MODELLING ADOPTION OF MEAT-REDUCED DIETS USING CELLULAR AUTOMATA, IN THE CONTEXT OF THE CLIMATE AND ECOLOGICAL EMERGENCY

by

Student number 100368172

Thesis presented in part-fulfilment of the degree of Master of Science/Master of Research in accordance with the regulations of the University of East Anglia

Word count:

9293

School of Environmental Sciences
University of East Anglia
University Plain
Norwich
NR4 7TJ

Abstract

Max 300 words

Table of Contents

1 List of Accompanying Material....1
2 Introduction....2
3 Literature Review....7
3.1 Meat and the Climate and Ecological Emergency....7
3.2 Reduced Impact Diets....7
3.3 System Change vs Individual Responsibility....9
3.4 Theories of Personal Change....9
3.5 Cellular Automata and Social Networks....11
4 Methods....13
4.1 Generation of Social Graph....13
4.2 Use of Data from Hielkema & Lund Study....17
4.3 Implementation of CA Model....19
4.4 Processing Model Runs Output....22
4.5 Sensitivity Analyses....23
4.5.1 Randomizing Coefficients in Upstream Model....23
4.5.2 Allowing Strong Ties Between Households....24
4.5.3 Feedback Model....24
5 Results....25
5.1 Observations....25
5.2 Linear Models....28
5.2.1 Percentage of Reducers Model....28
5.2.2 Percentage of Intenders Model....33
5.3 Estimates of awareness_pc and facility_pc....36
5.4 Sensitivity Analyses....36
5.4.1 Randomizing Coefficients in Upstream Model....36
5.4.2 Allowing Strong Ties Between Households....38
5.4.3 Feedback Model....39
6 Discussion....40
7 Conclusions....41
8 References....41
9 Appendices....42

1 List of Accompanying Material

git repo for scripts: https://github.com/alex-zeffertt/dissertation

2 Introduction

The early decades of the 21st century saw a profound change in many people's mental model of their relationship with the natural world. In western capitalist cultures in particular, the prevalent view throughout the 20th century was of natural abundance, that the planet would provide an inexhaustible supply of raw materials for our consumption and an inexhaustible sink for our waste. For many, it was unimaginable that human activity could ever significantly alter the natural world. This view has begun to change as the impacts of the climate and ecological emergency become impossible to ignore, as scientific understanding has grown, and as awareness is spread by cultural organisations and social movements.

The food people eat and the manner in which it is produced constitute one of the most significant impacts on the natural world. Food systems are responsible for 30% of anthropogenic greenhouse gas emissions, 70% of freshwater use, and 40% of land use (Willet et al., 2019). Livestock contributes a large part of this footprint with the majority of farmed land being used to grow animal feed (FAO, 2012). At sea, all but 10% of fisheries are overfished (Willet et al., 2019). A meta-analysis by Clune et al. (2016) showed that the global warming potential (GWP) of meat-heavy diets can be an order of magnitude greater than that of plant based diets. Figure 1, reproduced from their study, illustrates this point.

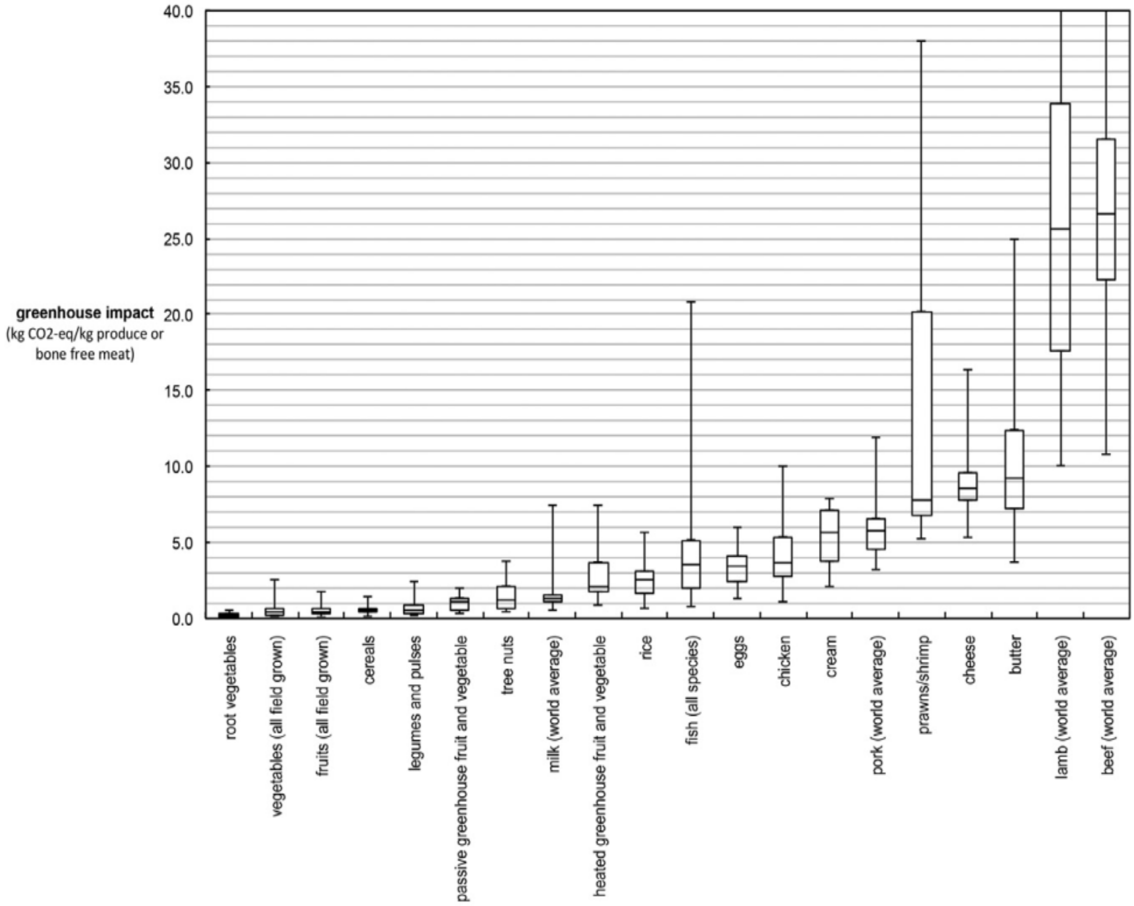


Figure 1: From Clune et al. (2016) - GWP ranges for different categories of food

There is a growing understanding that current patterns of consumption are unsustainable and that the existing economic system fails to consider damage done to shared assets, such as a stable climate and biodiversity. This is reflected in the fact that the term “externality” has migrated into common parlance from the field of economics. Economists have long argued that externalities should be priced in, for example via carbon pricing initiatives such as carbon taxes or emissions trading schemes (Baranzini et al., 2017). However, such system-led solutions have failed to become as widespread as hoped, or to set the price right (Boyce, 2018). In practice most solutions to the climate and ecological emergency have been framed in terms of individual choice. With better information, environmentally conscious individuals will, it is implied, choose lower impact goods and services, and peer pressure will accelerate that transition. There is some evidence for this in the context of food. For example Ploll et al. (2020) found that, in response to the climate and ecological emergency, many individuals are considering changes to their diets, becoming vegetarian, vegan, or simply reducing their meat intake. Meanwhile, Hielkema & Lund (2021) found that the behaviour of close family and friends had a strong influence on decisions to reduce meat consumption.

It is, however, unclear whether conscience-driven and demand-led changes alone are sufficient to create the "avalanche" effect required to produce the approximately 50% emissions reductions required this decade to keep warming below 1.5°C (IPCC, 2021), or to prevent further rainforest turning over to agriculture. Is there also a need for system change, such full pricing in of externalities, minimum standards, or even rationing? Agent based simulations provide one tool for answering this question.

Cellular Automata, or CA, are a type of agent-based simulation. These are computational models composed of a large number of cells (or "automata") each with a state and defined "neighbourhood". Every time step (also referred to as "iterations", or "generations"), each automaton state is updated by a rule based on the state of neighbouring automatons. This makes CA ideal for simulating systems such as the diffusion of human behaviours, in which the next state of each agent is influenced by the current state of its peers.

Conway's Game of Life (GoL) is an early example of CA, designed to explore the simplest physics able to support life-like behaviour (Gardner, 1970). Figure 2 illustrates some simple GoL sequences. However, GoL can also exhibit highly complex behaviour. In fact, Rendell (2011) showed that GoL is Turing Complete, meaning that it is capable of simulating computation and is therefore able to show behaviour as complex as any program conceivable.

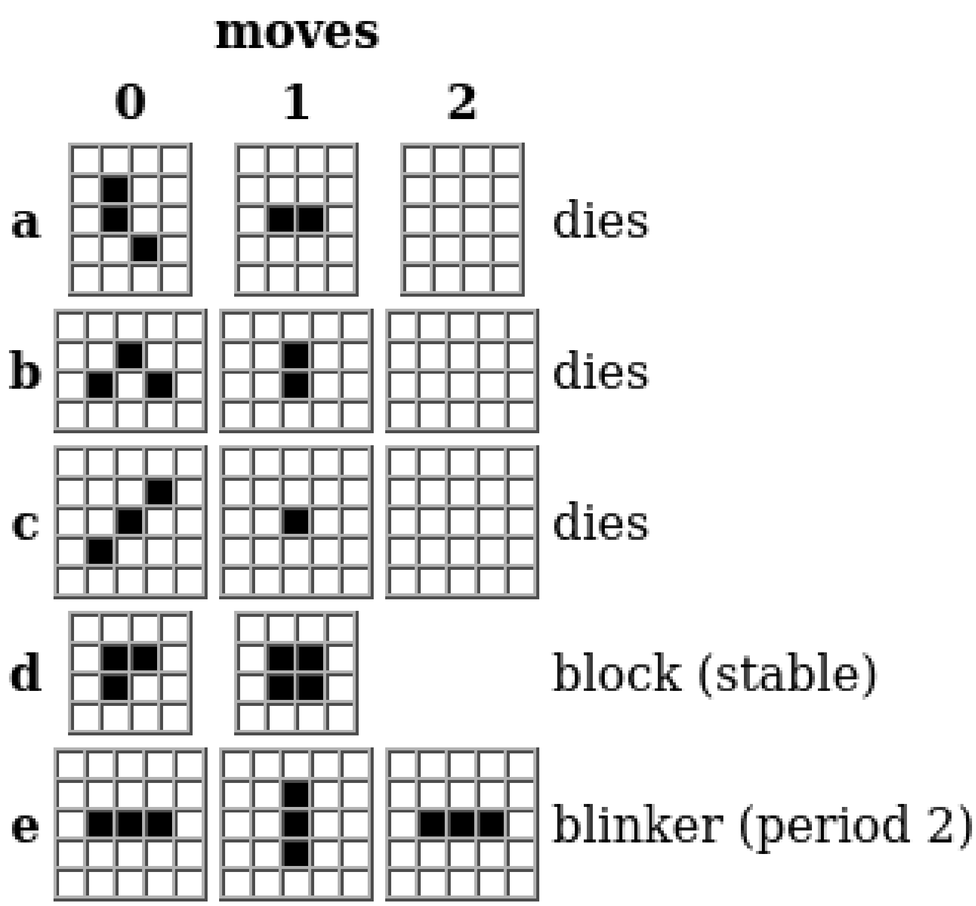


Figure 2: From Gardner (1970) - 3 generations of GoL from 5 alternative starting points. Neighbourhoods are 9-cell squares, states are absent (white) or present (black), and the rules are: 3 neighbours lead to cell-birth; 2 or 3 to cell-survival; and any other number to cell-death. More complex initial states generate apparently life-like behaviours.

In this study a model will be built to simulate the spread of meat-reduced diets through a population, in response to a society wide level of awareness of the climate and ecological emergency, and mediated by the influence of social ties.

It is clearly possible to build any number of contrived CA models for the spread of meat reduced diets. For example in one model produced by the author, each cell on a 200x200 grid is either

a) not considering reduction in meat consumption

b) considering reduction, or

c) reducing meat consumption.

The cells in this contrived model are initially 90% in state (a), 9% in state (b), and 1% in state (c), with the placing random. At each time step and for each cell, peer-pressure p is defined as the number of meat-reducer peers in the 4 adjacent cells plus half the number of meat-reducers in the 4 diagonally adjacent cells. If a cell is in state (a) the state is changed to (b) with a probability of $1 - e^{-0.3p}$; if it is in state (b) it is changed to (c) with probability $1 - e^{-0.1p}$; and if in state (c) it is changed to (a) with probability $1 - e^{-0.1(6-p)}$. This is intended to represent individuals changing their behaviour in response to the behaviour of their nearby peers. In this model every individual ultimately ends up a meat-reducer, and Figure 3 shows the iteration at which each cell reaches this state.

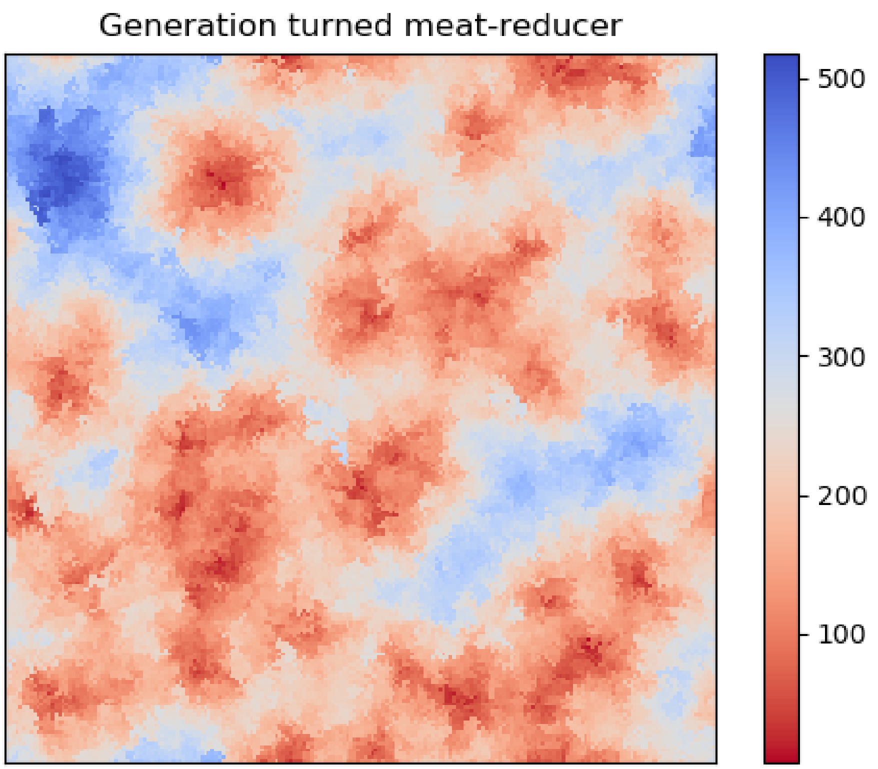


Figure 3: A contrived CA model for spread of meat-reduced diets.

Although such contrived models may produce interesting pictures, models built using real-world data produce more justifiable results. In order to produce a model from which conclusions can be drawn, a literature search is required. This may provide data with which to parametrize the model.

The first area for literature search relates to the domain in which the action takes place. Like many CA, the agents in the model described above are arranged in a 2D array. However, real social interactions occur in a social network, also known as a graph, in which individuals are represented by nodes connected to peers by edges. The likelihood of the presence of an edge between two nodes is obviously influenced by the physical proximity of the nodes, but the correlation is not as simple or deterministic as in naive CA models.

A second area for literature search pertains to the definition of the states and probabilities of transitioning between them. In the contrived model above, an individual transitions from “not considering” to “considering” or “considering” to “meat-reducer”, but only if neighbouring cells contain meat reducers. However, in real life there are non-zero probabilities for transition even amongst individuals with no ties to reducers. Data is required to furnish a CA model with justifiable transition probabilities.

In this study, peer-reviewed research and grey-literature were reviewed, and used to build a realistic CA model for the spread of meat-reduced diets in response to growing awareness of the climate and ecological emergency. The resulting model was run with various inputs and the results were analysed using statistical tools to produce a mathematical model. Finally the implications for the feasibility of demand-led solutions were determined.

3 Literature Review

3.1 Meat and the Climate and Ecological Emergency

In 2015, signatories at COP21 agreed to “limit global warming to well below 2, preferably to 1.5 degrees Celsius, compared to pre-industrial levels”. Six years later the IPCC identified that the remaining “1.5°C carbon budget” was likely to be in the range 300-900 $GTCO_2$ (IPCC, 2021). For context, the central estimate of 450 $GTCO_2$ will be exceeded in just 13 years, if emission rates remain unchanged. Agriculture is a major contributor, responsible for 30% of current emissions, with meat production directly or indirectly responsible for the majority of those, in as well as other greenhouse gases (FAO, 2012). In addition to climate, meat production negatively impacts ecosystems by requiring vastly more land and fresh water per food-calorie produced. In summary, our desire for meat significantly contributes towards the crossing of three of nine planetary boundaries (Rockström et al., 2009). These findings justify an exploration of factors that might facilitate the reduction of consumption levels.

3.2 Reduced Impact Diets

There is a broad consensus in the literature that a vegetarian diet usually has a lower carbon and ecological footprint than a meat-based diet. However, some authors have taken a more granular view. Clune et al. (2016) performed a systematic review of 369 studies to produce a comprehensive table of global warming potential (in terms of CO_2 or equivalent emissions per kg produce) for every food category. Whilst meat was generally found to have higher emissions than other foods, there was a large degree of overlap. For example the mean for rice (2.66 kg CO_2-eq/kg) exceeded that of mackerel (2.00 kg CO_2-eq/kg). For many foods there was a large variation depending on the method of production (if farmed) and type of transport (to distribution centre). For example, tomatoes grown in passively heated greenhouses had a mean of 1.02 CO_2-eq/kg, whereas those grown in natural gas heated greenhouses had a mean 2.59 CO_2-eq/kg, above that of many types of fish. Despite the efforts of Clune et al. to compile all of the information in one place, the complexity of determining which foods have the lowest impact make a piecemeal approach infeasible for ordinary consumers. It is therefore understandable that consumers who are concerned about their personal carbon footprints choose to adopt a shortcut. Meat-free or plant-based diets are simple to understand and to follow. However, the Clune et al. study suggests that a ruminant-free diet would reduce emissions by almost as much, and this would obviously be fairly straightforward to understand and follow too.

Whilst carbon footprints are important, there are many other environmental issues relating to food production not covered by the Clune et al. meta-analysis. Some consumers are also concerned about over-fishing, freshwater use, or biodiversity loss caused by the land use changes needed for animal rearing and feed. Cultured meat offers the potential to address these issues in addition to reducing greenhouse gas emissions. Although the technology to grow real meat in cell cultures is not yet scaled up, proof-of-concept trials do already exist (Stephens et al., 2018). Tuomisto & Joost Teixeira de Mattos (2011) estimated the carbon footprint, water use, and land use required to grow cultured meat at scale. They found that the carbon emissions of so-called “lab-grown” beef would be 1.9-2.4 CO_2-eq/kg, around a tenth of that of traditionally farmed beef. Water use would be around 367-521 litres/kg, or around 1/30th the usual figure, and land use would be negligible. The authors note that their estimates were based on existing mixes of energy and so emissions could be brought to zero, in a renewable-powered world. In fact, since switching to cultured meat frees up so much land, carbon emissions would be temporarily negative if that land were to be reforested.

If cultured, or “lab-grown”, meat can be produced at scale, many of the discussions in the literature surrounding the need to reduce meat intake for environmental reasons will become redundant. However, this all depends on whether cultured-meat can achieve widespread social acceptability, and shake off its image as “unnatural” (Siegrist & Sütterlin, 2017). This is difficult to predict, however, as Tuomisto & Joost Teixeira de Mattos point out, there is much about modern animal husbandry which is also “unnatural”.

3.3 System Change vs Individual Responsibility

Within the context of the climate and ecological emergency, proposed solutions fall into two categories: system level (or supply side) and individual responsibility (or demand side). An example of the former are international agreements which, if implemented, would likely see the price of emissions-intensive products such as meat rise. The latter is typified by the "Carbon Footprint Calculator", introduced by BP (2004), which aims to nudge consumers into lower emissions lifestyles. There is an active debate as to whether demand-led changes can feasibly impact widespread human behaviour. In fact, some commentators have claimed that BP's "Carbon Footprint" concept was introduced cynically in the belief that it would not reduce emissions. This theme is discussed by Baylor Johnson (2003) who invokes the "tragedy of the commons", arguing that calls for individual action detract from the system changes necessary. In response Marion Hourdequin (2011) counters that the desire to follow social norms may see an individual-led revolution sweep through society. Both authors make the case that their conclusions can be justified by game theory, the mathematical study of interactions between rational actors attempting to optimize some measure of personal utility. However, neither author attempts to build a model to test their claims.

3.4 Theories of Personal Change

A number of theories of individual change in the context meat-consumption are discussed in the literature. Ploll et al. (2020) look at the diffusion of vegetarian and vegan diets in Austria through a social innovation perspective, i.e. as a bottom-up social movement framed as a solution to a problem. Hodson & Earle (2018) consider political ideology as a predictor of lapses from vegetarian and vegan diets. The survey of attitudes towards meat-reduction in Denmark (Hielkema & Lund, 2019) adopts the transtheoretical model in which individuals pass through stages of change when modifying behaviour. All these studies are survey based, and can be used to produce a quantitative model for predicting meat-consumption based on demographics, political beliefs, and social ties. However, they all relate to just a snapshot in time. To quantify the potential of demand-led changes to reduce the environmental impact of diets it is necessary to create a dynamic model of meat-consumption, and let it run. This is because, as Hielkema & Lund demonstrate, social ties are an important factor, and therefore levels of meat-consumption will vary over time as new behaviours diffuse through society. This is true even if no system-level measures to combat climate change and ecological degradation are taken.

Although they do not build a dynamic model, Hielkema & Lund do provide information which could be incorporated into one. Firstly, the extent to which the likelihood of being a meat-reducer is influenced by peer behaviour is quantified. Second, they show that how the “carnivore” category can be broken down into sub-categories. Only 3.6% of the population surveyed were vegetarian or vegan, but there were many others reducing meat intake, for a number of reasons including the climate and ecological emergency. Therefore the obvious “carnivore” / ”vegetarian” / “vegan” split hides the area in which most of the meat reduction occurs, in addition to producing more noisy data. The sub-categories the authors used derive from the “Transtheoretical” or “Stages of Change” model developed by Prochaska & Di Clemente (1982). The “Stages of Change” model describes how individuals make behavioural changes, such as quitting smoking, by progressing and relapsing through stages nominally lasting 6 months. These stages are “Pre-contemplation”, “Contemplation”, “Preparation”, “Action”, and “Maintenance”. Hielkema and Lund determined the “stage of change” by asking respondents to select which of the following 5 options best described them.

1. *“no plan to reduce in the next six months”*
2. *“planning to reduce within the next six months”*
3. *“planning to reduce within the next month”*
4. *“already reduced within the last six months”*
5. *“reduced for longer than six months”*

(Hielkema & Lund, 2021)

The authors then identified 1 with a category they called “NO intention”, combined 2&3 into a category called “Intention”, and combined 4&5 into a category called “Reducer”.

The stages of change model has received a number of criticisms. For example, Bandura (1997) argues that there is insufficient support for the claim that individuals spend approximately 6 months in each stage, and Littell and Girvin (2002) question whether transitions really are constrained to single-step forward progressions, with relapses. This must be borne in mind when designing a dynamic model. A dynamic model should allow for resistance to change to vary across a population, and must not assume that individuals can only move between adjacent stages.

The existing studies chart the extent to which demographics (e.g. gender, conservatism, etc.) and personal beliefs (towards the environment, animal welfare, and health), predispose a person towards a meat-reduced diet, lowering the convenience and peer-influence thresholds at which a change occurs. However, no study was found that took an interdisciplinary approach of computer modelling the diffusion of meat reduced diets, although the concept of simulation in the broader context of sociology has been discussed (Halpin, 1999).

The survey by Hielkema & Lund (2019) shows peer-influence and climate awareness to be the most important factors in individual's decisions to consider or adopt a meat-reduced diet. This makes agent-based simulation such as cellular automata an ideal lens through which to study the problem. Although, other factors are almost as important they can be considered fixed, and modelled by making the automata demographics match those of a real population. Only peer-state and a global factors such as the society-wide level of climate awareness, and availability of non-meat options, need to be modelled as variables.

3.5 Cellular Automata and Social Networks

Griffeath (2015) describes multiple types of cellular automata based on geometric arrays, such as classic (deterministic) CA, and Probabilistic CA. For example, Figure 4 shows a PCA in which states represent one of two voting intentions (Democrat or Republican) and the update rule reflects influence by local peers.

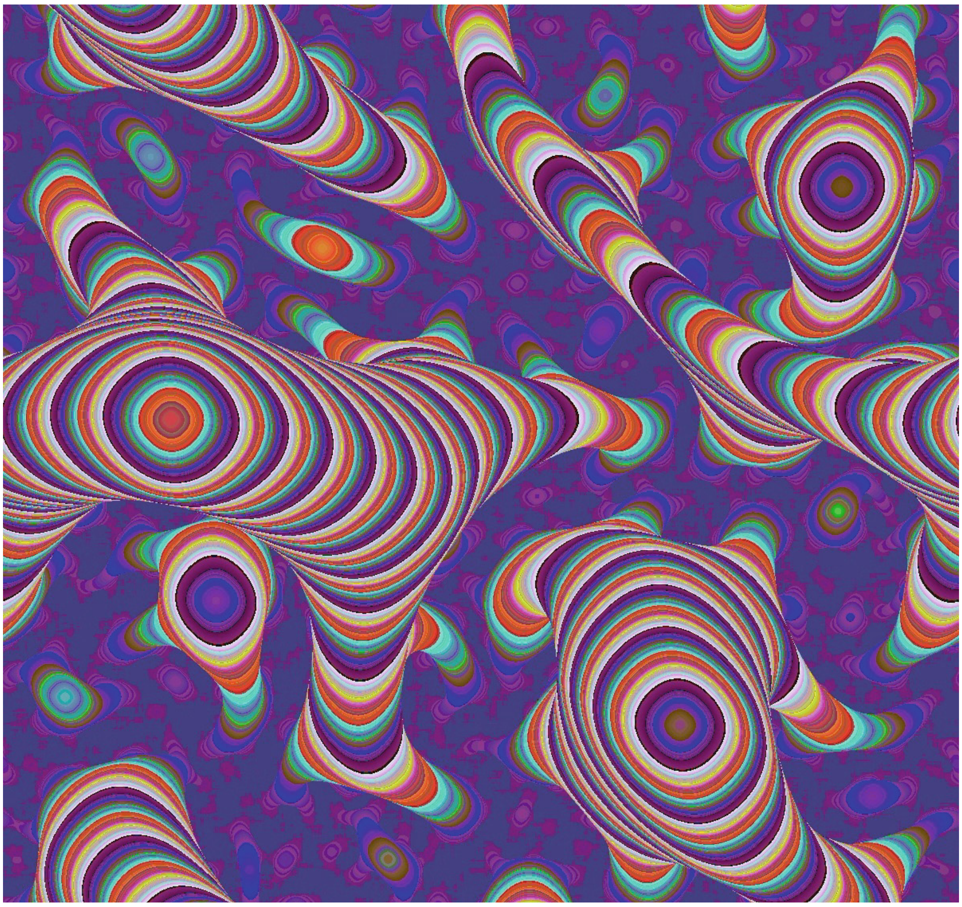


Figure 4: Adapted from Griffeath (2015), a PCA showing how a randomized population of Democrat and Republican voters converge towards a consensus given an update rule in which each cell switches allegiance to that of the majority of peers in its 21x21 cell neighbourhood. The colours indicate the generation in which the switch occurred.

The illustration shows that PCA are capable of simulating the type of non-linear phenomena, not amenable to normal mathematical techniques, that is required by this study. However the 2D arrays used by most classic CAs are not representative of real human social ties.

Other authors have replaced the classical grid array with nodal graphs in their studies (Małecki, 2017). Such relation based graph CAs, or “r-GCAs”, provide an ideal framework for considering the diffusion of peer-influenced human behaviour, as one’s peers are not simply those people who are geometrically closest. However, there is a debate over the best type of graph to use to represent social ties. Barabäsi et al. (2000) found WWW hyperlinks to be a Scale Free Network (SFN), in which the probability of having n links is proportional to $n^{-\gamma}$ for some constant $2 < \gamma < 3$. (Web page linking is a form of social network.) However, others have argued that SFNs are rare in real social networks (Broido & Clauset, 2019). An alternative that is widely used is the Small World Network (SWN) in which most links are geographically-local and there are a smaller number of long distance links (Huang et al., 2005). SWNs have average path lengths proportional to the log of the number of nodes in the network, a feature that has led to the idea that “six handshakes” link any two humans alive. In their study of urban disease outbreaks, Eubank et al. (2004) found the contact network to be similar to an SWN. The truth is probably that the model needed depends on the context, and a graph for describing social media links, or viral contagion, may look very different to one that only includes the sort of peers that might influence a dietary change. For that reason, in any given scenario it is important to determine if there is additional information that can be used to try and construct a social network relevant to the problem being considered, rather than simply adopting one or other pre-defined model.

4 Methods

The Hielkema & Lund study (2021) was used as a starting point for this investigation. The model they provide was found to be necessary but not sufficient for the creation of a CA model for the spread of meat-reduced diets, and therefore each run of the CA model required additional parameters to be supplied. These additional parameters will be described in the following sections.

The authors collected stratified data for the whole of Denmark via an online questionnaire (Hielkema & Lund, 2021b), and used the results to build a quantitative model. Their model predicts a categorical variable, called “stage of change” which comprises the categories “NO intention”, “Intention”, and “Reducer”. (The “stage of change” concept comes from Prochaska & Di Clemente’s “Transtheoretical” model (1982).) The explanatory variables Hielkema & Lund studied included a categorical variable describing the number of ties to meat-reducers and the strength of those ties. Their model enables the creation of a stochastic CA algorithm in which each agent’s “stage of change” can be updated every time step, using probabilities derived from real-world data.

4.1 Generation of Social Graph

Creation of large social graphs (or networks) is a slow process. To avoid the need to regenerate the same graphs multiple times, these were created in advance and saved to local files named after the parameters used.

Although the Hielkema & Lund model is for the whole of Denmark, creating social graphs for a country of 5.9 million people (Statistics Denmark, 2021) would result in excessive file sizes. Each edge requires 3 values (source, destination, and strength), and so if the average individual is involved in 10 ties, and values are stored in 64bit integers, a graph for 5.9 million individuals would require over 700MB of disk space. Therefore just one region was selected, specifically North Jutland, as this had approximately 1/10th of the population of the whole of Denmark. For convenience, it was assumed North Jutland was representative of the country as a whole.

Open source population density data collected by Meta (originally Facebook) was downloaded (Meta, 2020), and polygons for Danish administrative regions were downloaded from the Database of Global Administrative Areas (GADM, 2022). A python script was written to combine these to produce gridded population density data for North Jutland, with a small correction to scale to the correct population size provided by Statistics Denmark (2021). The grids divided the region (8.213°E-11.199°E) x (56.555°N-57.749°N) into 1000x1000 equal latitude and longitude cells. The resulting population density map is shown in Figure 5.

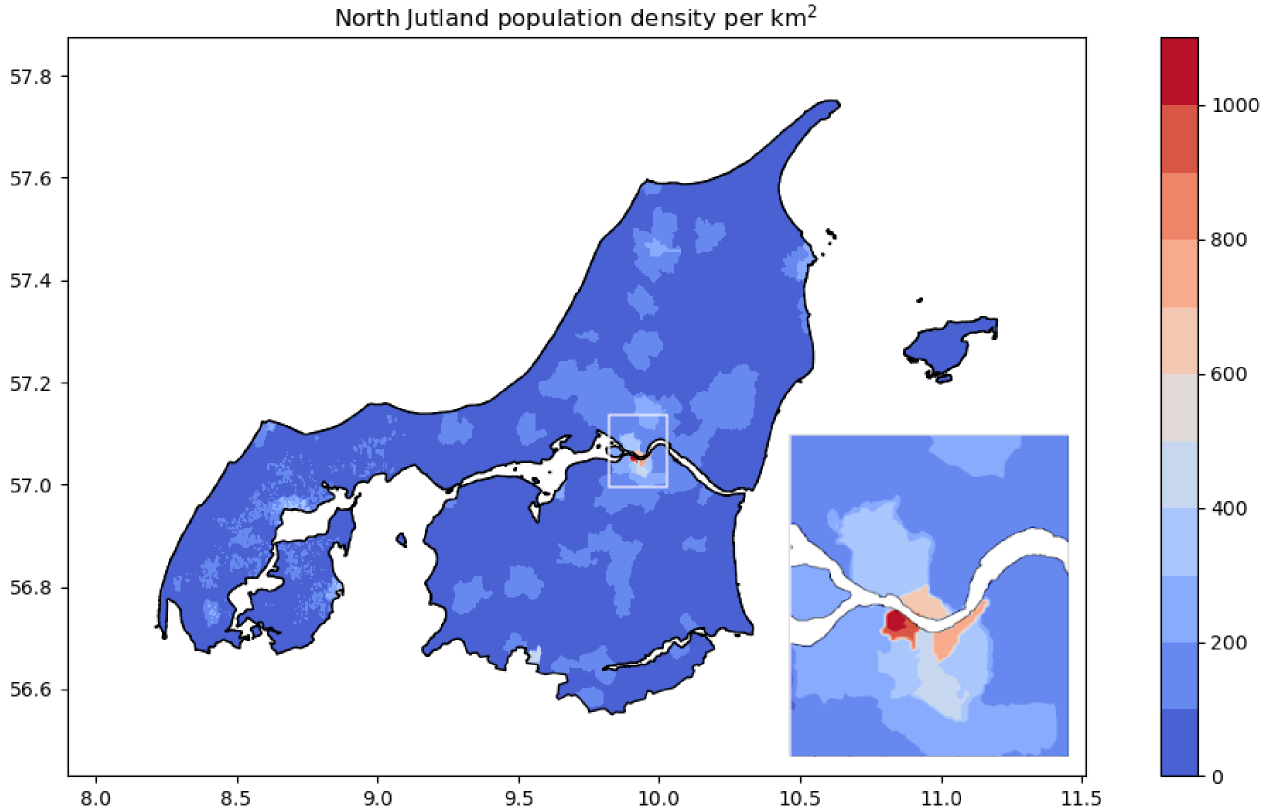


Figure 5: Population density for North Jutland, Denmark, with area around Aalborg expanded

Hielkema & Lund's questionnaire specifies two types of social tie, namely weak and strong, but does not define these terms (Hielkema & Lund, 2021b). Given the context is dietary influences, this study takes "strong tie" to indicate cohabitation (as cohabitants usually dine together every day), and "weak tie" to indicate any other person with whom one dines frequently.

The distribution of household sizes for Denmark was downloaded from Statistica (2021) and used in combination with the gridded population density data to randomly place households, and thereby individuals, adding strong links between household members in the process. A close up of the result of this placement algorithm is shown in Figure 6.

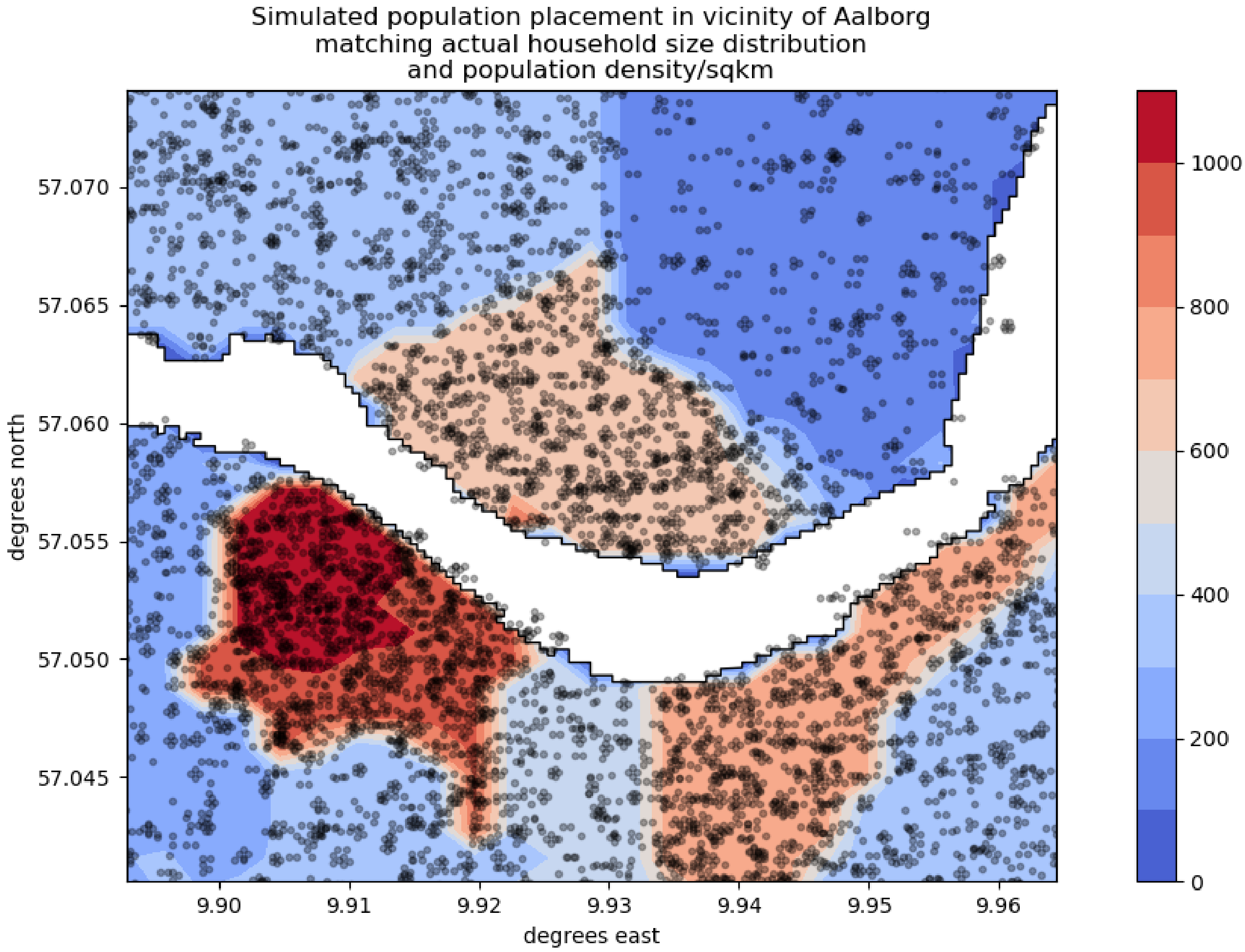


Figure 6: Population placement with multiple-occupant households clearly visible

No data supporting a specific allocation of weak links was found in the literature, so a novel algorithm was used, based on a method by Watts-Strogatz (1998) for building Small World Networks. In the Watts-Strogatz Model, nodes are arranged in a ring and each pair of nodes within a specified range form a link with a specified probability. This is often used to generate realistic social networks. However, using this without modification would ignore the geographical information previously obtained. The novel algorithm used here adapts Watts-Strogatz to a situation in which nodes come with 2D location data:

- The caller specifies the following values when building a social graph:
  - The mean number of weak ties per person, `mean_n_weak_ties`
  - The modal distance between individuals in a weak tie, `modal_weak_tie_km`
- Individuals are processed sequentially and added to `N/2` weak ties, where `N` is chosen from the Poisson distribution with mean `mean_n_weak_ties.` This simulates a situation where weak ties are assigned to individuals at random.
- The peer is chosen at random using a normal distribution, i.e. a probability proportional to

$$e^{-\frac{d^2}{2\sigma^2}}$$

where d is the distance to the peer, and σ is `modal_weak_tie_km.`

Mathematically this is expected to result in i) per-agent weak-tie counts following a Poisson distribution with mean `mean_n_weak_ties`; ii) weak-tie `x` and `y` displacements following centred normal distributions with standard deviation `modal_weak_tie_km`; iii) weak-tie distances having a modal value of `modal_weak_tie_km`. These were all confirmed by plotting the results.

An example of the resulting graph is shown in Figure 7, where the population density has been thinned by a factor of 100 to make the result more clear.

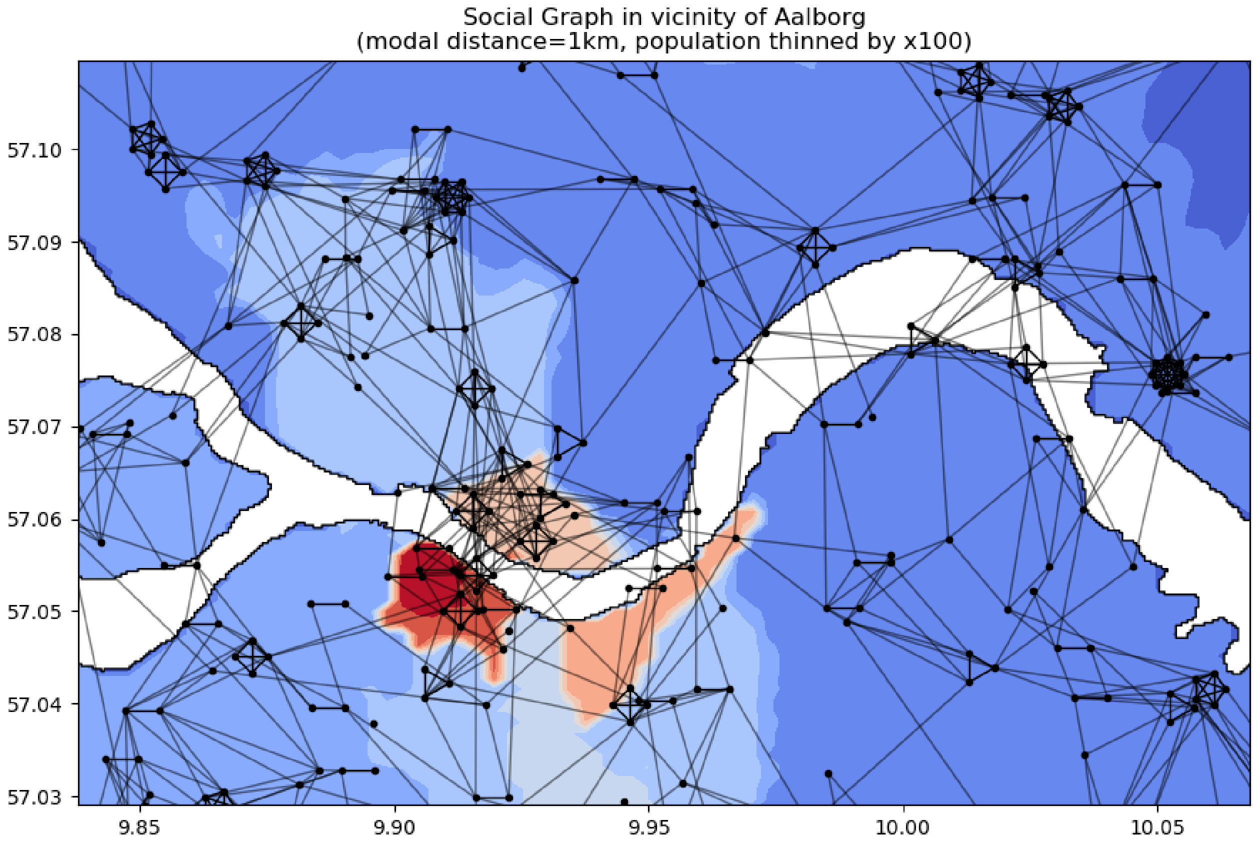


Figure 7: Example social graph with strong ties (intra-household) in bold

The script was invoked with every combination of `mean_n_weak_ties=2,3,4,5,6,7,8` and `modal_weak_tie_km=1,2,3,4,5` to produce 35 files varying between 28MB and 57MB in size.

4.2 Use of Data from Hielkema & Lund Study

The Hielkema & Lund model is a "multinomial logistic regression" model which predicts "stage of change", or Y. This study will focus on just one of their predictor variables, the categorical variable "social networks", or X. Table 1, is reproduced from Hielkema & Lund (2021) with additional annotations identifying the categories of X and Y.

Table 1: Social networks: Type and number of social ties with those who have reduced or stopped eating meat (reference group: knowing no one (X=0))

	Intention (*Y=1*) vs NO intention (*Y=0*)	**Meat reducers (*Y=2*) vs NO intention (*Y=0*)**
	OR (CI)	**OR (CI)**
(*X=1*): 1-2 weak ties	0.69 (0.27–1.81)	0.73 (0.36–1.45)
(*X=2*): 1-2 strong ties	2.78 (1.47–5.26)**	2.54 (1.55–4.16)**
(*X=3*): 1-2 strong and weak ties	3.81 (1.70–8.57)**	3.19 (1.65–6.17)**
(*X=4*): 3+ strong and/or weak ties	0.65 (0.16–2.65)	3.30 (1.56–6.98)**

** 10% significance (p < 0.10)*
*** 5% significance (p < 0.05)*

The multinomial logistic model uses the numbers above to calculate an expected probability, using the equation

$$P(Y=j|X=i) = P(Y=0|X=i)e^{C_j+\beta_{ij}}$$

where β_{ij} are the values in Table 1, where $P(A|B)$ should be read as "the probability of A given that B", and where the C_1, C_2 are some unknown constants. For example, if a person has "1-2 strong and weak ties" then $i=3$, then substituting $j=2$ and $\beta_{ij}=3.19$ produces an equation indicating the probability of them being a "Meat reducer" vs that of having "NO intention".

The above equation can be extended to work for $i=0$ and $j=0$. (Since knowing no one (X=0) is the reference group, $\beta_{0j}=0$ by definition, and substituting $i=0$ shows that $C_0=0$, which forces $\beta_{i0}=0$.) This enables an absolute form of the probabilities to be obtained

$$P(Y=j|X=i) = \frac{e^{C_j+\beta_{ij}}}{\sum_{k=0..2} e^{C_k+\beta_{ik}}}$$

In the formula above, there are two unknown values C_1, C_2. The information present in the paper was insufficient to to determine them. The authors were contacted for additional data but no response was obtained, so it was necessary for these values to be treated as parameters in the cellular automata model. Fortunately these two values are equivalent to (as in "may be derived from") two alternative parameters with a more a meaningful interpretation. These are shown in Table 2.

Table 2: Additional parameters needed to complete predictive "stages of change" model

Name	Description	Formula
`awareness_pc`	% of no-ties population either intending to reduce meat consumption, or actively reducing.	$100 \times P(Y > 0 \mid X = 0)$
`facility_pc`	% of “aware” no-ties population that are active reducers	$100 \times \dfrac{P(Y = 2 \mid X = 0)}{P(Y > 0 \mid X = 0)}$

An individual with no ties to meat reducers who is actively reducing their meat intake or contemplating doing so, cannot be doing so as a result of direct peer influence. It is therefore reasonable to attribute this behaviour to a society-wide level of awareness. For this reason the first of the parameters above is named `awareness_pc`.

The proportion of these individuals that then go on to become meat reducers, reflects the ease of that transition. Since these people have no meat-reducer peers it is reasonable to consider this as reflecting how easy society in general makes meat reduction (e.g. by providing meat alternatives in shops and restaurants). For this reason the second of the parameters above is named `facility_pc`.

The two parameters will be referred to frequently throughout this study. The above table serves as a reference for their technical definitions.

4.3 Implementation of CA Model

The model was implemented in `python`, and the `matplotlib` package was used for illustrations. It was originally intended to use the `networkx` package to support graph creation, but this was found to be slow and have a large memory footprint, so `numpy` arrays were used instead.

The CA model script took five parameters, the first four of which have already been described

1. `mean_n_weak_ties`
2. `modal_weak_tie_km`
3. `awareness_pc`
4. `facility_pc`
5. `p_update_logit_normal_sigma`

The parameters `mean_n_weak_ties` and `modal_weak_tie_km` selected a pre-built social graph to be loaded into memory from file. The parameters `awareness_pc` and `facility_pc` allowed the probabilities for each “stage of change” (Y) to be calculated given an individual’s “social network” (X). The last parameter will be described later.

The script initialized the Y-states using the probabilities for “no-ties”. It then updated all individuals’ Y-states in a loop, with iterations nominally representing 6 months, as in the Transtheoretical model of Prochaska & Di Clemente (1982). This involved identifying which of the 5 “X” categories each individual belonged to at the end of the previous iteration, and then using the probabilities to randomly assign a new “Y” value.

Unfortunately the Hielkema & Lund paper is ambiguous about the precise definition of each “social network” category, and it is difficult to reconcile the descriptions they give with mutually exclusive and exhaustive interpretations. For completeness, the interpretations used in this study are shown in Table 3.

Table 3: Names given by H&L and interpretations used in this study. The interpretations correspond to a strict ordering by number of strong ties, followed by number of weak ties.

H&L description	**Interpretation used in model**
(*X=0*): knowing no one	0 strong ties, 0 weak ties
(*X=1*): 1-2 weak ties	0 strong ties, 1+ weak ties
(*X=2*): 1-2 strong ties	1-2 strong ties, 0 weak ties
(*X=3*): 1-2 strong and weak ties	1-2 strong ties, 1+ weak ties
(*X=4*): 3+ strong and/or weak ties	3+ strong ties

Real individuals do not all re-assess their life choices equally frequently: some are more resistant to change than others. It was therefore decided to assign to each individual a “resistance to change” in the form of a number p. This defined the probability that a new Y-state would be chosen (at random) each iteration. For example, if $p = 0.25$ the individual would have their state updated on

25% of all iterations, on average. Since p is bounded by 0 and 1, a logit-normal distribution of p-values was assumed (i.e. it was assumed that $logit(p)$, or equivalently $log(p/(1-p))$, was normally distributed). It was also assumed that the mean value of p was 0.5, an assumption that can be justified by dropping the requirement that the iteration period represents exactly 6 months. The standard deviation of $logit(p)$ was set by the parameter `p_update_logit_normal_sigma`.

Figure 8 shows the distributions of p for different values of `p_update_logit_normal_sigma`. It is evident that at around `p_update_logit_normal_sigma`=1.5 the distribution changes from unimodal to bimodal (representing a population with a "conservative" peak and an "open-to-change" peak).

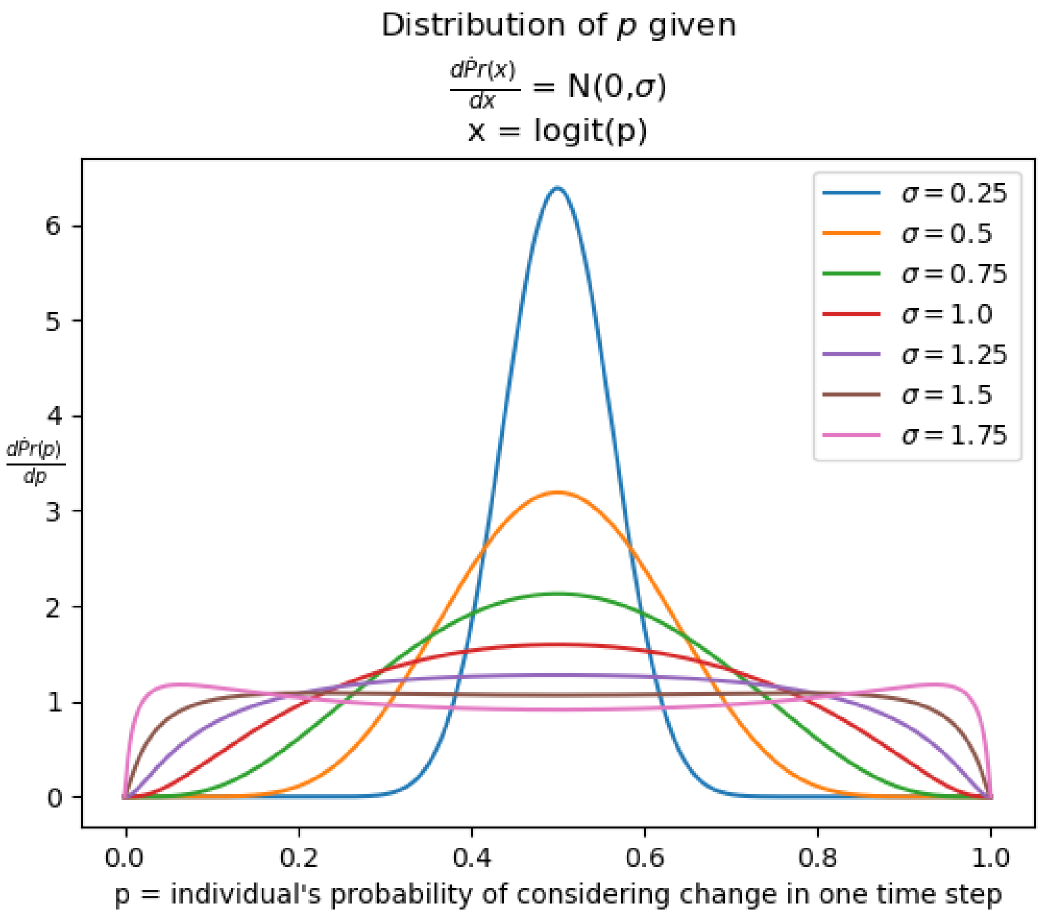


Figure 8: Logit normal distributions with mean 0.5

Finally a script was written to invoke the CA model script thousands of times with randomly selected values for each of 5 parameters, as shown in the code snippet below:

```
mean_n_weak_ties = np.random.choice(range(2,9))
modal_weak_tie_km = np.random.choice(range(1,6))
awareness_pc = np.random.uniform(10,90)
```

```
facility_pc = np.random.uniform(10,90)

p_update_logit_normal_sigma = np.random.uniform(0.25,3)
```

4.4 Processing Model Runs Output

It was observed that numbers of individuals in each “stage of change” converged rapidly to a limit as shown in Figure 9. This limit was usually reached within 5-10 “generations”. Since this was a stable feature of the model, irrespective of the parameters, the calling script was configured to request 20 iterations on each call of the model, and output the final values for the number of individuals in each category.

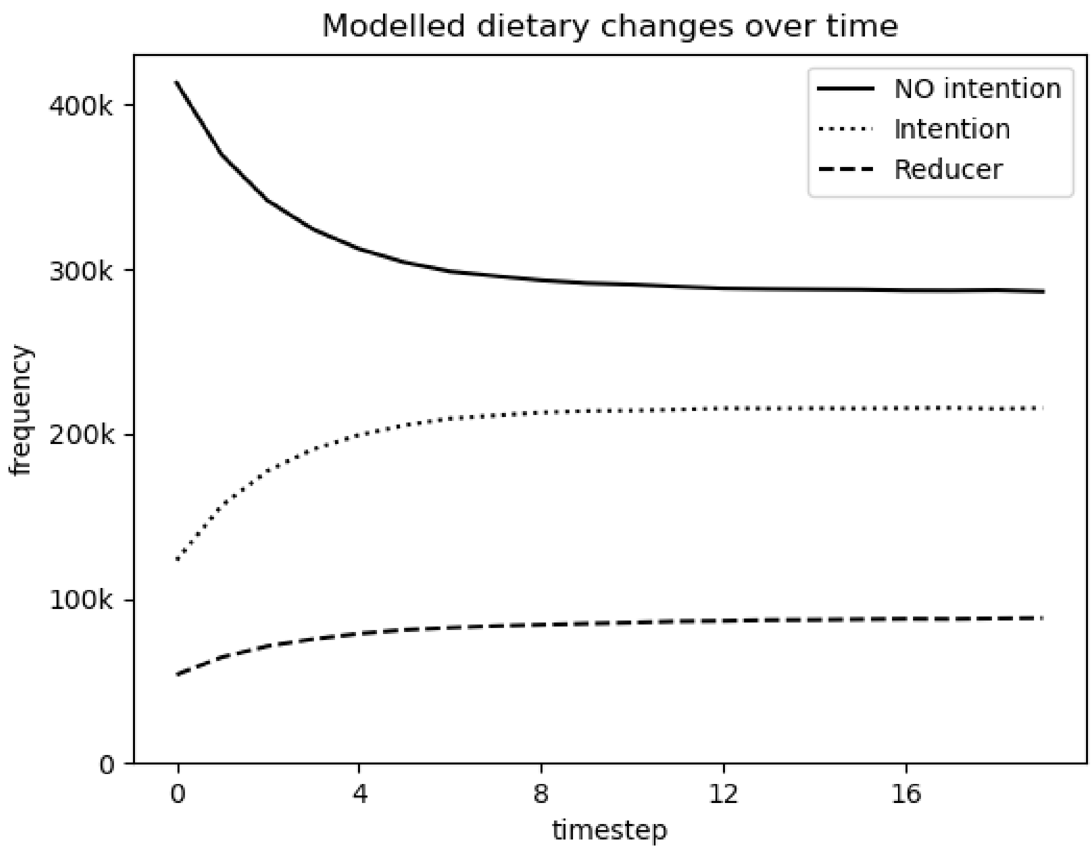


Figure 9: Typical convergence of "stage of change" distribution over time

The results of the model runs were collected into a `.csv` file alongside the parameters used. This was then loaded into the stats package R.

Within R the numbers that end up in each “stage of change” category were converted into percentages, and a model was constructed for how `percent_reducer`, and `percent_intention` depend on each of the 5 parameters.

Linear models

A linear model is a mathematical equation which predicts a response variable from one or more explanatory variables. An example of a simple linear model for the effect of the parameters on percent_reducer might be

$$\begin{aligned}\text{percent\_reducer} =&\alpha_0 + \\ &\alpha_1 \times \text{mean\_n\_weak\_ties} + \\ &\alpha_2 \times \text{modal\_weak\_tie\_km} + \\ &\alpha_3 \times \text{awareness\_pc} + \\ &\alpha_4 \times \text{facility\_pc} + \\ &\alpha_5 \times \text{p\_update\_logit\_normal\_sigma}\end{aligned}$$

Given a dataset and a formula like the one above, the R language routine `lm()` finds the α_i that minimize the sum of the squares of the errors.

Although the formula above is simple it is often better to transform variables prior to building a linear model. Variables v representing percentages were transformed into $logit(v/100)$ which is neither bounded above nor below, and therefore more appropriate for use in a linear model. After this step a visual inspection of relationships was performed to check for non-linear relationships, and square terms were added where necessary. Finally, terms that did not contribute significantly to the model were dropped using Akaike Information Criterion, a standard method for finding a balance between parsimony and explanatory power in a linear model.

This process produced a formula, which was passed to `lm()` to produce confidence intervals for the coefficients and an estimate for the amount of variation in the results “explained” by the model.

4.5 Sensitivity Analyses

4.5.1 Randomizing Coefficients in Upstream Model

In Table 1 the β-values are followed by 95% confidence intervals, for example “0.69 (0.27 – 1.81)”. The central estimate is the geometric mean of the upper and lower estimates. These upper and lower estimates can therefore be used to obtain the standard deviation of $log(\beta)$ which the multinomial logistic model assumes to be normally distributed. A sensitivity analysis was conducted by randomly perturbing each β_{ij} from its central estimate in each of several thousand runs to create a sensitivity analysis dataset. The model for `percent_reducer` was then regenerated using this new dataset.

4.5.2 Allowing Strong Ties Between Households

The social network graphs produced assume that all intra-household ties are strong and all inter-household ties are weak. This may be a reasonable approximation given the context. However, it does mean that the strong-tie graph is highly disconnected, and it is therefore possible that even a small proportion of strong ties between households may change outcomes dramatically.

A representative set of parameters was chosen and the CA model was re-run, but with varying proportions of the inter-household links randomly reconfigured as strong ties. This enabled the sensitivity of the model to be tested against the assumption that all inter-household ties were weak.

4.5.3 Feedback Model

In the CA model described earlier `awareness_pc` and `facility_pc` are fixed over time. In the real-world it is unlikely that society-wide measures of awareness (of the arguments for meat-reduction) or facility (of reducing meat consumption) remain constant. In particular there is likely to be some feedback causing these values to change. For example, increased numbers of meat-reducers might encourage supermarket and restaurant chains to offer meat alternatives, and to advertise their environmental benefits. Conversely, there could be a backlash as meat-producers use lobbying power to attempt to reverse such changes.

There are many ways in which feedbacks could be simulated. Choosing one method is difficult to justify, as it is equivalent to predicting the responses of a small number of influential individuals in business, media, and government. Nonetheless it is instructive to explore how a nominal feedback term might affect the output of the CA model.

A feedback parameter taking values between 0 and 1 was added to the CA model. The following calculation was performed at the end of each iteration:

$$\text{effective\_awareness\_pc} = \text{awareness\_pc} + \text{feedback} \times \frac{\text{percent\_reducer}}{100} \times (100 - \text{awareness\_pc})$$
$$\text{effective\_facility\_pc} = \text{facility\_pc} + \text{feedback} \times \frac{\text{percent\_reducer}}{100} \times (100 - \text{facility\_pc})$$

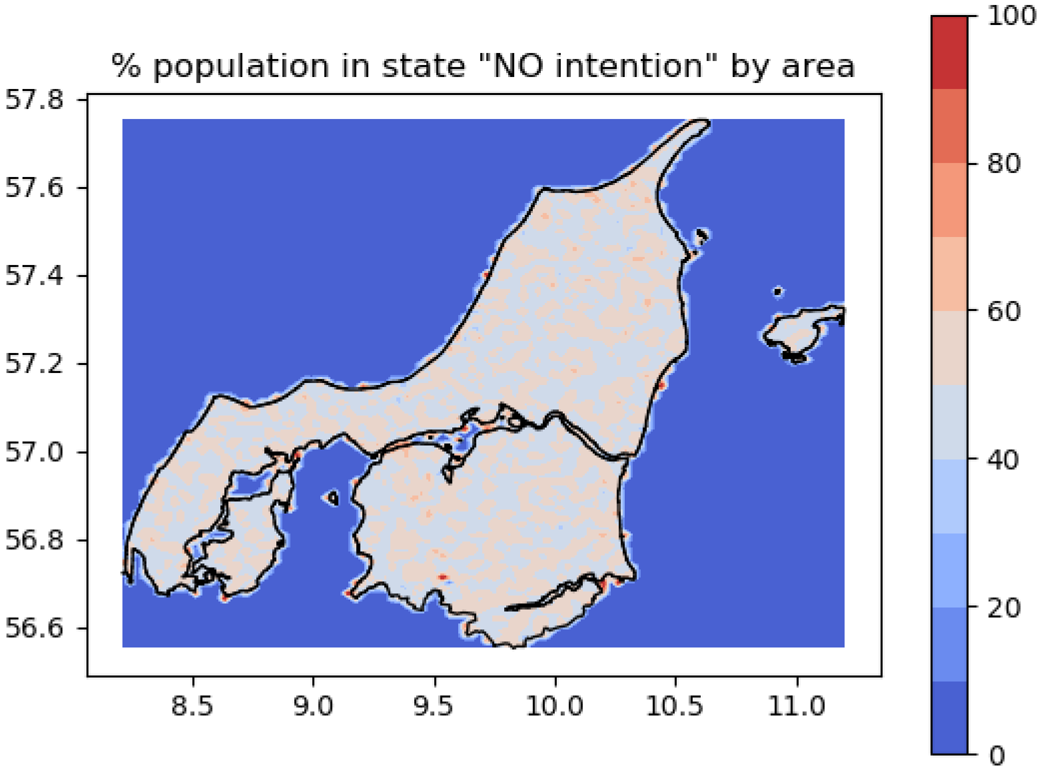
where `percent_reducer` represents the current proportion in state reducer, and where `effective_awareness_pc` and `effective_facility_pc` are used to recalculate the probabilities for the next iteration. A dataset was created in which `feedback` took random values and this was used to build a linear model with `lm()`.

5 Results

The CA runs were performed on a 2GHz quad core Intel PC with 8GB RAM. The script was initially run for about 24 hours which resulted in a `.csv` file containing the results of 2963 CA runs. This process was repeated for the randomized coefficient sensitivity analysis, resulting in a second `.csv` file containing the results of 4669 CA runs.

5.1 Observations

The final distribution of individuals in each stage of change appeared to be homogenous rather than showing a correlation to population density or geography. This homogeneity feature appeared to be independent of the parameters used. The results for the default parameters (`mean_n_weak_ties=6, modal_weak_tie_km=1, awareness_pc=30, facility_pc=30.33, p_update_logit_normal_sigma=0.5)` are shown in Figure 10.



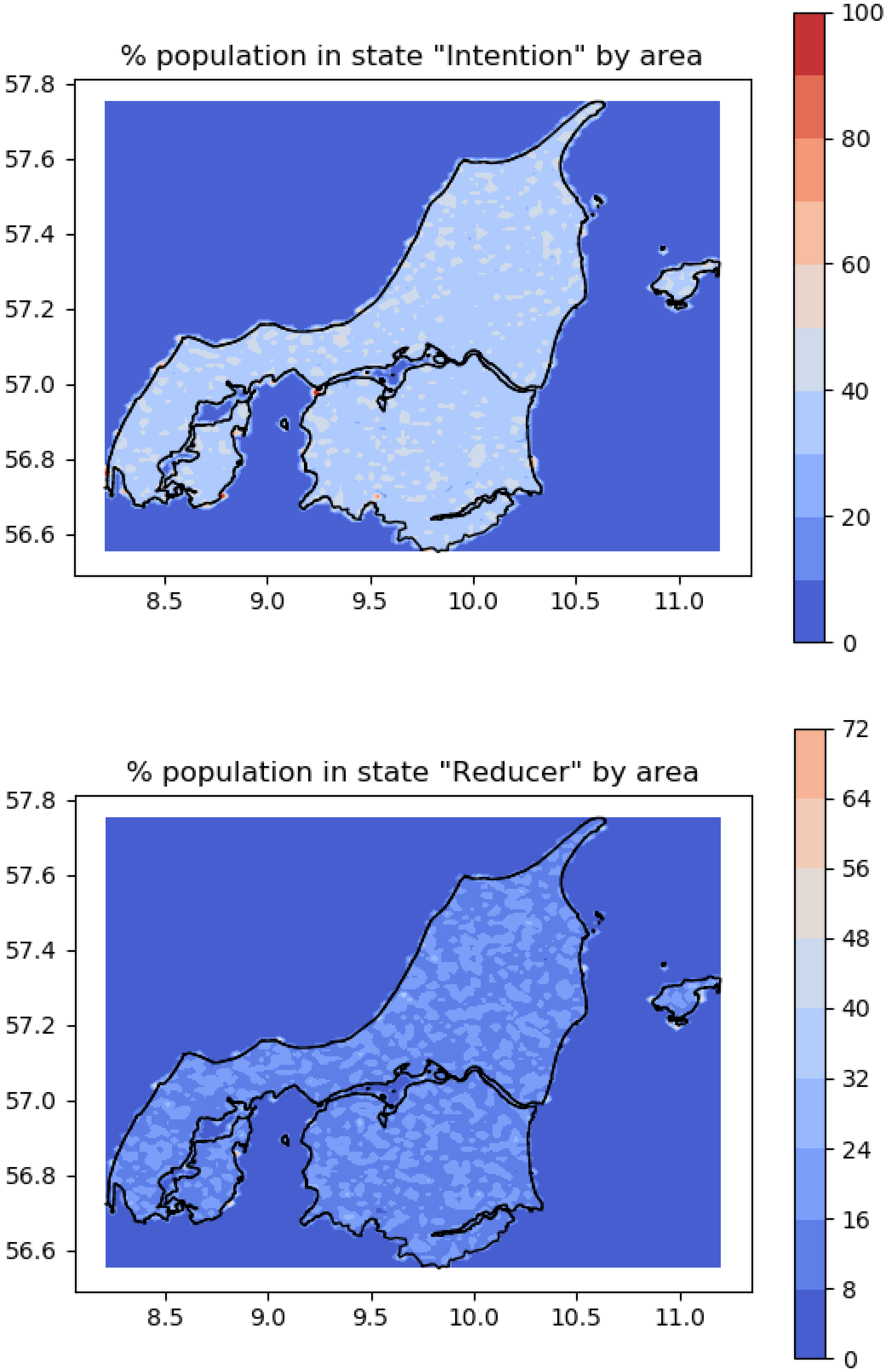


Figure 10: Distribution of individuals in each final state (using default parameters)

This homogeneity is supported by the fact that the proportions within grid squares appear to be normally distributed as shown in Figure 11.

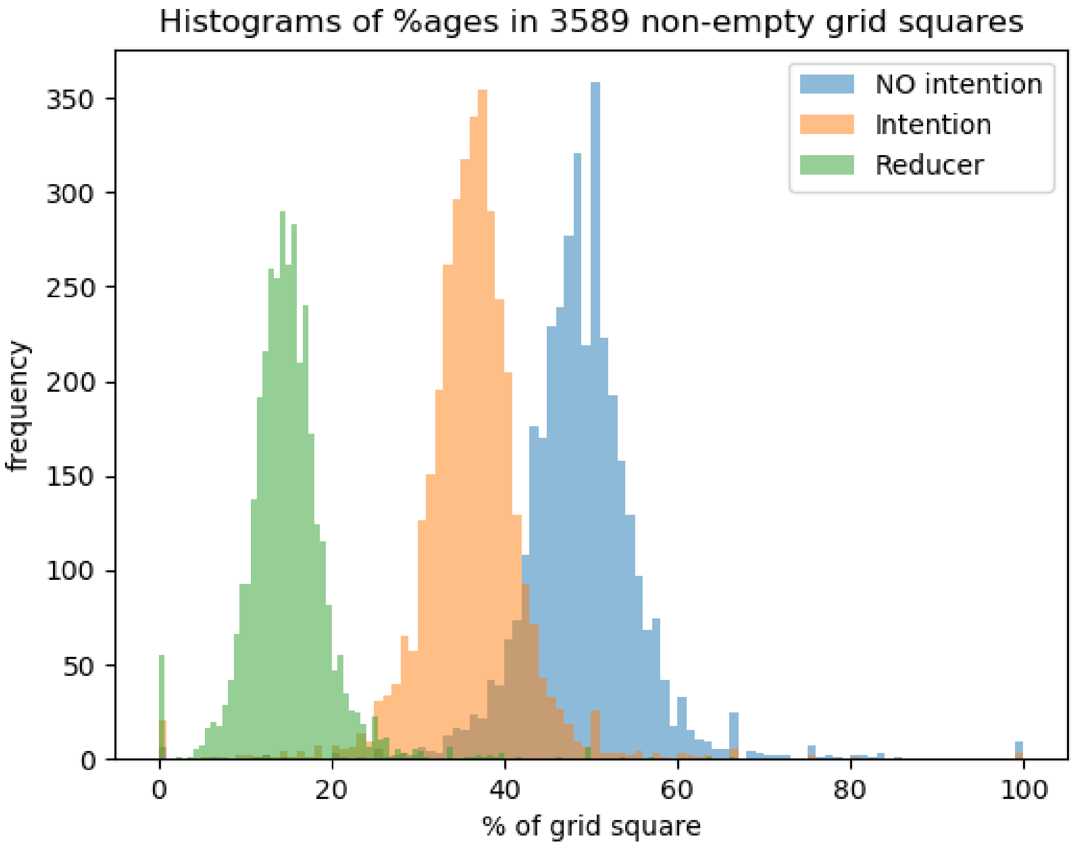


Figure 11: Proportions of each stage of change in grid squares seem normally distributed (default parameters)

Although homogeneity is the rule on the large scale, it is apparent from a visual inspection that households tend to divide into either meat-reducer, or non-meat-reducer households. Figure 12 shows an example, which appears to show that if a single person in a household becomes a meat reducer then the co-habitants will often follow suit. As a result the largest households tend to have far more meat-reducers than in the general population.

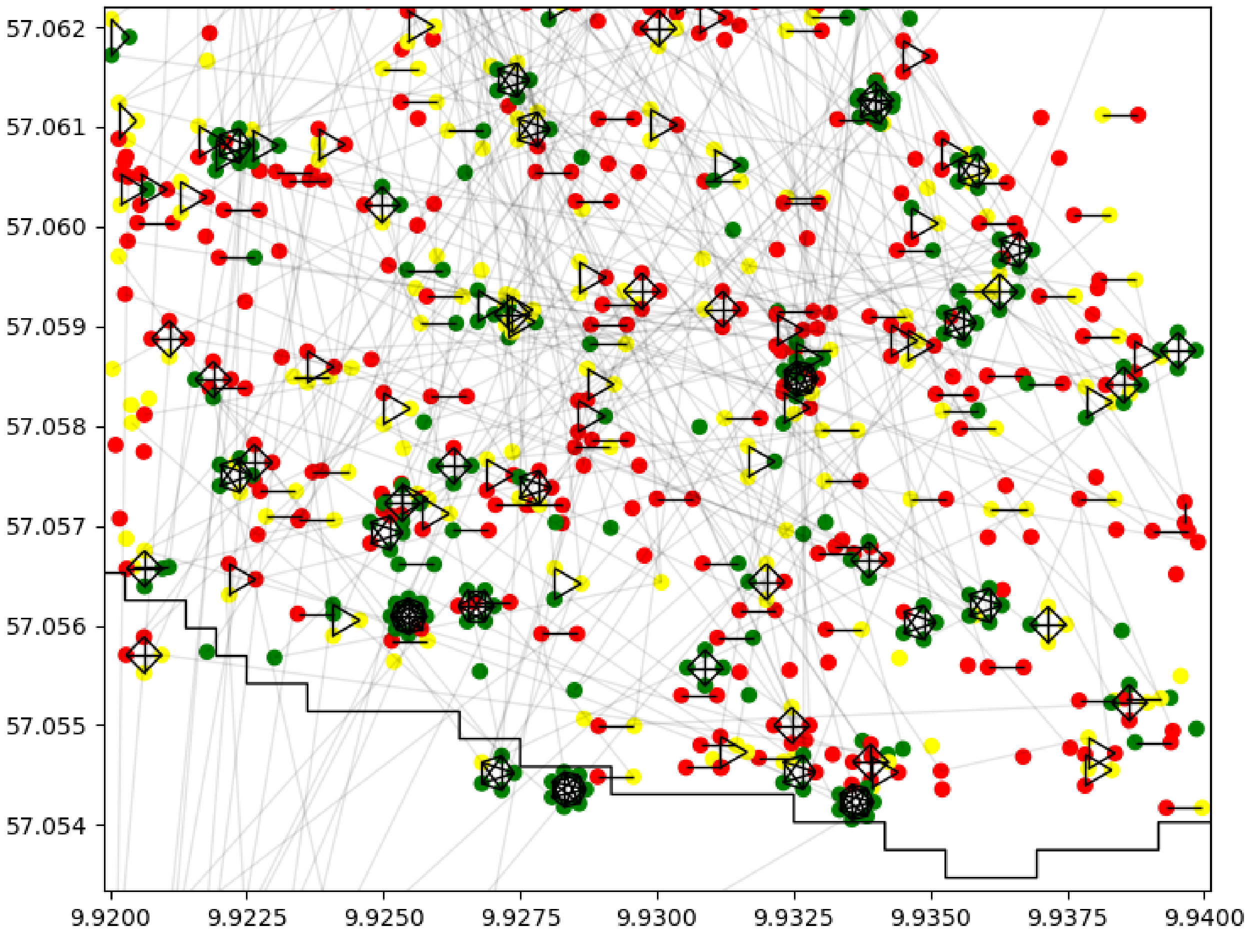


Figure 12: Close up of final stage-of-change distribution with Red="NO intention", Yellow="Intention", Green="Reducer". Large households appear to be more likely to be all reducer.

5.2 Linear Models

The results of the 2963 runs were loaded into R, and processed to obtain a model for the percentage of reducers after convergence, and the percentage of individuals intending to reduce, in terms of the explanatory variables.

5.2.1 Percentage of Reducers Model

The response variable `percent_reducers` was plotted against each explanatory variable and a histogram of `percent_reducers` was also plotted. The results are shown in Figure 13.

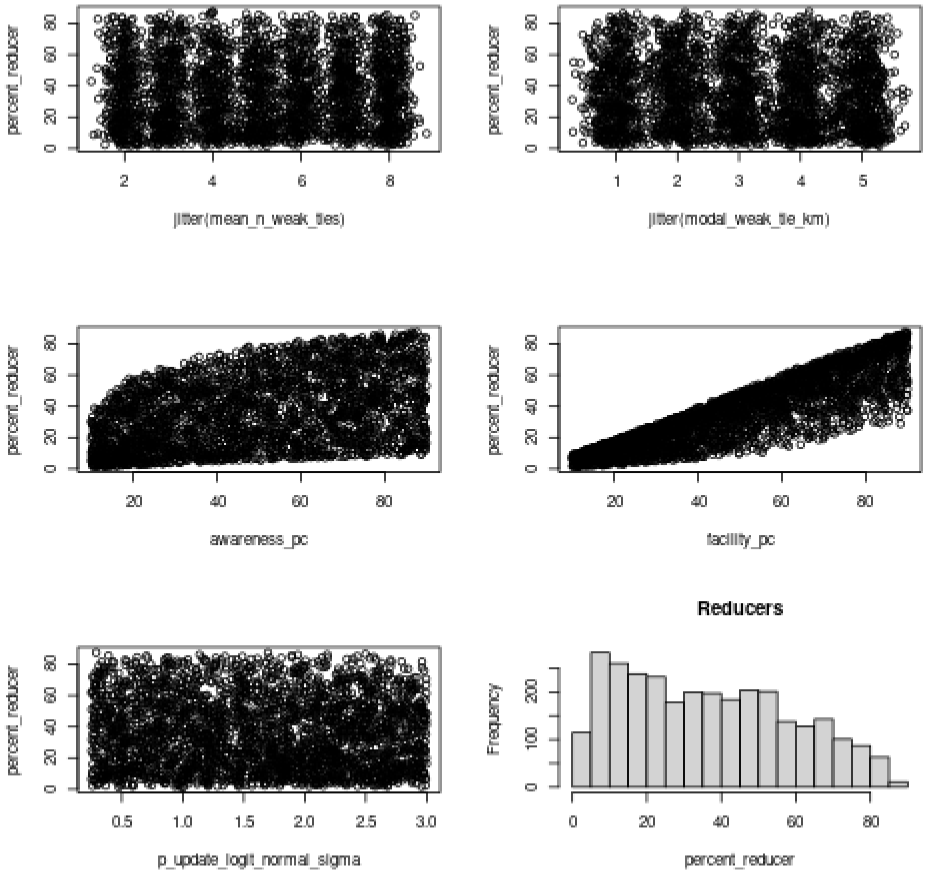


Figure 13: Histogram of percent_reducer and scatter plots against each explanatory variable

These charts show a clear positive relationship with `awareness_pc` and `facility_pc`, although the relationship with other parameters is less clear. The histogram shows a positive skew (the median is much higher than the mean) which suggests transforming the response variable. Since the response variable `percent_reducer` is a percentage it is appropriate to use a logistic transform to obtain a new response variable that is neither bounded above nor below. The new response variable was named `logit_reducer` and defined as $logit(\text{percent\_reducer}/100)$. The explanatory variables `awareness_pc` and `facility_pc` were likewise transformed into new explanatory variables `logit_awareness` and `logit_facility`. After transforming the variables the histogram and scatter plots were recreated, as shown in Figure 14.

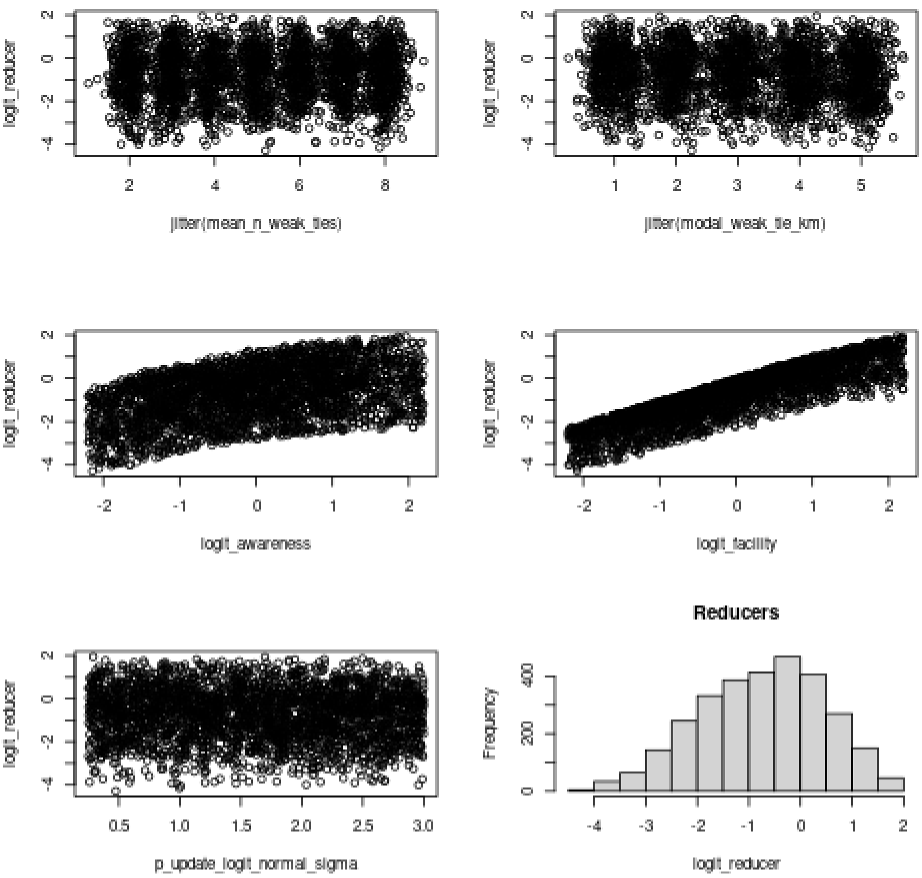


Figure 14: Histogram of logit_reducer and scatter plots against transformed explanatory variables

After transforming the response variable the histogram appeared much closer to a normal distribution, indicating that the transformation made the data more appropriate for use in a linear model. From visual inspection, it was apparent there were positive correlations with `logit_awareness` and `logit_facility`. It was also clear from the graphs that these relationships are non-linear, and so the 2^{nd} order terms `logit_awareness`^{2} and `logit_facility`^{2} were included in the model.

A linear model for `logit_reducer` was created using all 5 transformed explanatory variables including the two 2^{nd} order terms. Akaike Information Criterion was used to determine that `modal_weak_tie_km` could be dropped without reducing the explanatory power of the model.

The final model, generated by the R function `lm()`, is reproduced in Table 4.

Table 4: Results of model for logit_reducer

Coefficients:	Estimate	Std. Error	t value	Pr(>|t|)
(Intercept)	-0.4464706	0.0041644	-107.21	<2e-16 ***
mean_n_weak_ties	0.0105095	0.0005616	18.71	<2e-16 ***
logit_awareness	0.4697890	0.0010208	460.23	<2e-16 ***
I(logit_awareness^2)	-0.1230449	0.0008832	-139.31	<2e-16 ***
logit_facility	0.9568436	0.0010377	922.04	<2e-16 ***
I(logit_facility^2)	-0.0684336	0.0008842	-77.40	<2e-16 ***
p_update_logit_normal_sigma	-0.0594041	0.0014540	-40.85	<2e-16 ***

*Signif. codes: 0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1*
Residual standard error: 0.06239 on 2956 degrees of freedom
Multiple R-squared: 0.9973, Adjusted R-squared: 0.9973
F-statistic: 1.832e+05 on 6 and 2956 DF, p-value: < 2.2e-16

The R-squared value indicates that 99.7% of the variation in `logit_reducer` is explained by this model, making the model extremely accurate. Figure 15 shows that the residuals are more or less normally distributed, and not significantly correlated with the fitted values, which confirms the assumptions made by `lm()`. It also gives a visual indication of how close the predictions are to the actual values of `logit_reducer`.

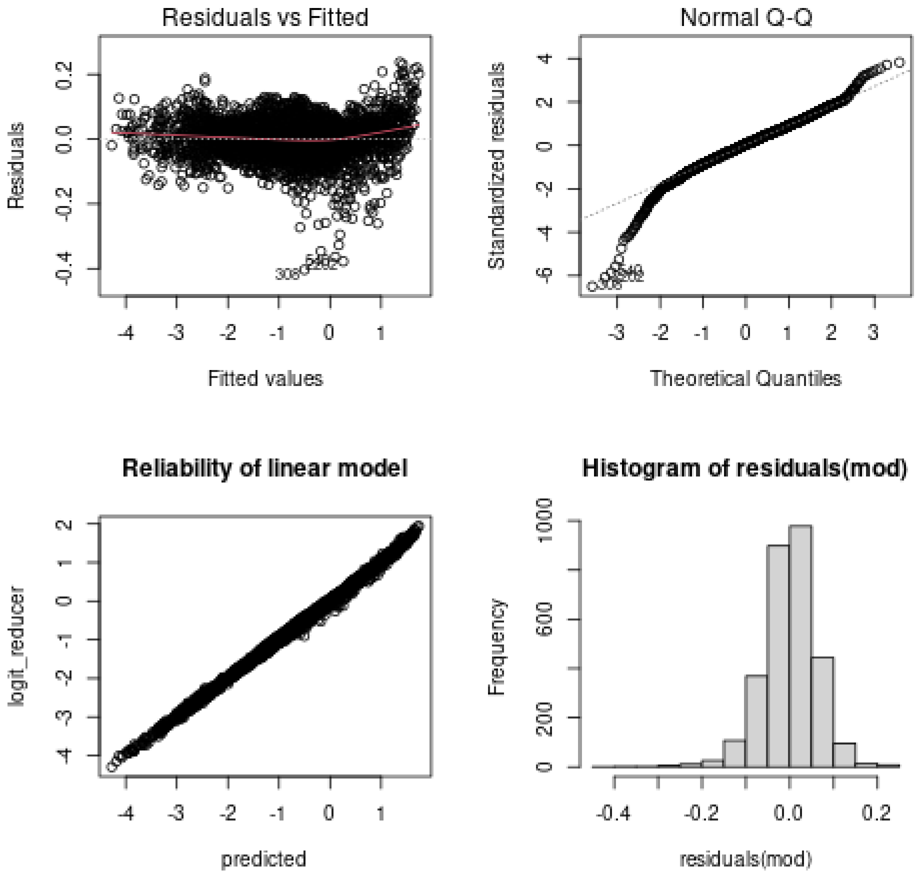


Figure 15: Visual confirmation of model assumptions

The model can be rearranged into an equation for `percent_reducer` as shown below

$$
\begin{aligned}
&\text{percent\_reducer} = \\
&\quad 100 \times logit^{-1}(\\
&\qquad -0.4464706+ \\
&\qquad\; 0.0105095 \times \text{mean\_n\_weak\_ties}+ \\
&\qquad\; 0.4697890 \times logit(\text{awareness\_pc}/100) - 0.1230449 \times logit(\text{awareness\_pc}/100)^2+ \\
&\qquad\; 0.9568436 \times logit(\text{facility\_pc}/100) - 0.0684336 \times logit(\text{facility\_pc}/100)^2+ \\
&\qquad -0.0594041 \times \text{p\_update\_logit\_normal\_sigma})
\end{aligned}
$$

The model was regenerated using scaled and centred explanatory variables in order to determine the relative importance of each. When this was done it was found that `logit_facility` was approximately twice as important as `logit_awareness` which was more than 10 times as important as any other variable.

5.2.2 Percentage of Intenders Model

A similar analysis was performed to produce a model for the percentage of the population intending to reduce meat consumption, following convergence. A histogram and scatter plots of the response variable `logit_intention` versus the explanatory variables is shown in Figure 16.

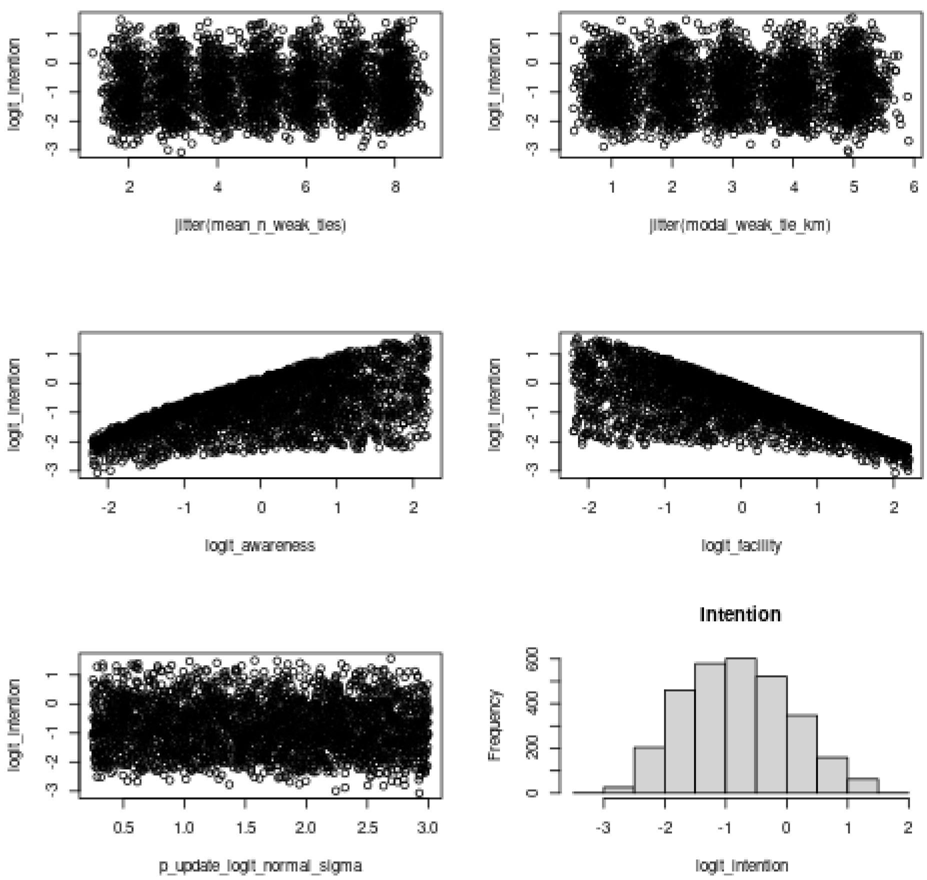


Figure 16: Histogram of logit_intention and scatter plots against transformed explanatory variables

As with `logit_reducer` the histogram shows that the response variable is more or less normally distributed and that there is a clear positive correlation with `logit_awareness`. However, unlike with `logit_reducer`, `logit_intention` appears to be negatively correlated with `logit_facility`. This is unsurprising as facilitating meat reduced diets means making it easier to transition from simply intending to reduce meat consumption to actually doing it.

A linear model was created for the response variable `logit_intention` using all 5 explanatory variables, including the squares of `logit_awareness`, and `logit_facility` (to account for the non-linear relationship apparent in the scatter plots). As with the model for `logit_reducer`, use of Akaike Information Criterion led to the explanatory variable `modal_weak_tie_km` being removed, as its presence did not increase the explanatory power of the model.

The final model, generated by the R function `lm()`, is reproduced in Table 5.

Table 5: Results of model for logit_intention

Coefficients:	**Estimate**	**Std. Error**	**t value**	**Pr(>|t|)**
(Intercept)	-0.571931	0.017466	-32.745	<2e-16 ***
mean_n_weak_ties	0.027138	0.002355	11.522	<2e-16 ***
logit_awareness	0.438523	0.004281	102.429	<2e-16 ***
I(logit_awareness^2)	-0.104790	0.003704	-28.288	<2e-16 ***
logit_facility	-0.557292	0.004352	-128.041	<2e-16 ***
I(logit_facility^2)	-0.124620	0.003708	-33.605	<2e-16 ***
p_update_logit_normal_sigma	-0.032487	0.006098	-5.327	1.07e-07

*Signif. codes: 0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1*
Residual standard error: 0.2617 on 2956 degrees of freedom
MMultiple R-squared: 0.9083, Adjusted R-squared: 0.9081
F-statistic: 4882 on 6 and 2956 DF, p-value: < 2.2e-16

The R-squared value indicates that 90.8% of the variation in `logit_intention` is explained by this model, making the model almost as accurate as the model for logit_reducer. Figure 17 shows that the residuals are more or less normally distributed, which confirms the assumptions made by `lm()`. However, there is some patterning of the residuals vs the fitted values. Nonetheless the visual indication of the model reliability shows that it makes good predictions of `logit_intention` based on the explanatory variables.

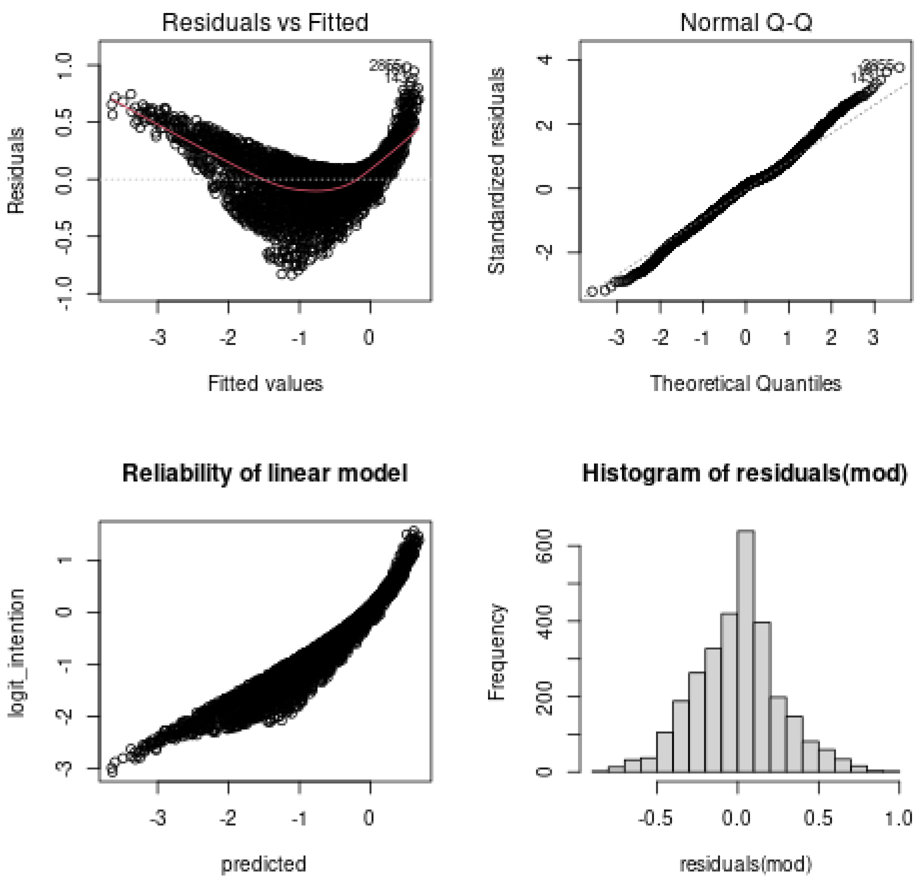


Figure 17: Visual confirmation of model assumptions

The model can be rearranged into an equation for `percent_intention` as shown below

$$
\begin{aligned}
&\text{percent\_intention} = \\
&\quad 100 \times logit^{-1}(\\
&\qquad -0.571931+ \\
&\qquad\; 0.027138 \times \text{mean\_n\_weak\_ties}+ \\
&\qquad\; 0.438523 \times logit(\text{awareness\_pc}/100) - 0.104790 \times logit(\text{awareness\_pc}/100)^2+ \\
&\qquad -0.557292 \times logit(\text{facility\_pc}/100) - 0.124620 \times logit(\text{facility\_pc}/100)^2+ \\
&\qquad -0.032487 \times \text{p\_update\_logit\_normal\_sigma})
\end{aligned}
$$

The model was regenerated using scaled and centred explanatory variables in order to determine the relative importance of each. When this was done it was found that `logit_facility` and `logit_awareness` were more or less equally important, and about 10 times as important as any other variable.

5.3 Estimates of `awareness_pc` and `facility_pc`

As stated earlier, the data from Hielkema & Lund was not sufficient to determine all of the probabilities for each stage of change (Y) given each social network (X). As a result, two additional parameters were specified, namely `awareness_pc` and `facility_pc` as defined in Table 2. These can be thought of as conveying the society-wide levels of awareness (of the need for meat reduction) and facility (of embarking on a meat-reduced diet). Although values for these parameters could not be inferred from the original study, the CA model produced here allows estimates of them to be made.

Since `awareness_pc` and `facility_pc` were by far the most significant parameters it is reasonable to ignore the other parameters in the model and assert that the `awareness_pc` and `facility_pc` values which generate the most realistic outcomes are closest to the correct values.

The Hielkema & Lund study stated that 57.3% of all respondents indicated no intention to reduce, whilst 11.4% indicated an intention and 31.2% indicated that they were actively reducing meat consumption. The model run that produced the results closest to this had an `awareness_pc` of 14% and a `facility_pc` of 76% with the ten closest results having `awareness_pc` between 11 and 16%, and `facility_pc` between 70 and 80%.

This rough estimate suggests that, in the population studied, the proportion of those at least intending to reduce despite no ties to reducers (`awareness_pc`) is low: between 1 in 10 and 1 in 5. However, amongst those the proportion actually reducing (`facility_pc`) is high (around three quarters) suggesting there are not too many barriers to meat reduction for those wishing to do so.

5.4 Sensitivity Analyses

5.4.1 Randomizing Coefficients in Upstream Model

The results of the 4669 sensitivity analysis runs were loaded into R. In these runs the central estimate values determined by Hielkema & Lund and shown in Table 1 were replaced by randomized values, using the distribution given by the 95% confidence intervals.

The analysis for `percent_reducer` was repeated using this new dataset. Again, `modal_weak_tie_km` was found to be an unnecessary explanatory variable, using AIC, and was removed from the model. The final model produced by the R function `lm()` is shown in Table 6.

Table 6: Results of model for logit_reducer in randomized coefficient sensitivity analysis

Coefficients:	**Estimate**	**Std. Error**	**t value**	**Pr(>|t|)**
(Intercept)	-0.283056	0.030611	-9.247	<2e-16 ***
mean_n_weak_ties	0.017156	0.004189	4.095	4.29e-05 ***
logit_awareness	0.460082	0.007589	60.624	<2e-16 ***
I(logit_awareness^2)	-0.123726	0.006485	-19.078	<2e-16 ***
logit_facility	0.866348	0.007772	111.469	<2e-16 ***
I(logit_facility^2)	-0.097502	0.006677	-14.602	<2e-16 ***
p_update_logit_normal_sigma	-0.101315	0.010678	-9.488	<2e-16 ***

*Signif. codes: 0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1*
Residual standard error: 0.5747 on 4662 degrees of freedom
Multiple R-squared: 0.7805, Adjusted R-squared: 0.7802
F-statistic: 2763 on 6 and 4662 DF, p-value: < 2.2e-16

The R-squared value indicates 78% of the variation in the response variable is explained by this model. This shows that the model is still very predictive despite the deliberate pertubations. All of the coefficients in the sensitivity analysis produced model match those from the normal model in terms of order of magnitude and sign, although `logit_facility`, `logit_facility`^{2}, and `p_update_logit_normal_sigma` in the normal model are outside 2 standard errors.

A visual check of the validity of the sensitivity analysis produced model is shown in Figure 18. It is clear that the model has a predictive value but that uncertainty in the Hielkema & Lund coefficients leads to the greatest level of uncertainty at the higher end predictions of `percent_reducer`.

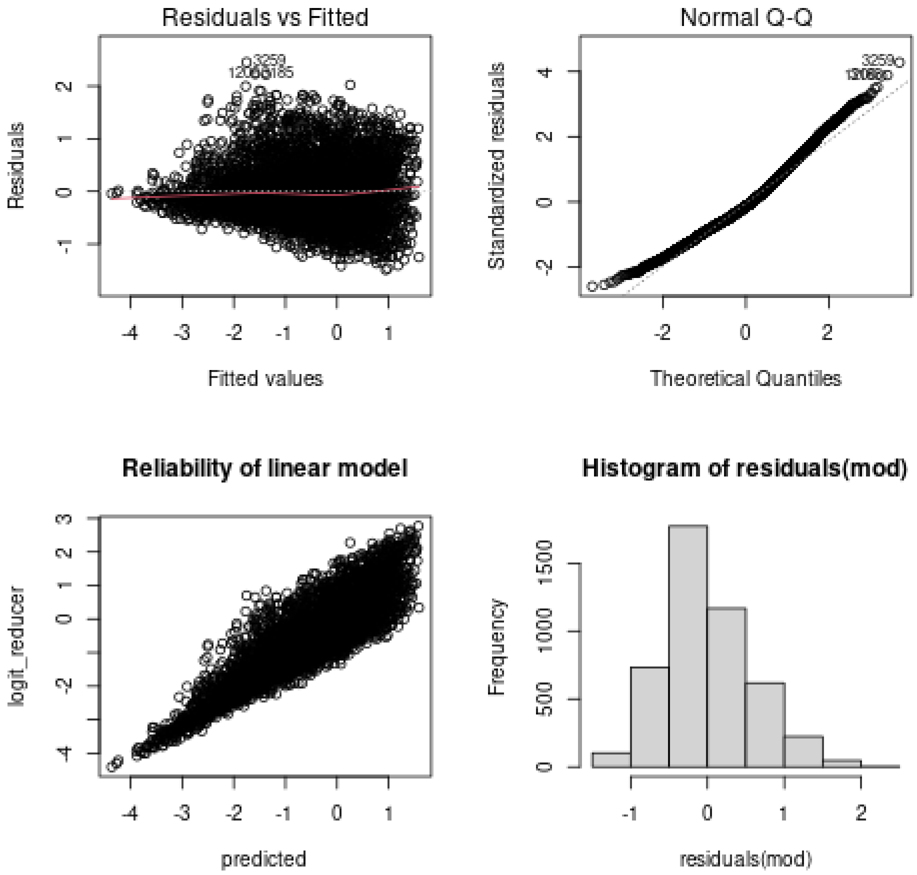


Figure 18: Visual confirmation of model assumptions

5.4.2 Allowing Strong Ties Between Households

The present CA model has no strong-ties between households, whereas every pair of cohabiting individuals is linked by a strong tie. A visual check of the results suggested it was far more likely for individuals in large households to become meat-reducers(see Figure 12). This opens up the possibility that even a small number of strong-ties between households could trigger a cascade effect, ultimately leading to the whole population becoming meat-reducers.

To test this hypothesis 5 additional runs were performed. The best estimates of `awareness_pc` and `facility_pc` (from 5.3) were used for all 5 runs. In each run a different proportion of inter-household ties were selected at random and replaced with strong ties. The results are shown in Table 7.

Table 7: Result of re-running CA model with some inter-household ties treated as strong. Other parameters used were: awareness_pc=14, facility_pc=76, mean_n_weak_ties=6, p_update_logit_normal_sigma=0.5. Percentages indicate post-convergence values.

Percent inter-household ties strong	Percent NO intention	Percent Intention	Percentage Reducer
0	50.4	13.2	36.4
5	44.9	14.6	40.5
10	40.1	15.7	44.2
15	35.9	16.3	47.8
20	32.6	16.3	51.1
50	24.0	10.6	65.4
100	23.4	3.0	73.6

Replacing inter-household ties clearly does have an impact, but the impact is gradual and there is no cliff edge threshold. This suggests that the CA model is not overly sensitive to the assumption that all inter-household ties are weak. Even if as much as one in five inter-household ties are strong the "percent reducer" prediction would only need to be adjusted upward by about 15 points.

5.4.3 Feedback Model

It was found that adding a large feedback value typically caused the length of time taken for convergence to double. Simulations involving feedback were run for five times as many iterations to guarantee convergence. A dataset with 1849 runs was collected and a linear model was created using the same technique described in 5.2.1. The result was a model explaining 93% of the variation in the response variable `logit_reducer`. The coefficients are shown in Table 8.

Table 8: Results of model for logit_reducer in feedback parameter sensitivity analysis

Coefficients:	Estimate	Std. Error	t value	Pr(>|t|)
(Intercept)	0.864862	0.026791	32.28	<2e-16 ***
logit_feedback	0.480340	0.005434	88.40	<2e-16 ***
logit_awareness	0.379694	0.009201	41.27	<2e-16 ***
I(logit_awareness^2)	-0.097686	0.007949	-12.29	<2e-16 ***
logit_facility	1.082118	0.009065	119.38	<2e-16 ***
I(logit_facility^2)	-0.124271	0.007665	-16.21	<2e-16 ***
p_update_logit_normal_sigma	-0.144507	0.012913	-11.19	<2e-16 ***

*Signif. codes: 0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1*
Residual standard error: 0.4285 on 1842 degrees of freedom
Multiple R-squared: 0.9311, Adjusted R-squared: 0.9309
F-statistic: 4147 on 6 and 1842 DF, p-value: < 2.2e-16

The coefficients in this model are similar to those in the non-feedback model except that: a) `mean_n_weak_ties` no longer improves the model sufficiently to be left in; b) the new parameter `logit_feedback` has appeared and its coefficient is of the same order of magnitude as that of `logit_awareness`.

There does not appear to be any kind of avalanche effect caused by adding in this type of feedback, instead `logit_reducer` appears to respond linearly to `logit_feedback` similarly to how it responds to `logit_awareness`.

6 Discussion

3000 words

definition of “reducer” vague and may reflect cultural desire to be seen to be doing the right thing

facility (as defined) twice as important as awareness (as defined)

How does novel algorithm for social network compare with Watts-Strogatz? Did it make much of a difference in the end? Why?

Denmark has a high proportion of single occupancy households. May be different in different cultures in which this is not true – individualism.

Other populations – different data but same approach?

Guess real facility_pc awareness_pc based on distribution of Y in H&L

Put in all the caveats

- resistance to change mean unknown. we’ve said we don’t need to know this as one generation no longer represents 6mo. This is fine if it is still the same order of magnitude. However, if it is much longer then the 5-10 generations needed for convergence may end up being longer than a person’s lifetime, and the model breaks down. Needs data.
- Number of ties may be an effect rather than a cause…. Only circumstantial evidence that its an effect
- unreliable survey responses – people trying to look good
- For some reason H&L ignored veggies in building their model

plot of theoretical model percent_reducer vs awareness_pc (several values of facility_pc, ignore other less important explanatory variables) and percent_reducer vs facility_pc (ditto).

In our model households were complete graphs of strong ties and this led to large households tending to be more likely to be mostly veggie. Is this representative of real life? Paint a scenario

Explanation why it becomes homogenous quickly.

Say that fact that it is relatively independent of number of weak ties and independent of weak tie distance is surprising but a good thing as these were unknowns.

Say it's all about making life easier for veggies and making people aware.

Killer to the BP argument that it's all about individual choice….!

7 Conclusions

No particular cliff edge found

Awareness and facility most important. Reference Sanchez-Sabate 2019.

Feedback important but an unknown

Individual choice with cascading unlikely to solve problem on its own

1000 words

8 References

Use all the below plus all from proposal plus scopus and then sort and filter

Griffeath D. (2015). Self-Organising Two-Dimensional Cellular Automata. In: Designing Beauty: The Art of Cellular Automata, ed. Springer, New York, pp.1-12.

BP (2004), "Drive down your carbon footprint", available at https://www.bp.com/en_gb/target-neutral/home/calculate-and-offset-your-emissions/travel.html>, last accessed 5th Mar 2022

Meta, 2020 https://data.humdata.org/dataset/c846d97b-4363-469a-b66b-1ac2674913b0/resource/cd2c4503-053a-4006-9d0f-efd985e2188f/download/dnk_general_2020_csv.zip

Statistics Denmark, 2021 https://www.statistikbanken.dk/FOLK1AM

GADM, 2022 https://gadm.org/download_country_v3.html

Hielkema & Lund, 2021b additional data

https://ars-els-cdn-com.uea.idm.oclc.org/content/image/1-s2.0-S0950329321001403-mmc1.pdf

Statistica (2021), "Number of households in Denmark as of January 2021, by household size", available at https://www.statista.com/statistics/582641/households-by-household-size-in-denmark>, last accessed 2nd Mar 2022

*Watts, D. J.; Strogatz, S. H. (1998). "Collective dynamics of 'small-world' networks" (PDF). Nature. **393** (6684): 440–442. Bibcode:1998Natur.393..440W. doi:10.1038/30918. PMID 9623998.*

IPCC (2021) Climate Change 2021: The Physical Science Basis. Contribution of Working Group I to the Sixth Assessment Report of the Intergovernmental Panel on Climate Change [Masson-Delmotte, V., P. Zhai, A. Pirani, S.L. Connors, C. Péan, S. Berger, N. Caud, Y. Chen, L. Goldfarb, M.I. Gomis, M. Huang, K. Leitzell, E. Lonnoy, J.B.R. Matthews, T.K. Maycock, T. Waterfield, O. Yelekçi, R. Yu, and B. Zhou (eds.)]. Cambridge University Press. In Press.

Gardner M. (1970) The fantastic combinations of John Conway's new solitaire game "life". Scientific American, 223:120-123.

9 Appendices