5-2018

Self-Organized Structures: Modeling Polistes dominula Nest Construction with Simple Rules

Matthew Harrison
East Tennessee State University

Self-Organized Structures: Modeling *Polistes dominula* Nest Construction with Simple Rules

---

A thesis

presented to

the faculty of the Department of Computing

East Tennessee State University

In partial fulfillment

of the requirements for the degree

Master of Science in Computer and Information Sciences

---

by

Matthew Steven Harrison

May 2018

Christopher Wallace, Ph.D., Co-chair

Istvan Karsai, Ph.D., Co-chair

Martin Barrett, Ph.D.

Keywords: Computational Modeling, Swarm Robotics, Self-organization, Stigmergy

ABSTRACT

Self-Organized Structures: Modeling *Polistes dominula* Nest Construction with Simple Rules

by

Matthew S. Harrison

The self-organized nest construction behaviors of European paper wasps (*Polistes dominula*) show potential for adoption in artificial intelligence and robotic systems where centralized control proves challenging. However, *P. dominula* nest construction mechanisms are not fully understood. This research investigated how nest structures stimulate *P. dominula* worker action at different stages of nest construction. A novel stochastic site selection model, weighted by simple rules for cell age, height, and wall count, was implemented in a three-dimensional, step-by-step nest construction simulation. The simulation was built on top of a hexagonal coordinate system to improve precision and performance. Real and idealized nest data were used to evaluate simulated nests via two parameters: outer wall counts and compactness numbers. Structures generated with age-based rules were not significantly different from real nest structures along both parameters.

DEDICATION

I dedicate this thesis to my parents, Michael and Suzan Harrison, for the unconditional love, encouragement, and support they have provided me throughout my life.

ACKNOWLEDGEMENTS

I wish to extend my sincerest thanks to my thesis advisor and committee co-chair Dr. Christopher Wallace for introducing me to this area of research and for encouraging me to pursue a thesis for my capstone project. Dr. Wallace's ideas on how to overcome hurdles encountered in the research process proved invaluable in the creation and engineering of the software simulations. I also thank Dr. Wallace for his continued attention, support, and patience in the completion of this thesis project.

I also wish to extend my deepest gratitude to my thesis advisor and committee co-chair Dr. Istvan Karsai for sharing his passion for research with me. Dr. Karsai's willingness to patiently and repeatedly explain the underlying biological principles to me was instrumental in the production of this research report. I am especially grateful for Dr. Karsai helping me maintain focus and direction during this research project and for his prompt responses and edits for any of my queries and draft documents.

Special thanks to Dr. Martin Barrett for agreeing to serve on my thesis committee and as an editor for this report. I appreciate Dr. Barrett's prompt responses and his making time during a busy semester to help revise this report.

Finally, thanks to all of my thesis committee for helping me keep a sense of perspective during this project.

TABLE OF CONTENTS

Page

LIST OF FIGURES

Figure Page

CHAPTER 1

INTRODUCTION

Swarm robotics is a branch of robotics that focuses on coordinating groups of simple robots with algorithms inspired by research performed on different species of social insects (Şahin 2004). One area of active research in swarm robotics is achieving group coordination using models based on self-organizing behaviors of ants (Rosalie, et al. 2017) via ant colony optimization (Dorigo 2011) and fireflies (De Rango, et al. 2015). Related research on termites (Mizumoto, Kobayashi and Matsuura 2015) and bees (Johnson 2009) focuses on self-organized construction behaviors performed by each type of insect. Other social insects, such as paper wasps, also demonstrate self-organized construction behaviors. A previous study (Karsai and Pénzes 1996) of nests built by European paper wasps (*Polistes dominula*) found *P. dominula* construct compact and complex structures with several properties that stay consistent across nests of varying size, including outer wall and buildable site ratios. Understanding how *P. dominula* coordinate construction behaviors without centralized control helps further knowledge of the biology behind other insect societies and could lead to further research into decentralized control algorithms in swarm robotics.

Mechanisms for nest construction by paper wasps are not fully understood. A previous investigation (Karsai and Pénzes 1993) into *P. dominula* nest construction created a three-dimensional simulation to model nest construction behaviors with probabilities assigned to different worker actions. A later two-dimensional model was built and investigated the role simple rules play in *P. dominula* workers selecting initiation locations (Karsai and Pénzes 2000).

The overall goal of this research was to develop a novel model for how *P. dominula* construct nests by combining aspects of the two previous models from (Karsai and Pénzes 1993)

and (Karsai and Pénzes 2000) into a stepwise, three-dimensional simulation. The simulation applied simple rules to a stochastic site selection model and two nest parameters, outer wall counts and compactness numbers, were recorded for all simulated nests for comparison against randomly-generated nests, idealized nests, and real nests from (Karsai and Pénzes 1996). Design and implementation of a hexagonal coordinate system was first completed to serve as the foundation of the three-dimensional simulation. Nests generated with the three-dimensional simulation were visually inspected to verify realistic structures; discovery of unnatural structures through visual inspection led to the development of building constraints applicable to site selection in all rules. Visual inspection also helped determine which rules generated structures like those found in nature, which helped narrow down rules for statistical analysis by outer wall counts and compactness numbers.

The rest of this report is divided into five chapters. Chapter 2 provides background information about self-organization, stigmergy, and *Polistes dominula*. Chapter 3 describes the foundational work on a hexagonal coordinate system for use with the new model. Chapter 4 details the methods for the creation of the three-dimensional simulation and stochastic model. Chapter 5 reports the results obtained from nests generated from the use of simple rules in the simulation. Chapter 6 discusses the results and suggests directions for future work.

CHAPTER 2

BACKGROUND

2.1: Self-Organization

Self-organizational behaviors are found in many types of social animals. Birds, insects, and mammals create and maintain systems to fulfill some purpose. While humans create plans, leadership hierarchies, and control systems, social insects, such as ants, termites, and wasps, build and maintain systems with no overall plan, no central leadership, and no understanding of the system as a whole (Middleton and Latty 2016). Instead, these insect species rely on individual workers making decisions and completing actions using self-organizing systems. Social insect workers follow a simple set of rules to make decisions. These rules start with some form of local input, such as the physical characteristics of a worker's location. Using this input, the worker then follows a simple algorithm to arrive at decisions. The workers take the action, which transitions the system from one state to another state. Worker insects do not keep any history of actions taken or states observed; workers act based only on the current state of the system. No recorded history, paired with local environmental cues, allow multiple workers to act in parallel without any sort of centralized leadership controlling the system (Karsai 1999).

Self-organization, as a modern concept in biology, was inspired by works from the Renaissance and the early modern era. Self-organization was first described from a theoretical point of view (Descartes 1637) and from the perspective of how disorder tends to increase in an isolated thermodynamic system (Clausius 1851). The modern definition of self-organization focuses on the underlying parts of a system: a dynamic system will tend towards an equilibrium determined by the environment generated by the system's own member subsystems (Ashby

1947). The subsystems of a system place constraints on the entire system and promote state transitions towards a point of equilibrium (Ashby 1947).

2.1.1: Positive Feedback Loops

An attractor is a system state that drives a system forward towards that attractor state, such as equilibrium. As the system's subsystems influence each other towards the attractor state, a positive feedback loop forms: changes of one subsystem to the attractor state causes other subsystems to also change to the attractor state.

Positive feedback loops take several forms, one of which is based on network effects. A network effect describes how the value of a system is impacted by the number of individuals using the system (Shapiro and Varian 1998). An example of the network effect is the telephone: as more users use the telephone system, the more valuable a telephone is to an individual user, which then attracts additional users to the telephone system in a positive feedback loop. A more modern example of this type of network effect is a social network gaining popularity as more users join the network.

As these positive feedback loops create synergistic dependencies between subsystems, new behaviors and system structures appear. This process is known as emergence and describes any system where some higher-level entity comes into being through the interactions of lower-level parts of the whole. One example of emergence is formation of snowflakes through the arrangement of ice crystals. A second example of emergence is termite colony building; a study of termites that modified two attractor state parameters resulted in emergent construction behaviors (Mizumoto, Kobayashi and Matsuura 2015).

2.2 Stigmergy

Stigmergy is a mechanism in which the results of prior actions are the stimulus for new actions. The term was coined to describe observations of termite workers stimulated to act by way of other termites' prior efforts (Grassé 1959). Stigmergy also serves as a mechanism for indirect communication of information. Indirect communication was observed among ant workers: as workers completed a task, other workers would observe the system's changed state and take further action based on those changes (Huber 1861). Other researchers noted the phenomenon in their own work: The terms "indirect social interactions" (Michener 1974) and "sematectonic communication" (Wilson 2000) describe the same phenomena as stigmergy.

2.2.1: Application

Researchers have applied stigmergy to robotic systems to do useful work. A set of experiments were performed where up to five robots, using simple algorithms, gathered randomly-distributed items into piles (Beckers, Holland and Deneubourg 1994). Compared to a single agent, gathering efficiency increased as up to two additional agents were added; however, efficiency decreased when a fourth or fifth agent were added due to additional time-intensive interactions between the robotic agents (Beckers, Holland and Deneubourg 1994).

Inspired by behavior observed in ants sorting their brood and the work in (Beckers, Holland and Deneubourg 1994), another experiment created robots that sort two different Frisbee types using algorithms to account for seven different possible stimuli (Holland and Melhuish 1999). These robots possessed only the ability to discern between the two Frisbee types and did not contain any memory or any ability to orient in space; it was determined that the algorithms,

as implemented, did not always provide good sorting but was a candidate for further investigations (Holland and Melhuish 1999).

Recent work used stigmergic feedback to enable agent communication and coordination in emergency situations, such as a bomb defusal scenario (Eleftherakis, et al. 2015). Additional work incorporated stigmergic models in ad-hoc communications between failure-prone agents conducting terrain exploration (Rodriguez, Gomez and Diaconescu 2015). Stigmergic round robin was utilized amongst swarming agents to establish Mobile Ad-hoc Networks across regular, irregular, and dynamic network topologies (Fraser and Hunjet 2016).

2.3 *Polistes dominula* Nest Construction

European paper wasps (*P. dominula*) are social wasps that build nest structures to rear offspring. Each nest consists of a comb composed of hexagonal cells and at least one petiole (stalk that anchors the nest to another surface). Individual hexagonal cells contain one offspring.

The stimuli that influence *P. dominula* nest construction behaviors are not fully understood. While earlier work on wasps proposed a plan-based building mechanism for wasp species that build with a single worker (Smith 1978), plan-based building will prove difficult in larger colonies that require coordination between workers (Camazine 2003), such as those constructed by *P. dominula* (Karsai 1999).

2.3.1: Stigmergic Influence

Stigmergy is one possible explanation for nest-building behaviors observed in wasp genus *Polistes* (Karsai 1999). However, not all the research agrees on the role stigmergy plays at various stages of nest construction. One proposal in (Downing and Jeanne 1988) stated that

Polistes nest-building behaviors are initially split into two phases: work completed before the nest reaches two cells in size and work completed after. Other research found that *Polistes* create two-celled nests via multiple pathways and nest structures: Different stimuli, such as gravity and wasp leg positioning on the petiole, appear to play a role in when workers switch from petiole building to cell construction (Karsai and Theraulaz 1995).

Research into termite construction behaviors proposed that stigmergy fails to describe when nest-construction behaviors by social insects should end (Stuart 1967), a view also put forth in (Downing and Jeanne 1988) regarding *Polistes*. However, later research showed social insect nests are not limited in shape, size, or age, but stigmergy cannot account for all nest-building processes, i.e., there are other factors unrelated to nest structure that have roles in nest-building behaviors (Karsai 1999).

Separate work on termites found that stigmergy does not account for nest repair and reconfigurations (Harris and Sands 1965), which was also applied to *Polistes* in (Downing and Jeanne 1988). More recent work investigating *Polistes* found that workers appear to behave in the same manner, regardless of the type of construction activity (Karsai and Theraulaz 1995).

2.3.2: Modeling Construction Behaviors

Researchers have worked towards a better understanding of *P. dominula* nest construction behaviors by building simulations to evaluate the effectiveness of nest construction models in producing lifelike structures. One such model was created as a complete simulation of building behavior from pulp foraging to depositing pulp on the nest (Karsai and Pénzes 1993). Static probabilities were used to represent stimulus for different wasp actions, e.g., the agent had a lower chance to deposit pulp at a cell site with two or fewer walls versus cell sites with three or

more walls (Karsai and Pénzes 1993). Later work used Markov chains to select cell initiation sites based on simple "rules of thumb" (Karsai and Pénzes 2000). An earlier model (Karsai and Pénzes 1998) found that individual construction rules can lead to well-formed nest structures; the two-dimensional model described in (Karsai and Pénzes 2000) determined the best "rule of thumb" for constructing naturally occurring nest forms with few non-natural forms was an age-based rule that summed neighboring cell ages where each eligible initiation site had a weight assigned to it based on the calculated summed ages of neighboring cells.

More recent work (Adoe 2010) extended the simple rules from (Karsai and Pénzes 2000) to a simulation that investigated all the possible nest forms producible with the use of different simple rules. Generated nests were evaluated across five variables; however, the limits of precision inherent to primitive data types caused difficulty in discerning between distinct nest configurations, requiring the use of arbitrarily large number libraries that significantly impacted simulation performance (Adoe 2010).

CHAPTER 3

HEXAGONAL COORDINATES

The overall goal of this research was to investigate the role simple rules play in a step-by-step, three-dimensional simulation of nest construction and evaluate if generated nest structures resemble naturally-occurring nest structures. However, the precision issues observed in previous work (Adoe 2010) necessitated the development of some alternative system that addressed precision issues inherent to Cartesian coordinates before building the simulation and running the experiment. Additionally, any alternate coordinate system to Cartesian coordinates should allow comparison of simulated data against real nest data. A hexagonal coordinate system addresses precision issues inherent to Cartesian coordinates and calculations derived from hexagonal coordinates, such as compactness, are directly comparable to calculations from Cartesian-mapped nests, e.g., real nest data obtained in (Karsai and Pénzes 1996).

3.1: Design

The core idea that led to an integer-based hexagonal coordinate system was eliminating as many irrational numbers as possible from the simulation. Irrational numbers were the root cause of the precision issue in (Adoe 2010): mapping hexagons to a Cartesian plane, which requires an irrational coefficient of $\sqrt{3}$ (Fig. 1), results in approximations that affect all calculations that involve coordinates, such as the sum of squared distances calculation. Floating point approximations further affect compactness number calculations, which requires taking the square root of the sum of squared distances. A hexagonal coordinate system addresses the precision issues by using integers in place of irrational numbers.

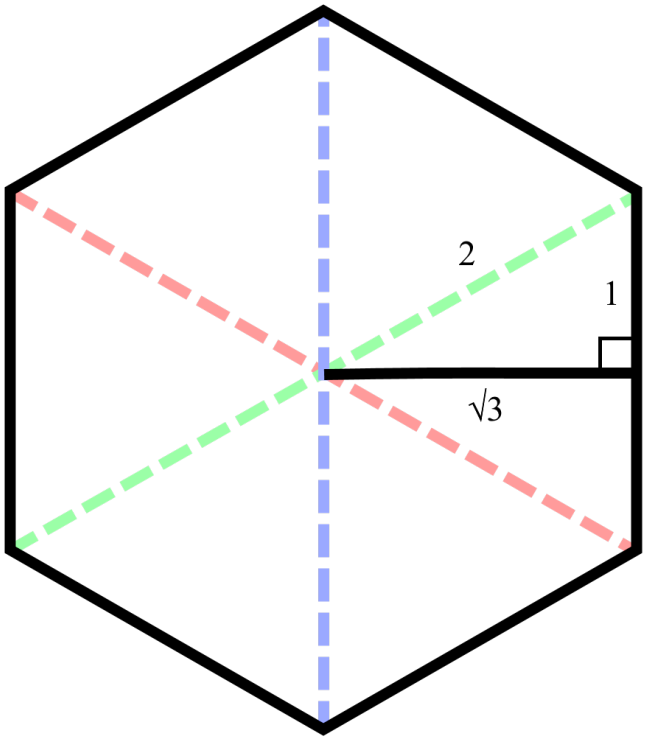

Figure 1: Irrational coefficient in Cartesian coordinates.

3.1.1: Hexagonal Coordinate System

The hexagonal coordinate system was designed to answer the following questions:

- Can the precision issue be addressed in a way that allows sufficiently sensitive nest comparison without resorting to arbitrarily large number libraries?
- Will switching to an integer-based system improve performance of a nest generating simulation?
- Can measurements and calculations made in hexagonal coordinates be directly compared to measurements and calculations made in Cartesian coordinates?

A hexagonal coordinate system uses three axes instead of two to map a hexagonal lattice to a two-dimensional plane, as seen in Figure 2. The Cartesian x-axis is split into two separate axes: a hexagonal x-axis that is 30° below the Cartesian x-axis and a hexagonal y-axis that is 30° above the Cartesian x-axis. The Cartesian y-axis becomes the hexagonal z-axis.

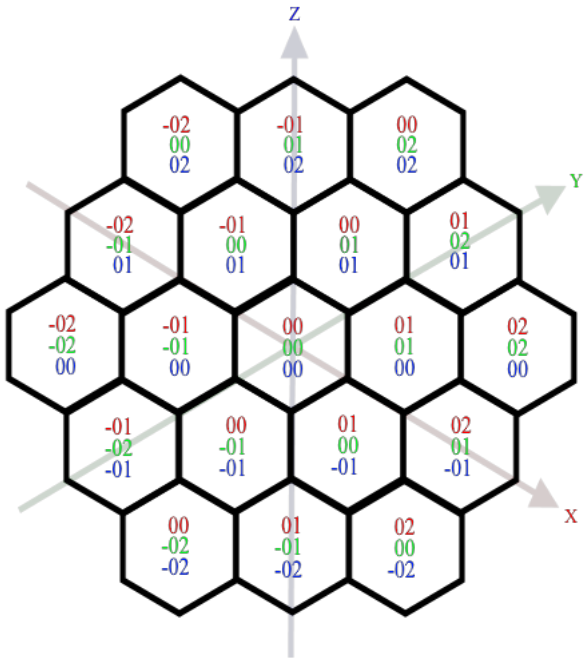

Figure 2: Hexagonal coordinate system where red represents the x axis and coordinate, green the y axis and coordinate, and blue the z axis and coordinate.

Hexagonal coordinates are represented by an ordered triple (x, y, z), where each element is an integer. Hexagonal coordinate representations eliminate the irrational coefficient present in hexagons mapped to Cartesian planes while retaining the same scale of the nest.

3.1.2: Cartesian Nest Calculations

The compactness number, one parameter used in (Karsai and Pénzes 1996) to evaluate nests of the same size but with a different arrangement of cells, measures the sum of the

distances from a cell's center to the geometric center of the nest. Equation 1 shows how to calculate a nest's compactness number in Cartesian coordinates (C_R). Equation 2 shows the Cartesian sum of squared distances (D_R), which is the calculation the compactness calculation is derived from. Equation 3 demonstrates the calculation of a nest's geometric center.

$$C_R = \sum_i \sqrt{(x_i - x_e)^2 + (y_i - y_e)^2} \quad (1)$$

$$D_R = \sum_i (x_i - x_e)^2 + (y_i - y_e)^2 \quad (2)$$

$$(x_e = \frac{1}{N} \times \sum_i x_i,\ y_e = \frac{1}{N} \times \sum_i y_i) \quad (3)$$

3.1.3: Hexagonal Nest Calculations

Converting between Cartesian coordinates and hexagonal coordinates requires a few simple equations. Averaging the hexagonal x and y coordinates results in the equivalent Cartesian x coordinate, while multiplying the hexagonal z coordinate by an irrational coefficient results in the equivalent Cartesian y coordinate, as shown in Eq. 4. Converting a Cartesian coordinate pair to a hexagonal coordinate triple requires solving a system of linear equations, shown in Eq. 5, 6, and 7.

$$(x,\ y, z) = (\frac{x+y}{2},\ \frac{\sqrt{3}}{2} \times z) \quad (4)$$

$$z_h = \frac{2\sqrt{3}}{3} y_r \quad (5)$$

$$x_h + y_h = 2x_r \quad (6)$$

$$x_h - y_h = -z_h \quad (7)$$

Calculating nest measurements in hexagonal coordinates necessitates changes to the Cartesian-based equations. Hexagonal coordinate calculations of geometric nest center (Eq. 8),

sum of squared distances (Eq. 9), and compactness number (Eq. 10) all have the z coordinate added to the equation.

$$(x_e = \frac{1}{N} \times \sum_i x_i,\ y_e = \frac{1}{N} \times \sum_i y_i,\ z_e = \frac{1}{N} \times \sum_i z_i) \tag{8}$$

$$D_H = \sum_i (x_i - x_e)^2 + (y_i - y_e)^2 + (z_i - z_e)^2 \tag{9}$$

$$C_H = \sum_i \sqrt{(x_i - x_e)^2 + (y_i - y_e)^2 + (z_i - z_e)^2} \tag{10}$$

3.2: Implementation

In this two-dimensional simulation, all nest permutations of 10 cells and fewer were procedurally generated and then compared against one another to detect distinct permutations. Nest generation was completed with Cartesian coordinates, floating point-based hexagonal coordinates, and integer-based hexagonal coordinates. Nests were compared with calculated compactness numbers as the primary comparer and nest eccentricity[1] along individual axes as the secondary comparer.

A nest visualization utility was constructed to help analyze nest data and evaluate specific nest statistics across multiple coordinate systems.

3.2.1: Software

This simulation was built in the C# programming language using Visual Studio 2015. The operating system environment was Windows 10 Professional 64-bit. All data generated was output to flat files for storage.

[1] Eccentricity is the distance the first initiated cell is from the two-dimensional geometric nest center.

3.2.2: Hardware

The test platform for the experiment was a desktop computer with an Intel Core i5-6600k CPU at 3.50 GHz ran on Intel's Z170 Express chipset. The test platform had 16 gigabytes of DDR4 RAM at 2133 MHz and 14/14/14/35 CAS timing.

3.3: Results

Counts of distinct permutations from nest simulation was compared against data generated in (Adoe 2010). Cartesian coordinate and floating point-based hexagonal coordinate compactness comparisons resulted in the same number of generated nests as those observed in (Adoe 2010), except for two duplicate nest forms, one at N=9 and one at N=10, for the Cartesian coordinates.

Table 1: Hexagonal coordinate system simulation results. * indicates duplicate nest structures.

Number of Cells	**Number of Unique Nest Configurations**		
	Rectangular	*Hexagonal Floating Point*	*Hexagonal Integer*
1	1	1	1
2	1	1	1
3	3	3	3
4	7	7	7
5	22	22	22
6	80	80	81
7	324	324	324
8	1377	1377	1389
9	6171*	6170	6192
10	28353*	28352	28416
Average Run Time:	**89.3 seconds**	**91.1 seconds**	**24.4 seconds**

The integer-based hexagonal coordinates located additional unique nest configurations at N=6 (Fig. 3), N=8, N=9, and N=10 cells. In addition, execution time for the integer-based hexagonal coordinates was, on average, three times faster than the execution time for Cartesian coordinates and floating point-based hexagonal coordinates.

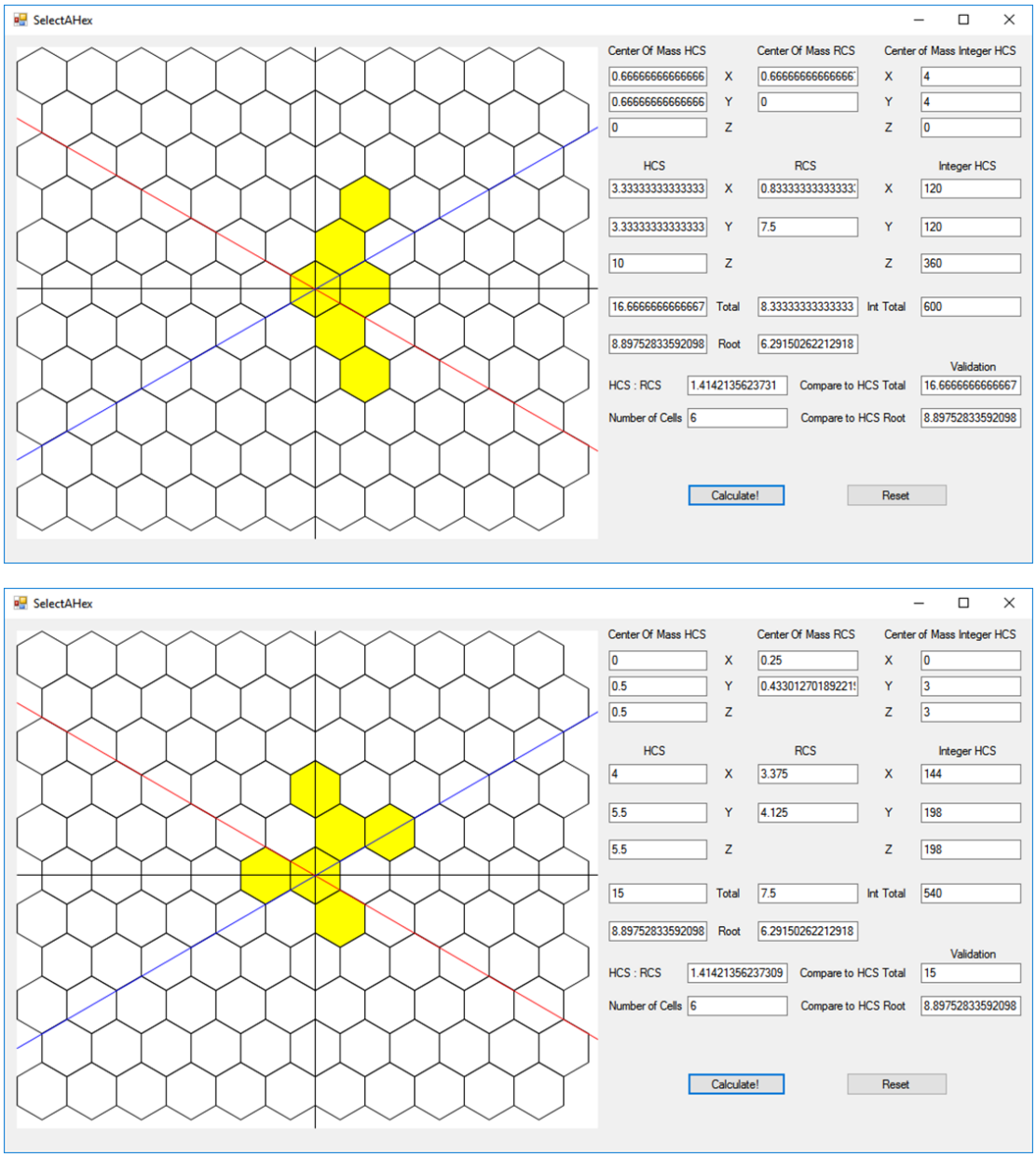

Figure 3: Two distinct nest forms with the same compactness number. Nests differ in eccentricity and sums of squared distances.

3.3.1: Ratios Between Compactness Numbers

In the generated nests, a relationship emerged between sums of squared distances and compactness numbers calculated in Cartesian coordinates and hexagonal coordinates: hexagonal sums of squared distances were twice as large as corresponding Cartesian sums of squared distances (Equation 11). In addition, hexagonal compactness numbers were $\sqrt{2}$ larger than Cartesian compactness numbers (Equation 12). After multiplying by the appropriate coefficient, hexagonal calculations were directly comparable to Cartesian calculations, and vice versa. See Appendix A for a direct proof of this relationship.

$$D_H = 2 \times D_R \tag{11}$$

$$C_H = \sqrt{2} \times C_R \tag{12}$$

CHAPTER 4

THREE-DIMENSIONAL NEST SIMULATION

The overall goal of this research was to investigate which simple rules generate nest structures that resemble those found in nature. This chapter discusses the design of the experiment, the stochastic model, the three-dimensional simulation, and the utilities necessary to implement the simulation and run the experiment.

4.1: Experimental Design

Nests were generated in the three-dimensional simulation and compared against real nest data across two parameters: outer cell walls / number of cells and compactness / number of cells. Additionally, generated nests were visually inspected with the 3D visualizer application to aid in understanding generated nest data.

Loads of paper pulp was used as the time metric in the simulation, where each load of pulp represents a wasp worker's successful gathering effort. A load of pulp is used at one cell site, either in site initiation or site lengthening. The simulation requires 25 loads of pulp for a cell to undergo initiation and lengthening to a mature height, e.g., a 100-cell nest of mature, maximum height cells would have 2500 loads of pulp, whereas a less-mature 100-cell nest would have fewer loads of pulp deposited. Pulp was deposited to the minimum wall(s) of a nest before any taller walls were lengthened. This type of lengthening allows for newly-initiated cells to "catch up" to the height of existing walls before those existing walls are lengthened.

Heights of walls were represented by an integer where each unit of height is equal to 100 microns, or one-tenth of a millimeter. For example, a wall height of 200 represents a wall that is 20 mm tall.

20,000 nests were generated per rule, with 200 nests generated per 10 loads of pulp in the range of 10-500, and 200 nests generated per 100 loads of pulp in the range of 500-5500. Additionally, for each pulp load target, 100 nests were generated without site selection constraints while 100 nests were generated with site selection constraints.

The hexagonal coordinate system designed and implemented in Chapter 3 was adapted for use in the simulation by adding a fourth (w) axis to represent height of a cell. A cell's coordinates were represented by a 4-tuple (x, y, z, w) where w was the minimum wall height of the cell. A cell with equal walls can be conceptualized as a hexagonal prism (see Figure 4).

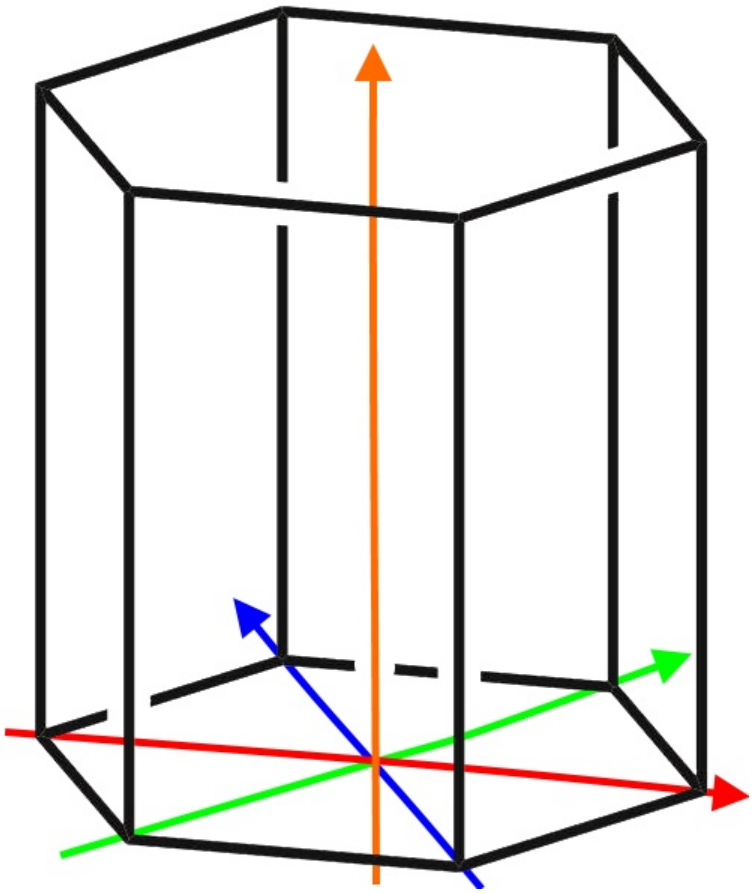

Figure 4: Hexagonal prism representation with hexagonal coordinate axes. Red is the x axis, green is the y axis, blue is the z axis, and orange is the w axis.

4.1.1: Statement of Questions

Three questions were asked for this experiment:

1. Do nests generated with non-random rules differ significantly from randomly-generated nests in outer wall counts and compactness numbers as a function of nest size?
2. Do generated nests differ significantly from ideal nests in outer wall counts and compactness numbers as a function of nest size?
3. Do generated nests differ significantly from real nest data in outer wall counts and compactness numbers as a function of nest size?

To determine if any differences exist between test data (generated nests) and control data (randomly generated nests, idealized nest data, and real nest data), scatterplots of generated nest outer wall counts and compactness numbers were created. Power functions were fit to the data: goodness of power function fit to the plotted data were evaluated through coefficients of determination. Each rule had two stratified random samples taken from the rule's simulated nest population, based on the pulp load counts between 10-500 and 100-5000. Heat maps of each rule's population were created to evaluate how representative the two samples taken were of the overall population. The power coefficient (b) for each fitted power function from experimental data (b_e) was then tested against power coefficients from control data (b_{random}, b_{ideal}, and b_{real}) by use of two-sample t-test, assuming unequal variances. The 99% confidence level was selected to evaluate the hypotheses.

4.1.2: Statement of Hypotheses

Question 1:

- Null hypothesis (H_0): $b_{random} - b_e = 0$
- Alternate hypothesis (H_a): $b_{random} - b_e \neq 0$

Question 2:

- Null hypothesis (H_0): $b_{ideal} - b_e = 0$
- Alternate hypothesis (H_a): $b_{ideal} - b_e \neq 0$

Question 3:

- Null hypothesis (H_0): $b_{real} - b_e = 0$
- Alternate hypothesis (H_a): $b_{real} - b_e \neq 0$

4.1.3: Model Assumptions

Several aspects of nest construction and worker action were simplified or abstractly represented:

- Pulp gathering was not modeled in this simulation: it was assumed that the agent had access to an unlimited supply of material for construction and always successfully gathered pulp.
- Each load of pulp brought back to the nest was assumed to be the same amount. Modeling varying-sized loads of pulp did not directly impact the experiment.

- Construction actions that are not initiation or lengthening were not modeled: it was assumed worker(s) will strengthen the petiole(s) or further build the base sheet of the nest as necessary.
- The base sheet is modeled as flat: real nests do not always have a flat base sheet.
- Cells are modeled at a constant diameter: real nest cells can be thinner at the base and wider at the mouth.
- A single agent performed all nest construction. Additionally, that single agent had the ability to evaluate every site on the nest. In real nests, multiple agents that check a limited number of sites before building (H. A. Downing 1994) would be concurrently acting on the nest.
- Brood were not directly modeled in the simulation; however, brood stimuli possibly play a role in both the maximum age rule and the maximum height rule (Karsai and Pénzes 2000).

4.1.4: Simulation Constraints

Three constraints were designed and implemented for use with the nest simulation: a maximum height constraint, a site initiation constraint, and a site lengthening constraint. All generated nests were subject to the maximum height constraint. Half of the generated nests, termed unconstrained nests, were generated without use of the site initiation constraint and the site lengthening constraint. The other half of generated nests, termed constrained nests, were subject to the site initiation and site lengthening constraints. Formal statements of the three constraints are below:

1. A maximum height of 30 mm, or 300 height units (one height unit equals 100 microns), was enforced on each wall: this constraint was placed on all generated nests and comes from the average height of mature nests observed in (Karsai and Pénzes 1996).
2. A site initiation constraint was implemented that limited site initiation to locations with two or more existing walls. This constraint was enforced on half of the dataset. Visual inspection of the unconstrained data set prompted creation of this constraint (see chapter 5, section 5.1.1).
3. A site lengthening constraint was implemented that limited lengthening to sites that had at least two neighboring cells with minimum walls taller than the candidate site. This constraint was enforced on half of the dataset. "Chimney-building" behavior observed in the unconstrained data set prompted creation of this constraint.

4.2: Site Selection Model

A stochastic model that assigns weights to each potential initiation site and existing cell was built for use in the three-dimensional simulation. Each step of the simulation represented the depositing of one load of pulp on the nest at a site selected by the model. After each load, the weights of each eligible build site were calculated based on the current state of the nest. The experimental model did not track previous states, i.e., the model has the Markov property.

4.2.1: Rules

Each rule represent one or more stimuli *P. dominula* workers may use to make decisions in real nest construction, such as chemical or mechanical stimuli (Karsai and Pénzes 2000).

Weights applied to each eligible site are integer calculations based on different properties of the nest, such as age, height, and existing walls.

Eligible sites for rules to evaluate were limited by the simulation constraint. For example, a constrained random rule nest will not weigh a site with only one existing wall since the site initiation constraint removes any one-cell initiation sites from the set of eligible sites. Similarly, the maximum height rule will only weigh eligible sites that have at least two taller neighbors due to the site lengthening constraint.

4.2.1.1: Random rule. Each eligible site has the same probability to be selected; eligible sites are assigned a weight of 1 (Eq. 13). The random rule serves as a baseline for analysis of the other rules.

$$Weight(random) = 1 \tag{13}$$

4.2.1.2: Maximum wall rule. Each eligible site's weight is the count of existing walls (Eq. 14). A site with six existing walls has a weight of 6 while a site with two existing walls has weight 2.

$$Weight(max\_wall) = \sum_i Exists(wall_i) \tag{14}$$

4.2.1.3: Maximum age rule. Each eligible site's weight is the sum of the ages of the site's walls (Eq. 15). Eligible sites have at least one wall and at most six walls. Age of a wall is calculated by subtracting the pulp load the wall was initiated on from the current pulp load. For

example, on the fourth pulp load in a three-celled nest, the oldest cell would have the largest probability factor of 18, followed by the second-oldest cell at 13, followed by the youngest cell at 9.

Existing cells tend to have higher probability factors than initiation sites; however, two older walls in an initiation site may have a higher probability factor than a newly-initiated cell on the frontier of the nest.

$$Weight(max\_age) = \sum_i Current\ load - Initiated\ load(wall_i) \quad (15)$$

4.2.1.4: Maximum height rule. Each eligible site's weight is the sum of the height of its walls (Eq. 16). A cell with six walls, each height 10, would have a weight of 60, while initiation sites with two of those walls would have a weight of 20.

$$Weight(max\_height) = \sum_i Height(wall_i) \quad (16)$$

4.2.1.5: Height difference rule. Each eligible site's weight is the difference between the site's tallest wall and the site's shortest wall (Eq. 17). For example, a site with a maximum wall height of 100 and a minimum wall height of 60 would have a weight of 40. Each of these integer units represents 100 microns: a cell wall modeled with a height of 100 represents a wall with a real height of 10 mm.

$$Weight(height\_diff) = MaxWall(site) - MinWall(site) \quad (17)$$

4.2.1.6: Hybrid rules. Rules that combine the maximum rules with the height difference rule were designed and implemented (Eq. 18). No scaling between the weights were applied to examine the effect each hybrid rule had on the nests generated.

$$Weight(hybrid) = Weight(max\_rule) + Weight(height\_diff) \quad (18)$$

4.2.2: Site Selection Algorithm

Stochastic selection of weighted sites was implemented by randomly selecting an integer between 0 and the sum of the eligible site weights. Site weights were then subtracted from the randomly generated integer until the difference became negative, indicating the selected site. Conceptually, this system draws inspiration from roulette wheel selection found in genetic algorithms.

4.3: Random Number Generation

Implementation of the stochastic model required a random number generator. Software random number generation is achieved through pseudorandom number generators. Pseudorandom number generators generate random numbers by applying a mathematical function to a previous state to generate the next state. The first state acted on by a pseudorandom number generator is the seed state.

The randomness of numbers generated by pseudorandom number generators is an area of ongoing research in fields that require high-quality random numbers, such as cryptography. Pseudorandom number generators possess properties that are necessary to pass testing developed in (Rukhin, et al. 2001), including truly random values for seeding a pseudorandom number

generator, uniformity of pseudorandom outputs, and consistency of generator behavior across different seeds.

The stochastic model does not require cryptographically-sound random numbers. However, the model does require a pseudorandom number generator with a sufficiently-large period, a low failure rate, and an ability to generate large quantities of pseudorandom numbers with minimal impact on performance in terms of compute time and memory space.

4.3.1: Pseudorandom Number Generator

The pseudorandom number generator chosen for this experiment is xoroshiro128+. Xoroshiro is a combination of the operations performed on the prior state to generate the successor state: XOR, rotate, shift, and rotate. Xorshiro128+ was designed as an updated version of xorshift generators (Vigna 2016), which was preceded by (Marsaglia 2003).

In (Vigna 2016) and (Vigna 2017), the TestU01 testing framework (L'Ecuyer and Simard 2007) and PractRand testing suite (Doty-Humphrey 2014) were employed to evaluate the quality of xorshift-based generators against other common pseudorandom number generators, such as Mersenne Twister (Matsumoto and Nishimura 1998). Testing showed that xoroshiro128+ produced fewer failures (where *p*-value statistics fell outside the test interval) in the BigCrush battery of tests than prior xorshift generators and Mersenne Twister while generating random 128-bit integers in less than one nanosecond and passing PractRand testing (Vigna 2017).

To verify implementation of xoroshiro128+ with the stochastic model, a simple binning test was created where 1 billion random integers in were generated from 0 to UINT64_MAX (18,446,744,073,709,551,615 on the test hardware) and counted in 10 bins representing

percentages of UINT64_MAX. This binning test was performed 100 times. Results from this test showed equidistributed xoroshiro128+-generated numbers across the 10 bins (see Table 2).

Table 2: Xoroshiro128+ binning test.

Bin (Percent)	Mean Count	Std. Deviation
0% - 10%	99999833.42	9054.123102
10% - 20%	99999657.49	8934.339742
20% - 30%	100000691	9581.776946
30% - 40%	100001208.4	8822.812362
40% - 50%	99999654.36	9412.002029
50% - 60%	99997438.87	8769.796361
60% - 70%	100000618.5	8697.225702
70% - 80%	100001916.7	8784.505183
80% - 90%	99999977.57	9791.111558
90% - 100%	99999003.67	9749.518619

4.3.2: Seed Source

Xoroshiro128+ requires a truly-random seed to generate high-quality pseudorandom numbers. Intel's Digital Random Number Generator (DRNG) was selected as the seed generation source. DRNG is a technology built into Intel processor chips that utilize three components to generate nondeterministic seeds: a hardware-based source of entropy, a conditioner that uses AES-CBC-MAC to output a 256-bit number, and a nondeterministic random number generator to scale the conditioner output to an arbitrary *n*-bit value (Intel Corporation 2014). Independent analysis found DRNG capable of generating cryptographically-sound random numbers and suitable seeds for pseudorandom number generators (Hamburg, Kocher and Marson 2012).

4.4: Software

The site selection model, three-dimensional simulation, and xoroshiro128+ were implemented in Visual C++ (C++11 standard) using Microsoft Visual Studio 2015. Additional libraries used in the simulation implementation were Boost v1.66, Mongocxx v3.2.0, and Bsoncxx v3.2.0. All simulated nests were stored in MongoDB Community Server v3.6.

The 3D visualizer utility was written in Python v3.6 in the Anaconda v5.1 distribution. Additional libraries used in the visualizer were PyMongo v3.6.1, Vpython v7.4.1, and GlowScript v2.7.

Heat maps were generated with the use of the MatPlotLib, PyPlot, and NumPy libraries. Statistical analysis was performed with a combination of NumPy and Microsoft Excel's data analysis module.

The operating system environment was Windows 10 Professional 64-bit.

4.5: Hardware

The test platform for the experiment was a desktop computer with an Intel Core i5-6600k CPU at 3.50 GHz ran on Intel's Z170 Express chipset. The test platform had 16 gigabytes of DDR4 RAM at 2133 MHz and 14/14/14/35 CAS timing.

CHAPTER 5

RESULTS

5.1: Visual Inspection

Generated nests were first inspected visually to look for any unexpected appearances or unanticipated model behavior. Each rule had a minimum of 100 nests visually inspected at load sizes 100, 500, 1000, 2000, and 4000.

5.1.1 Unconstrained Nests

Unconstrained nests are generated nests that were not constrained by the site initiation or site lengthening constraints described in Chapter 4, section 4.1.4. Unconstrained nests had the site selection rules weigh eligible sites that contained one existing wall and sites with zero or one taller neighbor. Unconstrained nests were still constrained by the maximum cell height constraint.

Random rule nests appeared as expected (Fig. 5); non-random rule nests did not appear as initially expected (Figs. 6, 7, and 8). These unexpected nest configurations were due to the large

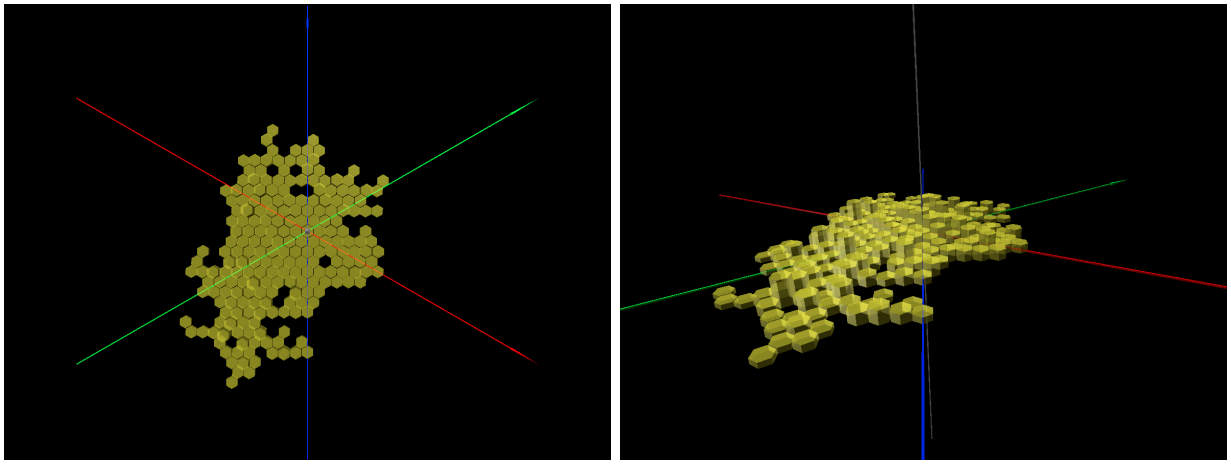

Figure 5: Random rule, unconstrained, 500 pulp loads.

number of one-wall initiation sites marked eligible by the unconstrained rule model. Each individual one-wall initiation site is not heavily weighted; however, the sum of those weights resulted in a disproportionate number of one-wall cell initiations.

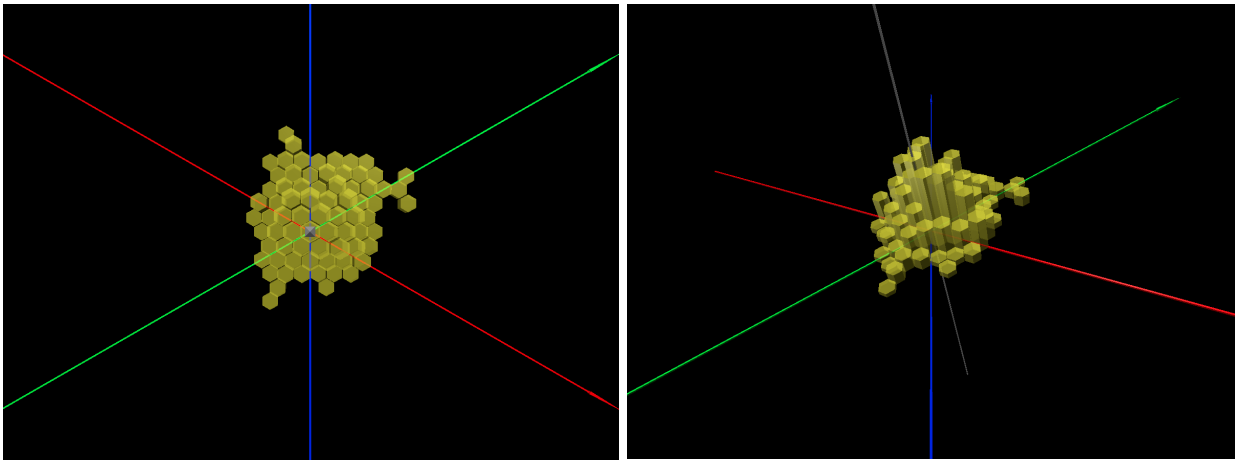

Figure 6: Maximum age rule, unconstrained, 500 pulp loads.

The maximum height rule tended to build "chimneys" away from any other structure. This behavior was due to the self-reinforcing nature of the maximum height rule: weights on individual cells become greater as the cell walls grow taller. These chimney-like cells would not

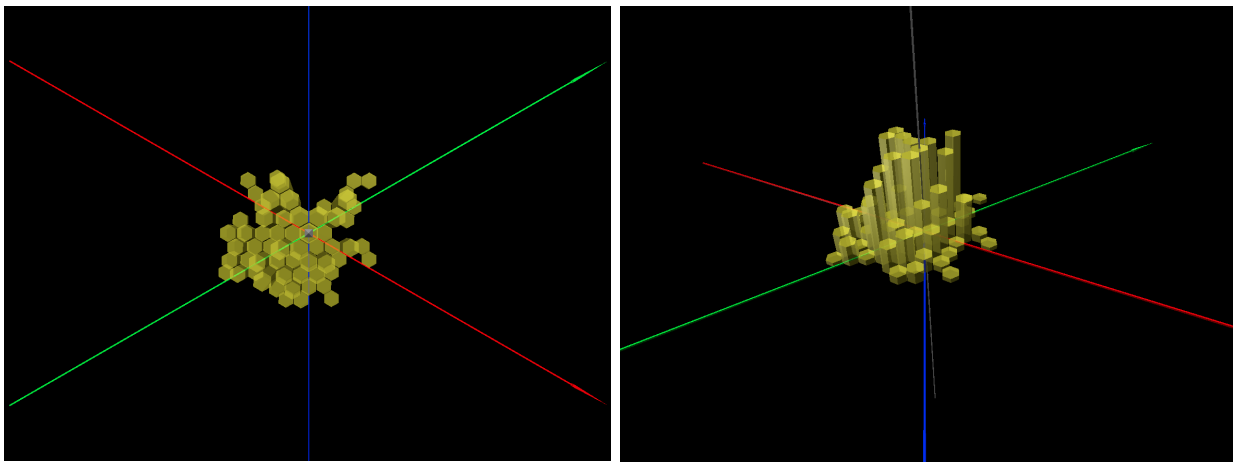

Figure 7: Maximum height rule, unconstrained, 500 pulp loads.

survive in nature due to the relative structural weakness of one freestanding paper cell. While the maximum age rule also resulted in some chimney-like cells, the maximum height rule was the primary source of this behavior.

The height difference rule resulted in nests with more in common with random nests than nests generated by other non-random rules. This was due to how the height difference rule was calculated. If an eligible initiation site had a maximum wall height of *n* and a minimum wall height of 0, then the weight of that site was *n*. Since evenly-built cells have smaller weights than uninitiated cells, the model selects more initiation sites than lengthening sites. The height difference initiation preference led to the creation of the hybrid rules to allow other rules, such as maximum age and maximum height, to be influenced by the stimulus provided by the height difference rule.

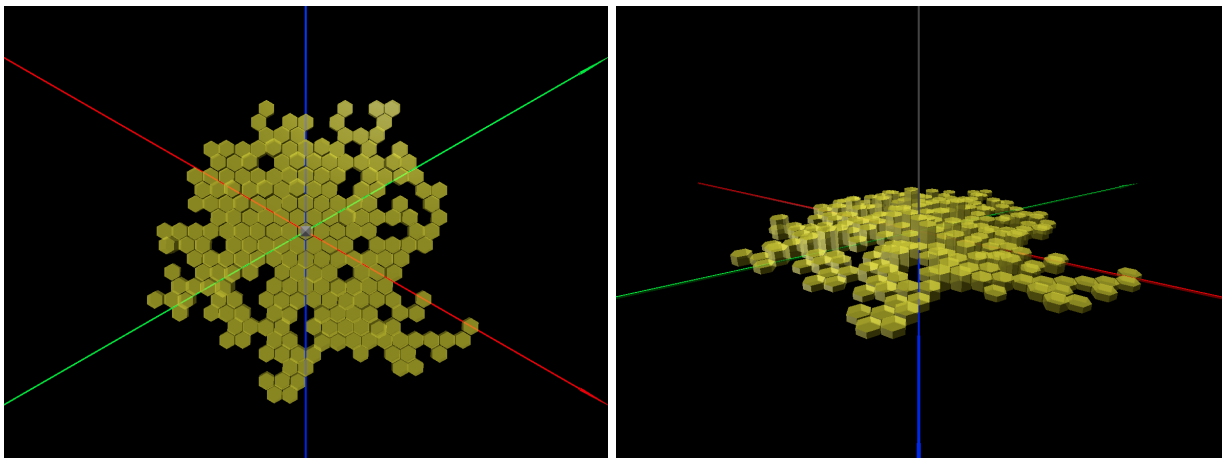

Figure 8: Height difference rule, unconstrained, 500 pulp loads.

The behavior of the unconstrained rules prompted the implementation of the site initiation and site lengthening constraints for subsequent nest generation. No further analysis was performed on unconstrained nests or the height difference rule as a standalone rule.

5.1.2 Constrained Nests

Generated nests that were constrained with the site initiation and site lengthening constraints do not exhibit the single-cell tails or chimneys observed in unconstrained nests due to the removal of single-walled initiation sites and sites without at least two taller neighbors. Nests generated by rules constrained by the constraints described in Chapter 4, section 4.1.4, appear to be more compact than those nests generated by the same rules without the constraints in place.

The random rule, with its equally-weighted sites, still generated unnatural nest forms, including missing initiation sites (Fig. 9). From a visual standpoint, the maximum age (Fig. 10) and hybrid age (Fig. 11) rules and the maximum height (Fig. 12) and hybrid height (Fig. 13) rules generated the most well-formed structures.

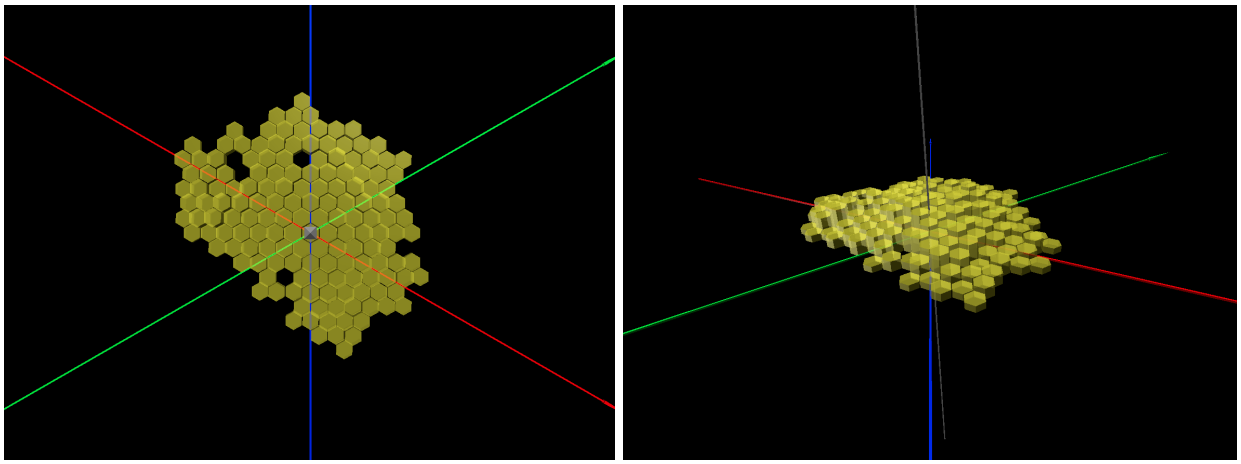

Figure 9: Random rule, constrained, 500 pulp loads.

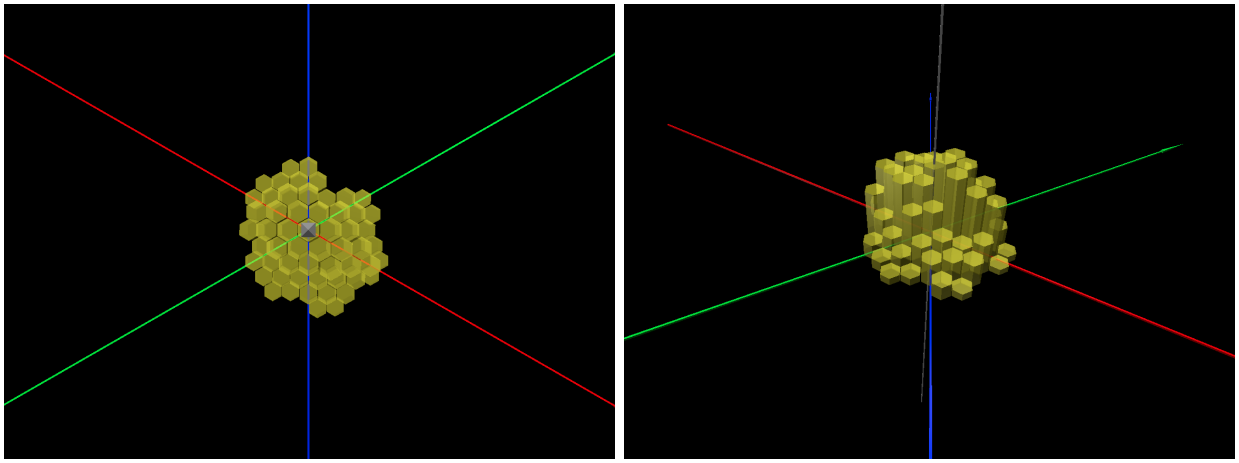

Figure 10: Maximum age rule, constrained, 500 pulp loads.

The hybrid version of the age rule does not appear to generate different nest structures from the maximum age rule. This observation may stem from how the weights are calculated. For example, a 1000-pulp load nest could have sites with weights in the 5000s for cell walls initiated early in construction ((1000 – 100) * 6 = 5400). In comparison, the height difference rule exerts a small fraction of that weight on the nest, e.g., a site with a maximum wall height of 250 and a minimum wall height of 50 has a height difference weight of 200. Future work (discussed in chapter 6, section 6.1) could apply either a static or dynamic weighting to the age component and the height difference component of the hybrid rule.

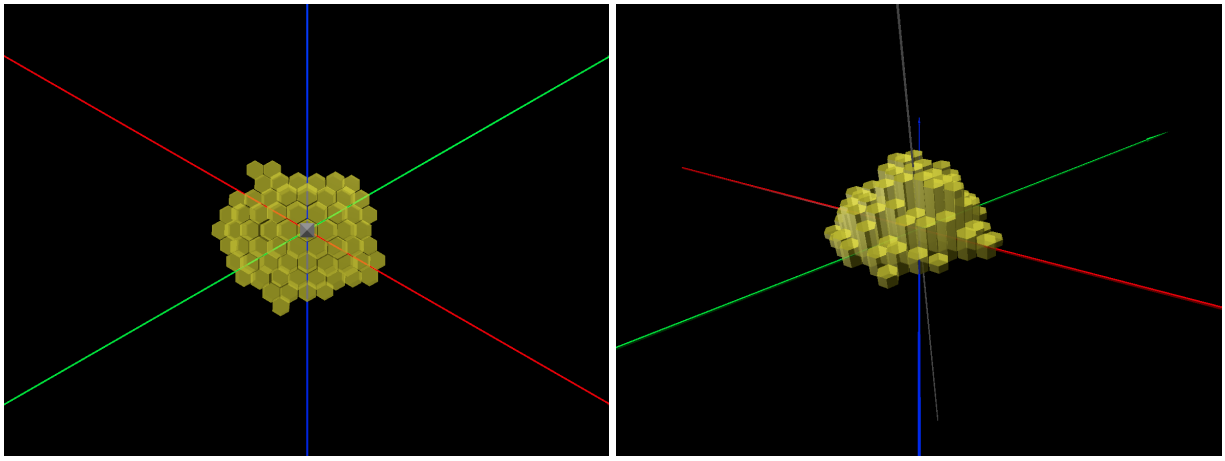

Figure 11: Hybrid age rule, constrained, 500 pulp loads.

Visual comparisons of the constrained age rules and height rules show that while the height rules generated fewer cells per pulp load target (compare Figs. 10 and 12), the age-based rules appear to generate nests with a smoother gradient of heights from the center of the nest towards the edges of the nest.

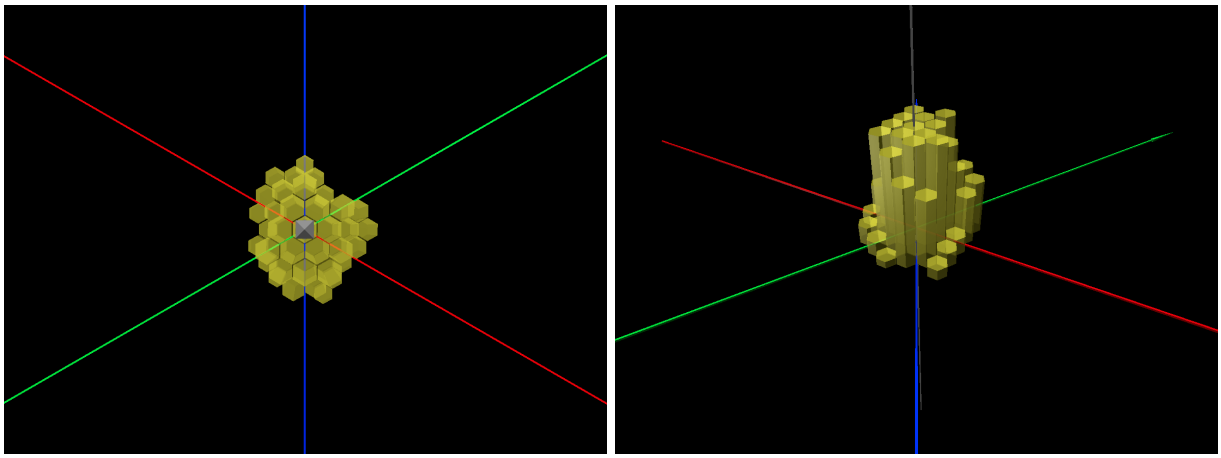

Figure 12: Maximum height rule, constrained, 500 pulp loads.

As observed with the age-based rules, the maximum height rule and the hybrid height rule generate visually similar structures. While not as drastic a ratio as that of the age-based rules, the ratio of weights between the two height-based rules still strongly favors the sum of

wall heights over the height difference: a site with five walls of 200 height units and one wall of 50 height units (summing to 1050 weight) will have a height difference weight of 150.

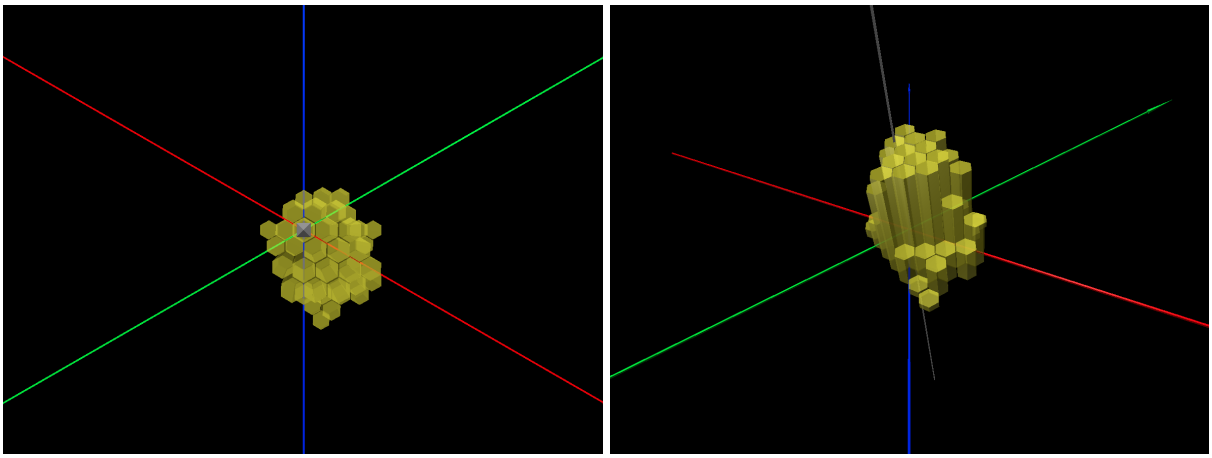

Figure 13: Hybrid height rule, constrained, 500 pulp loads.

The maximum wall and hybrid wall rules (Figs. 14 and 15) did not generate well-formed structures: wall-based structures visually appear to have more in common with random nests (Fig. 9) than nests generated with the age-based and height-based rules (Figs. 10 – 13). The weight exerted by sites with six existing walls is three times greater than the weight exerted by a site with two existing walls. In comparison, an age-based, center-nest site with six walls of total weight 5000 will outweigh an outer-nest initiation site with two walls of total weight 100 by a factor of 50.

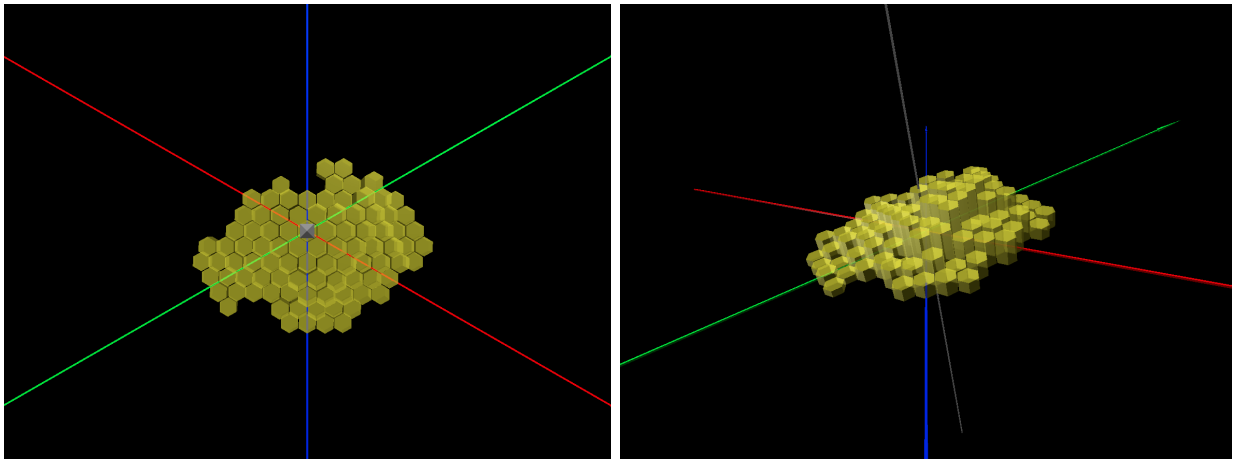

Figure 14: Maximum wall rule, constrained, 500 pulp loads.

The height difference portion of the hybrid height rule could exert more weight on site selection versus the maximum-wall portion of the rule. For example, a site with six walls, with maximum height 100 and minimum height 50, would have weight 6 from the maximum-wall portion of the rule and weight 50 from the hybrid height portion of the rule. However, there did not appear to be significant differences between the nests generated by the maximum wall rule (Fig. 14) and the hybrid wall rule (Fig. 15).

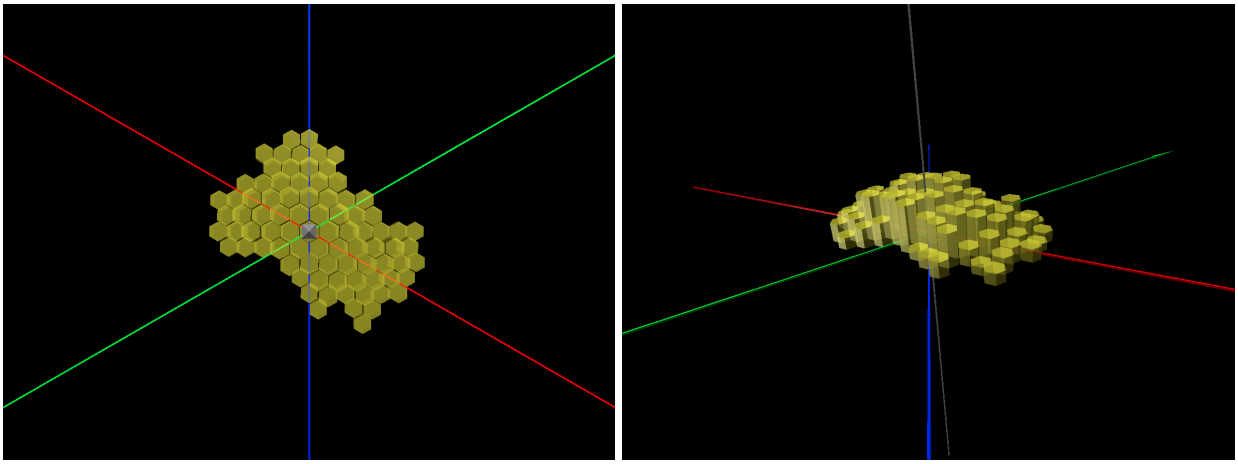

Figure 15: Hybrid wall rule, constrained, 500 pulp loads.

5.2: Quantitative Parameter Analysis

Data from (Karsai and Pénzes 1996) recommended that intra-species studies of nest variability in *Polistes* fit a curve between number of cells and two quantitative traits: outer wall count and compactness. The fitted curves for the nests generated with non-random rules were then compared against curves for each quantitative trait from control nests: random nests, ideal nests, and real nests. Fitted curves for number of cells versus outer walls and number of cells versus compactness for real nest data were found to have an $r^2 = 0.97$ and $r^2 = 1$, respectively (Karsai and Pénzes 1996); t tests found no significant difference between idealized and real nests data for outer walls ($p > 0.1$) but a significant difference between idealized and real nest data for compactness ($p < 0.01$) (Karsai and Pénzes 1996).

Analysis for constrained generated nests consisted of generating scatter plots for outer walls and compactness, each versus number of cells in a nest, and fitting power regressions to the data. Regression coefficients (b) were compared between data generated with one of the non-random rules (b_e) and data from a control source (b_{random}, b_{ideal}, and b_{real}) were tested for significance via two sample t-test with the assumption of unequal variances.

5.2.1: Sampling Strategy

For each rule of interest, 1000 nests were sampled in two different stratified random samples: 500 nests total were sampled from 10-500 loads of pulp, with 10 nests randomly sampled every 10 pulp loads, and 500 nests total from 100-5000 loads of pulp, with 10 nests randomly sampled every 100 pulp loads. 10% of the total constrained nest population was sampled by each stratified random sample.

MongoDB provides random sampling through use of the aggregation pipeline and the $sample command. Samples are randomly selected by MongoDB using a pseudorandom sort on a collection's documents and then selecting the first 10 randomly-sorted nest documents.

5.2.2: Sample Quality

Heat maps for each rule's population were created, based on sampling range (10-500 pulp loads or 100-5000 pulp loads) and analyzed parameter (outer wall count or compactness number). Figure 16 shows the key for interpreting the heat maps.

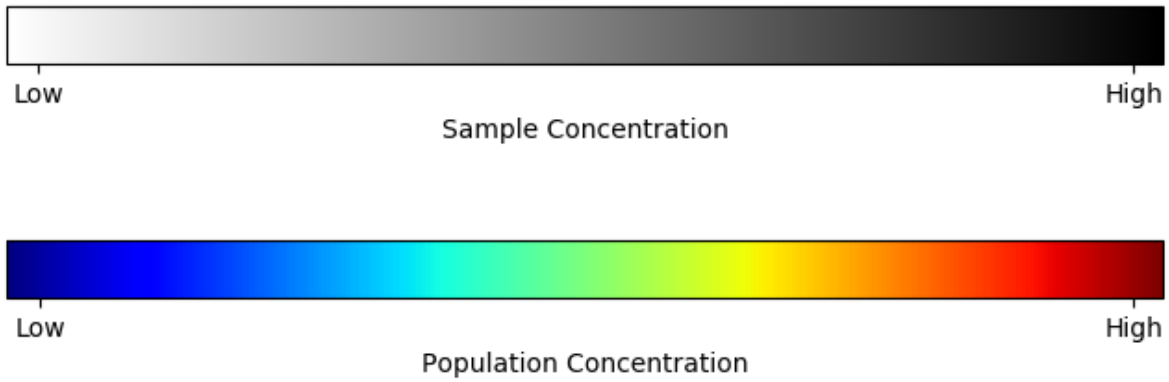

Figure 16: Color bars for heat map interpretation. The grayscale bar is for sampled points. The lighter the sample point color, the fewer number of samples taken at that point. The color bar is for the population of nests generated by a rule, separated by sample range and parameter (outer wall counts and compactness number). The cooler the color of the population points, the fewer individuals that had the parameter value at that number of cells.

The color points represent a rule's total population, plotted by parameter (outer wall counts or compactness numbers) as a function of number of cells. One point represents one or more nests, depending on the color of the point. Red (warm) points represent large concentrations of individuals while blue (cool) points represent sparse concentrations of individuals. The grayscale points plotted above the population heat map represent sampled points. Darker points had more samples taken than lighter points.

5.2.2.1: Outer Wall Samples. Figure 17 shows a heat map for random rule nest data from the broad (100-5000 pulp load) sample. Darker sample points are concentrated in warmer areas of the population heat map. The stratified random sample obtained from the random rule's population appears to be a representative sample.

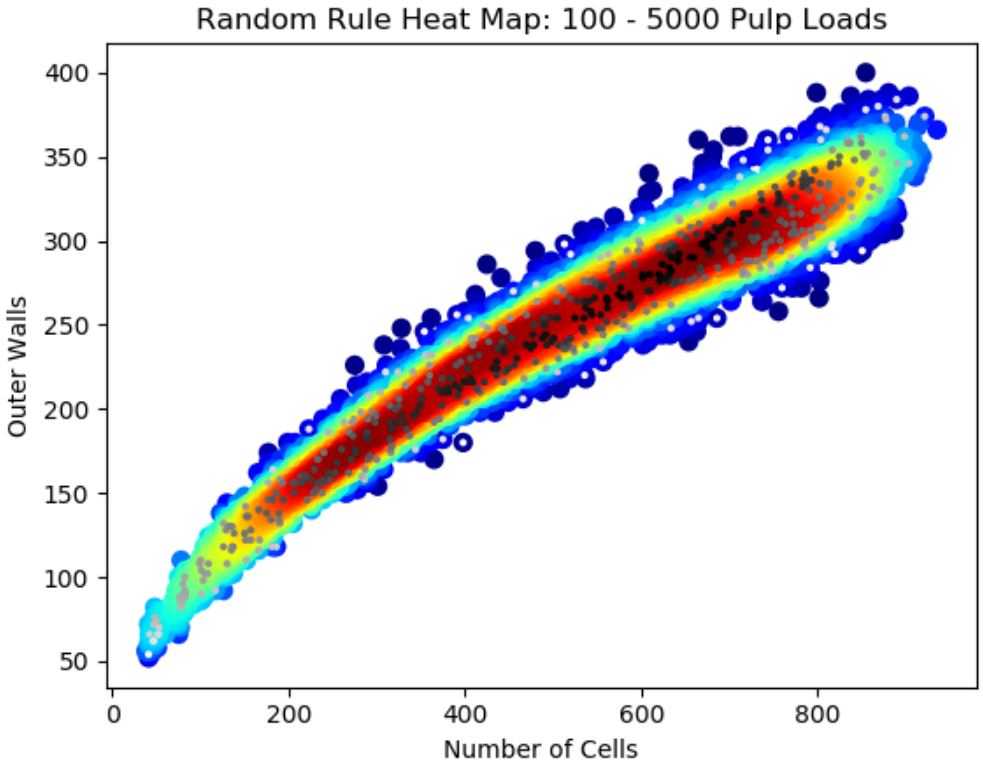

Figure 17: Random rule heat map of the outer walls parameter for the broad sample. Colorized points are the rule's generated nest population for the 100–5000 pulp load range while the grayscale points are for the sample taken from the random rule's population. Warmer / darker points represent higher densities.

Figure 18 shows a heat map for the maximum age rule from the broad sample. As with the random rule's heat map, it appears that the sample taken for the maximum age rule is representative.

Heat maps of outer wall data for the remaining rules in the broad sample (Figs. 36 - 40) and for all rules in the narrow sample (Figs. 41 – 47) also appear to be representative samples of their respective populations. The other rules' heat maps are in Appendix B of this report.

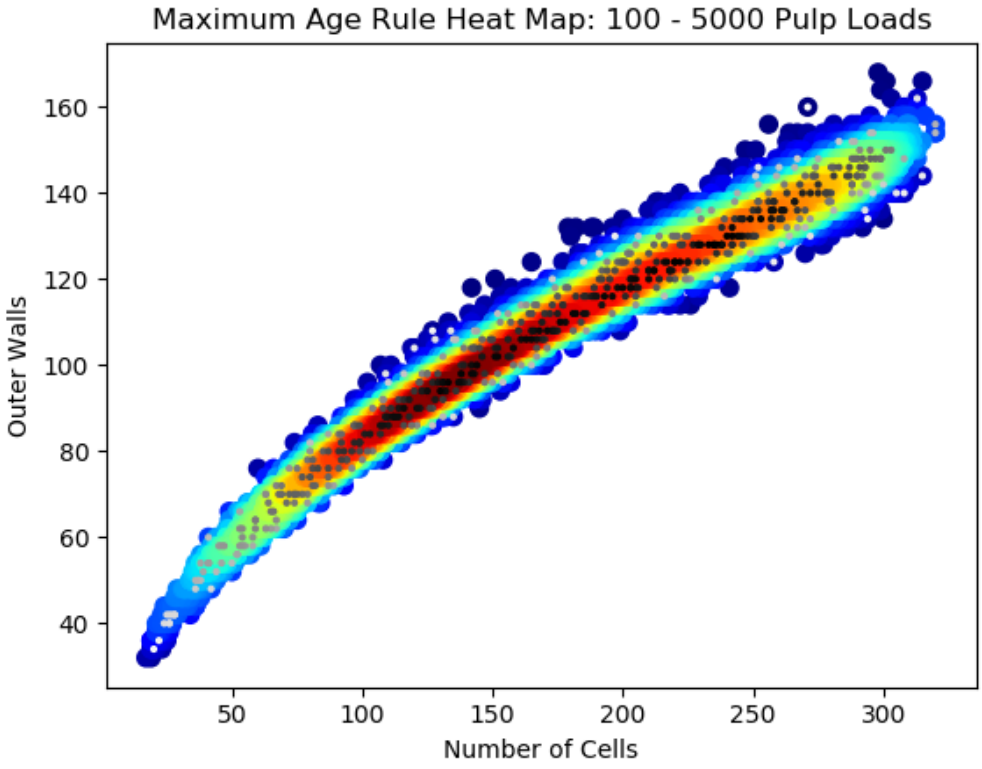

Figure 18: Maximum age rule heat map of the outer walls parameter for the broad sample.

5.2.2.2: Compactness Samples. Figure 19 is a heat map of the compactness data from the random rule's 100-5000 pulp load population data overlaid with the sample data. The same scaling for outer wall heat maps apply to the compactness heat maps. As with the outer wall data, the sample's compactness data appears representative of the random rule's population.

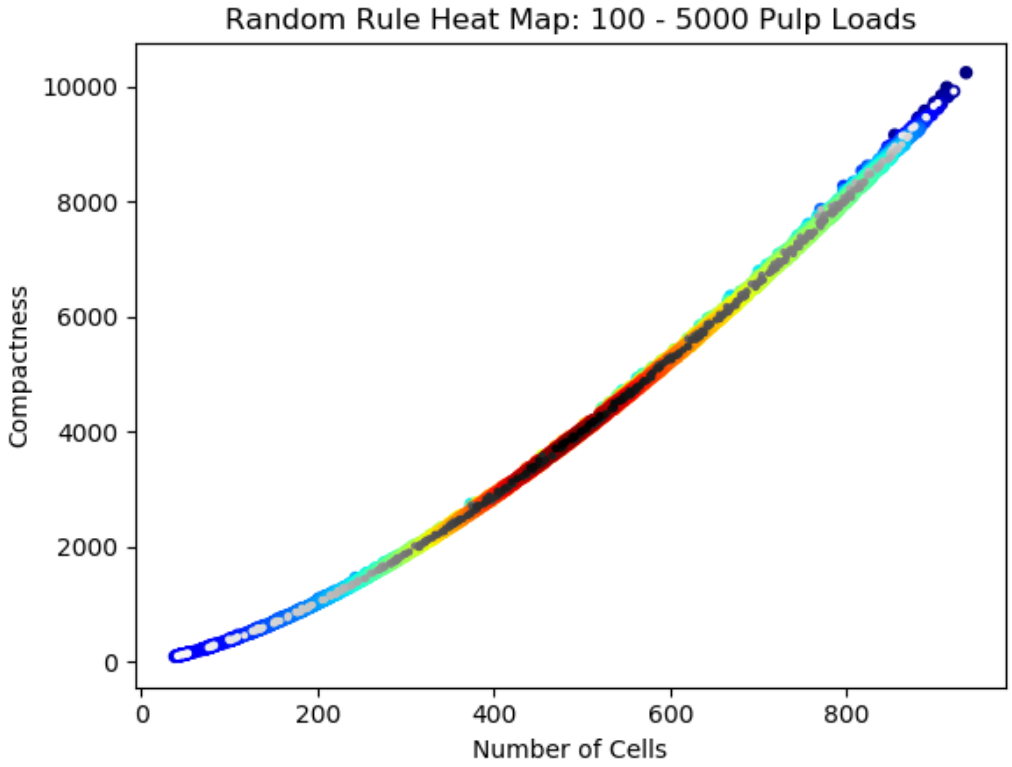

Figure 19: Random rule heat map of the compactness parameter for the broad sample. The same key from outer wall heat maps (Fig. 16) applies to compactness heat maps.

Figure 20 is a heat map of compactness data for the maximum age rule. As with the outer wall heat map, the compactness heat map for the maximum age rule shows similar distributions for the sampled points and the population points.

Appendix B contains the remaining rules' broad sample heat maps (Figs. 48 – 52) and all the rules' narrow sample heat maps (Figs. 53 - 59) for compactness numbers.

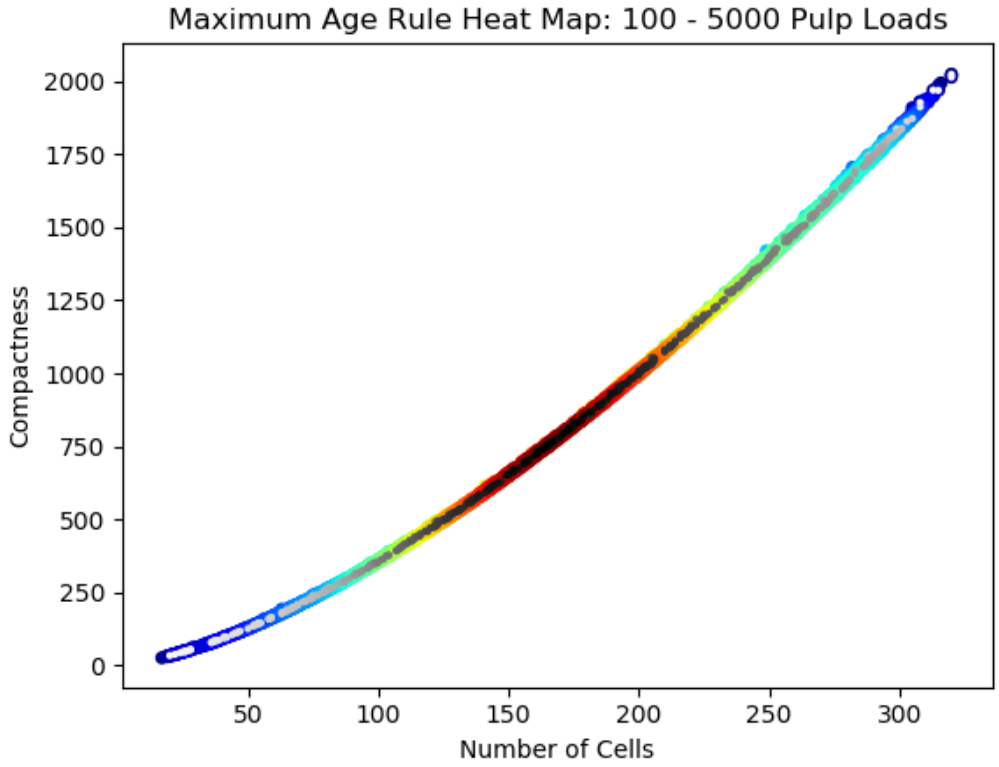

Figure 20: Maximum age rule heat map of the compactness parameter for the broad sample.

5.3: Real and Idealized Data

Data from (Karsai and Pénzes 1996) were used to generate comparison coefficients to test against generated nest parameters. Scatterplots of real nest data with fitted power functions were created for outer wall counts (Fig. 21) and compactness numbers (Fig. 22). Both parameters showed high coefficients of determination ($R^2 > 0.99$), indicating both parameters strongly depend on the number of cells in real nests. This correlation was also observed in (Karsai and Pénzes 1996).

For each parameter, the power function regression coefficient was compared against idealized nest data derived from minimally-compact nest forms. Ideal nest forms are the most compact nest forms for a given number of cells, which, in turn, have the fewest outer walls. Ideal calculations from (Karsai and Pénzes 1993) and data from (Karsai and Pénzes 2000) were used to generate idealized nest forms.

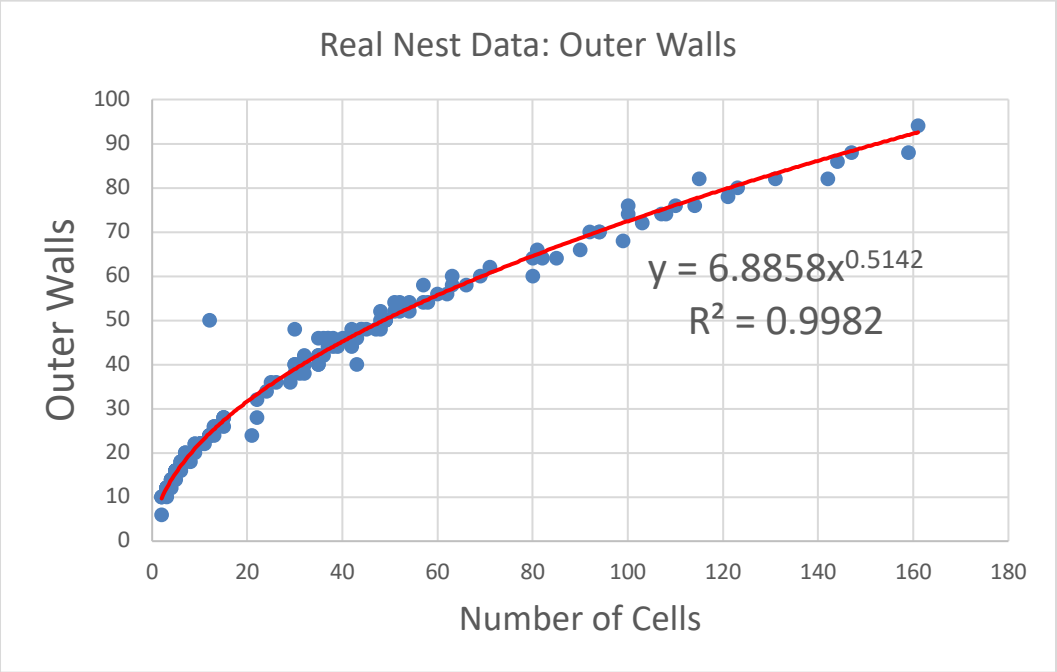

Figure 21: Real nest data - outer walls (Karsai and Pénzes 1996).

Scatterplots of idealized nest data with fitted power functions for outer wall counts (Fig. 23) and compactness numbers (Fig. 24) were created for comparison against real nest data. As with the real nest data coefficients of determination, values for both parameters of ideal nests strongly depend on the number of cells in the nest ($R^2 > 0.999$).

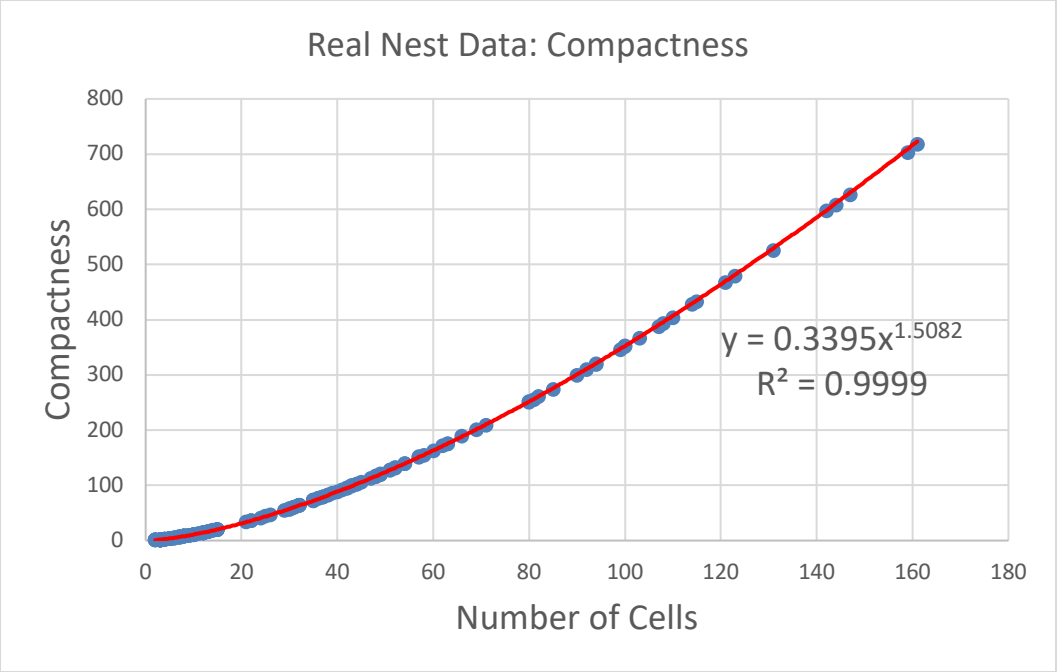

Figure 22: Real nest data - compactness (Karsai and Pénzes 1996).

For each parameter, t-tests were performed on the regression (*b)* coefficients of real vs ideal nest data as a control. There was no significant difference between idealized and real data ($p > 0.1$) for outer wall data; there was a significant difference in compactness data between idealized and real data (t-test, $p < 0.01$). The calculated *p*-values from the t-tests match the *p*-values calculated in (Karsai and Pénzes 1996).

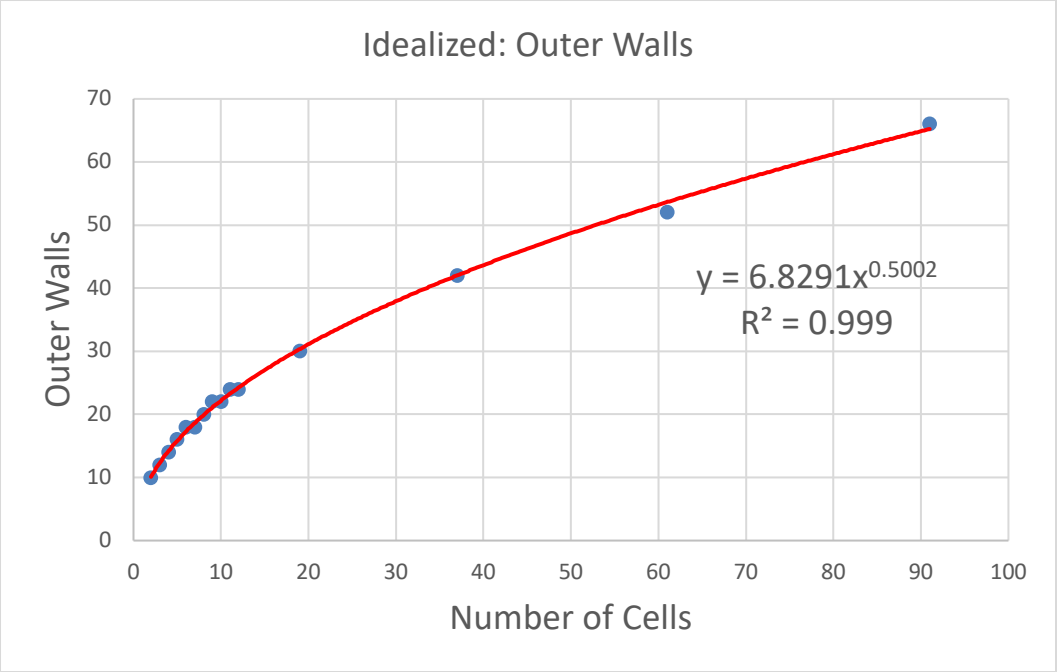

Figure 23: Idealized nest data – outer walls.

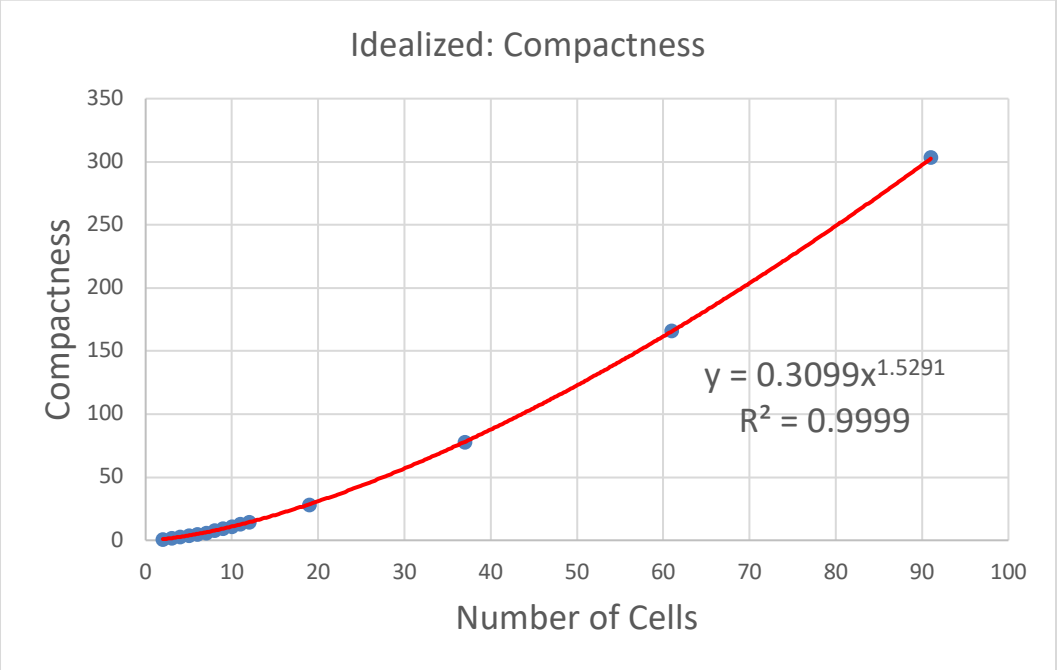

Figure 24: Idealized nest data – compactness.

5.4: Tests of Hypotheses: Outer Wall Counts

5.4.1: Random Rule

Randomly-generated nests did not result in lifelike structures. Figures 25 and 26 show scatterplots with similar fitted power functions ($R^2 > 0.96$) for the two samples, indicating that randomly-generated nests, while less dependent on number of cells than real and idealized nest data, still have high correlation between each nest parameter and the number of cells in a nest.

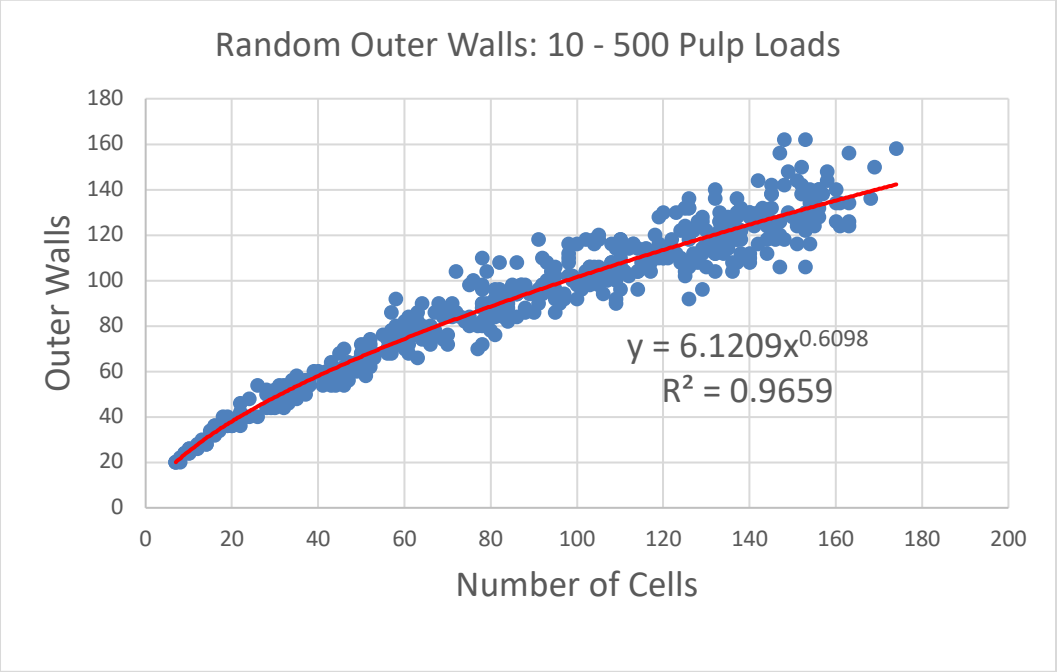

Figure 25: Random rule – outer wall data for 10 – 500 pulp loads.

For both the narrow and broad samples, calculated regression coefficients for random rule nests were significantly different from both real nests ($p < 0.01$) and ideal nests ($p < 0.01$). The significant difference between random data and control data by outer wall counts means that real nests and idealized nests are not likely to be randomly built.

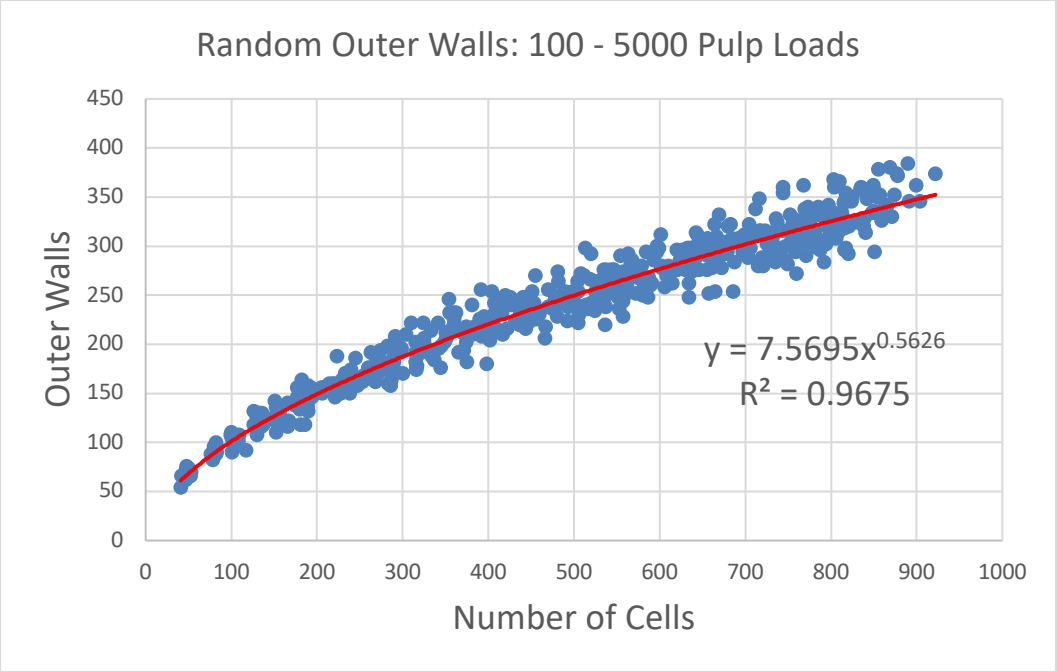

Figure 26: Random rule – outer wall data for 100 – 5000 pulp loads.

5.4.2: Wall Rules

The maximum wall rule and the hybrid wall rule generated similar nests in terms of outer wall counts. Figure 27 shows the broad-sample scatterplot for outer wall counts as a function of number of cells. A high coefficient of determination ($R^2 > 0.97$) was calculated for the fitted power function. Scatterplots for the maximum wall rule outer wall counts in the narrow sample (Fig. 68) and for the hybrid wall rule outer wall counts in both samples (Figs. 64 and 65) are in Appendix C.

Analysis of regression coefficients found different results based on the pulp load strata. For the narrow sample, the regression coefficients for the maximum wall rule were significantly different from random ($p < 0.01$), ideal ($p < 0.01$), and real ($p < 0.01$) regression coefficients. However, the broad sample for the maximum wall rule regression coefficients were not significantly different ($p > 0.01$) from the random coefficient. Significantly different regression

coefficients were found in the maximum wall rule broad sample from both idealized data ($p < 0.01$) and real data ($p < 0.01$).

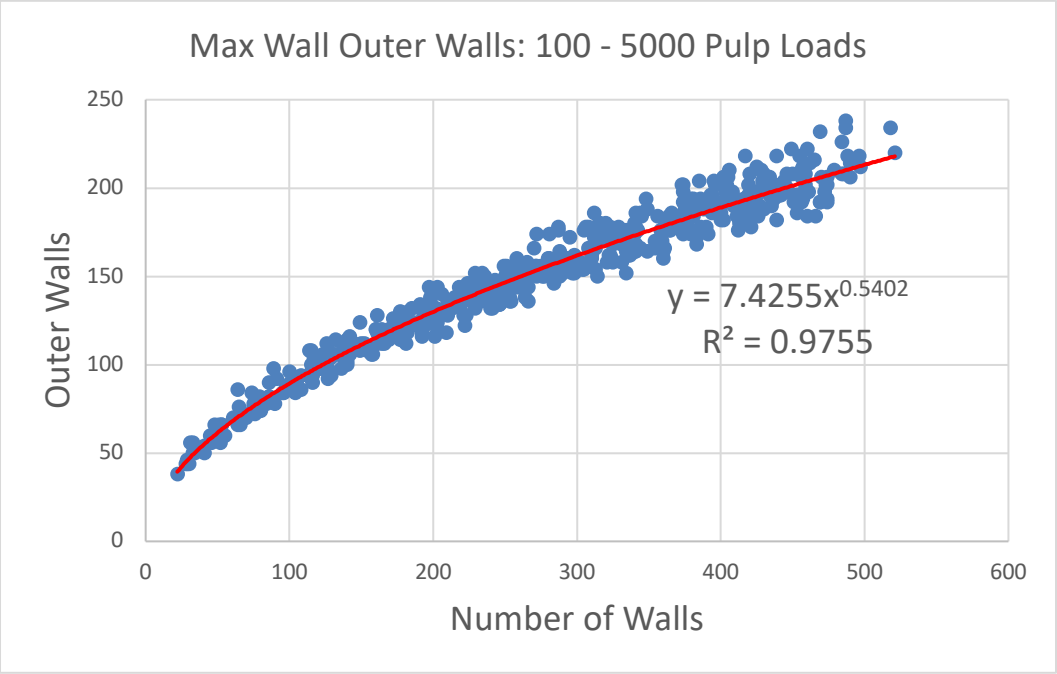

Figure 27: Max wall rule – outer wall data for 100 – 5000 pulp loads.

The hybrid wall rule showed identical t-test results to the maximum wall rule: significant differences from random ($p < 0.01$), ideal ($p < 0.01$), and real ($p < 0.01$) regression coefficients in the narrow sample and ideal ($p < 0.01$) and real ($p < 0.01$) data in the broad sample. No significant differences ($p > 0.01$) between the broad sample outer wall regression coefficients and the random rule regression coefficients were found.

5.4.3: Age Rules

The maximum age and hybrid age rules generated similar structures across the outer wall parameter. Power functions fit to the maximum age and hybrid age data sets had high

coefficients of determination ($R^2 > 0.97$). Figure 28 shows the scatterplot for the broad sample of the maximum age rule. Scatterplots for the narrow sample of the maximum age data (Fig. 66) and for both samples of the hybrid age data (Figs. 60 and 61) are in Appendix C.

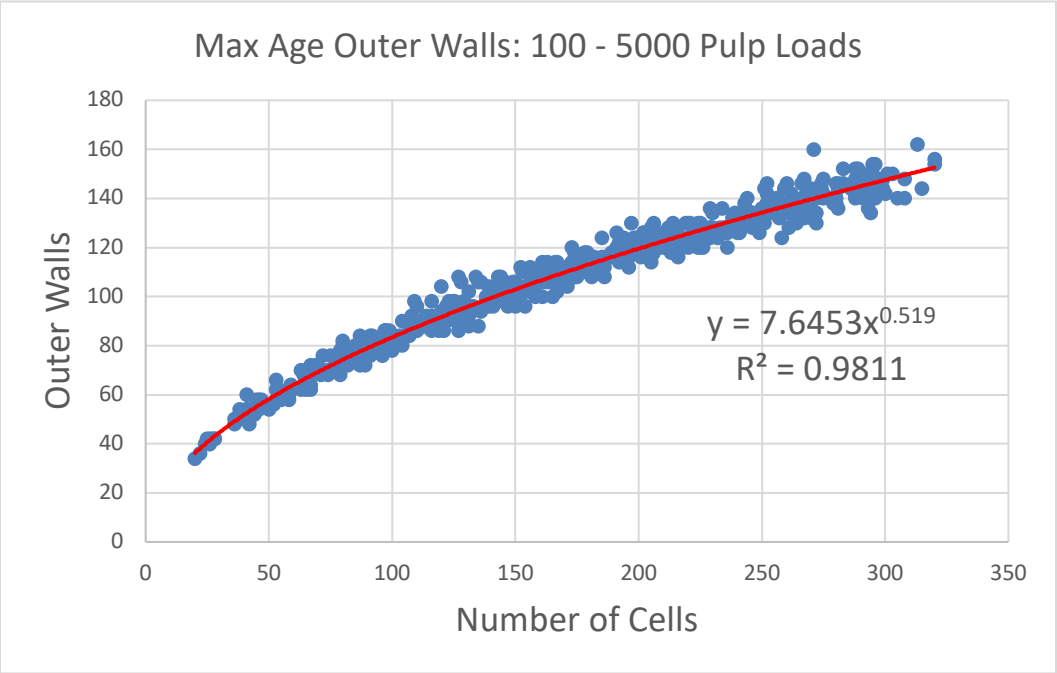

Figure 28: Max age rule – outer wall data for 100 – 5000 pulp loads.

Regression coefficients for the maximum age rule data in the narrow sample showed significant differences with random coefficients ($p < 0.01$), ideal coefficients ($p < 0.01$), and real coefficients ($p < 0.01$). The regression coefficients for the maximum age rule data in the broad sample showed significant differences with random coefficients ($p < 0.01$) and ideal coefficients ($p < 0.01$). However, there was no significant difference between the regression coefficients in the maximum age broad sample and real nest data ($p > 0.1$), indicating that the broad sample maximum age nests, in counts of outer walls as a function of number of cells, show an insignificant difference to real nest data along the outer wall parameter.

Hybrid age rule data differed from maximum age rule data along the outer wall parameter in the narrow sample. While the regression coefficients for the hybrid age rule in the narrow sample had significant differences with the regression coefficients of random nests ($p < 0.01$) and idealized nests ($p < 0.01$), no significant difference with the regression coefficients of real nests ($p > 0.01$) was found. The hybrid age rule data in the broad sample had the same test results as the maximum age rule: significant differences were found between hybrid age regression coefficients, random coefficients ($p < 0.01$), and ideal coefficients ($p < 0.01$), but no significant difference between hybrid age coefficients and real nest coefficients ($p > 0.1$).

5.4.4: Height Rules

The scatterplots for the maximum height and hybrid height rules outer wall counts were not as tightly packed as the other non-random rules. However, power functions fit to the data had high coefficients of determination ($R^2 > 0.95$). Figure 29 shows the scatterplot for the maximum height rule in the broad sample. Scatterplots of the maximum height rule in the narrow sample (Fig. 67) and they hybrid height rule in both samples (Figs. 62 and 63) are in Appendix C.

Analysis of the regression coefficients between the maximum height rule and the random rule in both the narrow sample and the broad sample found no significant difference ($p > 0.01$) between the data. Significant differences in regression coefficients were observed between the maximum height data, idealized data ($p < 0.01$), and real data ($p < 0.01$) in both the narrow and broad samples.

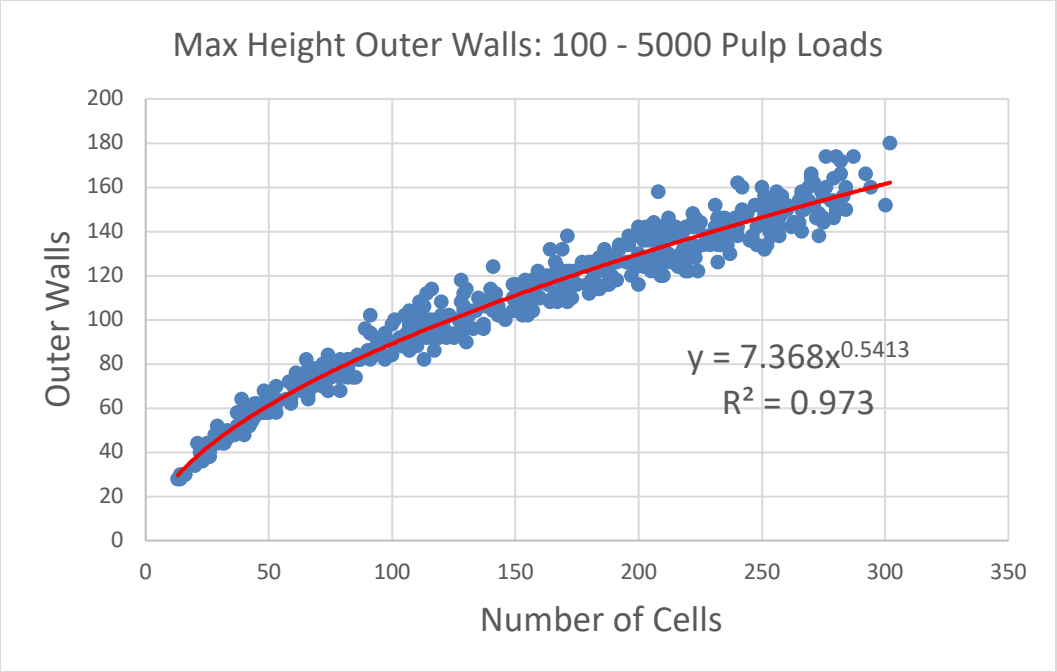

Figure 29: Max age rule – outer wall data for 100 – 5000 pulp loads.

The hybrid height rule analysis found analogous results to those found with the maximum height rule: no significant difference in regression coefficients between hybrid height data and random data ($p > 0.01$) was calculated across both samples. Significant differences were calculated between hybrid height regression coefficients, idealized coefficients ($p < 0.01$) and real coefficients ($p < 0.01$) over both samples.

5.5 Tests of Hypotheses: Compactness Numbers

5.5.1: Random Rule

The random rule did not generate compact nest structures, which provides a contrast to the compact nest structures of idealized and real data. Power functions fit to the random rule

compactness numbers in the narrow sample (Fig. 30) and the broad sample (Fig. 31) both have high coefficients of determination ($R^2 > 0.99$).

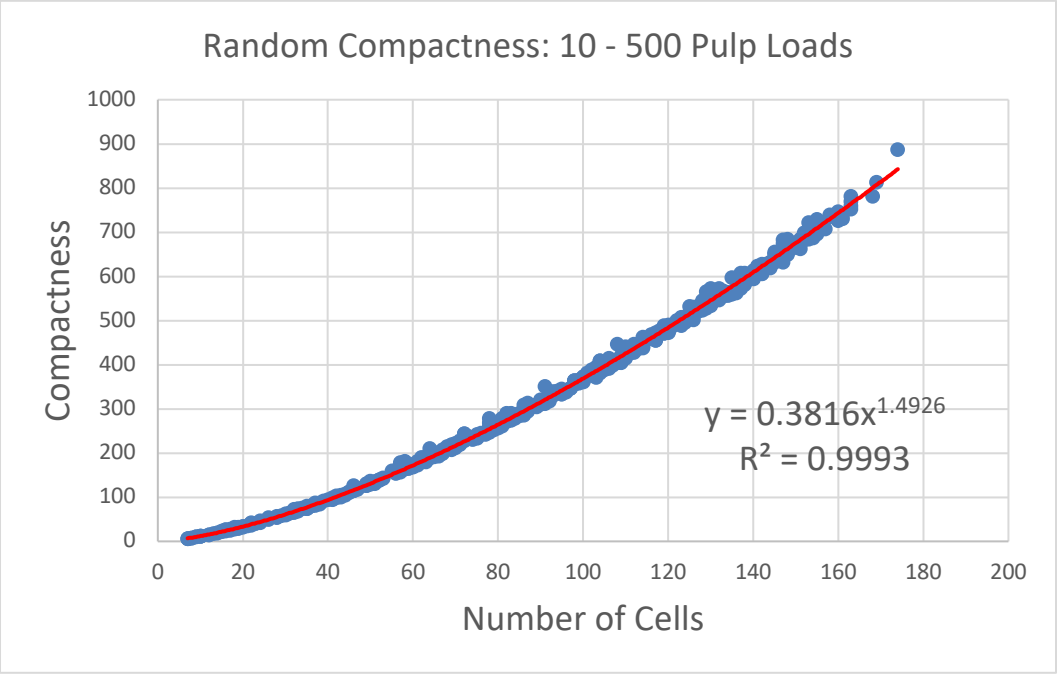

Figure 30: Random rule – compactness data for 10 – 500 pulp loads.

Statistical analysis of the random rule regression coefficients in the narrow and broad samples found significant differences between random data and ideal data ($p < 0.01$). Significant differences between regression coefficients of random data and real nest data ($p < 0.01$) were also found. The significant differences reinforce the observations from visual inspection of randomly generated nest structures—random nests are not compact, implying that the compact structures seen in idealized nests and real nests have a low likelihood of being the result of random construction behaviors.

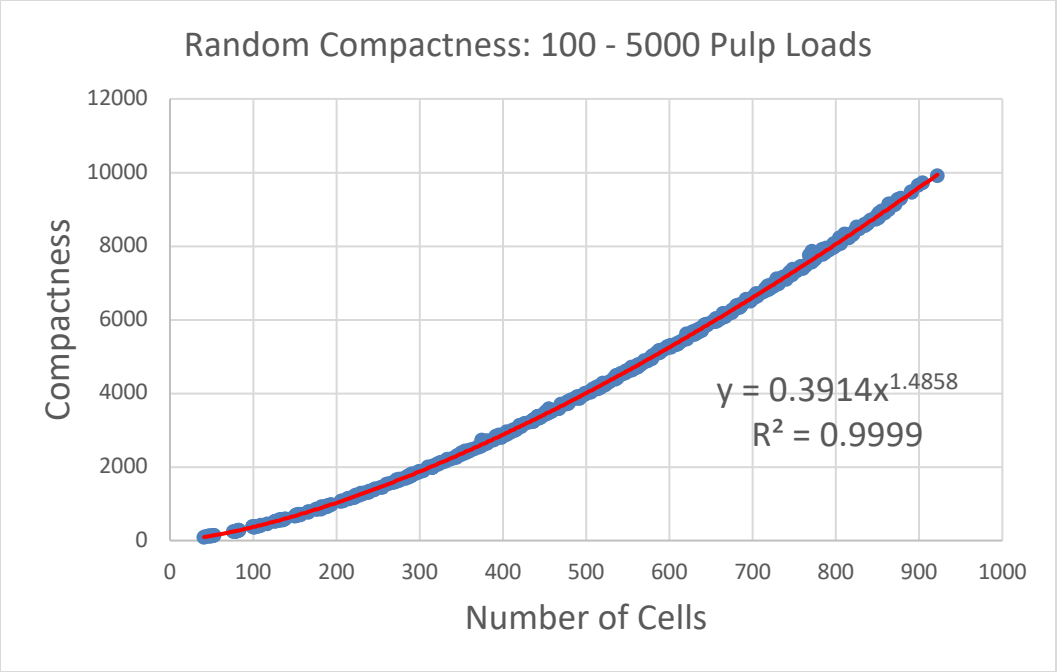

Figure 31: Random rule – compactness data for 100 – 5000 pulp loads.

5.5.2: Wall Rules

The wall-based rules generated the least compact nest structures of the non-random rules: visually, the wall-based nests appeared to have compactness numbers closer to random data than to data from the age-based and height-based rules. Number of cells strongly predicts compactness numbers of nests generated with wall-based rules ($R^2 > 0.99$). Figure 32 shows the scatterplot of the maximum wall rule in the broad sample. Scatterplots for the maximum wall rule in the narrow sample (Fig. 77) and for the hybrid wall rule in both samples (Figs. 73 and 74) are in Appendix C.

Statistical analysis of maximum wall rule regression coefficients confirms what was visually observed regarding compactness similarities to randomly-generated nests. There was no significant difference between regression coefficients of the maximum wall data and the random data in the narrow sample ($p > 0.1$). The similarity between random and maximum wall

compactness data increases in the broad sample: the *p*-value between random data and maximum wall data in the broad sample was greater than 0.5, indicating there is little difference in compactness between maximum wall rule nests and random rule nests. Significant differences between maximum wall data and idealized data were found in both the narrow sample ($p < 0.01$) and the broad sample ($p < 0.01$). There was no significant difference in regression coefficients between maximum wall data and real data in the narrow sample ($p > 0.01$); there was a significant difference in the broad sample ($p < 0.01$). However, the much-larger *p*-value between random data and maximum wall data suggests maximum walls have more in common with random nests than real nests.

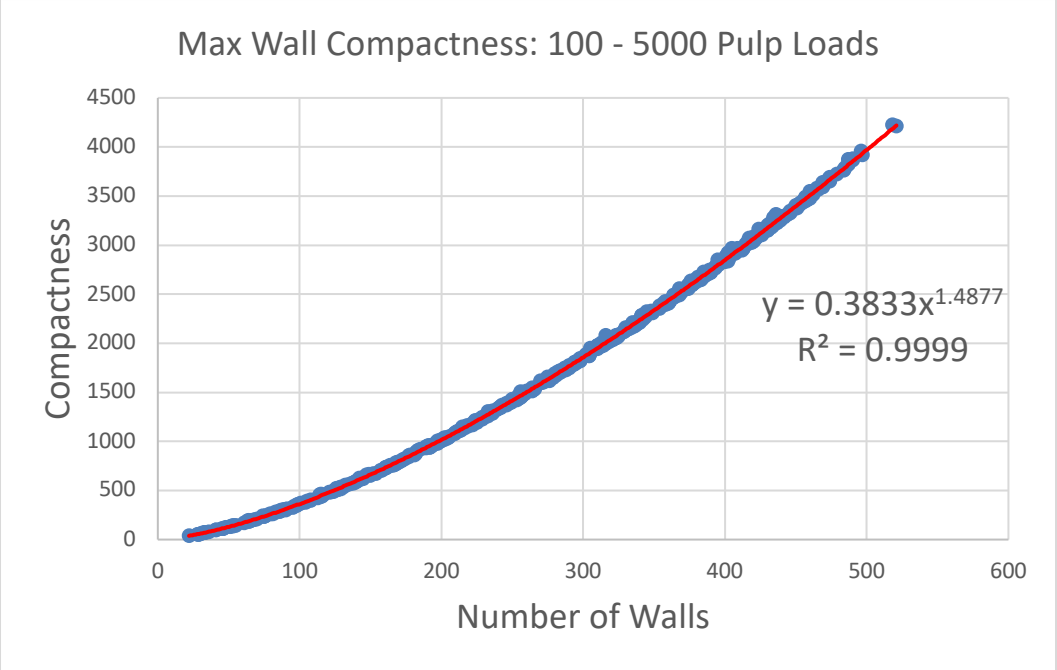

Figure 32: Max wall rule – compactness data for 100 – 5000 pulp loads.

Analysis of hybrid wall nest data along the compactness number parameter resulted in similar *p*-values as those from the maximum wall rule analysis. No significant difference was

found between regression coefficients of random data and hybrid wall data in the narrow sample ($p > 0.1$) and in the broad sample ($p > 0.5$). Significant differences in regression coefficients between maximum wall data and idealized data were found in both the narrow sample ($p < 0.01$) and the broad sample ($p < 0.01$). Regression coefficients between hybrid wall data and real data had no significant difference in the narrow sample ($p > 0.01$) but had a significant difference in the broad sample ($p < 0.01$).

5.5.3: Age Rules

The age-based rules, from a visual standpoint, generated relatively compact nest structures. Coefficients of determination were all high ($R^2 > 0.99$) for the scatterplots of the age-based rules. Figure 33 is the scatterplot for the maximum age rule in the broad sample. Appendix C contains figures for the maximum age rule in the narrow sample (Fig. 75) and for the hybrid age rule in both the narrow sample (Fig. 69) and the broad sample (Fig. 70).

Statistical analysis of the maximum age rule by the compactness number parameter yielded a mixed result. There was no significant difference in the narrow sample's regression coefficients between maximum age data, random data ($p > 0.05$), and real data ($p > 0.05$). There was a significant difference in the narrow sample's regression coefficients between maximum age data and idealized data ($p < 0.01$). For the narrow sample, it seems the maximum age nests occupy a middle ground of compactness between real nests and random nests since real nests and random nests significantly differ in compactness. The broad sample had no significant difference in the regression coefficients between maximum age data and random data ($p > 0.05$). There was a significant difference between coefficients in the maximum age data, idealized data ($p < 0.01$), and real data ($p < 0.01$).

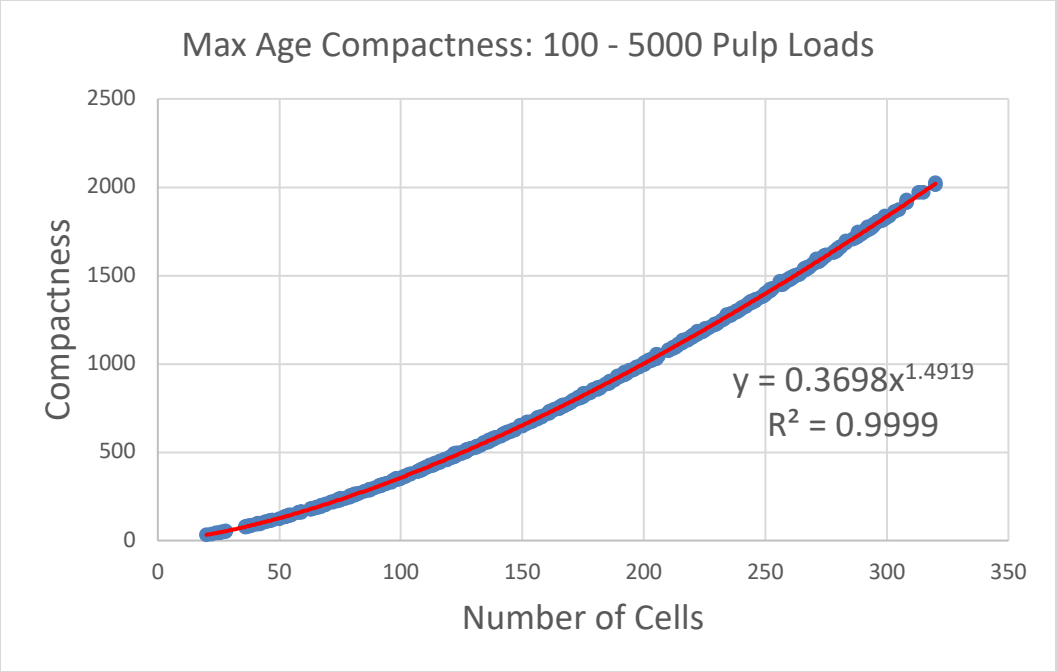

Figure 33: Max age rule – compactness data for 100 – 5000 pulp loads.

The hybrid age rule's compactness parameter also gave a mixed result; however, the hybrid age data tended more towards random data than real data. There was no significant difference in the narrow sample's regression coefficients between hybrid age data, random data ($p > 0.1$), and real data ($p > 0.01$). As evidenced by the p-values, the narrow sample's hybrid data had more in common with random data than real data. In the broad sample, there was no significant difference in the regression coefficients between hybrid age data and random data ($p > 0.05$); there was a significant difference between the coefficients in the hybrid age data and the real data ($p < 0.01$). Idealized data regression coefficients differed significantly from hybrid age data coefficients in both the narrow sample ($p < 0.01$) and the broad sample ($p < 0.01$).

5.5.4: Height Rules

The height-based rules, like the age-based rules, generated relatively compact structures from a visual standpoint. The fitted power functions for the height-based scatterplots all had high coefficients of determination ($R^2 > 0.99$). Figure 34 is a scatterplot of the maximum height rule data from the broad sample. Scatterplots for the maximum height rule data from the narrow sample (Fig. 76) and hybrid height rule data for both samples (Figs. 71 and 72) are in Appendix C.

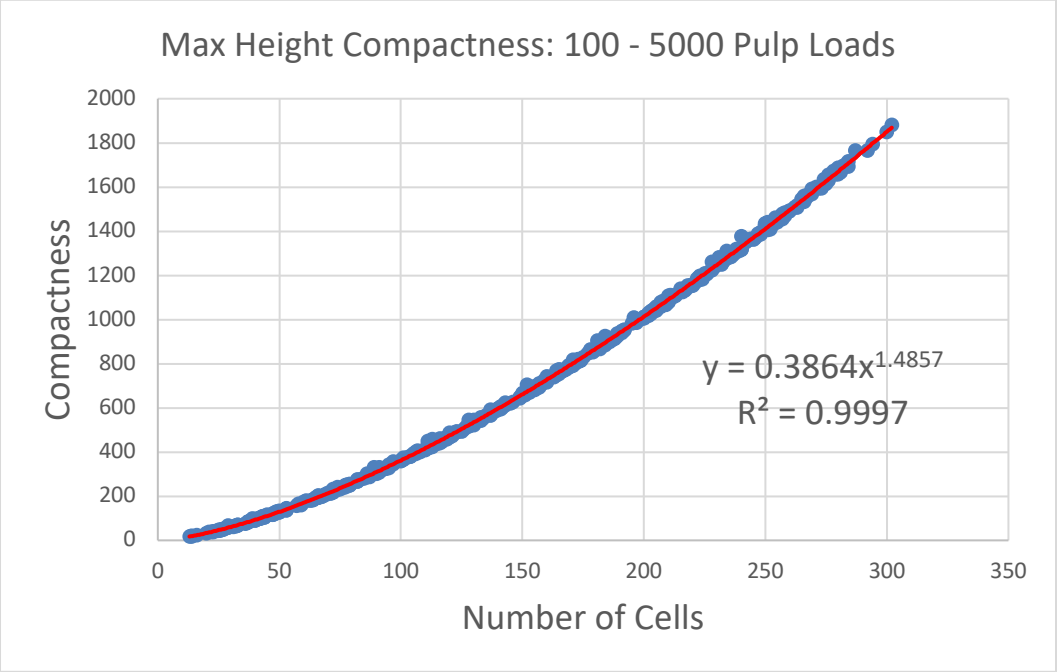

Figure 34: Max height rule – compactness data for 100 – 5000 pulp loads

The statistical analysis of the maximum height rule in the narrow sample yielded an interesting result. Significant differences were found between regression coefficients of the maximum height data, random data ($p < 0.01$), and idealized data ($p < 0.01$). Real nest compactness data and maximum height compactness data were remarkably similar: no

significant difference in regression coefficients were found ($p > 0.9$). In contrast, the broad sample's maximum height data had no significant difference in regression coefficients with random data ($p > 0.5$) but had significant differences with both idealized data ($p < 0.01$) and real data ($p < 0.01$).

Analysis of the hybrid height rule revealed a similar result to that of the maximum height rule, albeit with less extreme *p*-values. The same statistically significant similarity to real nest data was found in the hybrid height rule's narrow sample ($p > 0.7$) while significant differences were seen in the regression coefficients of random data ($p < 0.01$) and idealized data ($p < 0.01$). The broad sample data found no significant difference in regression coefficients between the hybrid height data and the random data ($p > 0.4$). Significant differences were found in the regression coefficients of hybrid height data with idealized data ($p < 0.01$) and real data ($p < 0.01$).

CHAPTER 6

DISCUSSION

This report described two separate pieces of work: the foundational design and implementation of a hexagonal coordinate system and the design, implementation, and analysis of a novel three-dimensional nest construction simulation that stochastically modeled *P. dominula* decision making with simple rules. To perform this work, an applications suite of seven different utilities, including two-dimensional and three-dimensional visualization utilities, simulations, and analysis tools were designed and implemented across three programming languages. The simulations written in the foundational work and the experimental work accomplished the goals set out for each in their respective designs. The hexagonal coordinate system improved simulation performance by implementing integer-based coordinates, which allowed direct comparison of hexagonal coordinate calculations and Cartesian coordinate calculations. Hexagonal coordinates also provided a greater than three times improvement in execution time in the procedural nest generation utility.

The design and implementation of the novel three-dimensional simulation was completed successfully with nearly a million nests generated using simple rules. Visual analysis of the generated nest structures gave a good baseline impression of each rule and helped shape the direction of the experiment by demonstrating the unnatural nest forms resulting from unconstrained nest generation. Visual inspection also helped narrow down which rules were more likely to generate natural nest forms, shaping the statistical analysis. The statistical analysis found that while no single rule could always generate realistic nest data across both parameters, the age-based rules, paired with the site selection constraints, were able to consistently generate

nests that resembled real nest data in terms of outer cell wall counts and compactness numbers. The age-based rules make a good basis for future work based on stochastic site selection.

6.1: Analysis of Hypotheses

The statistical analysis was performed to evaluate the null and alternate hypotheses proposed in Chapter 4, section 4.1.2. The first question asked for this experiment was whether randomly-generated nests differed from non-randomly generated nests. The null hypothesis for random nests stated that there was no difference between random nests and non-random nests along two parameters: outer wall counts and compactness numbers (both as a function of the number of cells in the compared nests). The alternate hypothesis stated there was a difference between random nests and non-random nests along the two parameters. In comparison to real and idealized nest data, the null hypothesis was rejected: there are statistically significant differences between random data, real nest data, and idealized data. The significant differences found when calculating the regression coefficients between random nests and real nests in both parameters agrees with results seen in prior research (Karsai and Pénzes 1993; Karsai and Pénzes 2000).

Analysis of the random rule against the wall-based, age-based, and height-based rules showed no clear answers across both parameters and both samples. Using outer wall counts as the nest parameter, the null hypothesis was rejected for the age-based rules, accepted for the height-based rules, and, depending on which sample, rejected (narrow sample) and accepted (broad sample) for the wall-based rules. When compactness is the nest comparison parameter, the null hypothesis is accepted for both the age-based rules and the wall-based rules across both samples, meaning no significant difference was found between those non-random data and random data in terms of nest compactness as a function of number of cells. In the narrow sample,

the height-based rules rejected the null hypothesis, i.e., were significantly different from random data; however, the null hypothesis was accepted in the broad sample.

The failure to reject the null hypothesis in the compactness parameter comparisons across all three rules may point to some factor all the rules had in common: the site selection constraints. The enforced selection of initiation sites that have at least two existing walls and lengthening sites with at least two taller neighbors may exert more pressure on the shaping of the nest than the individual rules do in terms of overall nest compactness. However, this trend did not appear in the analysis of outer wall counts. Analysis of the unconstrained nests may help shed some light on the impact of the constraints on overall site selection weighting.

6.1.1: Idealized and Real Data Analysis

The second experimental question was whether idealized nest data differed significantly from generated nest data. Across all rules, both random and non-random, and across both samples, the null hypothesis was rejected. Idealized nest forms likely differ significantly from all constrained nest data. This is not surprising: real nest data is significantly different from idealized nest data using the compactness parameter as comparison. Some nests, such as narrow-sample height-based nests, did come closer to the 99% confidence level threshold for significance; however, no generated nest data was ever less than a power of 10 away from crossing the significance threshold.

The third experimental question was whether real nest data differed significantly from generated nest data. Results of this hypothesis test depended on the rule, the sample, and the parameter. The wall-based rules, with one exception, all rejected the null hypothesis: wall rule nests were significantly different from real nests. The one occurrence where the null hypothesis

was accepted was in the narrow sample along the compactness parameter; however, *p*-values were not far above 0.01.

The age-based rules often failed to reject the null hypothesis, meaning the age-based nest structures had both parameters in common with real nests at equivalent numbers of cells. The cases where the null hypothesis was rejected are the broad sample compactness and the maximum age rule's narrow outer wall counts. The maximum age rejection was close to the threshold; a different random sample may result in failing to reject the null hypothesis. The compactness broad sample rejections are part of a larger trend observed across all the non-random generation rules. There are several possible causes for this trend. One possible explanation is the inability of these rules to predict nests at larger sizes. Another possible explanation is the limitation of the model as a single-agent, fully observable state model. Large *P. dominula* nests have multiple agents with partial environmental knowledge of the nest. Those wasp agents do not wander indefinitely; another species of *Polistes* averages 38 seconds and 18 site checks before making a final building decision (H. A. Downing 1994). A third explanation may be the previously-described constraint weighting forcing nests down a certain path, regardless of rule influence.

The height-based rules, with one exception, led to the rejection of the null hypothesis in comparison to real nest data. The exception was the narrow sample for the compactness number parameter, which found a remarkable similarity in compactness between real nests and height-based nests. One reason for this similarity may be how real nests grow at this phase. Once real nests reach 12-16 cells, they undergo an extended lengthening stage to grow the cells tall enough to accommodate the hatched larvae (Karsai and Pénzes 1996). After this lengthening phase,

further site initiations resume to continue growing the nest; nest lengthening does not continue much further past the point reached during the larval stage (Karsai and Pénzes 1996).

6.2: Study Limitations

This study was not without limitations. One limitation was the lack of characterization of nests in three dimensions versus two-dimensional analysis. Three-dimensional compactness of simulated structures was calculated but had two issues: it was not directly comparable to real nest compactness numbers and three-dimensional compactness was not at the same scale as the other three axes of the nest, limiting its usefulness. A second limitation was the debatable usefulness of the hybrid rules. Analysis showed little difference between hybrid and maximum versions of the wall, age, and height rules. Resolving the lack of three-dimensional characterization may help address the effect hybrid rules have on nest construction. Adding scaling coefficients to the stochastic site selection model to increase or decrease the influence of each rule may also help differentiate the hybrid rules from the maximum rules.

Another limitation was the role constraints played in site selection. Unconstrained nests were subjected to visually inspected and the unconstrained rules were discarded after discovery of chimney elongation and single-wall initiations. Further investigation into the difference between fully constrained, partially constrained, and unconstrained rules may shed light on how much constraints affected nest parameters. Finally, none of these findings directly address the underlying stimuli of how *P. dominula* construct nests. The age-based rule, for example, could be influenced by multiple stimuli, including the nest structure itself, pheromones deposited on the structure after various worker actions (H. Downing 1991), or pupae in the cells (Karsai 1997).

6.3: Future Work

One avenue for future projects to pursue is removing abstractions and simplifications from the model and attempting to separate as many stimuli into discrete variables as possible. For example, the age-based rule could be limited to site initiations, as in the model from (Karsai and Pénzes 2000) and another mechanism designed for cell lengthening, such as implementing realistic brood mechanics as a stimulus for cell elongation. The design of the stochastic site selection model could serve as a good basis for implementation of discrete stimuli and less generalized models, e.g., a site lengthening model weighted by brood mechanics and a site initiation model weighted by wall ages or pheromone deposits.

A second direction for future work is the implementation of a multi-agent simulation that utilizes workers with limited knowledge of the nest structure. Additional features for this future work could include worker collisions (and how those collisions affect site selection) and real-time simulation of pulp gathering to examine how timing affects the limited decision-making windows *Polistes* workers possess.

BIBLIOGRAPHY

Adoe, Fadel. *Generating Compact Wasp Nest Structures via Minimal Complexity Algorithms.* MS Thesis, Johnson City: East Tennessee State University, 2010.

Ashby, W. R. "Principles of the Self-organizing Dynamic System." *The Journal of General Psychology*, 1947: 125-128.

Beckers, Ralph, O. E. Holland, and Jean-Louis Deneubourg. "From Local Actions to Global Tasks: Stigmergy and Collective Robotics." 1994.

Camazine, Scott. *Self-organization in Biological Systems.* Princeton: Princeton University Press, 2003.

Clausius, Rudolf. "I. On the Moving Force of Heat, and the Laws Regarding the Nature of Heat Itself Which are Deducible Therefrom." *The London, Edinburgh, and Dublin Philosophical Magazine and Journal of Science*, 1851: 1-21.

De Rango, Floriano, Nunzia Palmieri, Xin She Yang, and Salvatore Marano. "Bio-inspired Exploring and Recruiting Tasks in a Team of Distributed Robots over Mined Regions." *Proceedings of the International Symposium on Performance Evaluation of Computer and Telecommunication Systems.* Society for Modeling & Simulation International (SCS), 2015. 1-8.

Descartes, René. *Discourse on the Method of Rightly Conducting One's Reason and of Seeking Truth in the Sciences.* 1637.

Dorigo, Marco, and Mauro Birattari. "Ant Colony Optimization." In *Encyclopedia of Machine Learning*, by Claude, and Geoffrey Webb Sammut, 36-39. Boston: Springer, 2011.

Doty-Humphrey, Chris. *PractRand.* 2014. http://pracrand.sourceforge.net/ (accessed 2018).

Downing, H. A. "Information Analysis by the Paper Wasp, Polistes fuscatus, During Nest Construction (Hymenoptera, Vespidae)." *Insectes Sociaux*, 1994: 361-377.

Downing, Holly A., and Robert L. Jeanne. "Nest Construction by the Paper Wasp, Polistes: A Test of Stigmergy Theory." *Animal Behaviour*, 1988: 1729-1739.

Downing, Holly. "The Function and Evolution of Exocrine Glands." In *The Social Biology of Wasps*, by K.G. Ross, & R.W. Matthews, 540-569. Ithaca: Comstock, 1991.

Eleftherakis, George, Milos Kostic, Konstantinos Rousis, and Anca Vasilescu. "Stigmergy Inspired Approach to Enable Agent Communication in Emergency Scenarios." *Proceedings of the 7th Balkan Conference on Informatics.* ACM, 2015. 22-29.

Fraser, Bradley, and Robert Hunjet. "Data Ferrying in Tactical Networks Using Swarm Intelligenze and Stigmergic Coordination." *26th International Telecommunication Networks and Applications Conference (ITNAC).* IEEE, 2016. 1-6.

Grassé, Plerre-P. "La Reconstruction du Nid et les Coordinations Interindividuelles chez Bellicositermes natalensis et Cubitermes sp. la Théorie de la Stigmergie: Essai d'interprétation du Comportement des Termites Constructeurs." *Insectes Sociaux*, 1959: 41-80.

Hamburg, Mike, Paul Kocher, and Mark E. Marson. *Analysis of Intel's Ivy Bridge Digital Random Number Generator.* San Francisco, March 12, 2012.

Harris, W. Victor, and W. A. Sands. "The Social Organization of Termite Colonies." *Symposia of the Zoological Society of London.* London: The Society, 1965. 113-131.

Holland, Owen, and Chris Melhuish. "Stigmergy, Self-organization, and Sorting in Collective Robotics." *Artificial Life*, 1999: 173-202.

Huber, Pierre. *Recherches sur les Moeurs des Fourmis Indigènes.* Paris: Libr. Fischbacher, 1861.

Intel Corporation. *Intel® Digital Random Number Generator (DRNG) Software Implementation Guide.* 2014. https://software.intel.com/en-us/articles/intel-digital-random-number-generator-drng-software-implementation-guide (accessed 2017).

Johnson, Brian R. "Pattern Formation on the Combs of Honeybees: Increasing Fitness by Coupling Self-organization with Templates." *Proceedings of the Royal Society of London B: Biological Sciences.* London: Royal Society Publishing, 2009. 255-261.

Karsai, István. "Brood Patterns in Wasp Combs: The Influence of Brood on Egg-Laying and Building by Adults." *Ethology Ecology & Evolution*, 1997: 27-44.

Karsai, István. "Decentralized Control of Construction Behavior in Paper Wasps: An Overview of the Stigmergy Approach." *Artificial Life*, 1999: 117-136.

Karsai, István, and Guy Theraulaz. "Nest Building in a Social Wasp: Postures and Constraints (Hymenoptera: Vespidae)." *Sociobiology*, 1995: 83-86.

Karsai, István, and Zsolt Pénzes. "Comb Building in Social Wasps: Self-organization and Stigmergic Script." *Journal of Theoretical Biology*, 1993: 505-525.

Karsai, István, and Zsolt Pénzes. "Intraspecific Variation in the Comb Structure of Polistes dominulusi Parameters, Maturation, Nest Size and Cell Arrangement." *Insectes sociaux*, 1996: 277-296.

Karsai, István, and Zsolt Pénzes. "Nest Shapes in Paper Wasps: Can the Variability of Forms be Deduced from the Same Construction Algorithm?" *Proceedings of the Royal Society of London B: Biological Sciences.* London: Royal Society, 1998. 1261-1268.

Karsai, István, and Zsolt Pénzes. "Optimality of Cell Arrangement and Rules of Thumb of Cell Initiation in Polistes dominulus: a Modeling Approach." *Behavioral Ecology*, 2000: 387-395.

L'Ecuyer, Pierre, and Richard Simard. "TestU01: A C Library for Empirical Testing of Random Number Generators." *ACM Transactions on Mathematical Software*, 2007: 22.

Marsaglia, George. "Xorshift RNGs." *Journal of Statistical Software*, 2003: 1-6.

Matsumoto, Makato, and Takuji Nishimura. "Mersenne Twister: A 623-dimensionally Equidstributed Uniform Pseudo-random Number Generator." *ACM Transactions on Modeling and Computer Simulation*, 1998: 3-30.

Michener, Charles Duncan. *The Social Behavior of the Bees: A Comparative Study.* Cambridge: Harvard University Press, 1974.

Middleton, Eliza JT, and Tanya Latty. "Resilience in Social Insect Infrastructure Systems." *Journal of the Royal Society Interface*, 2016: 20151022.

Mizumoto, Nobuaki, Kazuya Kobayashi, and Kenji Matsuura. "Emergence of Intercolonial Variation in Termite Shelter Tube Patterns and Prediction of its Underlying Mechanism." *Royal Society Open Science*, 2015: 150360.

Rodriguez, Arles, Jonatan Gomez, and Ada Diaconescu. "Foraging-inspired Self-organisation for Terrain Exploration with Failure-prone Agents." *9th International Conference on Self-Adaptive and Self-Organizing Systems (SASO).* IEEE, 2015. 121-130.

Rosalie, Martin, Matthias R. Brust, Gregoire Danoy, Serge Chaumette, and Pascal Bouvry. "Coverage Optimization with Connectivity Preservation for UAV Swarms applying Chaotic Dynamics." *International Conference on Autonomic Computing (ICAC).* IEEE, 2017. 113-118.

Rukhin, Andrew, Juan Soto, James Nechvatal, Miles Smid, and Elaine Barker. "A Statistical Test Suite for Random and Pseudorandom Number Generators for Cryptographic Applications." Mclean: Booz-Allen and Hamilton, Inc., 2001.

Şahin, Erol. "Swarm Robotics: From Sources of Inspiration to Domains of Application." *International Workshop on Swarm Robotics.* Berlin: Springer, 2004. 10-20.

Shapiro, Carl, and Hal R. Varian. *Information Rules: A Strategic Guide to the Network Economy.* Watertown: Harvard Business Press, 1998.

Smith, Andrew P. "An Investigation of the Mechanisms Underlying Nest Construction in the Mud Wasp Paralastor sp. (Hymenoptera: Eumenidae)." *Animal Behavior*, 1978: 232-240.

Stuart, Alastair M. "Alarm, Defense, and Construction Behavior Relationships in Termites (Isoptera)." *Science*, 1967: 1123-1125.

Vigna, Sebastiano. "An Experimental Exploration of Marsaglia's Xorshift Generators, Scrambled." *ACM Transactions on Mathematical Software*, 2016: 30.

Vigna, Sebastiano. *Xoroshiro+ / xorshift* / xorshift+ Generators and the PRNG Shootout.* 2017. http://xoroshiro.di.unimi.it/ (accessed 2018).

Wilson, Edward O. *Sociobiology.* Cambridge: Harvard University Press, 2000.

APPENDICES

Appendix A

Proof of Direct Comparison of Compactness Numbers

This section contains a direct proof of the relationship observed between hexagonal coordinate sum of squared distances equation and Cartesian coordinate sum of squared distances equation. Equations 19, 20, and 21 show the initial steps to the proof.

$$D_H = 2\times D_R \tag{19}$$

$$2D_R = 2\sum_i (x_{ri} - x_{re})^2 + (y_{ri} - y_{re})^2 \tag{20}$$

$$= 2\sum_i x_{ri}^2 - 2x_{ri}x_{re} + x_{ri}^2 + y_{ri}^2 - 2y_{ri}y_{re} + y_{ri}^2 \tag{21}$$

Using Eq. 3, substitute hexagonal coordinates for Cartesian coordinates.

$$= 2\sum_i \left(\frac{x_{hi}+y_{hi}}{2}\right)^2 - 2\left(\frac{x_{hi}+y_{hi}}{2}\right)\left(\frac{x_{he}+y_{he}}{2}\right) + \left(\frac{x_{he}+y_{he}}{2}\right)^2 \tag{22}$$
$$+\left(\frac{\sqrt{3}}{2}z_{hi}\right)^2 - 2\left(\frac{\sqrt{3}}{2}z_{hi}\right)\left(\frac{\sqrt{3}}{2}z_{he}\right) + \left(\frac{\sqrt{3}}{2}z_{he}\right)^2$$

Expand and combine like terms.

$$= \frac{1}{2}\sum_i (x_{hi} - x_{he})^2 + (y_{hi} - y_{he})^2 + 2(z_{hi} - z_{he})^2 \tag{23}$$
$$+(z_{hi}^2 - 2z_{hi}z_{he} + z_{he}^2) + 2(x_{hi} - x_{he})\,(y_{hi} - y_{he})$$

Substitute hexagonal x and y into expanded hexagonal z terms using Eq. 4, 5, and 6.

$$= \frac{1}{2}\sum_i (x_{hi} - x_{he})^2 + (y_{hi} - y_{he})^2 + 2(z_{hi} - z_{he})^2 \quad (24)$$
$$+(y_{hi} - x_{hi})^2 - 2(y_{hi} - x_{hi})(y_{he} - x_{he}) + (y_{hi} - x_{hi})^2$$
$$+2(x_{hi} - x_{he})\,(y_{hi} - y_{he})$$

Expand and combine like terms.

$$= \frac{1}{2}\sum_i 2(x_{hi} - x_{he})^2 + 2(y_{hi} - y_{he})^2 + 2(z_{hi} - z_{he})^2 \quad (25)$$
$$-2(x_{hi} - x_{he})\,(y_{hi} - y_{he}) + 2(x_{hi} - x_{he})\,(y_{hi} - y_{he})$$

Hexagonal sum of squared distances equation results from substitution into Cartesian sum of squared distances equation.

$$= \sum_i (x_{hi} - x_{he})^2 + (y_{hi} - y_{he})^2 + (z_{hi} - z_{he})^2 \quad (26)$$

■

Appendix B

Heat Maps of Stratified Random Samples

This section contains heat maps not shown in Chapter 5's discussion of sampling quality. See Fig. 35 for the heat map key.

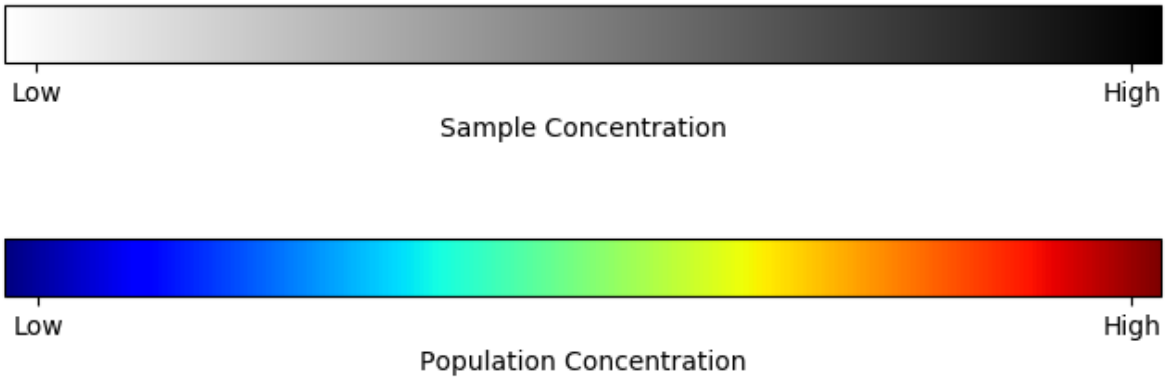

Figure 35: Color bars for heat map interpretation. The grayscale bar is for sampled points. The lighter the sample point color, the fewer number of samples taken at that point. The color bar is for the population of nests generated by a rule, separated by sample range and parameter (outer wall counts and compactness number). The cooler the color of the population points, the fewer individuals that had the parameter value at that number of cells.

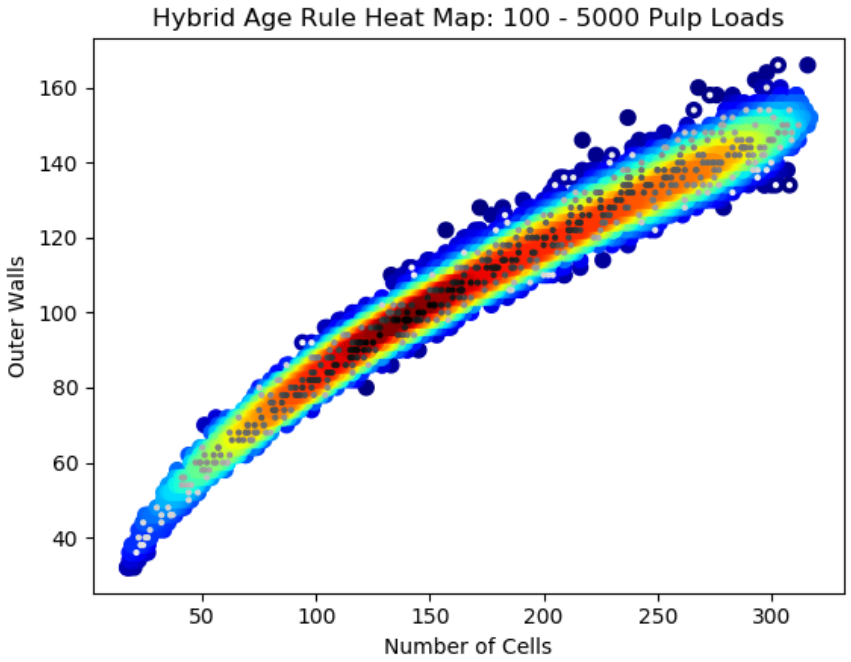

Figure 36: Hybrid age rule heat map of the outer walls parameter for the broad sample.

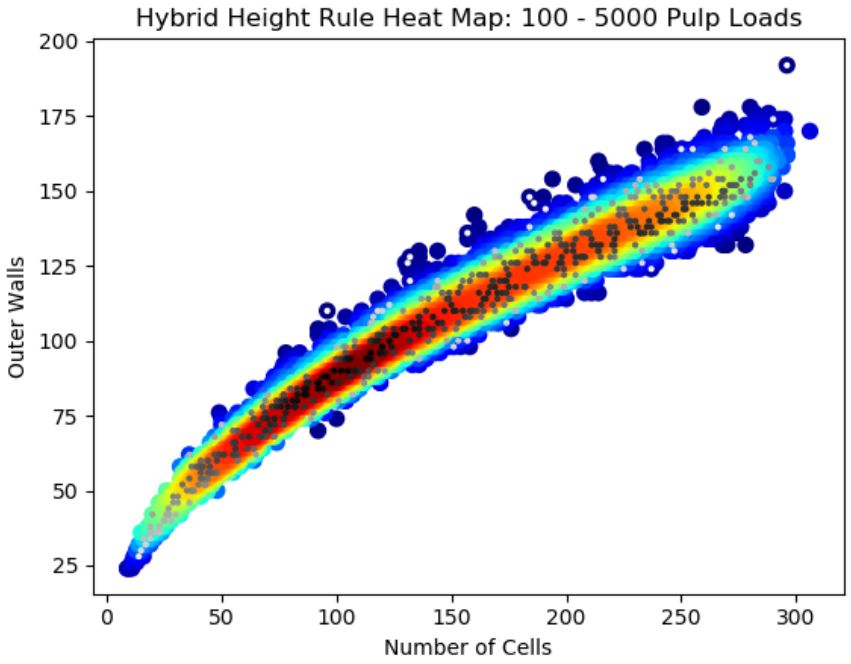

Figure 37: Hybrid height rule heat map of the outer walls parameter for the broad sample.

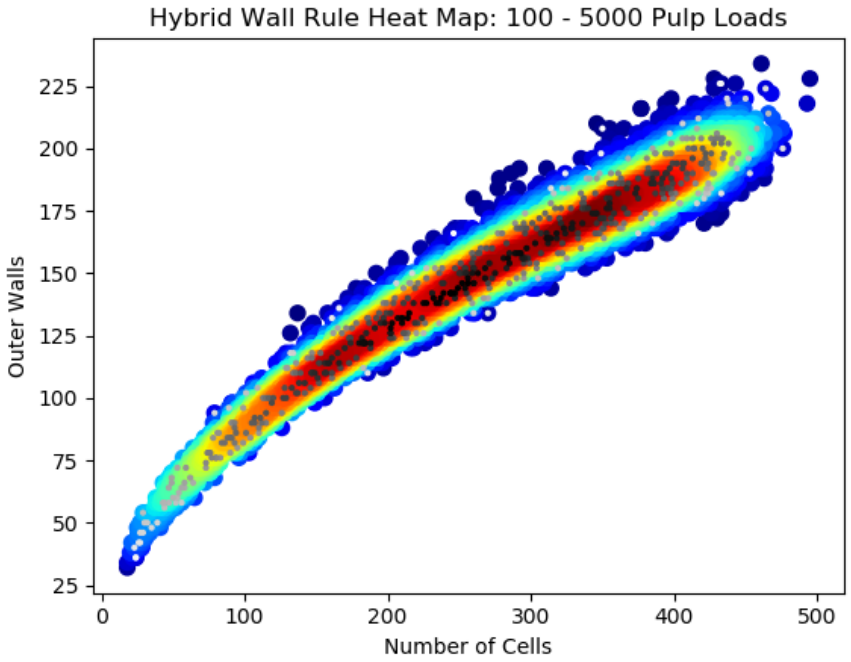

Figure 38: Hybrid wall rule heat map of the outer walls parameter for the broad sample.

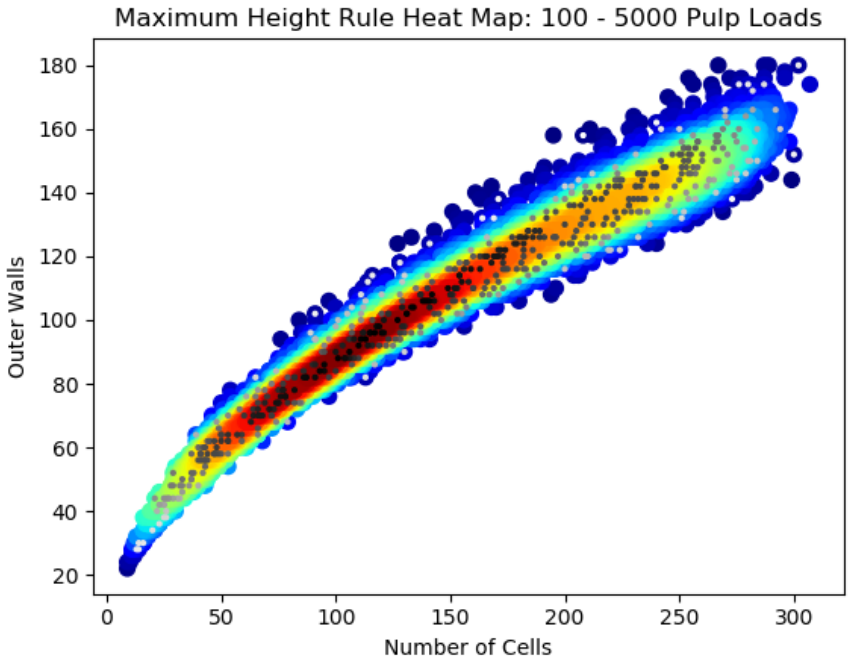

Figure 39: Maximum height rule heat map of the outer walls parameter for the broad sample.

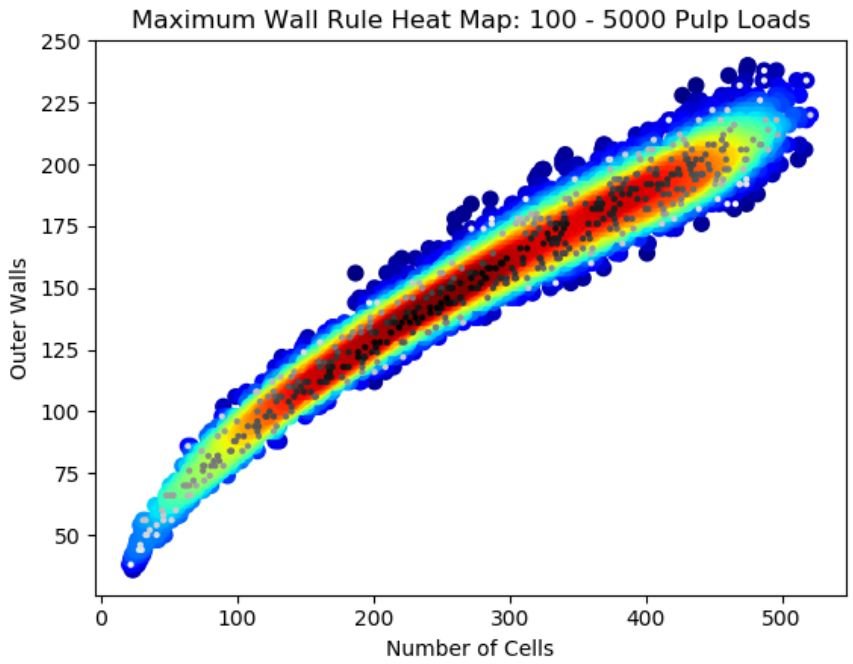

Figure 40: Maximum wall rule heat map of the outer walls parameter for the broad sample.

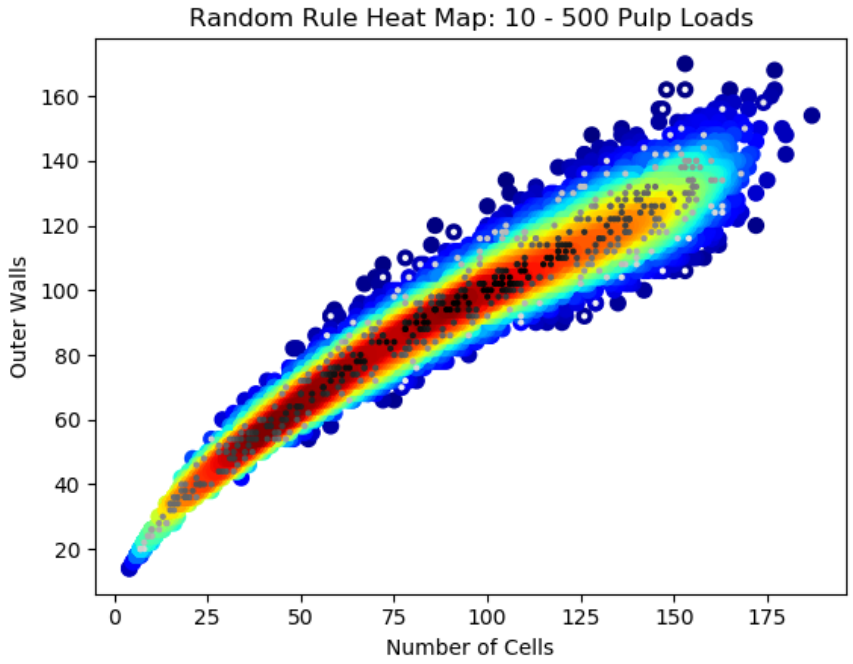

Figure 41: Random rule heat map of the outer walls parameter for the narrow sample.

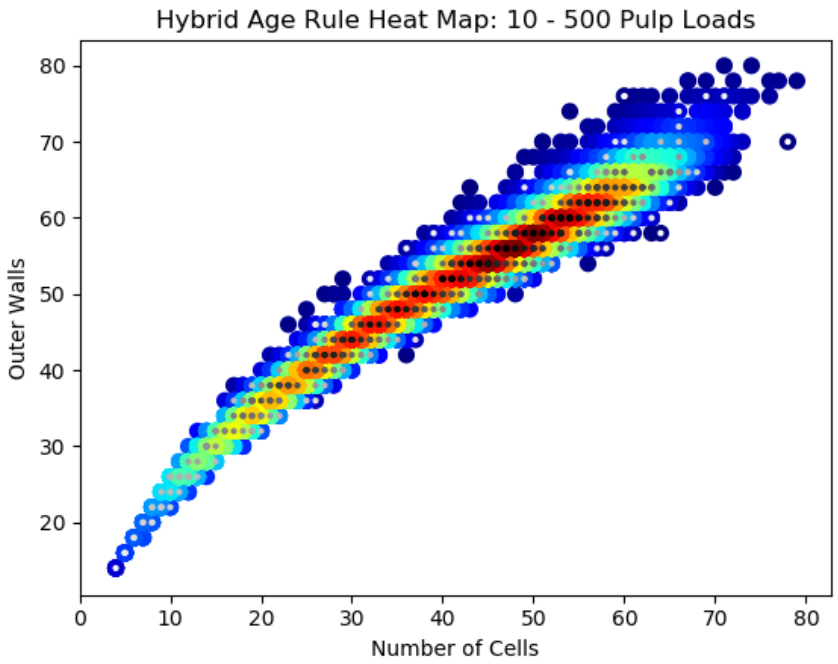

Figure 42: Hybrid age heat map of the outer walls parameter for the narrow sample.

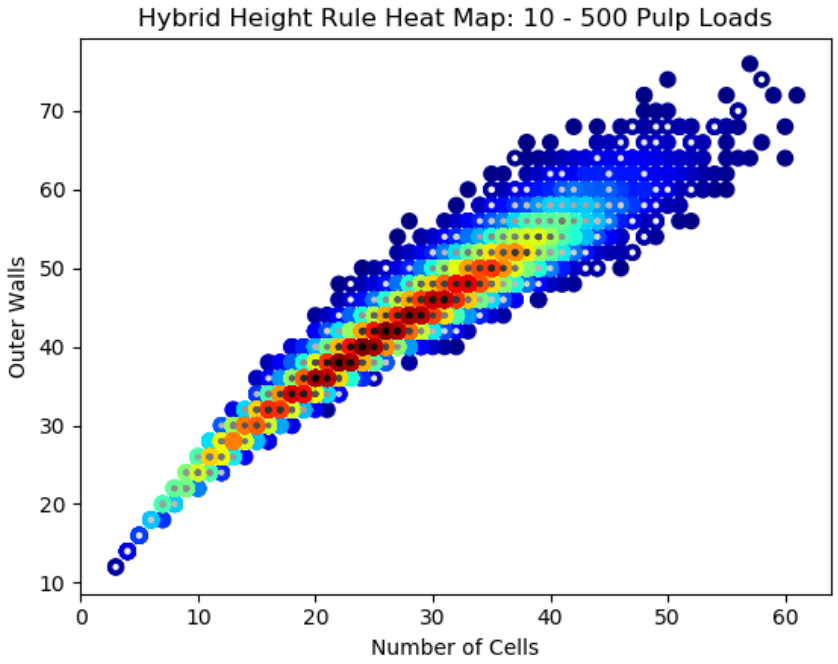

Figure 43: Hybrid height heat map of the outer walls parameter for the narrow sample.

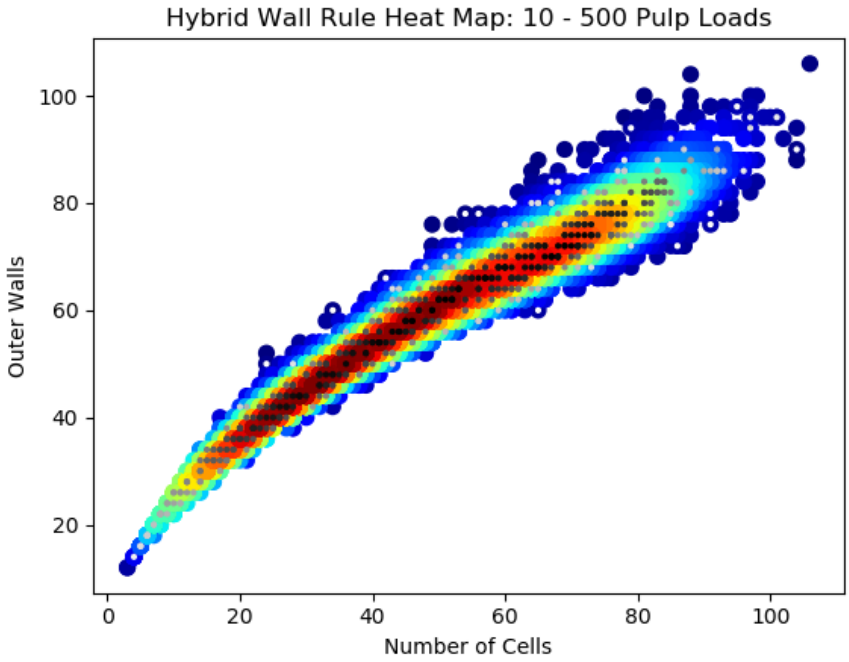

Figure 44: Hybrid wall heat map of the outer walls parameter for the narrow sample.

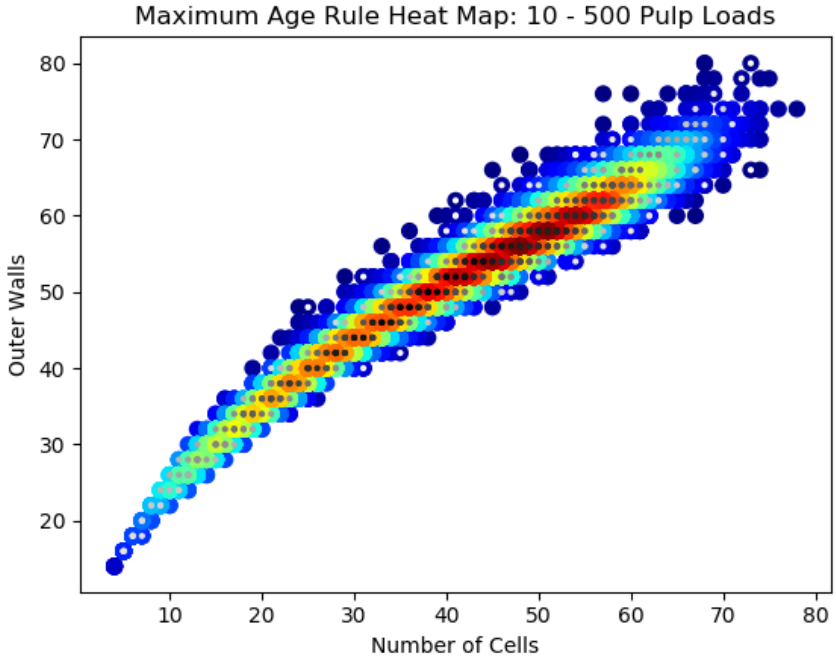

Figure 45: Maximum age heat map of the outer walls parameter for the narrow sample.

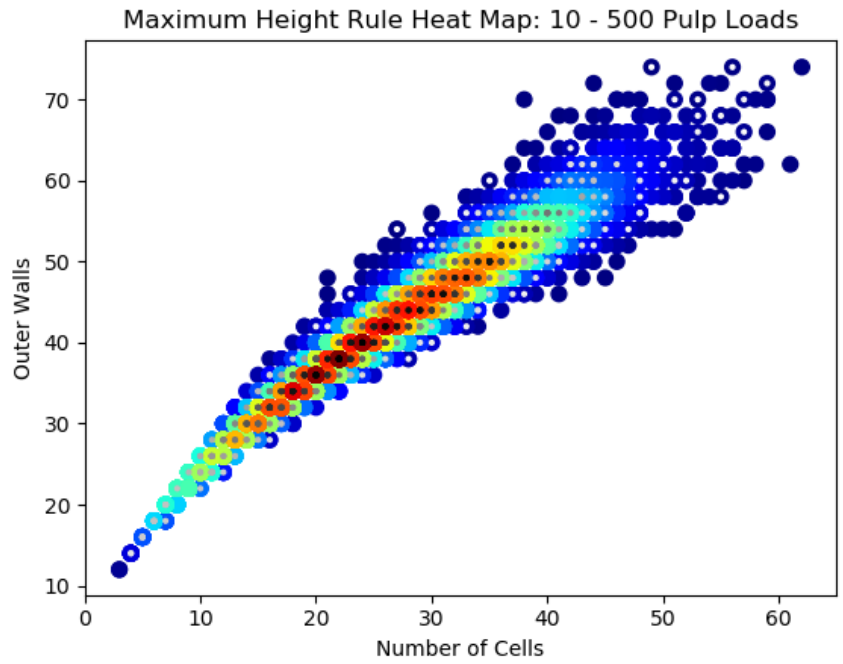

Figure 46: Maximum height heat map of the outer walls parameter for the narrow sample.

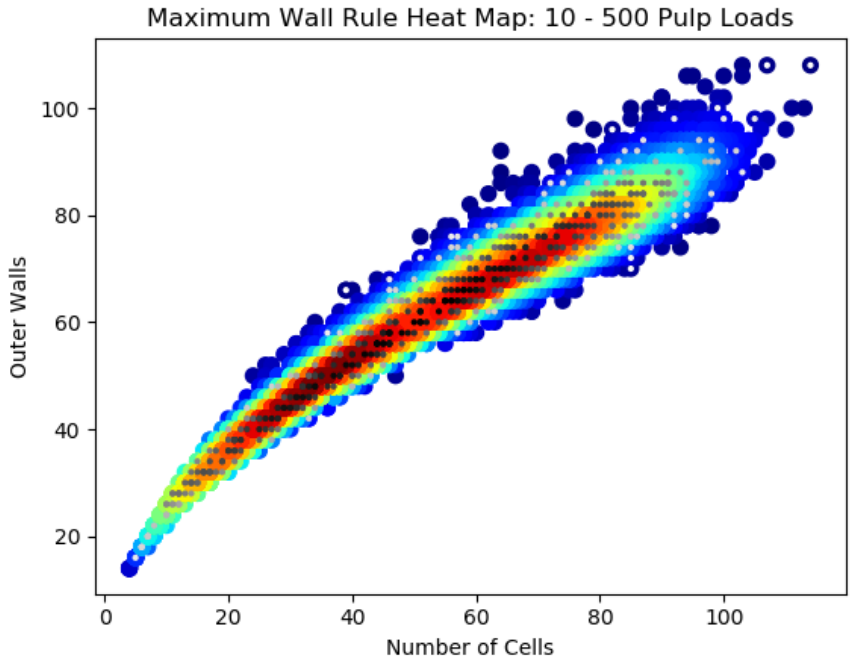

Figure 47: Maximum wall heat map of the outer walls parameter for the narrow sample.

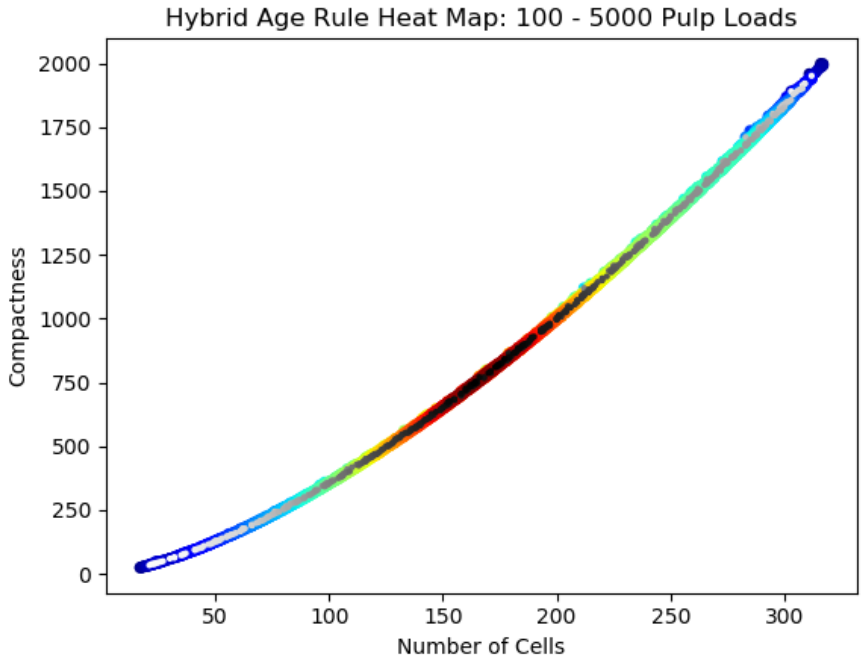

Figure 48: Hybrid age heat map of the compactness parameter for the broad sample.

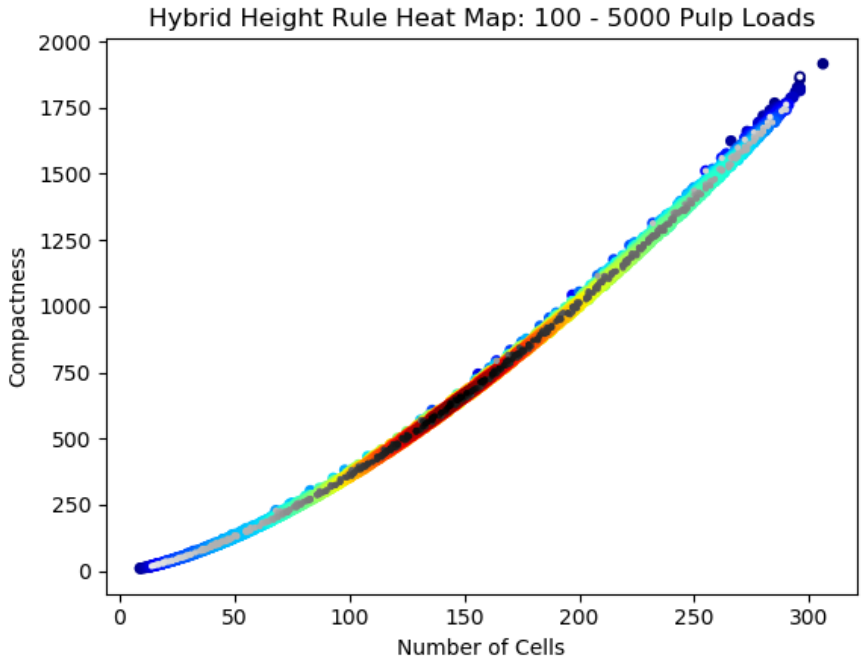

Figure 49: Hybrid height heat map of the compactness parameter for the broad sample.

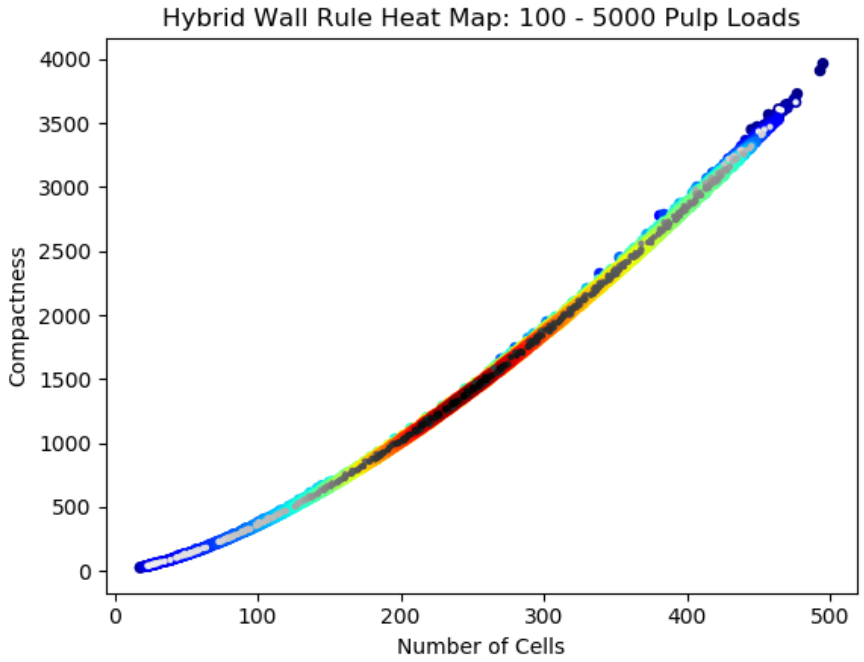

Figure 50: Hybrid wall heat map of the compactness parameter for the broad sample.

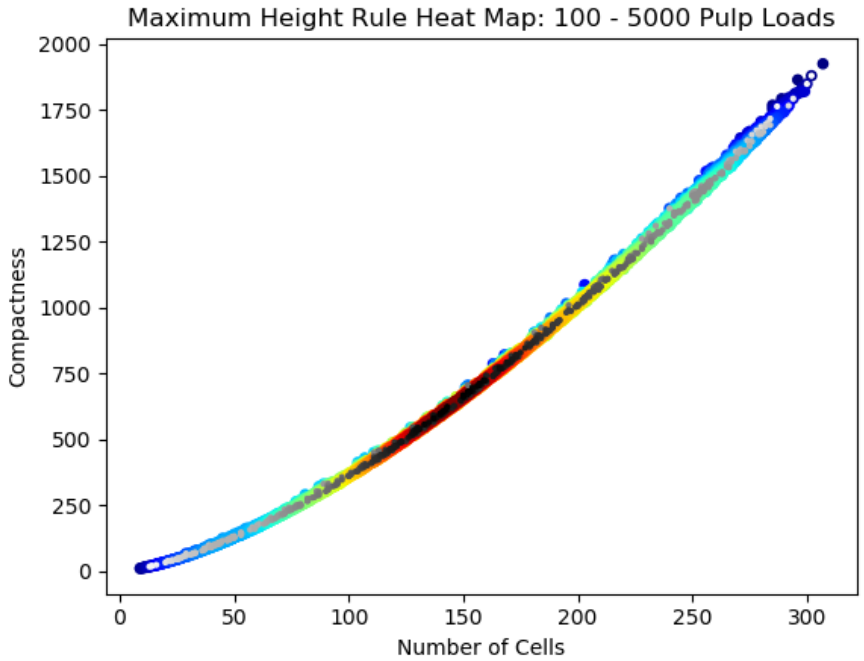

Figure 51: Maximum height heat map of the compactness parameter for the broad sample.

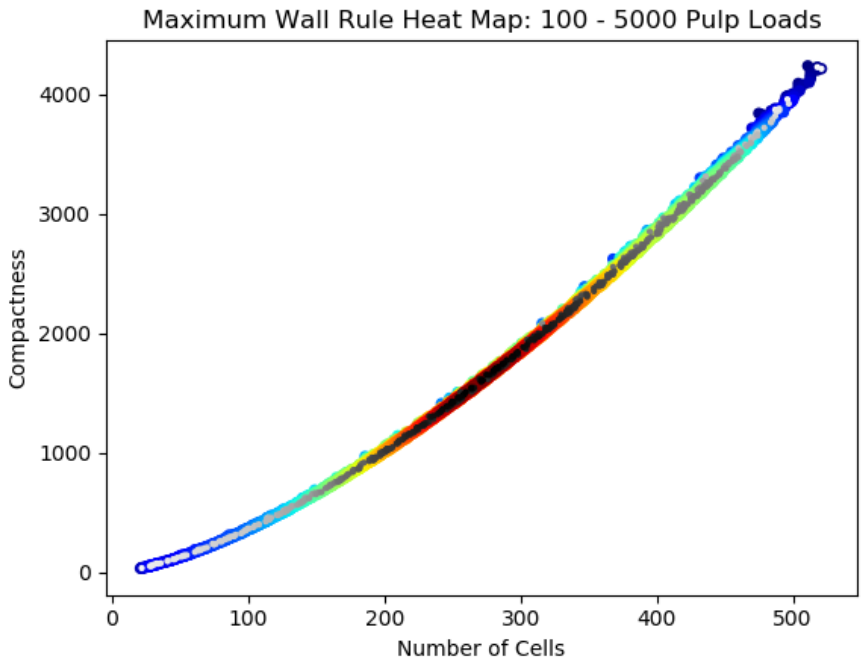

Figure 52: Maximum wall heat map of the compactness parameter for the broad sample.

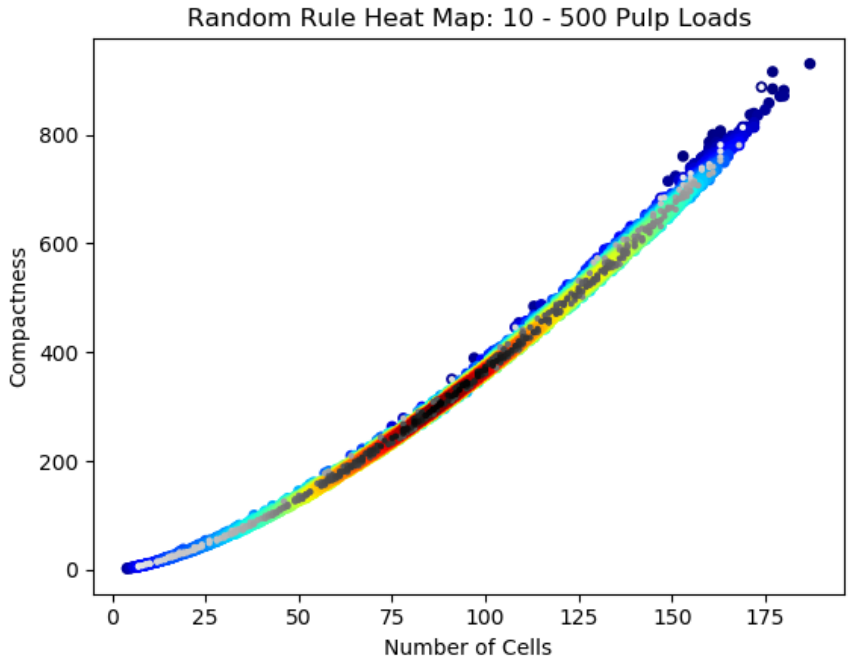

Figure 53: Random heat map of the compactness parameter for the narrow sample.

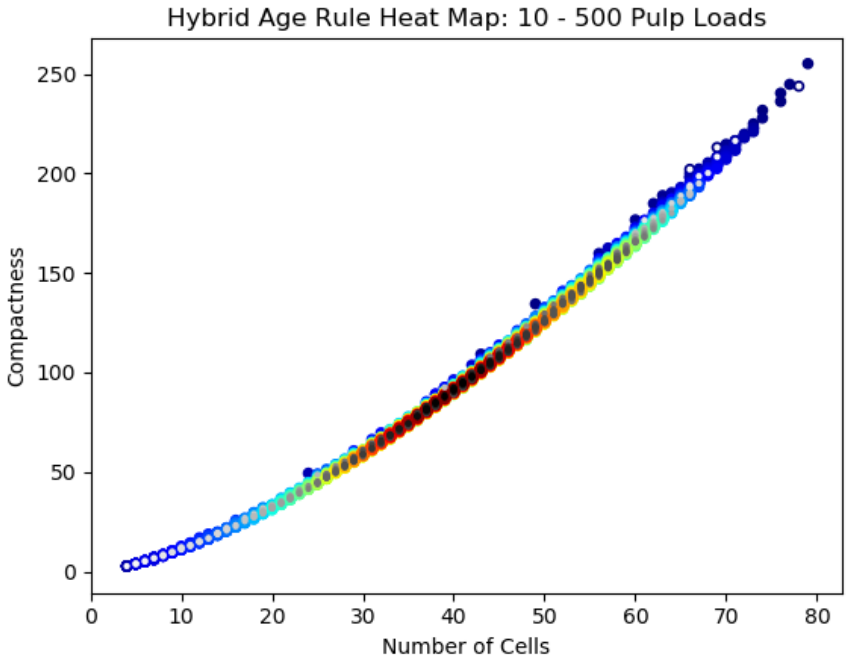

Figure 54: Hybrid age heat map of the compactness parameter for the narrow sample.

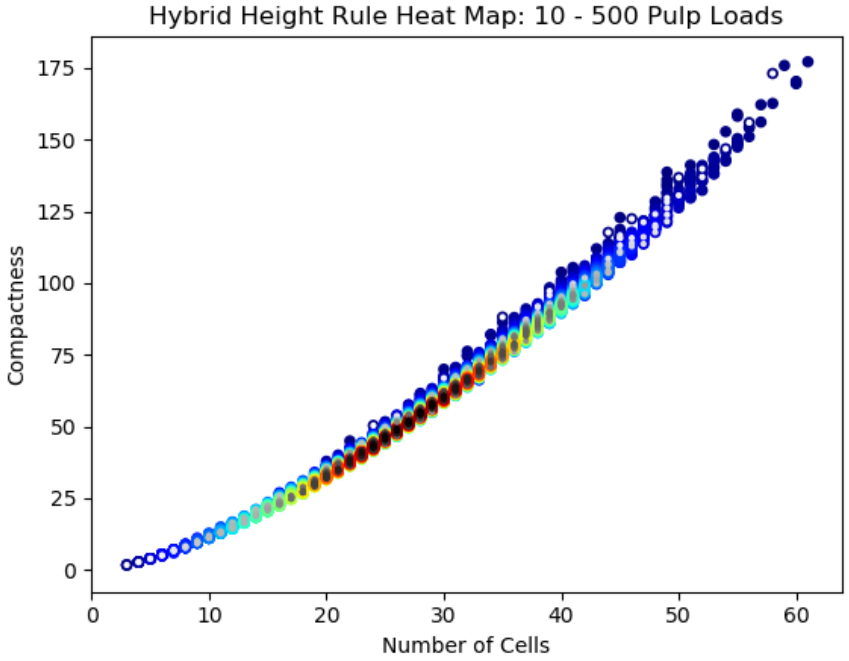

Figure 55: Hybrid height heat map of the compactness parameter for the narrow sample.

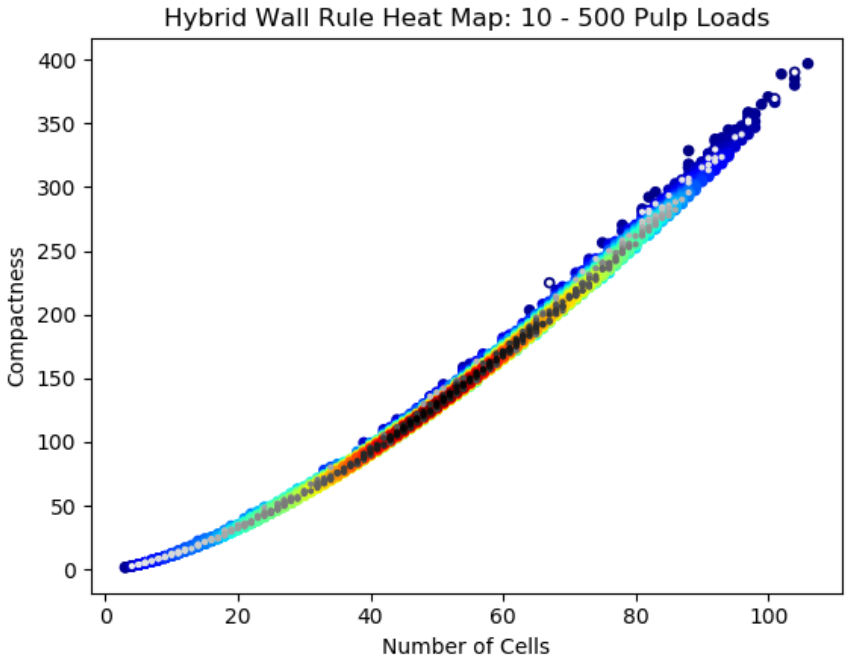

Figure 56: Hybrid wall heat map of the compactness parameter for the narrow sample.

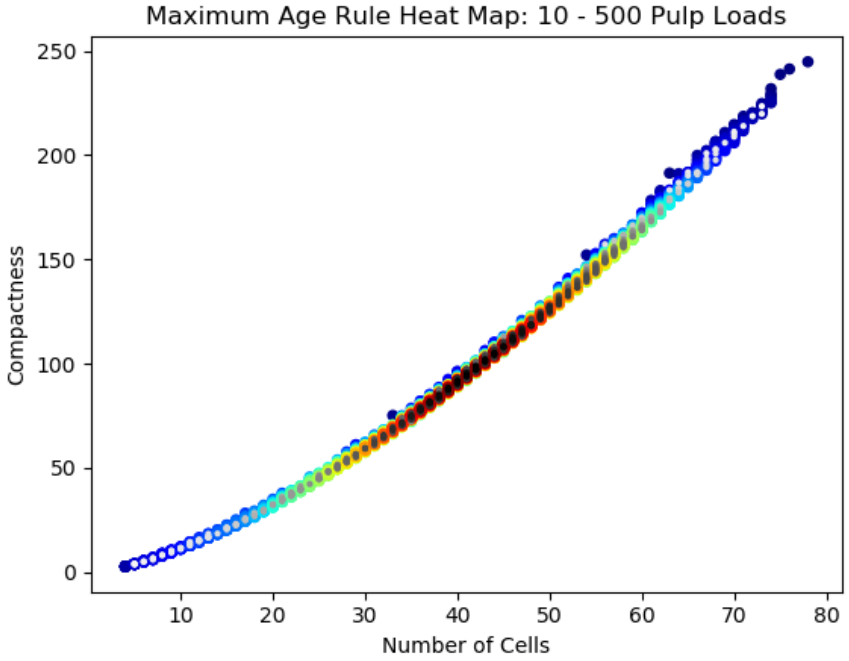

Figure 57: Maximum age heat map of the compactness parameter for the narrow sample.

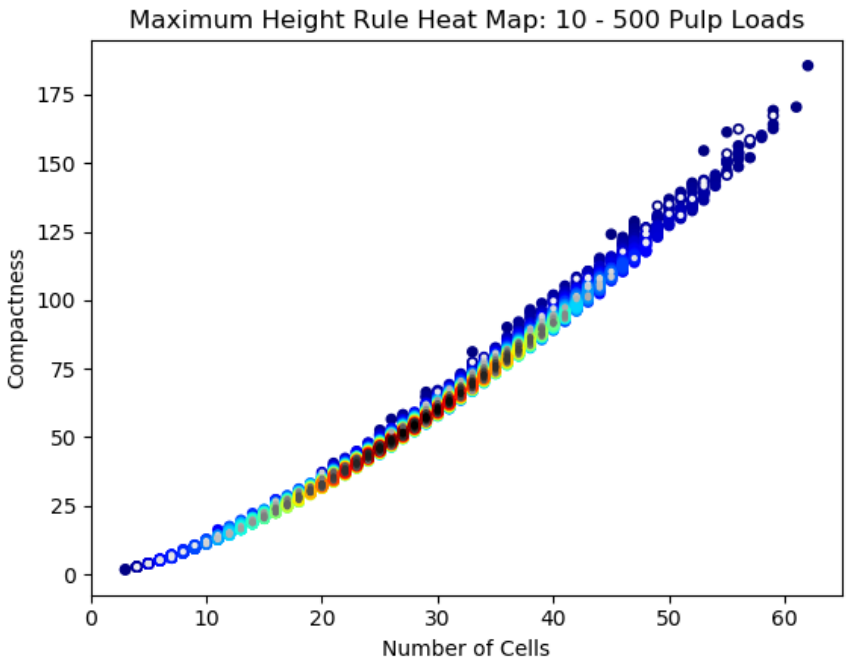

Figure 58: Maximum height heat map of the compactness parameter for the narrow sample.

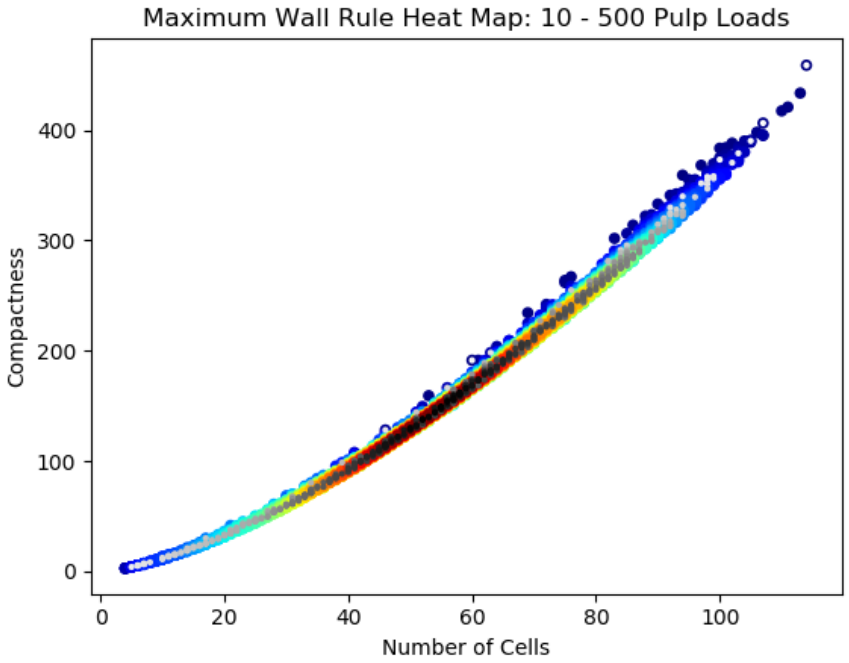

Figure 59: Maximum wall heat map of the compactness parameter for the narrow sample.

Appendix C

Scatterplots of Nest Parameters

This section contains the scatterplots of outer wall counts and compactness numbers not shown in Chapter 5's statistical analysis.

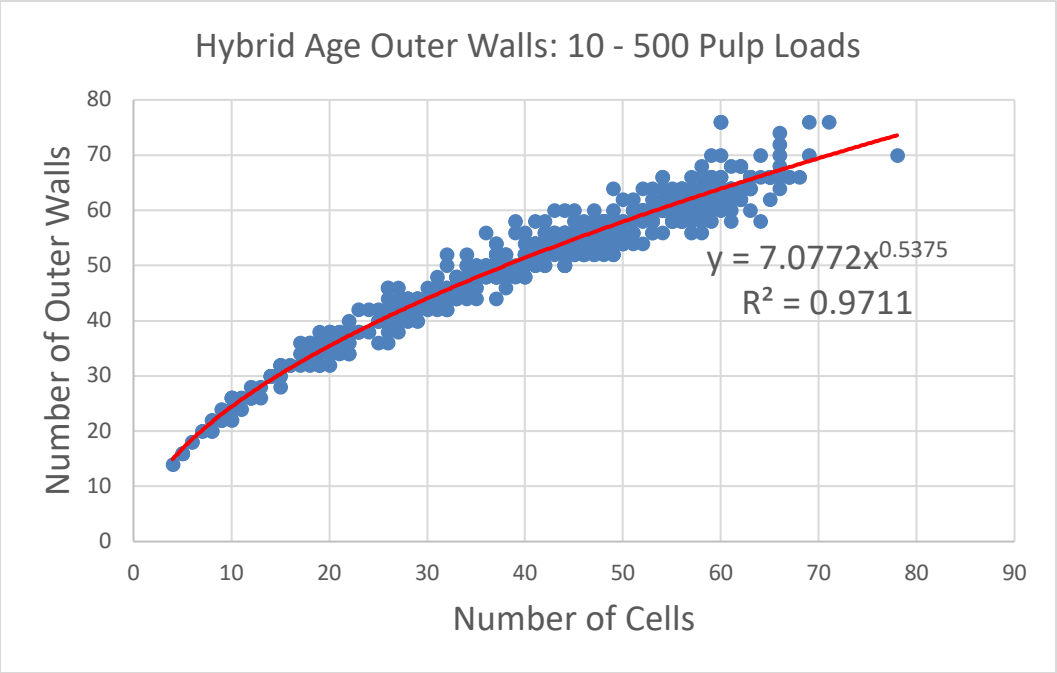

Figure 60: Hybrid age rule – outer wall data for 10 – 500 pulp loads.

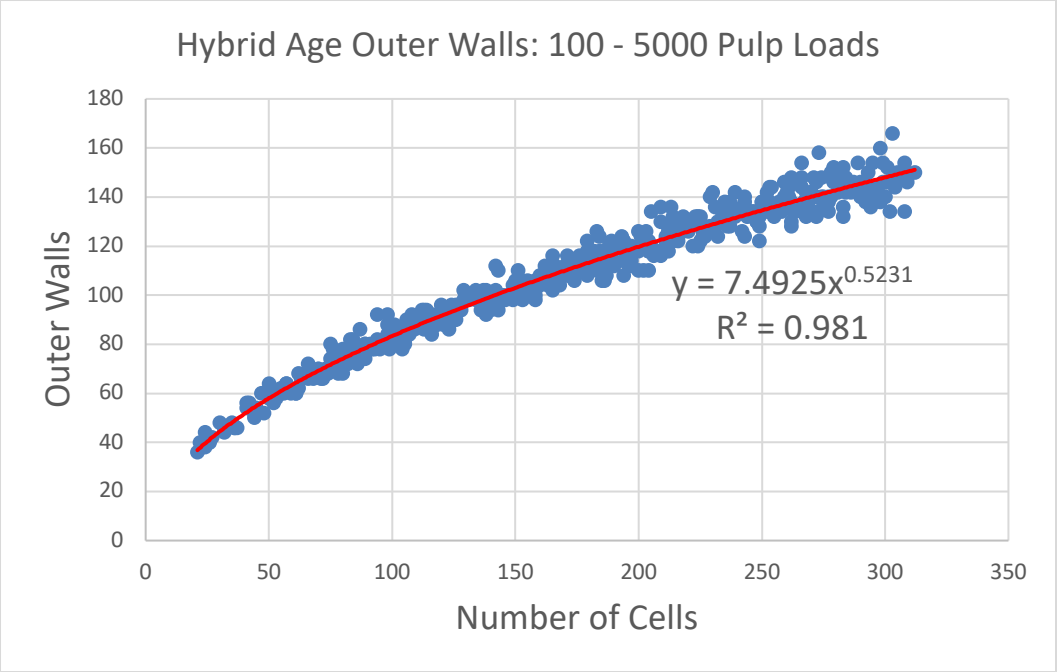

Figure 61: Hybrid age rule – outer wall data for 100 – 5000 pulp loads.

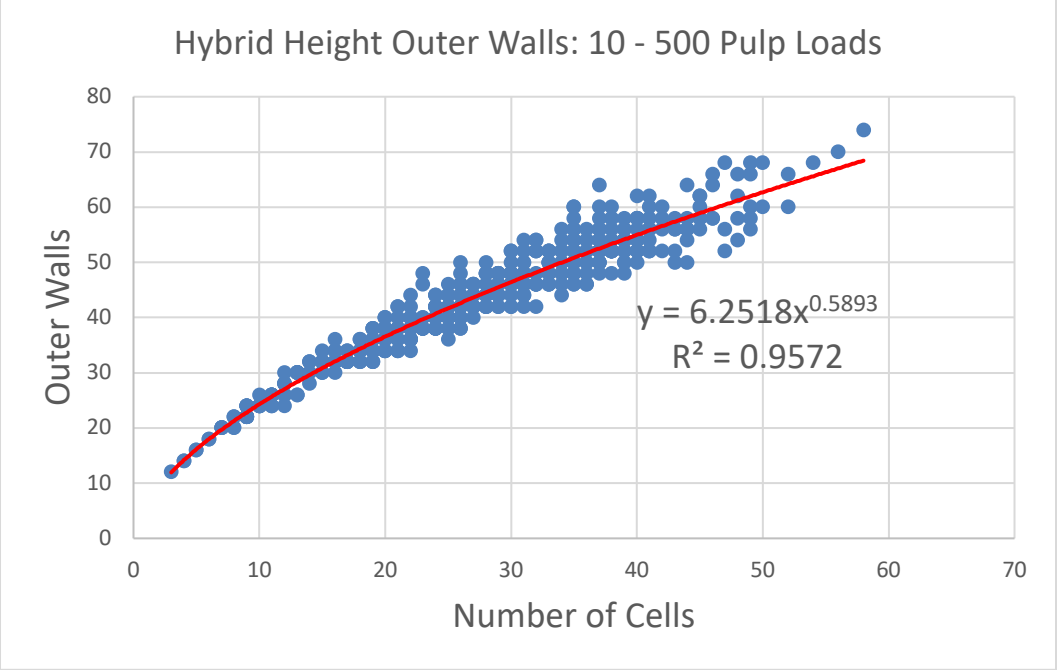

Figure 62: Hybrid height rule – outer wall data for 10 – 500 pulp loads.

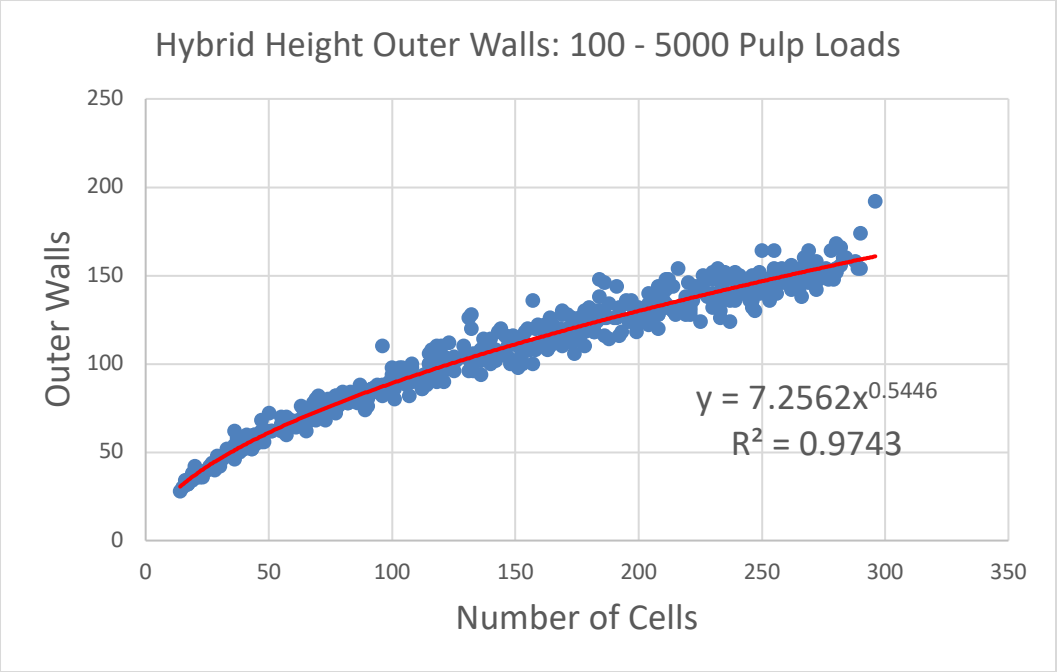

Figure 63: Hybrid height rule – outer wall data for 100 – 5000 pulp loads.

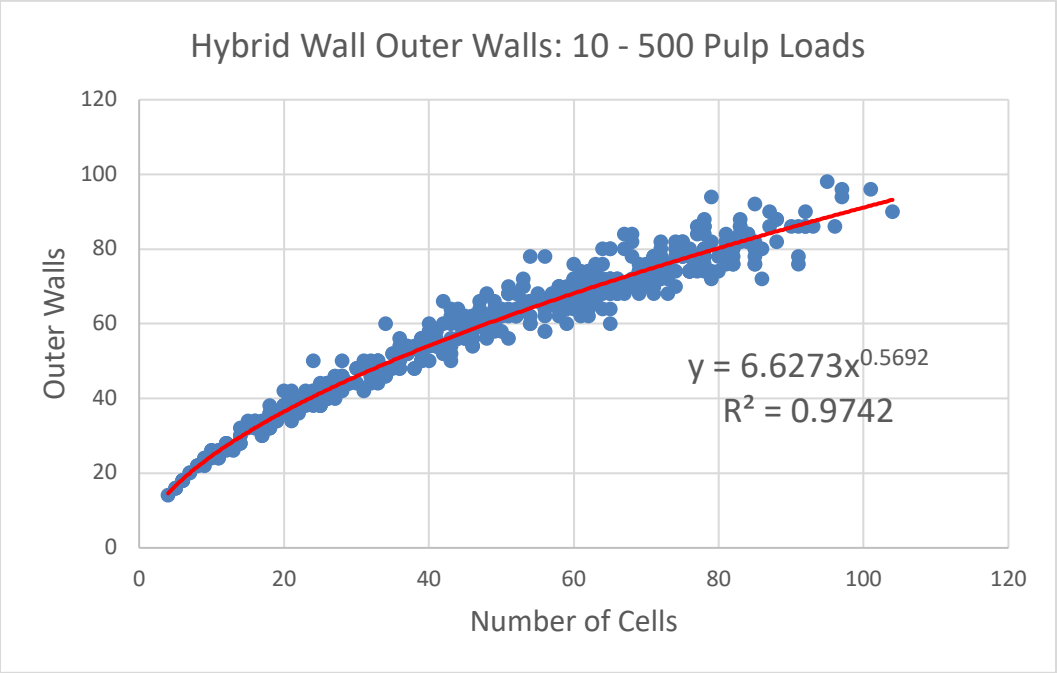

Figure 64: Hybrid wall rule – outer wall data for 10 – 500 pulp loads.

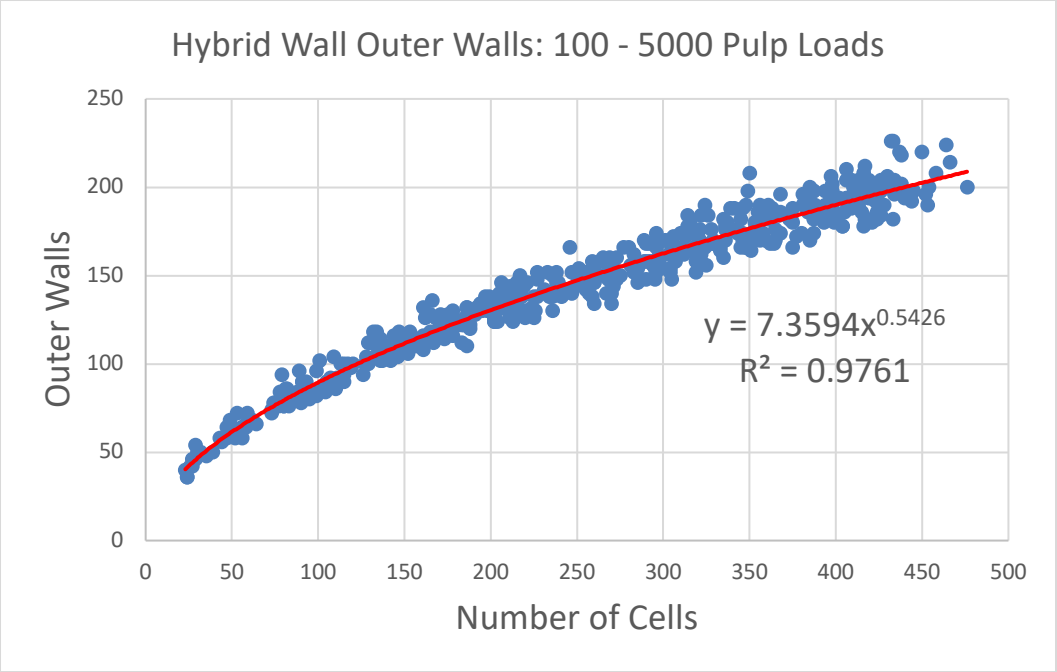

Figure 65: Hybrid wall rule – outer wall data for 100 – 5000 pulp loads.

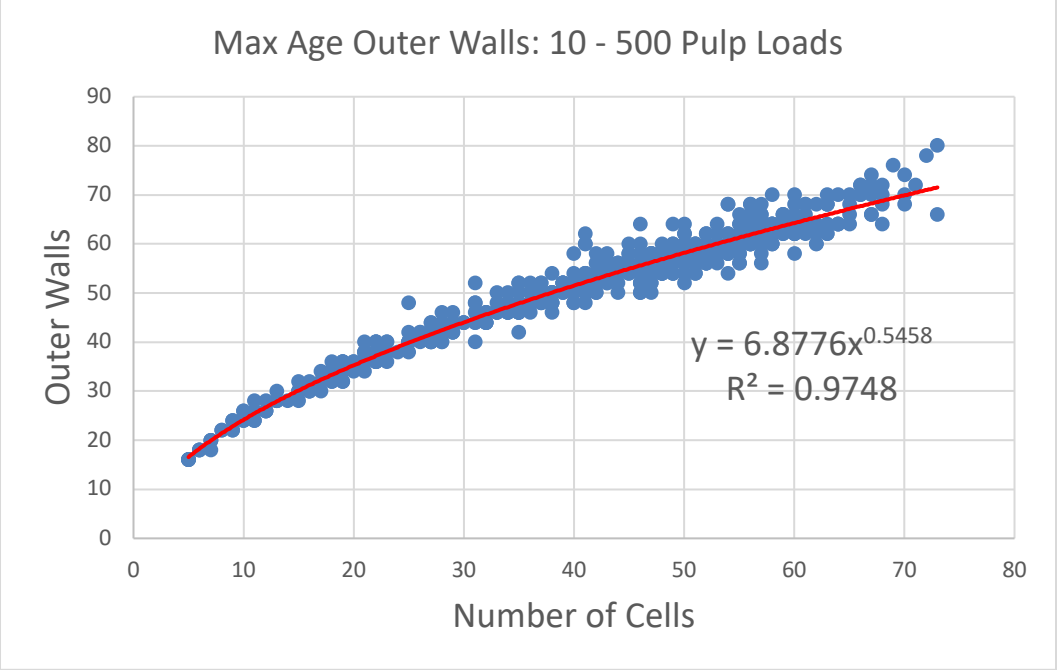

Figure 66: Max age rule – outer wall data for 10 – 500 pulp loads.

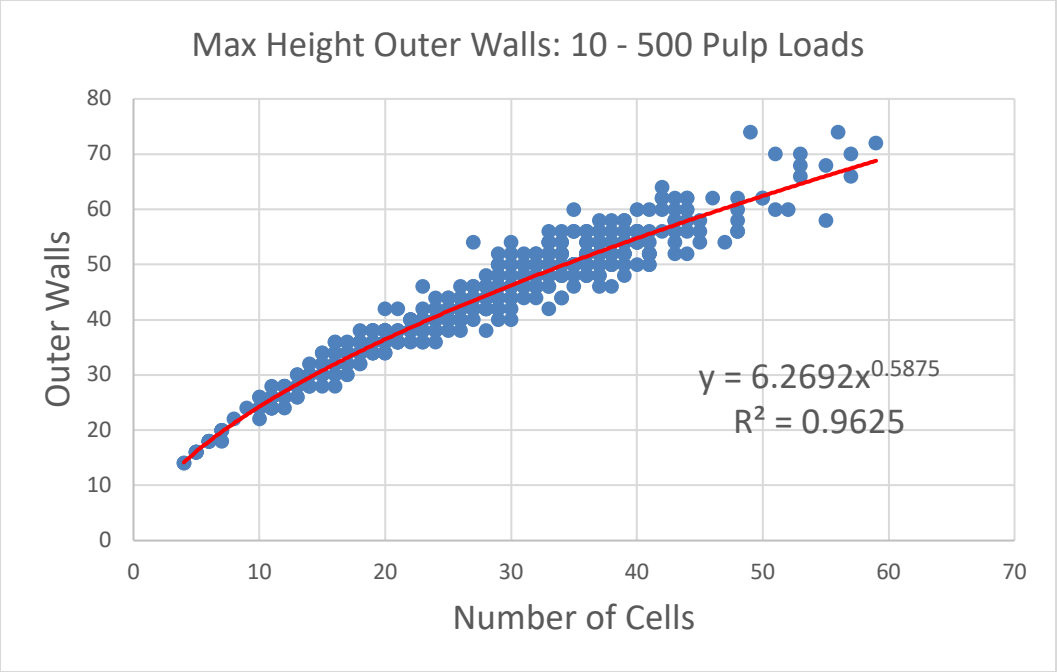

Figure 67: Max height rule – outer wall data for 10 – 500 pulp loads.

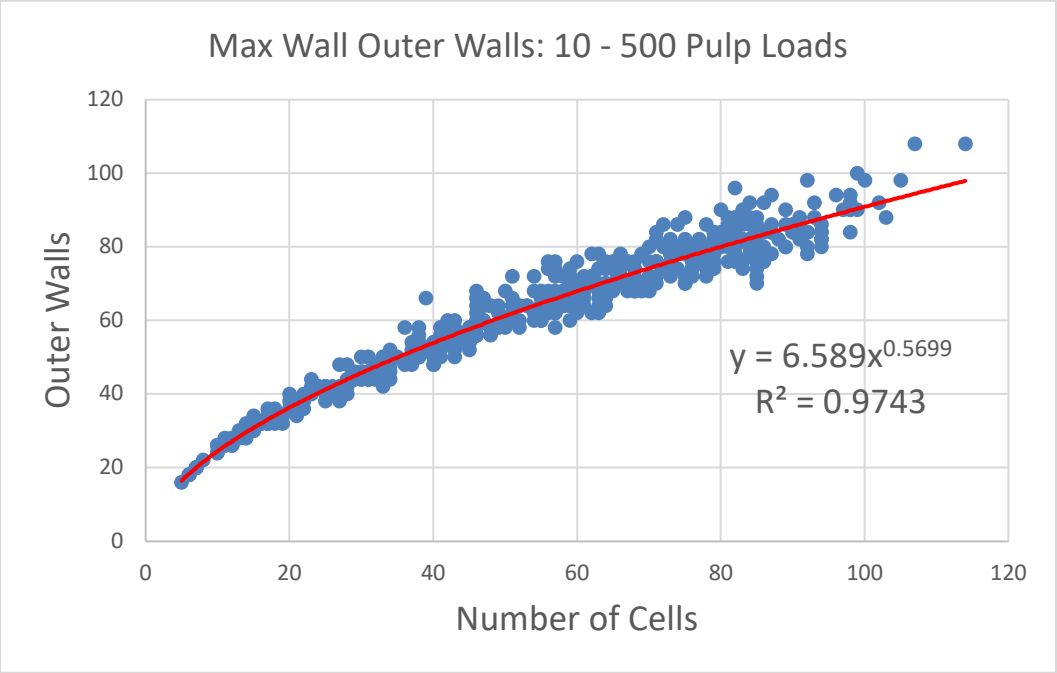

Figure 68: Max wall rule – outer wall data for 10 – 500 pulp loads.

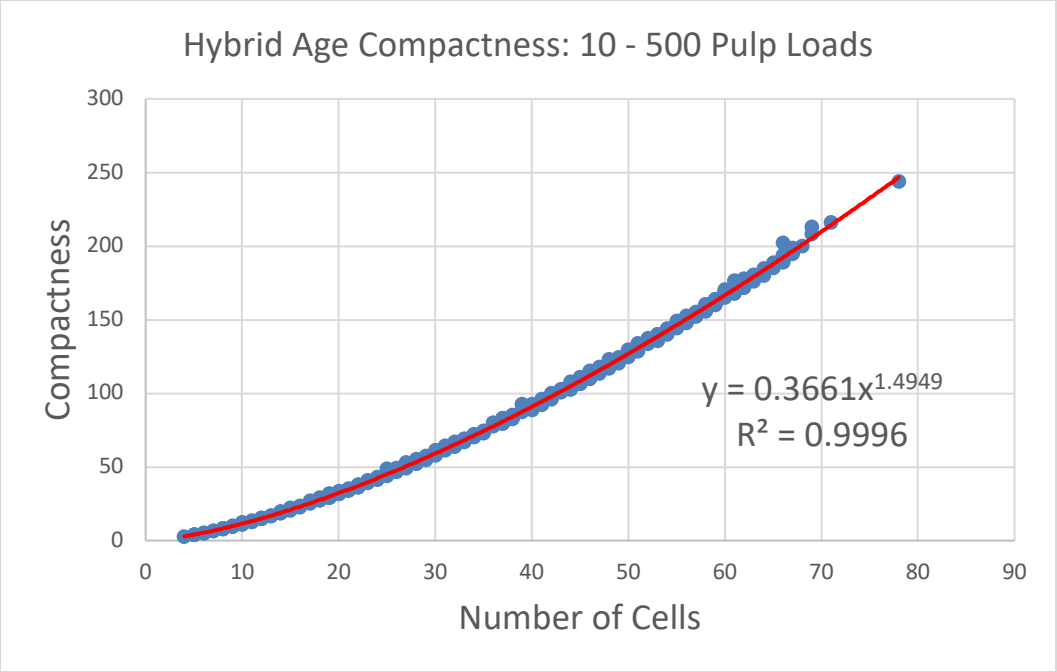

Figure 69: Hybrid age rule – compactness data for 10 – 500 pulp loads.

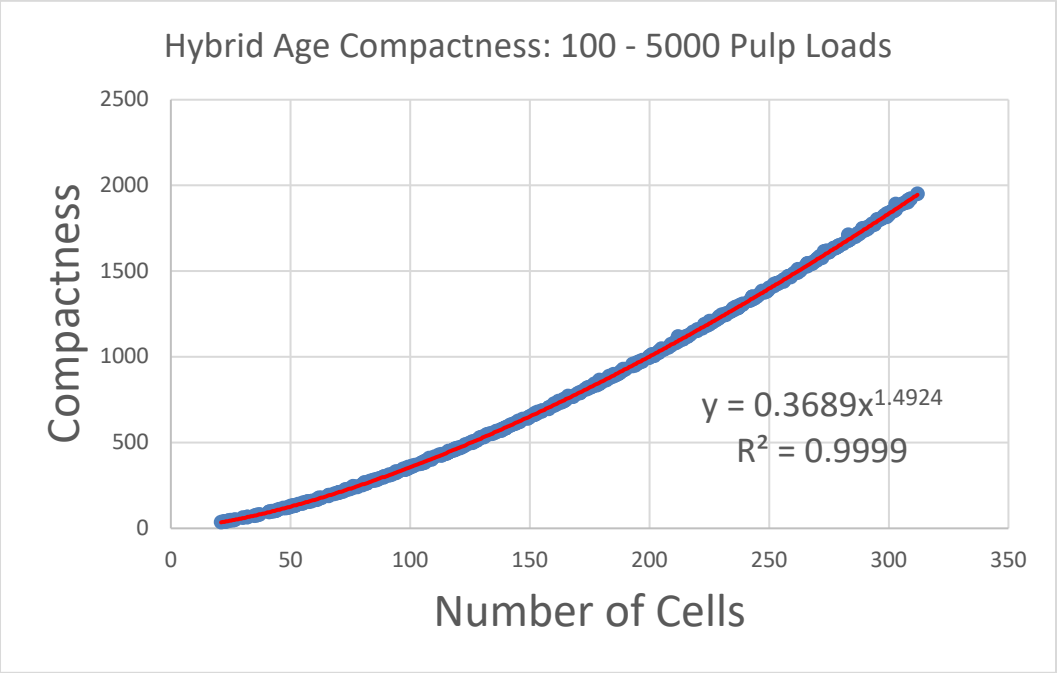

Figure 70: Hybrid age rule – compactness data for 100 – 5000 pulp loads.

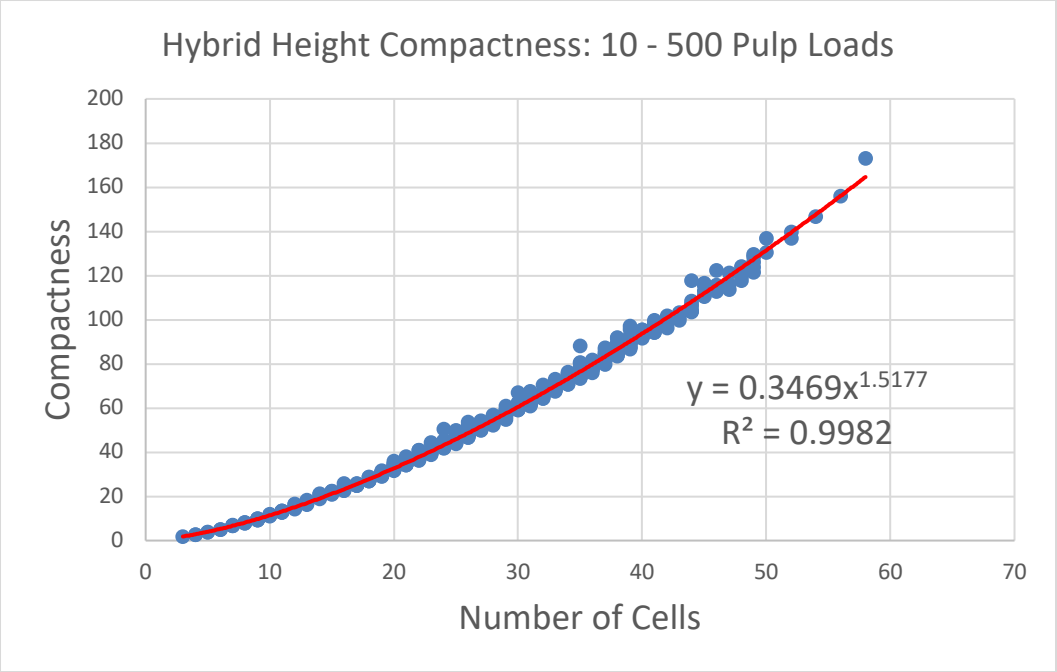

Figure 71: Hybrid height rule – compactness data for 10 – 500 pulp loads.

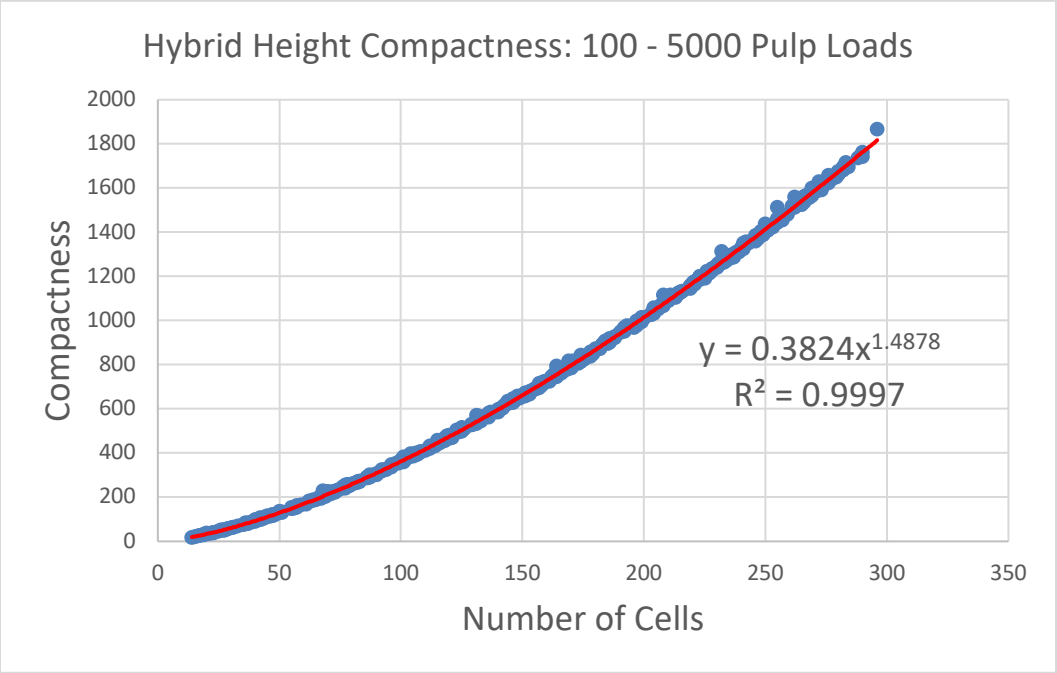

Figure 72: Hybrid height rule – compactness data for 100 – 5000 pulp loads.

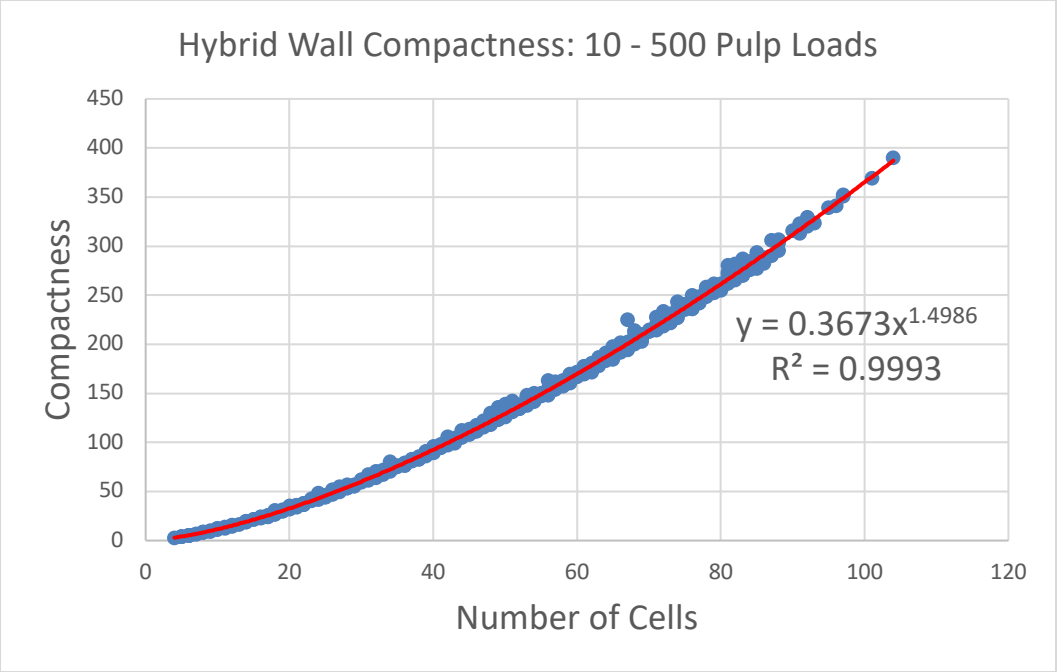

Figure 73: Hybrid wall rule – compactness data for 10 – 500 pulp loads.

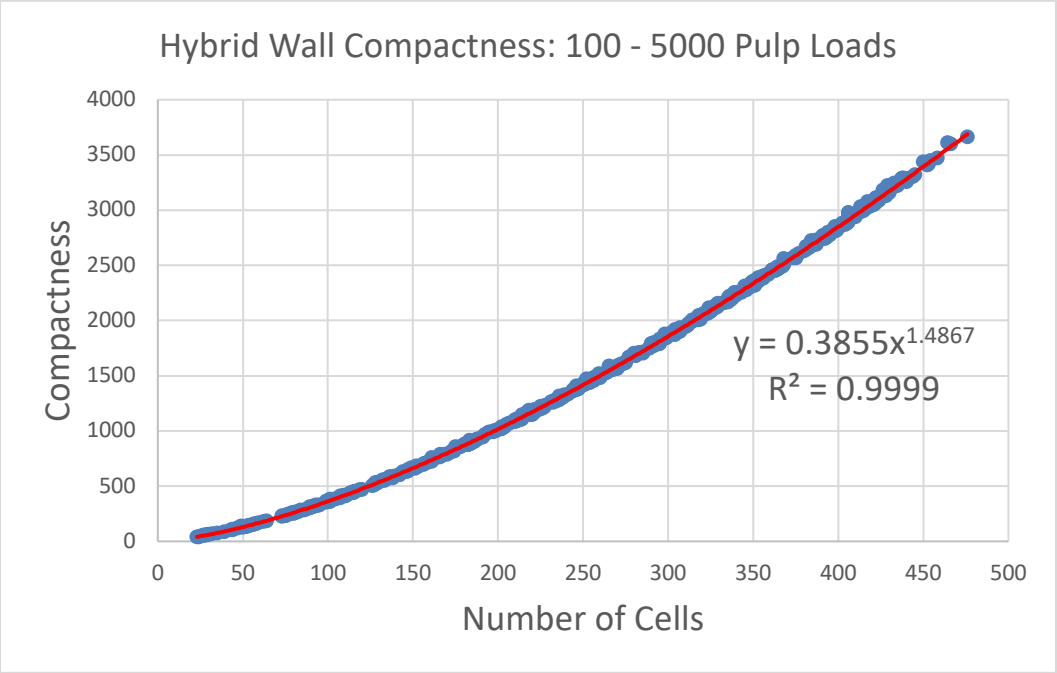

Figure 74: Hybrid wall rule – compactness data for 100 – 5000 pulp loads.

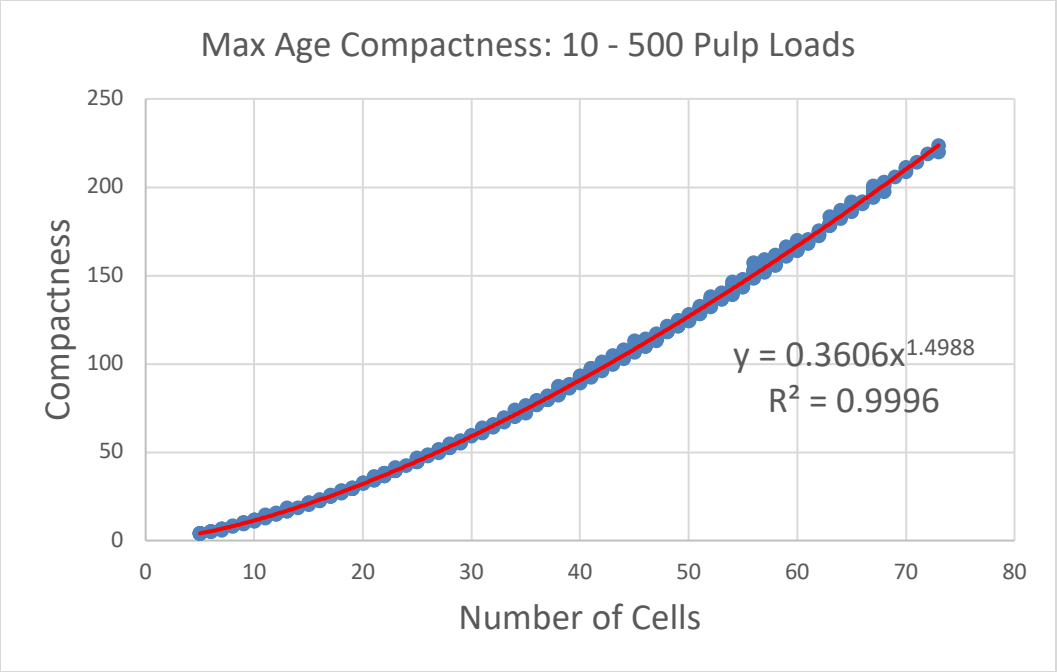

Figure 75: Max age rule – compactness data for 10 – 500 pulp loads.

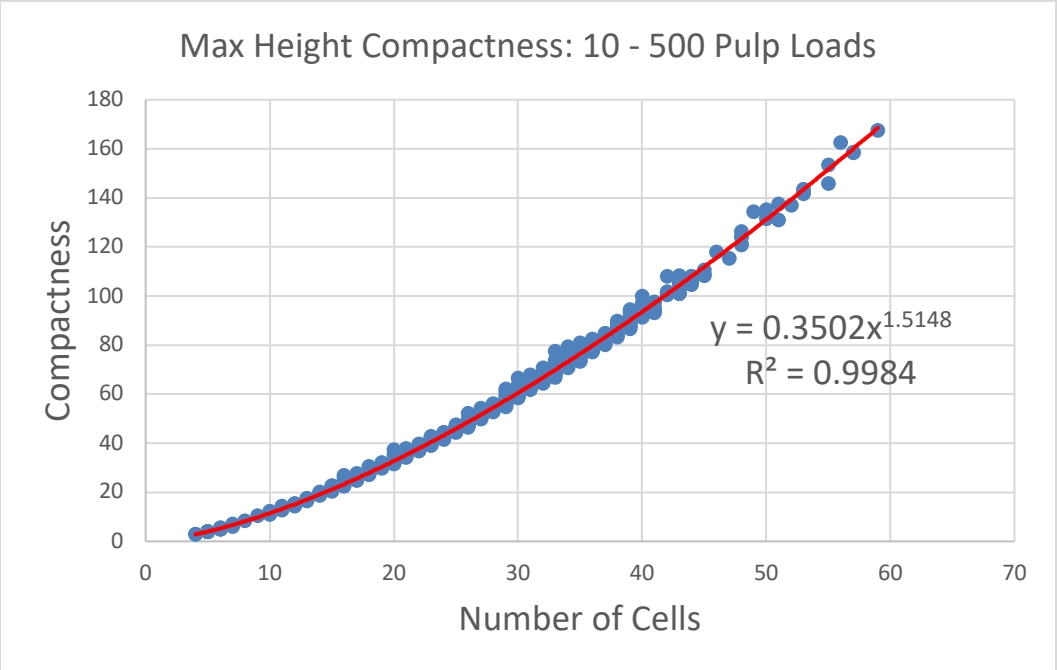

Figure 76: Max height rule – compactness data for 10 – 500 pulp loads.

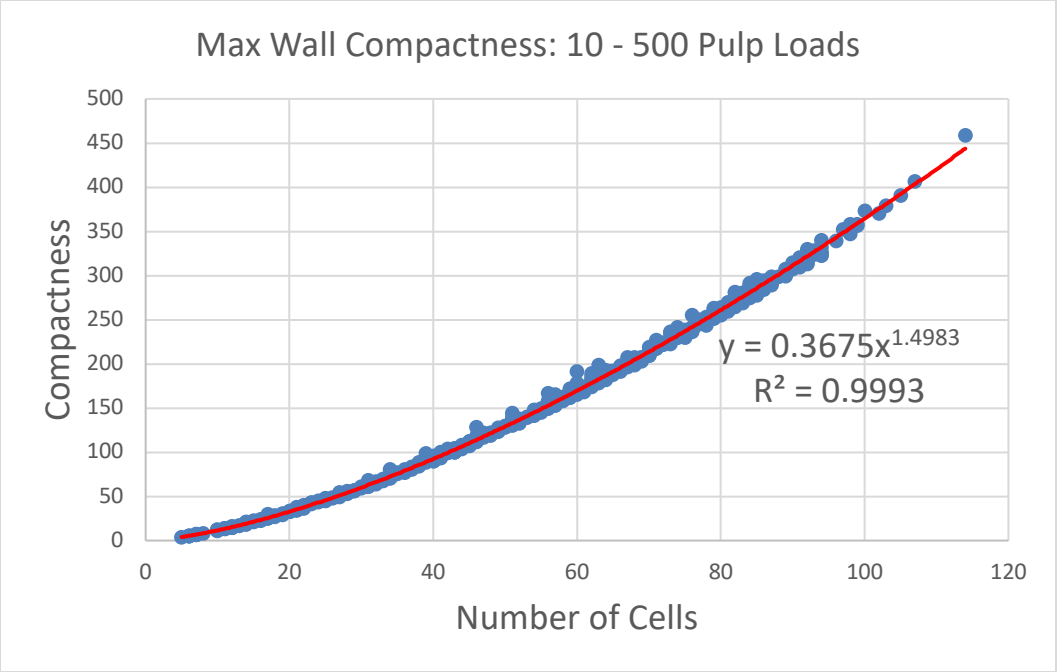

Figure 77: Max wall rule – compactness data for 10 – 500 pulp loads.

VITA

MATTHEW S. HARRISON

Education:	B.A. Biology, University of Virginia, Charlottesville, Virginia 2005 M.S. Computer and Information Science, East Tennessee State University, Johnson City, Tennessee 2018
Professional Experience:	Lecturer, East Tennessee State University, Department of Computing, Johnson City, Tennessee 2018 Graduate Assistant, East Tennessee State University, Information Technology Services, Johnson City, Tennessee 2016 – 2018 Telecommunications Technician, East Tennessee State University, Information Technology Services, Johnson City, Tennessee 2016 – 2018 Tutor I / II, East Tennessee State University, Center for Academic Achievement, Johnson City, Tennessee 2015 – 2016 Lab Specialist I, University of Virginia Health System, Center for Research in Reproduction, Charlottesville, Virginia 2005 – 2007
Publications:	Harrison, Matthew S., Istvan Karsai, and Christopher Wallace. "Self-Organized Structures: Modeling *Polistes dominula* Nest Construction with Simple Rules." Presentation, Appalachian Student Research Forum. Johnson City, Tennessee: East Tennessee State University, April 4, 2018. A-1. Harrison, Matthew S., Istvan Karsai, and Christopher Wallace. "Wasps 2.0: Modeling *Polistes dominula* Nest Construction." Poster, Appalachian Student Research Forum. Johnson City, Tennessee: East Tennessee State University, April 12, 2017. 146.
Honors and Awards:	Outstanding Computing Graduate Student, East Tennessee State University, Department of Computing, 2018