, where the variance is calculated as the average squared distance to the mean position. I can then compare the variance distributions coming from other grid types and, together with the 95% confidence interval for the

variance, evaluate whether the variances are significantly different.

The variance will be in a standard distance unit (sdu), where the distance between the center of two cells sharing a face is 1 sdu.

When comparing different grids, I ensure that the height of the simulation boxes match, as that is the distance that the lightning has to travel.

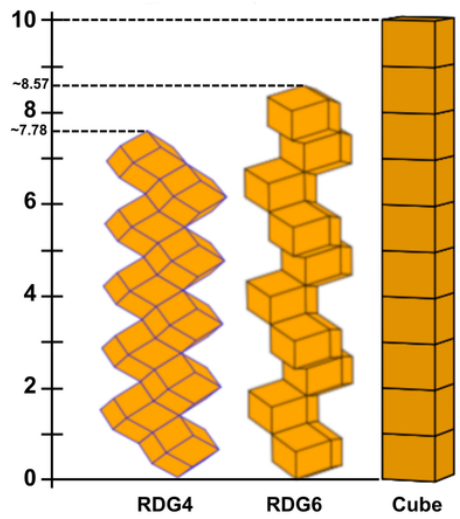


Fig. 14: The height (in sdu's) of 10 cells vertically stacked, for each grid type.

As discussed in section 6, and visualized in Figure 14, it matters whether we control for the number of cells in the vertical axis, or for the distance in the vertical axis. If we consider the number of cells, all three stacks in this example are equally tall. If we instead consider physical distance, we would need more cells for RDG4 and RDG6 to make them approximately as tall as the Cube stack.

In Figure 14 I measure the distance from the bottom vertex to the top vertex in the stack, however in the implementation of the simulation, I use the distance between the center of the top cell and the center of the bottom cell. We can describe this distance using the following functions, see Figure 6 for intuition.

$$
\begin{aligned}
h_{\mathrm{CG}}(z) &= z - 1, \\
h_{\mathrm{RDG4}}(z) &= (z-1)\frac{\sqrt{2}}{2}, \\
h_{\mathrm{RDG6}}(z) &= (z-1)\frac{\sqrt{6}}{3}.
\end{aligned}
$$

Two simulation boxes of grid types A and B are fixed to the same *physical height* when the distance according to the equations above is as similar as possible. Specifically, for a target integer physical height L, I select the number of cells in the vertical direction (z_{CG}, z_{RDG4}, z_{RDG6}) such that the physical height is as close to L as possible. I can describe the selected height dimension for a given target height L as follows:

$$
\begin{aligned}
z^*_{CG} &= L + 1 \\
z^*_{RDG4} &= \arg\min_z |h_{RDG4}(z) - L| \\
z^*_{RDG6} &= \arg\min_z |h_{RDG6}(z) - L|
\end{aligned}
$$

Two simulation boxes are fixed to the same *cell height* when they have the same number of vertical cells, $z_{CG} = z_{RDG4} = z_{RDG6}$. For a given target *cell height* C,

$$
z^*_{CG} = z^*_{RDG4} = z^*_{RDG6} = L
$$

Given lightning strike data from each grid type (CG, RDG4, and RDG6) that is conditioned on either the same cell height or physical height, I will compute the overall variance from the mean strike location and the 95% confidence interval of the variance.

The variance of a strike distribution will be calculated as the mean squared error from the mean strike position, where the error is the Euclidean distance. Where $\vec{\mu}, \vec{d}_i \in \mathbb{R}^2$ is the mean strike location and strike location of data point i respectively, and n is the number of data points, then the variance of the strike distribution is

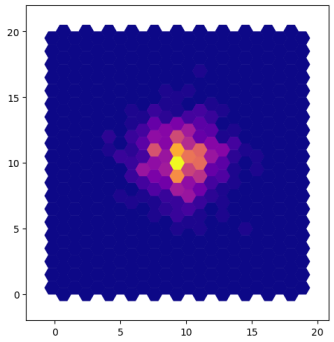
$$
\begin{aligned}
\sigma^2 &= \sum_{i=1}^{n} \frac{||\vec{\mu} - \vec{d}_i||^2}{n-1} \\
&= \sum_{i=1}^{n} \frac{(\vec{\mu}_x - \vec{d}_{i,x})^2 + (\vec{\mu}_y - \vec{d}_{i,y})^2}{n-1}
\end{aligned}
$$

I expect the strike distribution to be a bivariate normal distribution with no covariance $D \sim \mathcal{N}_2(\vec{\mu}, \Sigma = \sigma^2\mathbb{I})$, since there is no systematic directional bias in the horizontal plane, and since each strike is produced independently and with identical method I expect them to be i.i.d and follow a normal distribution. This, in turn, leads to the expectation that the variance will be distributed according to a central chi-squared distribution $\sigma^2 \sim \chi^2_n$, central since I manually center the distribution around 0. I use the central 95% probability region of the χ^2_n distribution to quantify the uncertainty in the sample variances. Note that this is not the same as a confidence interval as the interpretation is that 95% of values from the distribution fall within this range and not that 95% of such ranges contain the true population variance.

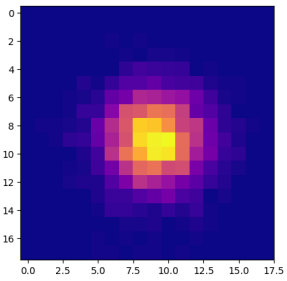
The central 95% probability region is also not the highest density interval for non-symmetrical distributions like χ^2_n. However, χ^2_n is approximately symmetric for large datasets like I will have (sample size larger than approximately 50).

8.3 Strike Distribution Results

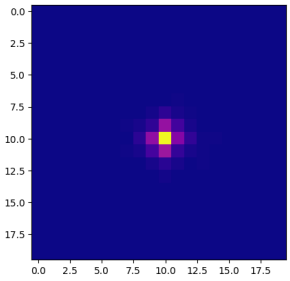
We can already see in the strike distribution plots in Figure 15 how it seems like the variance of RDG4 and RDG6 are significantly larger than the variance of the strike distribution for the CG. We can also visually get a sense that the distributions seem to have no covariance, that is, that there's no linear correlation. We can quantify this by calculating Pearson's r and it is corresponding p-value for each strike distribution. Using a threshold of $p < 0.0005$, no grid type has statistical significance in their linear correlation, with a p-value of 0.12, 0.47, and



(a) Strike distribution for RDG6 grid, 606 datapoints, box height is 9.8 sdu's (12 cells)



(b) Strike distribution for RDG4 grid, 3184 datapoints, box height is 9.9 sdu's (13 cells)



(c) Strike distribution for cube grid, 1344 datapoints, box height is 10 sdu's (10 cells)

Fig. 15: Strike distributions for each grid type, condition on the same physical height

0.31 for CG, RDG4, and RDG6, respectively. This supports the assumption that there would be no covariance in the strike data.

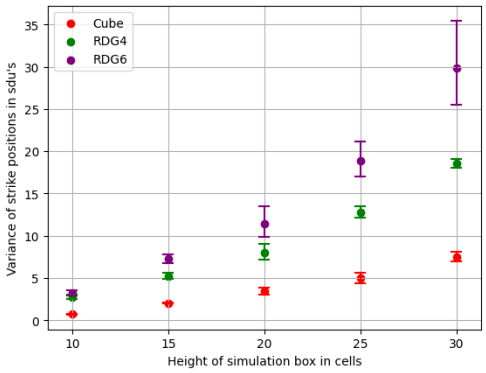


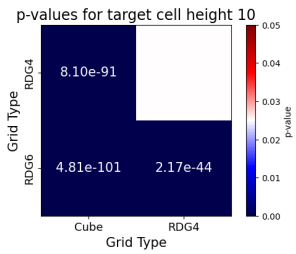
Fig. 16: Variance of lightning strike distributions for each grid type, given different heights the lightning needs to travel, in number of cells. Whiskers are central 95% probability regions over the variances.

In Figure 16 we see a distinct difference in variance across all three grid types, RDG6 having consistently the largest variance, and CG having consistently the lowest. The significance of the difference is supported by the non-overlap between the central 95% probability regions. Crucially, we observe that the difference in variance seems to increase, in absolute terms, as we increase the size of the box and hence the resolution of the simulation, which goes against the potential hypothesis that if the strike distributions have different variances at a low resolution, the difference would approach zero as we increase the resolution. This result is very significant; not only is there a distinct difference in variance between the different grids, but it seems like this difference doesn't diminish as we increase the resolution; the different grids have inherently different variances in them.

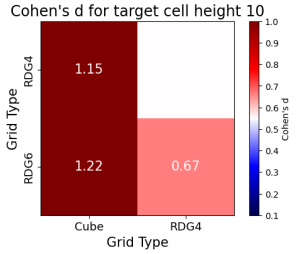
We can affirm that this difference is statistically significant ($p < 0.0005$) and practically significant by looking at Figure 17, displaying the significance values specifically for when the simulation box is of height 10, the leftmost points in Figure 16. We see in Figure 17a that for a height of 10 the p-value is, at its largest, around 40 orders of magnitude smaller than my selected threshold of 0.0005, meaning we have very high certainty that the variance is different between each grid. Remembering that Cohen's d is in standard deviation units and that it is common practice to consider the range $0.5 < d < 0.8$ to be a moderate effect size, and $0.8 < d$ to be a large effect size, we see in Figure 17b that the difference is moderate between RDG4 and RDG6, and large for the other two pairings.

Very similar results can be seen for all other simulation heights in Figure 35, p-values are extremely small, and only in three pairwise comparisons is Cohen's d smaller than 0.8, being either 0.74 or 0.76.

There is strong evidence for variance changing as we change grid type, and it seems to be at least of moderate practical significance.



(a) p-values from pairwise t-tests when target cell height is 10



(b) Cohen's d from pairwise t-tests when target cell height is 10

Fig. 17: Statistical and practical significance in the difference between each grid type, when simulation box height is 10 cells.

One may notice that the variance over increased simulation box height looks to follow an exponential curve; in fact, exponential curves fit the data very well (Figure 18). However, it is important to note that I do not have any strong prior belief that any given curve should fit the data, so there is no analytical support for the choice of exponential curves and it may not be the true relationship at play.

The exponential functions of best fit are:

$$\sigma^2_{CG} \approx 2.60e^{0.0476x} - 3.39$$
$$\sigma^2_{RDG4} \approx 7.20e^{0.0439x} - 8.45$$
$$\sigma^2_{RDG6} \approx 4.20e^{0.0705x} - 5.09$$

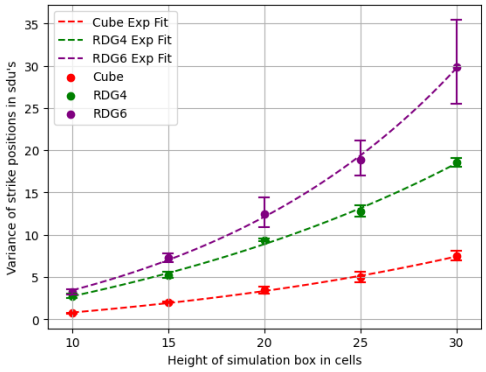


Fig. 18: Variance of lightning strike distributions for each grid type, with an exponential function fit to the variance from each grid type.

Here we need to be mindful of the indicated uncertainty for the last variance datapoints from RDG6, the exponential fit is especially sensitive to this last datapoint and hence we can't have too much confidence in the predictive power of the fitted exponential functions. However, it is noteworthy that we get approximately the same exponent for the two grid types where we have higher certainty in the variances, CG and RDG4, with a difference in the leading coefficient.

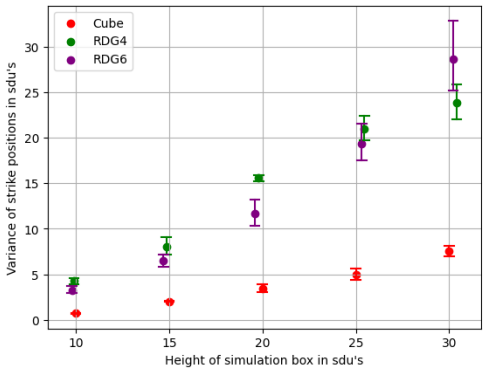


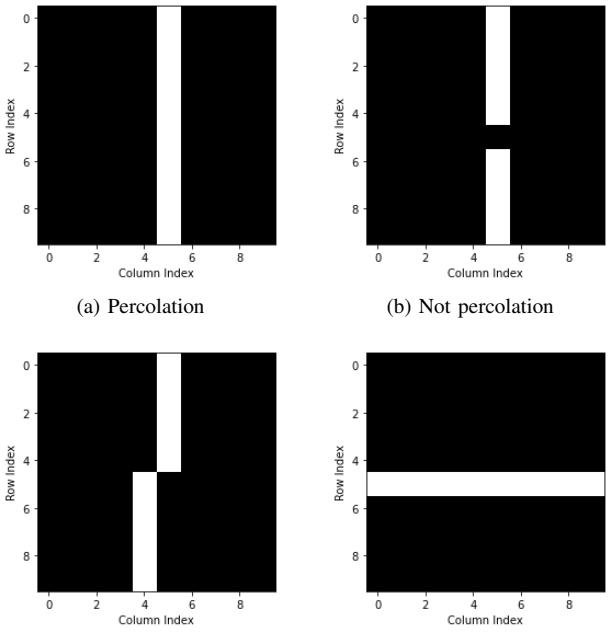
Fig. 19: Variance of lightning strike distributions for each grid type, given different heights the lightning needs to travel, in physical height (sdu). Whiskers are 95% confidence intervals over the variances.

We see in Figure 19 that there is a clear distinction between the variances in the cube grid and the two grid types using the rhombic dodecahedron cell shapes (RDG4 & RDG6), with there being much more similarity between the latter pair.

9 Percolation analysis

9.1 Method

I will produce simulation boxes where each cell is active with a probability p and inactive with a probability $1 - p$. Percolation occurs if there is a non-broken path using only "active" cells that travel from the top of the grid to the bottom. I will run the percolation models in 3D simulation boxes, but I use the 2D examples below to illustrate what percolation is.



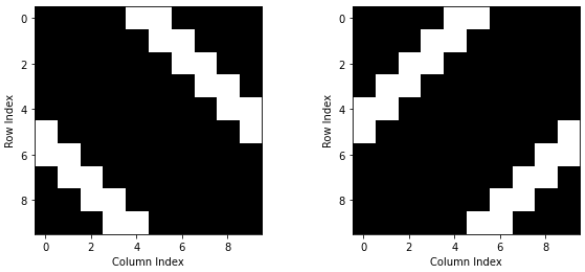
(a) Percolation (b) Not percolation

(c) Not percolation (d) Not percolation

Fig. 20: A 10×10 square grid, white cells are "active". Percolation is only considered from top to bottom, and there has to be an unbroken path of active cells.

I consider two cells to be neighbors if they share a face (edge in 2D), not if they only share a vertex; hence, Figure 20c is not an example of percolation. For the square and cube grid, this corresponds to using the Von-Neumann neighborhood rather than the Moore neighborhood.

I use periodic boundaries for all grids, as displayed in Figure 21.



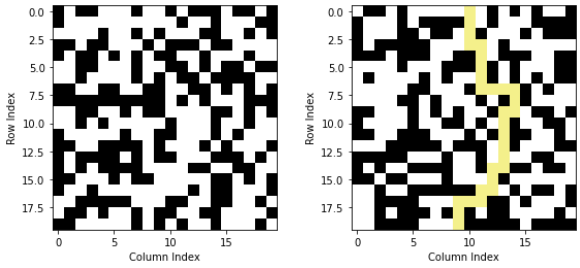
(a) Percolation only if periodic (b) Percolation only if periodic

Fig. 21: A 10×10 square grid, white cells are "active". Percolation occurs in both these cases if we let the sides be periodic, but not otherwise.

A simulation run consists of choosing a type of grid, a size N, and a probability of activation p. Each cell will randomly and independently be active with probability p; then we observe whether percolation has occurred.

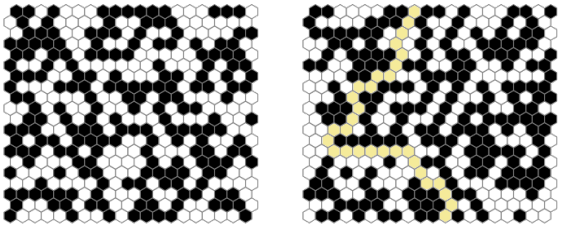
See examples of percolation on randomly generated grids on a square grid in Figure 22, and a hexagon grid in Figure 23.

The interesting outcome metric in these simulations is the critical probability p_c. it is the probability at which the percolation probability is independent of scale.



(a) Percolation did not occur (b) Percolation did occur

Fig. 22: 20×20 square grid ($N = 20$), $p = 0.55$ (white is active), and periodic boundaries.



(a) percolation did not occur (b) percolation did occurred

Fig. 23: 20×20 hexagon grid ($N = 20$), $p = 0.3$ (white is active), and periodic boundaries.

That is when the site activation probability $p = p_c$, the probability for percolation in a 4×4 grid is the same as in a 32×32 grid and, in general, an $N \times N$ grid. Hence this metric is agnostic of simulation resolution and instead only dependent on the specific graph or grid type used. The critical probability is also referred to as the site percolation threshold.

Another intuitive metric is the 50%-rate threshold ($p_{50\%}$), which is the activation probability p such that the probability of percolation is 50%. This threshold probability can be useful in practice, however it depends on the size of the simulation, which raises the question of which grid dimensions we should compare to each other. Regardless, as we increase the size of our simulation box, the 50%-rate threshold will approach the critical probability, hence it is the more useful metric to compare the intrinsic properties of the grid-types.

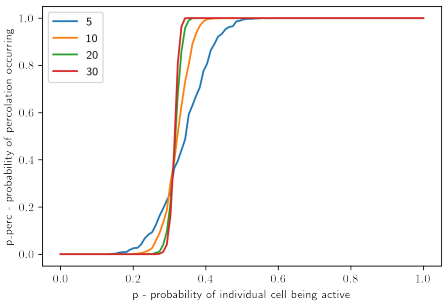
9.2 Percolation results

In Figure 24 we can see the results of the percolation simulations for the cube grid. Notice in Figure 24b that the critical probability does not occur when the probability of percolation surpasses 0.5, but rather when the probability of percolation intersects across different grid sizes, which occurs here at a cell activation probability of around $p_c \approx 0.355$.

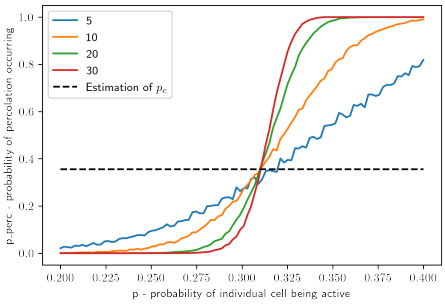
For RDG4 the critical probability is $p_c \approx 0.680$ (Figure 33) and for RDG6 it is $p_c \approx 0.417$ (Figure 34), as displayed in Figure 25.

10 Discussion

I have found that the choice of grid type, specifically between a cube grid and two orientations of the rhombic



(a) 100 different values of p in the range $(0, 1)$, 2050 simulation-runs per p-N combination



(b) 100 different values of p in the range $(0.2, 0.4)$, 5800 simulation-runs per p-N combination

Fig. 24: The probability of percolation p_c over the probability of individual cell activation p for a cube grid, with periodic boundaries. The critical percolation probability p_c is highlighted. Each line is a different size of grid, where if $N = 10$ it is a $10 \times 10 \times 10$ grid.

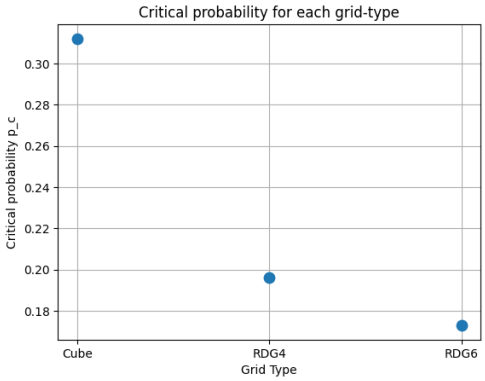


Fig. 25: The critical activation probability where the percolation probability becomes independent on grid size.

dodecahedron grid, has a significant impact on both strike distribution variance and critical percolation probability. The variance in lightning strike distribution for the DBM increases as we go from CG to RDG4 to RDG6. In the percolation model, we find that the critical percolation probability decreases as we go from CG to RDG4 to RDG6.

Importantly, these differences do not diminish as we increase the resolution of our simulation, and they are of relative practical significance. For example, the variance

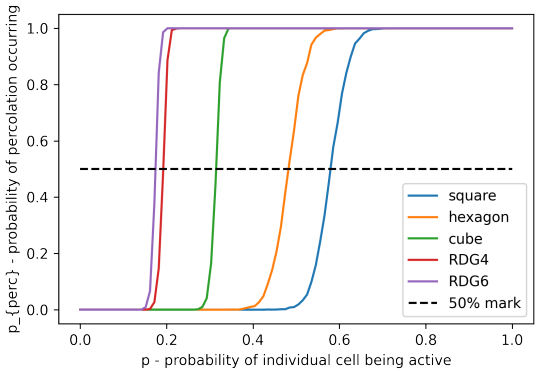


Fig. 26: Percolation probability over cell activation probability for each type of grid, at $N = 30$.

of the strike distribution in the DBM is at least 2.5 times larger for RDG4 as it is for CG, and then in turn about 1.3 times larger for RDG6 compared with RDG4. The fact that we do not see the differences decrease as we increase the resolution indicates that this is not a matter of resolution, but rather an inherent difference between the grid types. I also identified that this difference does not relate to a difference in average fractal dimensionality between the lightning structures, as this was held constant when $\eta = 6$. These results were all found with very strong statistical significance and at least moderate practical significance.

This confirms that researchers need to take much more care, discuss, and justify their grid choices when working with the DBM since it is a choice that may impact the given outcome metric they're investigating.

I also found that the grid choice impacts the critical percolation probability for the PM, a model that is, contrary to DBM, time-independent and has no global interactions. This enforces the duty of researchers to consider grid decisions and shows that grid type can be an extraneous variable in models with both local and global interactions and in either time-dependent or time-independent models.

An alternative explanation could be that it is in the nature of some models to be dependent on the grid choice when undergoing spatial discretization, and there could perhaps exist accurate models for the same tasks that do not rely on such spatial discretization. For example, the approach by L. Niemeyer in [27] uses a stepped leader model with an adapted coordinate system so as not to rely on any particular spatial discretization. Perhaps these are the types of models we need to rely on to avoid grid types as an extraneous variable.

The results of this project are very significant. I have compared the cube grid and two orientations of the rhombic dodecahedron grid under the dielectric breakdown model and the percolation model. The results show with high confidence that the choice of spatial discretization does indeed matter for physical simulations.

This prompts further research into what other models are grid-dependent, what makes a model grid-dependent, and which methods we can use to avoid the influence of grid choice.

References

[1] T. Kim and M. C. Lin, "Physically based animation and rendering of lightning," *CiteSeer X (The Pennsylvania State University)*, 11 2004.

[2] N. I. Petrov, G. N. Petrova, and F. D'Alessandro, "Quantification of the probability of lightning strikes to structures using a fractal approach," *IEEE Transactions on Dielectrics and Electrical Insulation*, vol. 10, pp. 641–654, 08 2003.

[3] R. Jiang, W. Lyu, B. Wu, Q. Qi, Y. Ma, Z. Su, S. Wu, Z. Xie, and Y. Tan, "Simulation of cloud-to-ground lightning strikes to structures based on an improved stochastic lightning model," *Journal of Atmospheric and Solar-Terrestrial Physics*, vol. 203, 07 2020.

[4] J. He, X. Zhang, L. Dong, R. Zeng, and Z. Liu, "Fractal model of lightning channel for simulating lightning strikes to transmission lines," *Science in China Series E: Technological Sciences*, vol. 52, pp. 3135–3141, 11 2009.

[5] M. Games, "The fantastic combinations of john conway's new solitaire game "life" by martin gardner," *Scientific American*, vol. 223, pp. 120–123, 1970.

[6] P. D. Beale and P. M. Duxbury, "Theory of dielectric breakdown in metal-loaded dielectrics," *Physical Review B*, vol. 37, pp. 2785–2791, 02 1988.

[7] D. R. Bowman and D. Stroud, "Model for dielectric breakdown in metal-insulator composites," *Physical Review B*, vol. 40, pp. 4641–4650, 09 1989.

[8] A. A. Dul'zon, V. V. Lopatin, D. Noskov, and O. I. Pleshkov, "Modeling the development of the stepped leader of a lightning discharge," *Technical Physics*, vol. 44, pp. 394–398, 04 1999.

[9] L. Niemeyer, L. Pietronero, and H. Wiesmann, "Fractal dimension of dielectric breakdown," *Physical Review Letters*, vol. 52, pp. 1033–1036, 03 1984.

[10] S. Satpathy, "Dielectric breakdown in three dimensions: Results of numerical simulation," *Physical Review B*, vol. 33, pp. 5093–5095, 04 1986.

[11] S. Doorwar and K. K. Mohanty, "Extension of the dielectric breakdown model for simulation of viscous fingering at finite viscosity ratios," *Physical Review E*, vol. 90, 07 2014.

[12] J. Vicente, A. Razzitte, and E. Mola, "Fractal characteristics of dielectric breakdown," *Proceedings of IEEE Conference on Electrical Insulation and Dielectric Phenomena - (CEIDP'94)*, pp. 524–531, 10 1994.

[13] C. P. Birch, S. P. Oom, and J. A. Beecham, "Rectangular and hexagonal grids used for observation, experiment and simulation in ecology," *Ecological Modelling*, vol. 206, pp. 347–359, 08 2007.

[14] V. E. Brimkov and R. P. Barneva, ""honeycomb" vs square and cubic models," *Electronic notes in theoretical computer science*, vol. 46, pp. 321–338, 08 2001.

[15] C. Wüthrich and P. Stucki, "An algorithmic comparison between square- and hexagonal-based grids," *CVGIP: Graphical Models and Image Processing*, vol. 53, pp. 324–339, 07 1991.

[16] J. A. Kimerling, K. Sahr, D. White, and L. Song, "Comparing geometrical properties of global grids," *Cartography and Geographic Information Science*, vol. 26, pp. 271–288, 01 1999.

[17] H. Miura and M. Kimoto, "A comparison of grid quality of optimized spherical hexagonal–pentagonal geodesic grids," *Monthly Weather Review*, vol. 133, pp. 2817–2833, 10 2005.

[18] L. Wang and T. Ai, "The comparison of drainage network extraction between square and hexagonal grid-based dem," *The International Archives of the Photogrammetry, Remote Sensing and Spatial Information Sciences*, vol. XLII-4, pp. 687–692, 09 2018.

[19] A. T. Nugraha, B. J. Waterson, S. P. Blainey, and F. J. Nash, "On the consistency of urban cellular automata models based on hexagonal and square cells," *Environment and Planning B Urban Analytics and City Science*, vol. 48, pp. 845–860, 01 2020.

[20] B. Hamilton and S. Bilbao, "Hexagonal vs. rectilinear grids for explicit finite difference schemes for the two-dimensional wave equation," *The Journal of the Acoustical Society of America*, vol. 133, pp. 3532–3532, 05 2013.

[21] T. Mir, P. Yao, K. Duranceau, and I. Prémont-Schwarz, "Are grid cells hexagonal for performance or by convenience?" *arXiv (Cornell University)*, 10 2024.

[22] Cyp, Wikipedia, 06 2004. [Online]. Available: https://en.wikipedia.org/wiki/File:Rhombicdodecahedron.jpg

[23] T. Ruen, Wikipedia, 05 2016. [Online]. Available: https://en.wikipedia.org/wiki/Bilinski_dodecahedron#/media/File:Bilinski_dodecahedron.png
[24] R. Itza Balam, M. Uh Zapata, and U. Iturrarán-Viveros, "Hexagonal finite differences for the two-dimensional variable coefficient poisson equation," *Examples and Counterexamples*, vol. 5, p. 100144, 06 2024.
[25] T. Kim, J. Sewall, A. Sud, and M. C. Lin, "Fast simulation of laplacian growth," *IEEE Computer Graphics and Applications*, vol. 27, pp. 68–76, 03 2007.
[26] J. Sañudo, J. B. Gómez, F. Castaño, and A. F. Pacheco, "Fractal dimension of lightning discharge," *Nonlinear processes in geophysics*, vol. 2, pp. 101–106, 06 1995.
[27] L. Niemeyer, "A stepped leader random walk model," *Journal of Physics D: Applied Physics*, vol. 20, pp. 897–906, 07 1987.

11 Numerical Appendix

General discrete Laplacian operator

When I approximate the solution to the Laplace equation for the RDGs in this paper, I let each cell equal the average of their neighborhood. In doing this I rely on the claim that this notion holds true for a general graph, not only the regular square case. I use this section to motivate this claim.

When working with functions that are differentiable and defined on a continuous, smooth space (like $\mathbb{R}^n$), we can apply the Taylor series to isolate the second derivative term and approximate it using nearby samples from the function. Letting the given grid define the samples of this space.

However, when we wish to move on to general graphs, this doesn't work anymore since graphs are not smooth manifolds. There is no guarantee of a consistent coordinate system, nor a notion of the infinitesimal displacement, h, used in the Taylor series.

We must hence use the discrete Laplacian $\Delta^2(\cdot)$, which is analogous to the continuous Laplacian $\nabla^2(\cdot)$ found using the Taylor series, except instead of letting our "samples" exist on a coordinate system with infinitesimal displacement h, we let the samples be neighboring nodes on a grid. One may also consider the discrete Laplacian the "coordinate-free" analog of the continuous Laplacian.

The finite difference approximation of the continuous Laplacian operator,

$$\nabla^2 u(x,y) \approx \frac{1}{h^2}(u_{i+1,j} + u_{i-1,j} + u_{i,j+1} + u_{i,j-1} - 4u_{i,j}),$$

is analogous to the discrete Laplacian operator Δ^2, which is simply using a different set of "samples", now on a non-smooth space,

$$\Delta^2 u_i = \sum_{j \in \mathcal{N}(i)} (u_j - u_i) = \sum_{j \in \mathcal{N}(i)} u_j - \deg(i) \cdot u_i.$$

u_i is the value at node i, and $\mathcal{N}(i)$ are all nodes connected to node i by an edge.

Notice how the interpretation of the Laplacian as a measure of deviation from the local neighborhood average, or equivalently, curvature at this point/node, is maintained. It is also easy to see how the finite difference approximation is a special case of the discrete Laplacian, one where the graph happens to be a square lattice.

Also, notice how the continuous Laplacian operator is only approximated by the finite difference formula since the domain is continuous, and we only reach the exact value for the second derivative as the displacement h approaches 0. While the discrete Laplacian operator is described exactly as the sum of the differences in its neighborhood since the "local neighborhood" is now defined exactly.

We now have a notion of the Laplacian operator on a general graph, which is equal to the finite difference estimations if the graph is a regular grid. We can now also conclude that on a general undirected graph G, the solution to the Laplace equation is still that every node is equal to the average of its neighborhood,

$$\begin{aligned}
\Delta^2 u_i &= 0 \\
\sum_{j \in \mathcal{N}(i)} u_j - \deg(i) \cdot u_i &= 0 \\
\deg(i) \cdot u_i &= \sum_{j \in \mathcal{N}(i)} u_j \\
u_i &= \frac{1}{\deg(i)} \sum_{j \in \mathcal{N}(i)} u_j
\end{aligned}$$

This is enough to convince ourselves that there is a corresponding operator on a general undirected graph, which supports the notion of finding the values making each cell equal to the average of its neighborhood.

One can also express the operator in matrix form, using the Graph Laplacian Matrix.

Let A be the adjacency matrix for a general undirected graph G: $A_{ij} = 1$ if nodes i and j are connected. Let D be the diagonal degree matrix: $D_{ii} = \deg(i)$, 0 on off-diagonals. $\vec{u} \in \mathrm{R^n}$ be the vector of values at each node. Then the graph Laplacian matrix $L = D - A$, $Lu = (D - A)u$, and

$$(Lu)_i = \sum_{j \in \mathcal{N}(i)} (u_i - u_j).$$

We can then set the Laplacian matrix to 0 to extract the Laplace equation:

$$Lu = 0 \iff u_i = \frac{1}{\deg(i)} \sum_{j \in \mathcal{N}(i)} u_j$$

For further curiosity, there are limits at which a graph may approximate a smooth manifold, under which the graph Laplacian converges to the Laplace-Beltrami operator, which *is* the continuous Laplacian on the manifold.

12 Figure Appendix

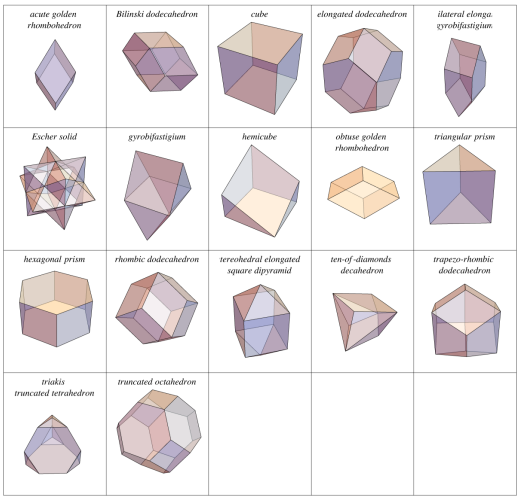
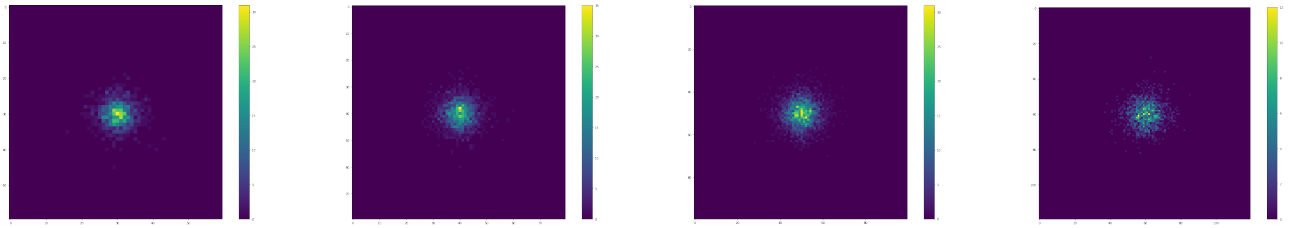


Fig. 27: The set of all polyhedra which are space-filling https://mathworld.wolfram.com/Space-FillingPolyhedron.html

η	p-value	Cohen's d
0	$2.69 \cdot 10^{-8}$	3.26
1	$1.87 \cdot 10^{-7}$	4.39
2	0.490	0.26
3	0.865	0.07
4	0.315	0.26
5	0.252	0.47
6	0.445	0.32

TABLE I: Statistical significance and effect size in difference of mean tests for average fractal dimensionality.



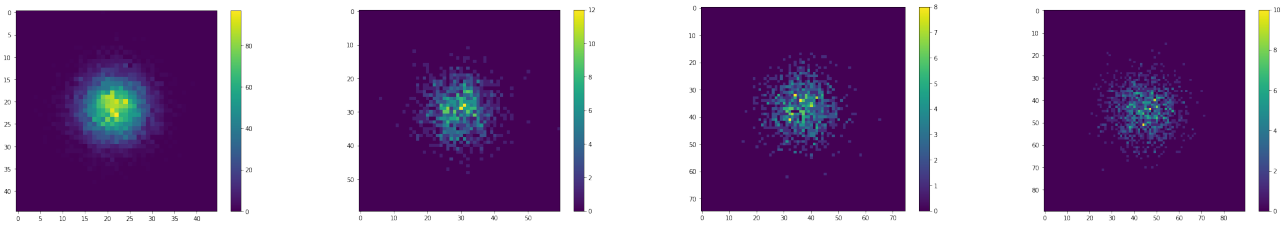
(a) $60 \times 60 \times 30$ box, $\eta = 6$, $\alpha = 0.001$, 1156 samples

(b) $80 \times 80 \times 40$ box, $\eta = 6$, $\alpha = 0.001$, 2130 samples

(c) $100 \times 100 \times 50$ box, $\eta = 6$, $\alpha = 0.001$, 3579 samples

(d) $120 \times 120 \times 60$ box, $\eta = 6$, $\alpha = 0.001$, 1456 samples

Fig. 28: Lightning strike location distributions for different simulation resolutions, using CG



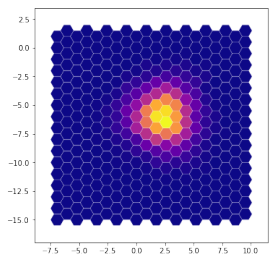
(a) $45 \times 45 \times 30$ box, $\eta = 6$, $\alpha = 0.001$, 10467 samples

(b) $60 \times 60 \times 40$ box, $\eta = 6$, $\alpha = 0.001$, 1295 samples

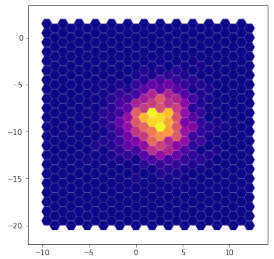
(c) $75 \times 75 \times 50$ box, $\eta = 6$, $\alpha = 0.001$, 1185 samples

(d) $90 \times 90 \times 60$ box, $\eta = 6$, $\alpha = 0.001$, 1516 samples

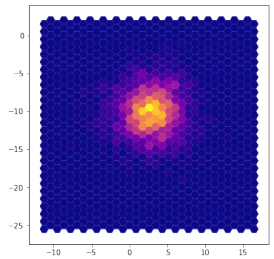
Fig. 29: Lightning strike location distributions for different simulation resolutions, using RDG-4



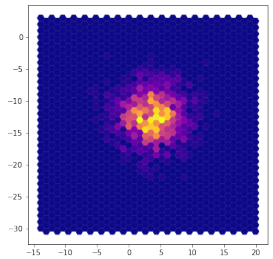
(a) $30 \times 30 \times 11$ box, $\eta = 6$, $\alpha = 0.001$, 9819 samples



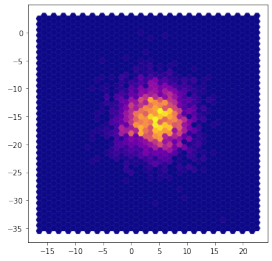
(b) $40 \times 40 \times 14$ box, $\eta = 6$, $\alpha = 0.001$, 2874 samples



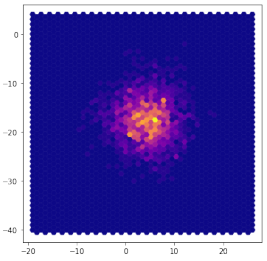
(c) $50 \times 50 \times 18$ box, $\eta = 6$, $\alpha = 0.001$, 2631 samples



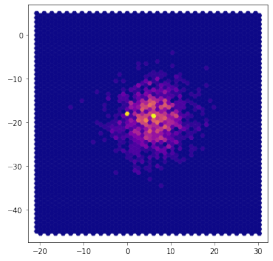
(d) $60 \times 60 \times 21$ box, $\eta = 6$, $\alpha = 0.001$, 1465 samples



(e) $70 \times 70 \times 25$ box, $\eta = 6$, $\alpha = 0.001$, 2842 samples

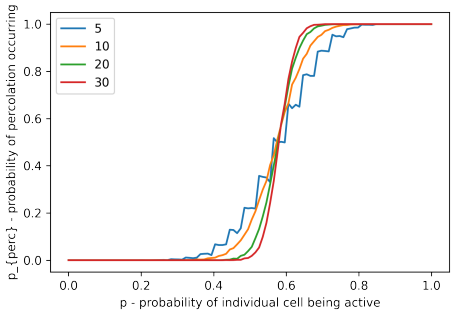


(f) $80 \times 80 \times 28$ box, $\eta = 6$, $\alpha = 0.001$, 1770 samples

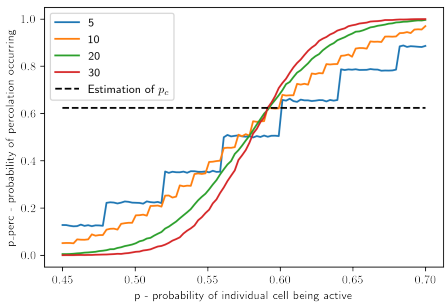


(g) $90 \times 90 \times 32$ box, $\eta = 6$, $\alpha = 0.001$, 1112 samples

Fig. 30: Lightning strike location distributions for different simulation resolutions, using RDG-6

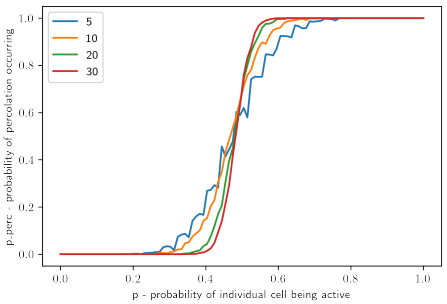


(a) 100 different values of p in the range $(0, 1)$, 3800 simulation-runs per p-N combination

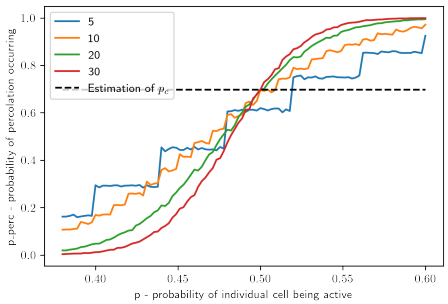


(b) 100 different values of p in the range $(0.45, 0.7)$, 30100 simulation-runs per p-N combination

Fig. 31: The probability of percolation p_{perc} over the probability of individual cell activation p for square grid, with periodic boundaries. The estimation for the critical probability p_c is highlighted. Each line is a different size of grid, where if $N = 10$ it is a 10×10 grid.

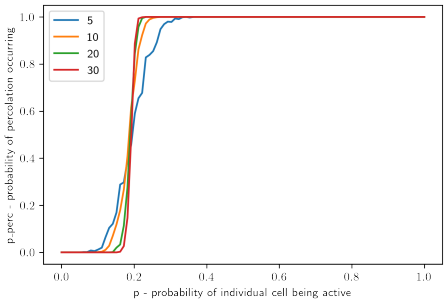


(a) 100 different values of p in the range $(0, 1)$, 1100 simulation-runs per p-N combination

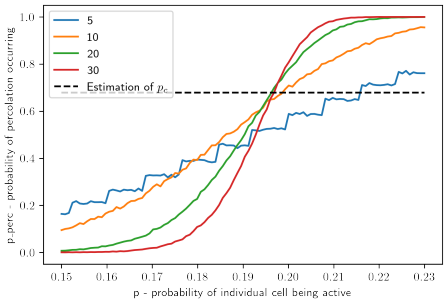


(b) 100 different values of p in the range $(0.38, 0.6)$, 6200 simulation-runs per p-N combination

Fig. 32: The probability of percolation p_{perc} over the probability of individual cell activation p for a hexagonal grid, with periodic boundaries. The estimation for the critical probability p_c is highlighted. Each line is a different size of grid, where if $N = 10$ it is a 10×10 grid.

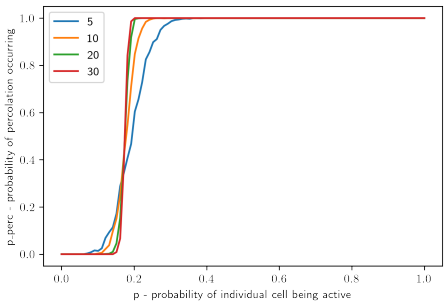


(a) 100 different values of p in the range $(0, 1)$, 2050 simulation-runs per p-N combination

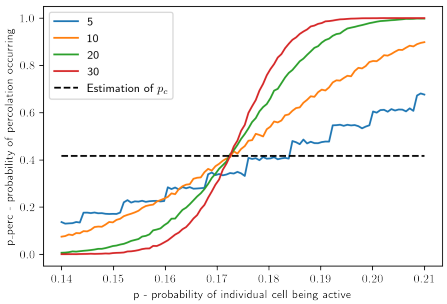


(b) 100 different values of p in the range $(0.15, 0.23)$, 11500 simulation-runs per p-N combination

Fig. 33: The probability of percolation p_{perc} over the probability of individual cell activation p for a RDG-4 grid, with periodic boundaries. The 50% line is highlighted. Each line is a different size of grid, where if $N = 10$ it is a $10 \times 10 \times 10$ grid.

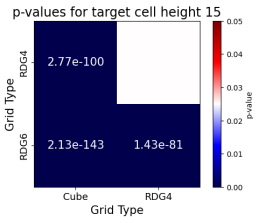


(a) 100 different values of p in the range $(0, 1)$, 1000 simulation-runs per p-N combination

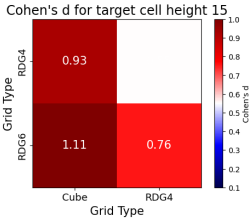


(b) 100 different values of p in the range $(0.14, 0.21)$, 15000 simulation-runs per p-N combination

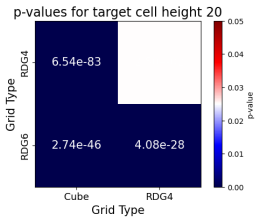
Fig. 34: The probability of percolation p_{perc} over the probability of individual cell activation p for a RDG-6 grid, with periodic boundaries. The 50% line is highlighted. Each line is a different size of grid, where if $N = 10$ it is a $10 \times 10 \times 10$ grid.



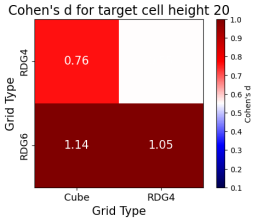
(a) p-values from pairwise t-tests for a target height of 15



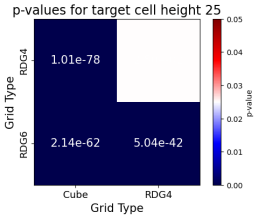
(b) Cohen's d from pairwise t-tests for a target height of 15



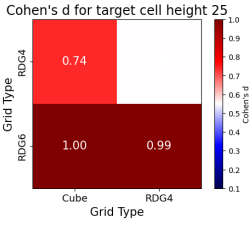
(c) p-values from pairwise t-tests for a target height of 20



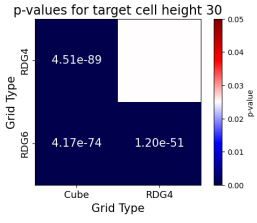
(d) Cohen's d from pairwise t-tests for a target height of 20



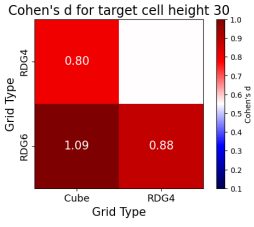
(e) p-values from pairwise t-tests for a target height of 25



(f) Cohen's d from pairwise t-tests for a target height of 25



(g) p-values from pairwise t-tests for a target height of 30



(h) Cohen's d from pairwise t-tests for a target height of 30

Fig. 35: All p-values and Cohen's d from each simulation box height visualized in Figure 16.