UNIVERSITY OF AMSTERDAM

MASTER THESIS

Complexity in Depression: From Causal Diagram Development to Topological Sensitivity Analysis

Author:
Bas CHÂTEL

Supervision:
Prof. Claudi BOCKTING
Dr. Rick QUAX

Examiner:
Prof. Peter Sloot

A thesis submitted in fulfillment of the requirements for the degree of Master of Science

in the

Institute of Advanced Study

November 18, 2019

"Did you ever stop to think, and forget to start again?"

Winnie the Pooh

UNIVERSITY OF AMSTERDAM

Abstract

Faculty of Science
Institute of Advanced Study

Master of Science

Complexity in Depression: From Causal Diagram Development to Topological Sensitivity Analysis

by Bas CHÂTEL

Despite years of research, knowledge about the causal factors of Major Depressive Disorder (MDD) remains sparse. Treatment efficacy has not increased over at least four decades; around 50% of patients that are treated are unresponsive and prevalence rates remain stable. 93% of studies only investigate one key variable versus 7% of studies that study two or more. These findings warrant a system dynamics approach as it appears that the reductionist way of solving depression one factor at a time is not fruitful. Here we present a methodology for building and quantifying a system dynamics model (SDM). The method is then partly adopted to create a database containing structured and categorized causal factors and their respective relationships in depression. This database is acquired by performing structured interviews with experts in the field of depression. To assess the sensitivity of such expert knowledge-based systems, we also present a type of topological sensitivity analysis (TSA) by increasingly introducing combinations of mistakes in weight polarity and relationship topology within the network. We found that faulty network topologies converge towards being a random graph after around 3 mistakes in either direction and the addition of extra relations impedes fitting to data more than removing relations. These results provide a method of creating and testing SDMs and give MDD researchers a head start by providing a network topology database. We anticipate that our database can function well as a template for further MDD research that employs the network perspective. Furthermore, upon using our method for developing an SDM the sensitivity of the resulting topology can be tested using our TSA method.

Acknowledgements

All the experts in appendix A are gratefully acknowledged for participating in the interview round, providing us with an invaluable pool of information. Also, everyone working at the Institute of Advanced Study is appreciatively acknowledged for providing insightful discussions. Furthermore, I want to thank my colleagues and guidance in the research group during this project (Julia Anten, Claudi Bockting, Denny Borsboom, Sacha Epskamp, Ria Hoekstra, Rick Quax) for the interesting discussions and always be willing to engage in meaningful conversation. And lastly, my gratitude goes out to Rianne Schoon for always being supportive and providing a critical perspective at times when most needed.

Contents

List of Figures

List of Tables

List of Algorithms

List of Abbreviations

CLD	**C**ausal **D**iagram
DSM	**D**iagnostic and **S**tatistical **M**anual of Mental Disorders
GA	**G**enetic **A**lgorithm
MAD	**M**ean **A**bsolute **D**ifference
MDD	**M**ajor **D**epressive **D**isorder
ODE	**O**rdinary **D**ifferential **E**quation
SDM	**S**ystem **D**ynamics **M**odel
TSA	**T**opological **S**ensitivity **A**nalysis

Chapter 1

General introduction

Globally depression affects more than 300 million people of all ages making it the leading cause of disability worldwide (Day and Director-general, 2018). The global burden of Major Depressive disorder (MDD) was estimated in 2015 to account for 44 million disability-adjusted life-years, representing the sum of the years of life lost due to premature mortality and years lived with disability (Kassebaum et al., 2016). Furthermore, the economic burden of individuals with MDD amounts to $210.5 billion in 2010 in the United States (Greenberg et al., 2015). Even though depression research has progressed swiftly, effect sizes for MDD treatment efficacy have not increased over at least four decades; around 50% of patients that are treated are unresponsive and prevalence rates remain stable (Frodl, 2016, Ferrari et al., 2013). It seems that developing more techniques for treating MDD is redundant as this plurality of existing treatment paradigms does not result in higher remission rates. Evidence for leading theories that explain the onset and maintenance of depression is fragmented (Siegle, 2019). Cognitive therapy seems to be as effective as medication for as long as the treatment continues; moreover, upon stopping, cognitive therapy has a more enduring effect than medication (Dobson et al., 2008, Kupfer and Frank, 2001, Haman et al., 2005).

The majority of research that has been done on MDD with regards to treatment (Chukhraev et al., 2017, Smith et al., 2018, Hetlevik et al., 2019), prevention (Dias et al., 2019, Freeman, 2019, Bockting et al., 2015) and causal factors (Valles-Colomer et al., 2019, Zimmermann and Papa, 2019, Libuda et al., 2019). However, 93% of studies only investigate one key variable versus 7% of studies that study two or more (Wittenborn et al., 2016). Furthermore, the investigation of key causal variables has not been particularly fruitful. To examine the evidence for current leading theoretical models that explain MDD onset and relapse, a combined $\pm$140.000 articles were reviewed by Brouwer, 2019 on onset of MDD, Fu, 2019 on relapse of MDD, Kennis, 2019 on both onset and relapse. Out of this total of $\pm$140.000 articles considered over these three reviews, only 121 papers were found eligible for further study as they fitted the criteria of having a longitudinal prospective design and to measure a minimum of one theory-driven vulnerability factor. This approach brought forth a total of 9 factors where causal inferences could be made, distributed over the three papers, as can be seen in table 1.1. With this, only some evidence could be found for leading approaches and theories regarding MDD. For most theories, the evidence was either not decisive or was even missing. These findings pose a clear problem in that there has been very little support for actual causal inference within the body of literature of MDD. This is an indication that the search for individual causal constituents may not be the optimal way to approach the problem, but a multivariate system dynamics approach could be better suitable for finding the causal factors of MDD. The notion of expanding focus from a single variable within a research design has been postulated and explored by several researchers (Borsboom and Cramer, 2013,

TABLE 1.1: Findings of three meta-analyses analyzing prospective studies in search for causal factors regarding Major Depression. A total of 9 individual factors were found over these three studies for onset and relapse.

Brouwer et al., *	**Fu et al., ****	**Kennis et al., *****
Negative Attributional Styles	Dysfunctional beliefs	Genetics
Neuroticism	Negative emotionality	
Residual Symptoms	Dysfunctional attitudes	
	Response style	
	Neuroticism in negative emotionality	

* Relapse ** Onset *** Onset and Relapse

Herrera and Bleijenbergh, 2016, Susta and Bizik, 2011, Hassan, 2008).

Depression is seen as a disorder that is caused by the interaction between mental, biological, stress-related and societal factors that can change over time characterized by large individual differences. One of the main research challenges is to understand the causal interplay between these factors. An integrated systemic approach is the best step needed to generate insights regarding the development of innovative, more effective treatments. The need for a holistic view of MDD is becoming more and more prominent. This need can be captured by applying complexity science. A complex system can be described as an entity or phenomenon which is coherent in some recognizable way, but whose elements, iterations, and dynamics generate structures that transcend the sum of their parts and whose behavior is intrinsically difficult to model (Batty and Torrens, 2019, Churchman, 1968).

Though quite new, the idea of approaching depression from a complex systems perspective has been coined several times before. For example, Wittenborn et al., 2016 have formed quite an extensive causal map defining 13 key reinforcing feedback loops that involve nine candidate drivers of depression. This model was built upon earlier efforts to map the causal mechanisms of depression (Kendler, Gardner, and Prescott, 2002, Kendler, Gardner, and Prescott, 2006, Borsboom and Cramer, 2013). Furthermore, Cramer et al., 2016 have conducted a complex network analysis of MDD where an intra-individual symptom-based process model to investigate two phenomena. Firstly, the model aimed to examine the vulnerability hypothesis (i.e. natural connectivity within the system that contributes to a degree of vulnerability) in which high connectivity between nodes within the system leads to high activity within the system. Secondly, increasing external stress to a system that starts with a low internal stress state causes an earlier tipping point then decreasing external stress from a highly stressed system.

Although MDD has been subject to the field of complexity science, research has not always been done thoroughly as research often stops at creating a causal diagram (CLD) of the complex system (Wittenborn et al., 2016). CLDs can be viewed as a conceptual blueprint of a system and will be explained in further detail later on in this paper. To be able to properly explore dynamics, an extra step must be made to turn these conceptual models into computational simulations. These steps are, however, poorly documented, incomplete, and no comprehensive protocol has been developed to create and turn these CLDs into a quantifiable computational model enabling the exploration of system dynamics. Documenting this conversion of CLDs to a functional system dynamics model (SDM) will be one of the aims this paper will pursue. Its contribution will be used in another paper by Crielaard et al., that will be published later on.

This methodology will then partly be adopted to create a causal mapping of MDD that is at the interplay of knowledge postulated from literature and gained from expert knowledge using structured interviews. As findings from the meta-analyses suggest that there is little hard evidence for causal relations in MDD, the intuition of experts in the field is used to create a holistic overview of causal relations in MDD over three temporal stages; onset, maintenance, and relapse. This causal mapping can be used as a blueprint for modeling at a later stage in the project, outside the scope of this paper.

As the CLDs are comprised out of information gained from expert knowledge, the resulting model is made under the premise that the collective knowledge of the expert pool will cover the entire underlying network of MDD. However, there is no literature on how the performance of expert knowledge-driven models is impacted by faults in the underlying relations within the CLD. This gives rise to questions that this paper will aim to address by investigating the effects of increasingly incomplete and wrongly wired system dynamics models using a type of topological sensitivity analysis (TSA). This method leads to an understanding of the influence that topological errors in networks have and on what order they operate on. Furthermore, the addition or removal of relations between factors are predicted to affect favorably in the latter case as removing relations will create less uncertainty in the system.

Lastly, for future reference, as we have had some difficulty in planning interviews and feedback moments with experts around the world, a web-based template will be presented to facilitate projects that pursue a similar research design. Here CLDs are extracted from a collaborative online platform where experts can work on these complex causal networks with shared ownership as an incentive. Even though this exceeds the scope of this paper, these ideas will be presented in the discussion under "future work".

1.1 Aims and hypotheses

This paper will henceforth be split into four chapters. Each chapter will address one of several aims of this paper and elements posed that will be used in the following chapter. Each chapter will entail a short introduction, hypothesis, results, and discussion.

1. Methodology - Building a CLD and converting it into a system dynamics model
2. Constructing a causal diagram based on Experts and Literature
3. Quantifying reliability of System Dynamics topology
4. Feedback platform prototype: Internet-based collaboration platform for causal diagram formation and rudimentary system dynamics simulation

Aim one - Methodology - Building a CLD and converting it into a system dynamics model

As a start, a methodology is developed to address how a CLD can be procured optimally and iteratively. It serves as a guideline in constructing a CLD and subsequently convert it into an SDM. This chapter will address questions such as

- What are the shortcomings in literature for creating a CLD/SD model?
- How can a usable CLD consistently be created by iterating through available data?
- How can a conceptual CLD be turned into a functioning system dynamics model?
- How can the system dynamics model be quantified?

This chapter will lay the groundwork of systems thinking and poses methods that are adopted in addressing the next aim.

Aim two - Constructing a causal diagram based on Experts and Literature

Using the iterative nature en methods, with some deviations, given in addressing the first aim of this paper an overview of the efforts within our research group regarding gathering expert information through structured interviews will be given. The resulting CLD will be presented and our early findings regarding our systems approach regarding depression will be elaborated upon. This chapter provides a blueprint of factors and relations of MDD which can be used in future research.

Aim three - Quantifying reliability of System Dynamics topology

As the network resulting from aim two is based largely on expert knowledge, this chapter will be focusing on developing a method to assess whether expert knowledge-driven networks perform better than random networks. It will then move on to evaluating the sensitivity of said networks with regards to two types of mistakes (reversing polarity, rewiring relations between factors) within the topology of the network (i.e. a causal diagram). Lastly, an assessment is made to explore the effect of missing links altogether to be able to give a coherent recommendation as to whether to overshoot or undershoot the number of possible relations within a system and build intuition around the sensitivity of networks.

Aim four - Feedback platform prototype: Internet based collaboration platform for causal diagram formation and rudimentary system dynamics simulation

This aim will be discussed in the general discussion as a future research suggestion. A proposition is made for software that could solve certain problematic aspects that were raised in aim two during the interview process. This software would be made general as to be compatible with all research that follow the same methodology, it would automate part of the process of incrementally creating and quantifying a CLD and let experts that cannot be approached through group modeling sessions still provide their insights and promote cooperation between experts.

Chapter 2

Methodology - Building a CLD and converting it into a system dynamics model

2.1 Introduction

Systems thinking is used in a plethora of disciplines (e.g. Economics and Management (Long and Chen, 2019, Yang et al., 2019), public health (Luke and Stamatakis, 2012), psychology (such as with depression (Wittenborn et al., 2016), geography education (Cox, Steegen, and Elen, 2018), etc.). However, although creating CLDs and SDMs seems to be becoming increasingly popular, a clear protocol for this type of research is not forthcoming. Therefore, this chapter aims to provide a clear step-by-step guide for creating an executable SD model through the creation of a CLD. Firstly, we will begin by specifying elements that are currently not represented in literature but, to our understanding, should be included when creating a CLD or SDM. Secondly, we provide an algorithmic approach to building a CLD up from a list of causal factors (henceforth referred to as nodes). Thirdly, the contraction phase is discussed to move from a CLD to an SDM through labeling and rewriting the CLD. And lastly, a method for turning the system into equations and quantifying the SDM will be provided.

2.2 What are the shortcomings in literature for creating a CLD/SD model?

In our experience, current literature on how to actually form these CLDs are lacking important ingredients that allows for conversion to a (validated) executable SD (Levine, 2000, Lai and Wahba, 2001, Binder et al., 2004, Sterman, 2000). This mostly occurs when the goal is stopping at the CLD level. Some properties that are, in our understanding, currently missing in current methods and why these methods should be included are the following:

Ownership and consensus When creating a causal model, it is important to keep in mind who created the model. The result will not necessarily be a reflection of the truth, but rather a reflection of the collective knowledge of the group that created the model. In this sense, it is imperative to document ownership of the model. This way, it is clear from which experts the model is a reflection, but also where consensus comes from within the model. Recording consensus can also have computational implications as experts can disagree about certain nodes or edges, which would raise uncertainty within the model and might even lead to an ensemble of different models. This can bring about some challenges as performing analyses on an ensemble of models is more difficult than on a single model.

Quantifiable variables Some variables are inherently not quantifiable. This can have several reasons, for example, a variable is not defined precisely enough (e.g. environment) or there are simply no good, unambiguous ways to make measurements (e.g. measuring emotions (Mauss and Robinson, 2009)). It is significant to keep these unquantifiable variables into account as they will pose problems later on, and will need to be approximated by proxy variables as will be discussed later on in this paper.

Kind of relation In System Dynamics, it is often the case that relation within a system has an unknown functional form. This can be linear, in which case it can easily be computed, but it can also be non-linear which can bring a lot of uncertainties to the table (Coyle, 2000). For the sake of completeness, it is useful to try and describe these relational functions (e.g. quadratic, sigmoid, etc.) as they are needed for creating a dynamic model. This will be further elaborated upon later in this paper.

Justification of nodes and edges using literature During the process of creating a CLD, or further down the line an SD, it is often overlooked whether the nodes and edges within the system are supported by the literature. Though many CLDs do this implicitly as their models are informed by existing literature such as experimental evidence (Wittenborn et al., 2016, Vins et al., 2015, Tucker Lima et al., 2017), it is wise to make these sources explicit for each element in the model by providing their respective sources. This also applies when making use of expert knowledge. Each node or edge given by an expert explicitly requires a statement containing a reference to literature or lack thereof, and corresponding units accompanied by a statement of what the link entails (Burns and Musa, 2001). This is, as it is with recording consensus within the model, important for recording uncertainty. Higher degrees of uncertainty within a model will bring about difficulties for later fitting to data (Chatfield, 1995).

Non-edges Current methods focus on which causal links can be added. However, if a certain link is not added, it is typically unrecorded whether this is because the link is 'unknown' (or not discussed) or that it is 'known to be absent'. As such, it is implicitly assumed that unrecorded edges do not exist in reality, in contrast to the edge not being empirically examined in the first place. This is nevertheless an important distinction for the model space as this brings about a difference between Monte Carlo sampling from models with and without these links, versus only sampling models without them. Here the latter leads to underestimated uncertainty in the model output, i.e. when a link is 'unknown' you could sample from models with and without this particular link. If you don't record whether or not a link is 'unknown' or 'known to be absent', you miss a chance to reduce parameter space which would lead to less uncertainty within the system.

2.3 Algorithmic approach to building a CLD

An algorithmic approach to building a CLD will be provided in this section. At this point, we assume a full set of nodes through which can be iterated through and, as such, expand into a network with a coherent set of edges. The algorithm is written such that it performs in $O(N^2)$ time. However, in practice, this will turn out to be faster as some edges, or absence of edges, are known in advance thus speeding up the process (e.g. 'Variable A does not influence anything', 'Variable B is only a mediator for C and D but nothing else', etc.). As these relations are already processed, the algorithm can start as follows:

- Iterate through the 'unprocessed' set of pairs of variables:
  1. Establish whether there is a direct and complementary causal link (i.e., other than what is already encoded in the model thus far).
  2. If deemed so:
     (a) Ask what the (hidden/omitted) intermediary processes are (e.g. sleep could mediate between weight and stress) such that each intermediate step makes sense (is it considered a 'direct' effect). Store this with the link. The intermediary processes should not already be variables within the model as defined in the previous section (if so, start over with that pair of variables, leave the current pair as 'unprocessed').
     (b) In line with this, the causalities mapping should ensure that the causal pathways within the system dynamics model represent mutually exclusive mechanisms.
     (c) Ask based on which knowledge (literature, experiments) this link is based. Store this with the link.
     (d) Ask whether the type of relation between the two nodes is known (linear, quadratic, sigmoid, etc.). This is a difficult question but will provide useful upon implementing dynamics. If an expert is unsure about the exact mathematical function (as people usually don't think in these ways), try to draw the relationship (i.e. what happens to B if A increases?). If the type of relationship is known, store it and its information source. Also, when possible, try to validate this relationship through the use of data (regression, curve-fitting, etc.).
     (e) Verify that this causal link is 'direct'. Its intermediary processes should not already be encoded in the model (e.g. sleep should not already be a variable in the model). If so, add it, otherwise, change to the new pair of variables in question (weight and sleep) and start over the iteration.
     (f) If the link is added, verify that it does not form an intermediary process of an already existing causal link (e.g. weight $\rightarrow$ eating, mediated by 'stress'). If so, break up the existing link and add the remainder link (stress $\rightarrow$ eating). Transfer the relevant literature/experiment references and intermediary (hidden) variables that were previously stored with this existing link, if any.
     (g) Verify that this edge does not violate any existing non-edge (See step 3).
  3. If deemed not, then:
     (a) Store the fact that this is a non-edge, i.e., the claim that no causal effect is present between these variables. Establish the reasoning behind this decision (literature, experiments) and store this with the non-edge.

(b) Verify that no existing edges form a connected causal pathway between these variables, which would violate the claim.

(c) If deemed uncertain, then change the ’unprocessed’ status to ’unknown’. This is important later for quantifying the model space and hence the uncertainty of its predictions.

This process provides a clear overview of nodes, intermediary processes, edges, edge-types, non-edges, and unknowns. It provides a body of information which can, with some steps, be transformed in a fully executable system dynamics model. Yet, it is wise to first follow some validation steps which will be discussed in the next section.

2.4 Contraction phase and annotation into system dynamics

To move from a (validated) conceptual CLD into a functioning system dynamics model, the use of data is highly advised. Preceding this process, it is important to remain skeptical about what is actually being measured in any quantitative data that might be available. To properly be able to use this data to fit the CLD and obtain quantitative models, the variables within the data must be able to represent variables contained by the CLD. Here, however, some interaction with the experts still is needed as we do not initially create the CLD, go to the data and then never go back to the experts. Interacting with the experts needs to remain open to find out whether certain data can be used to represent concepts within the CLD. It is important to not divulge in advance what data is available as this risks an "inside the box" type of thinking. An expert thought-process might, as such, be unconsciously limited by the presented data. This section will elaborate on the choices one can make regarding the availability of data and the fitting thereof.

2.4.1 Proxies

Whenever there is a variable that is represented within the CLD but not in the data, the use of proxies can be considered. For instance, if ‘Social Economic Status’ is not measured but ‘Income’ and ‘Level of Education’ are measured we can deliberate whether the latter two could be used as a proxy for the former, or that the expert audience is willing to make this assumption. If so, change it and review the incoming and outgoing links.

One can also cluster variables in categories (i.e. ’Social Economic Status’ is a combination of, among others, ’Income’, ’Education’, and ’Occupation’). If data is only available for a subset of the variables within these categories, could the data from the rest within the cluster accurately represent that of the missing subset? This way we can replace variables that are missing by ones that we do have.

2.4.2 Unobserved variables

An unobserved variable can be added, but beware that their parameter values cannot be perfectly identified. This creates more computing time later for quantifying uncertainty. In the case of a feedback loop, there can only be one free parameter that should be fitted in each feedback loop, where a loop starts and ends in the same stock. Otherwise, we get expressions in which we define something by multiplying two unknowns, which means that we do not know the value of each unknown separately, we only know their multiplication.

2.4.3 Mediators

If a variable is a mediator on a single pathway (or a fork), then it could be removed and the incoming and outgoing links could be aggregated. For instance, if 'sleep' is only mediating in 'weight $\rightarrow$ sleep $\rightarrow$ stress' (Tuomilehto et al., 2009), then create a new edge 'weight $\rightarrow$ stress' and store with this edge the fact that 'sleep' is considered an intermediary process. This notion of not explicitly modeling of the intermediary process can be of use when we don't have data on the intermediary process itself while also cutting down in the dimensionality of the 'solution space'. If a variable is, however, a mediator on multiple pathways then it is advisable to keep the variable in the model. For instance, if in addition to the previous point 'sleep' would also influence 'exercise' through 'exercise $\rightarrow$ sleep $\rightarrow$ exercise' (so 'sleep' forms a crossroad) then 'sleep' creates a correlation between 'stress' and 'exercise'. Removing 'sleep' would disconnect these two pathways, which would (falsely) encode independence and potentially lead to false results.

In the case of a feedback loop, there can only be one free parameter that should be fitted in each feedback loop, where a loop starts and ends in the same stock. Otherwise, we get expressions in which we define something by multiplying two unknowns, which means that we do not know the value of each unknown separately, we only know their multiplication.

2.4.4 Shortage of data

During the inventory of nodes and matching data, we can also find out that there is simply a shortage of data. This may indicate that it is a possible starting point for new research.

2.4.5 Sensitivity analysis

In the occurrence of not having any data to work with, a sensitivity analysis can be performed. This can be used to find out whether certain nodes are at all important/influential in your model. To achieve this, one must first know the units and interactions of the nodes and edges.

2.4.6 Fitting data

Some relationships may be fitted by linear regression if the true underlying relationship is unknown. This will be discussed later on in this paper in greater detail.

2.5 Labelling the causal loop diagram

Now that we have made an inventory of the possibilities regarding the available data concerning the model, we can convert the CLD to a system dynamics model (SDM). For each variable, establish whether it is a constant, a stock, a flow, or an auxiliary. To do this, the temporal scale of the model should be considered. All variables that act and change on a larger temporal scale can be considered constants for this SDM. All variables that do change within the model's temporal scale will therefore not be constants. Quantities or variables that change faster than a model time step, can be considered to update instantly.

A stock is a quantity that changes on the temporal scale of the model, but for which the rate of change, instead of the quantity directly, is influenced by other variables, including possibly itself. For instance, 'body fat' could be a stock variable in a model that has a temporal scale of weeks or months. The most important quantity, the outcome variable of the model, is normally a stock variable. A flow is connected to a stock variable and determines how the stock changes over time: it can be perceived as a rate. An auxiliary can be either an exact re-definition of existing stocks/auxiliaries or it can be a quantity that changes on a significantly smaller temporal scale than the scale of the model. For instance, 'BMI' would be a re-definition of 'weight' and 'height' through a specific formula. It serves only to simplify representation, especially when e.g. 'BMI' feeds into multiple other variables. The influence from the variables 'depression' and 'physical activity' on the variable 'sleep' in a model for weight gain, on the other hand, is not a re-definition but the influence (within days) acts on a smaller temporal scale than the model (weeks/months). The 'sleep' variable can, therefore, be considered an auxiliary in this model. More information about this stock- and- flow system dynamics models can be found at for example (Sterman 2001).

2.6 Equations

The result of labeling a causal loop diagram would be a fully annotated SDM. In these models, the underlying relations can be represented by a set of equations that take the variable values as input and generate the corresponding variable values for the next time point as output. These equations simulate the behavior of the complex system according to the sequence as outlined in the SDM (following the curved arrows). Accordingly, for each causal link, the functional form of the relationship between the two respective variables needs to be estimated, i.e. whether the relationship between the two variables is linear or takes any other functional shape (this will be elaborated on further in the section 3d). During this process, dimensional consistency between the units of the left- and right-hand side of each equation should always be preserved.

An example of such equations representing a system can be a system of ordinary differential equations (ODE). As a very simplistic conceptual example, we can take the following system displayed in figure 2.1. On the left, there is a system of ODE's where the change in sadness (S) is defined as the sum of guilt (G) times parameter alpha minus avoidance (A) times parameter beta. The change in Avoidance is defined as gamma times Sadness and guilt is epsilon times Avoidance. These greek letters are the parameters within the system and these can be viewed as the weights of the relations in the model. If we were to increase beta to a very high degree, the great inhibitory strength would cause the system to converge towards zero, as sadness will be inhibited greatly. However, if we were to have a low beta that is close to zero and increase alpha, the system will grow to infinity as sadness is not inhibited and avoidance and guilt will keep increasing. On the right the system is depicted graphically consisting of three stocks; Sadness, Avoidance, and guilt. Four flows of which one inhibitory between Avoidance and Sadness. In this example, we will assume that sadness leads to an avoidance of emotions and this avoidance will consequently inhibit feelings of sadness in a short period. Avoidance, however, in the long run, would also increase feelings of guilt which, in turn, has an excitatory effect on sadness.

$$\frac{dS}{dt} = \alpha G - \beta A$$
$$\frac{dA}{dt} = \eta S$$
$$\frac{dG}{dt} = \epsilon A$$

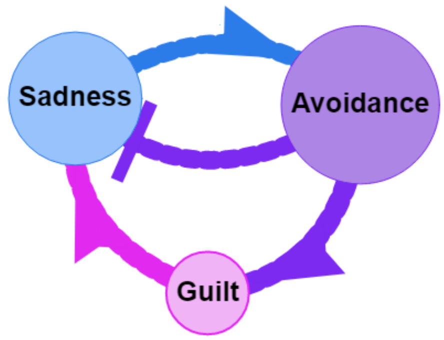


FIGURE 2.1: Small network in both a mathematical representation (left) and a graphical representation (right). Sadness (S) has an excitatory effect on Avoidance (A) which in turn inhibits S. A then excites Guilt (G) which in turn excites A on a slower time scale.

2.7 Quantification of a system dynamics model

Once a causal loop diagram has been contracted and labeled, the optimal parameter values are to be found. This can be done utilizing an optimization algorithm that aims to approximate the data by cleverly sampling/adjusting parameter values in the system. How well these parameters fit can be quantified by a cost function (also known as an objective function). The cost function is a measure of fit in the model in terms of its ability to estimate the relationship between the parameters. An example that comes to mind can be distance measures like Mean Absolute Difference (MAD) between the predicted value (output of the model) and the true value (from data). By constantly slightly adjusting the parameters within the system the fitness landscape is revealed (e.g. figure 2.3. When using MAD the distance between the predicted value and true value should be minimized, so a minimum sought within this fitness landscape. Cost functions are different in each individual problem, so it is essential to first clarify what it is exactly what we are trying to optimize and create a cost function accordingly. For further information about cost functions (Diewert, 2008) can be of use.

As earlier stated, optimization algorithms aim to approximate given data by tuning parameters based on the cost function. To give a quick and easy intuition of how this might be done we will shortly elaborate on one of the easiest optimization algorithms: The Hill Climber. In pseudocode algorithm 1 the hill climber in combination with the MAD cost function is given. Here we see that we initialize the system with a random set of parameters and then run the simulation and take those as the starting points. From there on with each iteration, the parameter values are slightly perturbed to see if that change changes the score for the better. If the new score is better, we take that as the best score

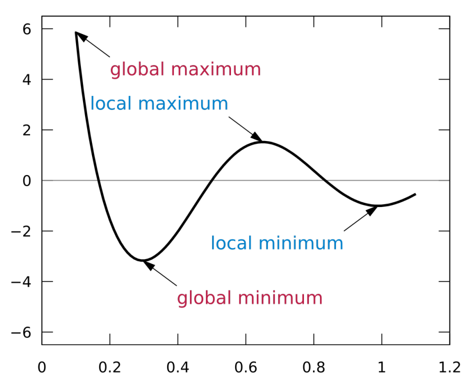


FIGURE 2.2: Example of a function displaying both a local and global minimum and maximum.

Algorithm 1 Example of Hill Climber pseudocode

```
Result: Convergence towards (local) optimum
Initialize simulation and costFunction
  finalParameterSet = Random set of parameter values
  Run simulation
  finalScore = Calculate output of simulation through costFunction
  while stopping criterion has not been met do
    tmpParameterSet = Slightly perturbed finalParameterSet
    Run simulation
    score = Calculate output of simulation through costFunction
    if score < finalScore then
      finalScore = score
      finalParameterSet = tmpParameterSet
    else
      pass;
    end
end
```

and in it's not, we try another perturbation and evaluate whether that score is better. Even though this method is very simple and has a good chance to converge towards a local minimum, it does illustrate nicely how the algorithm can incrementally converge towards a solution. Other optimization algorithms take this same idea and apply it in a more advanced manner.

It should also be noted that as an optimization algorithm converges towards a solution, it can happen that a parameter is being set to zero in the process. This can be an indication that the importance of the corresponding relationship between nodes is non-existent. In the occasion of ensemble modeling where there is disagreement in some edge and the optimization algorithm converges to zero in that same parameter, it could indicate the edge can be left out. This could, however, also indicate an inconsistency in the data. In these cases, it would be best to return to the experts.

In table 2.1 more examples of optimization algorithms are provided with some benefits, disadvantages, and literature for further reference for a few optimization algorithms. As we do not aim to be exhaustive, these are mere suggestions as there is a plurality of other possible methods out there.

Now that intuition has been developed for optimizing parameter values in a system, it is necessary to look at the functional form of the relations (i.e., linear,

TABLE 2.1: Examples of optimization algorithms with some benefits and downfalls. Also some literature is provided for each type.

	Pro's	Con's	Literature
Hillclimber	Very simple to make	Can get stuck in local optimum	(Hernando, Mendiburu, and Lozano, 2018)
Simulated Annealing	Can sometimes climb out of local optimum to find global optimum	Somewhat harder as there is a heat function that needs to be optimized	(Kirkpatrick, Gelatt, and Vecchi, 1983)
Metropolis-Hastings	Obtains a sequence of random samples from a probability distribution from which direct sampling is difficult	Works best with high dimensionality. For lower dimensionality, other methods can be more useful	(Kirkpatrick, Gelatt, and Vecchi, 1983, Chib and Greenberg, 1995)
Evolutionary Algorithm	Tends to converge faster by creating a population of solutions that search the parameter space.	Can have a lot of parameters in the algorithm itself that needs to be optimized (mutation chance, cross-over, immigration, etc.).	(Coello, Lamont, and Van Veldhuizen, 2007)

sigmoidal, quadratic, etc.). One of the simplest methods can be applied by linear regression which can be used to see if there are linear dependencies in the data. This method essentially tries to fit a straight line on the data and calculates how good the data can be predicted by this line. An example of how this can be done (there are several other ways) is by use of the least-squares approach. This simply computes the distance between each data point and the linearly fitted line and minimizes this distance iteratively.

There is, however, a downfall to linear regression as all individuals go to the same line after a single time step. If you model individuals (instead of at population level) then you must add a "personal bias". This personal bias entails moving the line to each individual in the model, thus eliminating that the whole population suddenly converges to a single line.

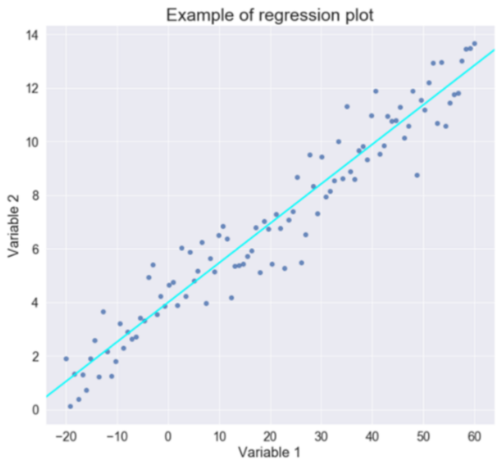


FIGURE 2.3: Example of a function on which linear regression has been performed. A straight line is fitted such that the distance between the line and data points are minimized.

In terms of the model, this means that for each causal link, decide on its functional shape and any free parameters it may contain if any. This is a difficult step. For predictions of a small number of time steps (for instance 6 months, if the model time step is 1 month) you may consider applying the 'locally linear approximation'. This means that the progression curves (e.g. of weight) are assumed to progress roughly linear in this short period. This makes the model deterministic: at this point, there are no random effects in the model, as we assume linearity. This does not matter so much, because the causal relationships between the variables do not change. This means that each causal link becomes a linear dynamic. For instance, if 'energy intake' increases 'weight' then the increase can be modeled as "a * [energy intake]" where 'a' is a constant (kg/kcal). If this constant is known from literature then it is a fixed parameter; otherwise, it is a free parameter and must be estimated in a later phase. Try to minimize the number of free parameters.

If the nature of relations between two nodes in the model is unknown and it does not immediately make sense to use linear approximation then elastic net regression can be applied. This technique has been applied by Schmelzer, Dwight, and Cinnella, 2019 where they constructed a large library of nonlinear candidate functions to regress data. From this library, the most relevant candidates are identified so that the functional form of a relationship can be identified. Elastic net regression is good at dealing with situations when there are correlations between parameters, but Computationally costly and has a probability of overfitting as the estimator is flexible. More information about this technique can also be found at (Zou and Hastie 2005).

Chapter 3

Constructing a causal diagram based on experts and literature

3.1 Introduction

According to estimates from the World Health Organization, more than 300 million people are living with depression. The number of depressed people increased by more than 18% between 2005 and 2015 (Day and Director-general, 2018). Even though it is the leading cause of ill health and disability worldwide, the collective knowledge about the causal factors (nodes) is surprisingly modest. As a threefold of meta-analyses suggest by Brouwer, 2019, Fu, 2019, and Kennis, 2019 current literature concerning true causal knowledge regarding Major Depressive Disorder (MDD) is lacking. Out of a total of ±140.000 articles only 121 papers were found eligible for further study as they fitted the criteria of having a longitudinal prospective design and to measure a minimum of one theory-driven vulnerability node. This approach brought forth a total of 9 nodes distributed over the three papers, as can be seen in table 3.1.

Furthermore, Wittenborn et al., 2016 suggests that 93% of studies only investigate one key variable versus 7% of studies that study two or more. This indicates that there is a lack of research from a systems perspective within the field of MDD. There are, however, several papers that already adopted this systems perspective and have had some interesting results. For example, Wittenborn et al., 2016 has already formed a causal through the use of literature. One problem is though, since literature is not sufficient enough (see meta-analyses), a causal diagram based purely on literature does not seem a complete option. That is why we propose to create a causal node model based not only on literature but also based on the intuition of

TABLE 3.1: Findings of three meta-analysis analyzing prospective studies in search for causal nodes regarding Major Depression. A total of 9 individual nodes were found over these three studies for onset and relapse.

Brouwer et al., *	**Fu et al., ****	**Kennis et al., *****
Negative Attributional Styles	Dysfunctional beliefs	Genetics
Neuroticism	Negative emotionality	
Residual Symptoms	Dysfunctional attitudes	
	Response style	
	Neuroticism in negative emotionality	
* Relapse	** Onset	*** Onset and Relapse

some of the top experts in the field of MDD. These experts should come from a diverse pool to sample expert knowledge from all types of disciplines.

One of the main purposes of this paper is to map out the causal nodes and to uncover its node-to-node relationships through the use of structured interviews with experts in the field. Furthermore, a temporal categorization will be attempted to prepare for the system dynamics model. To keep within the conventions of the field these nodes will henceforth be referred to as nodes, and the node-to-node relations will be referred to as edges. This process of mapping a causal diagram of MDD through structured interviews will also uncover answers to the following question; After assembling the node list and categorizing the timescales, how do the different timescales compare to one another in terms of the number of nodes and edges? This resulting multi-scale causal node model could serve as a template for MDD research in future endeavors.

Each expert will be asked to give a personal definition of MDD as upon discussing mental illness, one cannot dive into meaningful conversation without touching the subject of definitions and categorization. Although MDD is, as described in the introduction, of great significance both economically and sociologically, no clear and single definition has hitherto been posed to describe and/or explain the phenomenon fully and with full consensus in the scientific community Berrios, 1999. Currently, the Diagnostic and Statistical Manual of Mental Disorders (DSM-5) produced by the American Psychiatric Association (Association, 2013) is one of the most widely used categorization systems. The DSM requires five or more symptoms to be present for 2 weeks. Each symptom can either be present or not. One of the symptoms should, at least, be either a depressed mood or anhedonia. The secondary symptoms entail: appetite or weight changes, sleep difficulties (insomnia or hypersomnia), psychomotor agitation or retardation, fatigue or loss of energy, diminished ability to think or concentrate, feelings of worthlessness or excessive guilt and lastly suicidal ideation. As the DSM is the most used standard, experts will be asked to relate their vision of what MDD entails, to the definition posed by the DSM.

3.2 Methods

3.2.1 Timeline

Due to the extensive nature of this project, figure 3.1 presents a timeline regarding the time it took for several of this chapter to occur. As the main ideas were already clear, we started by reading literature and iteratively creating a list of potential expert candidates which would fill some aspect of MDD related knowledge. During this process the interview template was written (see appendix C for full transcript) that would capture the information needed to develop a causal network of MDD. After little more than a month, the list of potential candidates grew, and additional information needed to be gathered regarding their contact information, the disciplines and specializations they would represent, etc. This eventually led to the fulfillment of a list of potential experts that were important in the field of MDD and a period of planning and performing interviews could start. After each interview, the information was extracted within a week (details will be discussed later on), and further processing was done thereafter. Finally, after extracting all data, group modeling sessions were performed within the research group to perform data cleanup, categorizations and relational inference between nodes.

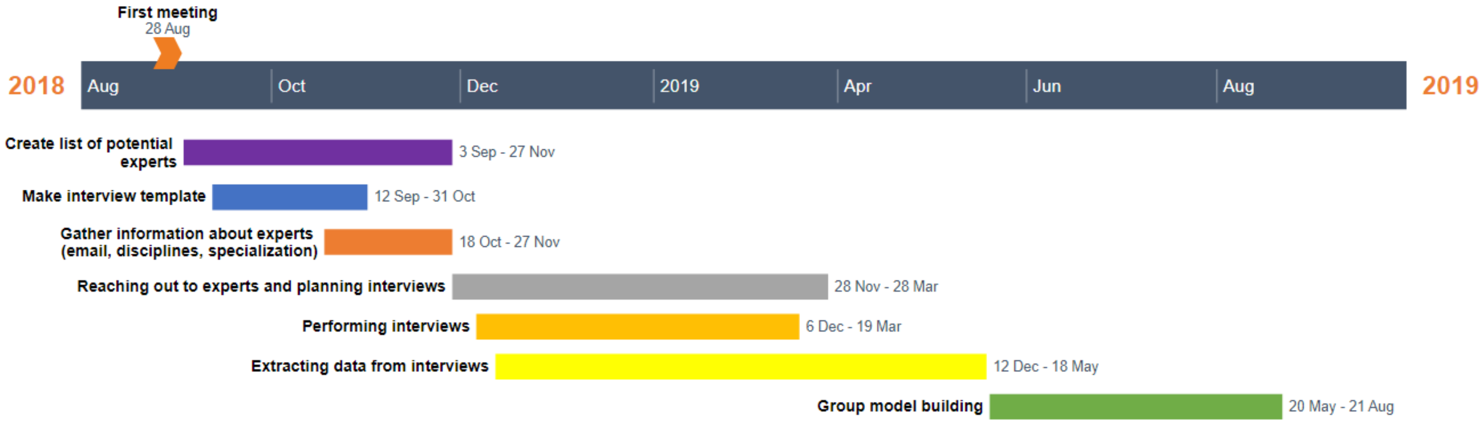


FIGURE 3.1: Timeline of the project.

3.2.2 Recruiting experts

The major objective of this study was to create an overview of causal nodes and edges of MDD. This has been done using structured interviews of 20 experts with various backgrounds, yet all relevant to MDD. For a full list of the expert list, refer to appendix A. Decisions regarding which experts to invite were made with the help of an interactive visualization made in JavaScript with the D3.js library (see figure 3.2 for reference). The visualization portrayed experts in distinct colors where they could represent multiple positions in domains (researcher, experience, policy maker) and different disciplines (e.g. biology, board member in the house of representative, board member of mental health institute, economics, lived experience, psychiatry, psycho-biology, psychology and sociology, philosophy, etc.) which entailed multiple specializations (e.g. cross-cultural mental health, epidemiology, computational modeling, etc.). This was done to assure the greatest diversity in expert knowledge that was possible without having to interview too many experts.

3.2.3 Interviews

Interviews were conducted over a span of three months and were performed using Skype, over the phone or in person. All interviews were recorded both audio-visually and with an audio-only backup. The interviews were continually performed by two native Dutch-speaking students and took around an hour. These students had a moderate knowledge of MDD. English was used as the lingua franca where native language didn't match up (13 out of 20 interviews). Interviews not performed in English were performed in Dutch (7 out of 20 interviews). During every interview, there was one interviewer and at least one secretary making notes and could step in when needed.

The main goal of the interviews was to gather nodes which the experts regarded as important to include in our causal node model. These nodes were divided into four categories, namely general, onset, maintenance, and relapse. Furthermore, they were enquired with regard to their personal definition of MDD, temporal scales with regard to nodes, treatment nodes, interesting articles, personal questions regarding their own specialization with depression, and lastly to specify phenomena that the simulation model should be able to explain (e.g. spontaneous remission, relapse, individual differences, etc.). For the complete script of the interview see appendix

C. Each expert was interviewed using the same script, although deviations from the script were allowed when the conversation took an interesting/useful direction.

3.2.4 Workflow

These interviews yield an additional database outside the scope of literature. This new database required some cleaning up, stripping excess information down to its core resulting in what you could call a "fat" excel file. The data would need to be compressed into something that is machine-readable to perform computations and simulations on which could be called a "slim" excel file. From this, an iterative process within the research group started by engaging in group model building sessions. Then Python was used to read and process data and yield a comprehensive image of the network. In future advancements of this study, this image will then be used to gain feedback from experts regarding the nodes and edges, and the iterative process will start again. This feedback moment is, however, outside of the scope of this thesis. This workflow is captured in figure 3.3.

3.2.5 Data filtering and compression

Interview data were scored independently by two members of the research team in an excel file. These two members adhered to the same method in scoring interviews. nodes, phenomena, and other observations were extracted and shortly explained. The resulting excel file can be viewed as a "fat" excel file as all details and nuances are recorded in it. The output of these two individuals was then compared and, upon discrepancy, corrected by a third person who is trained in researching, and working with, depression. However, disagreements were settled after careful deliberation between the two scorers and no third person was needed. When the results

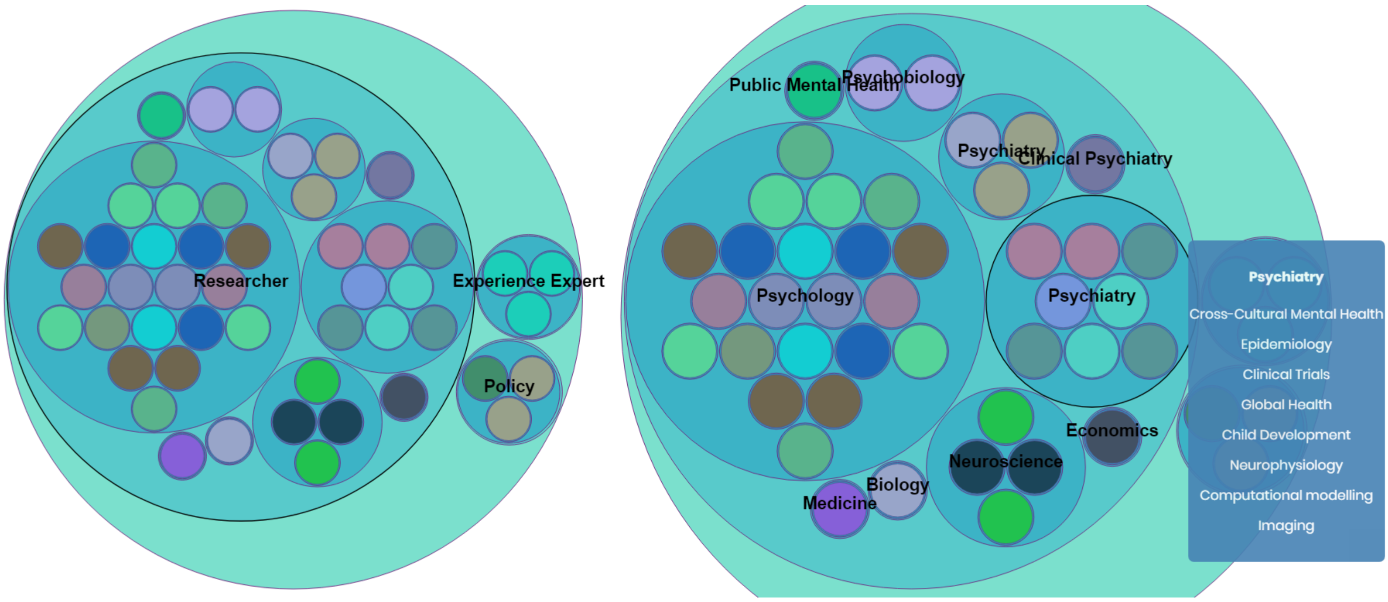


FIGURE 3.2: Interactive visualization made for inventory of different domains in which experts were represented. This has helped in making choices regarding the expansion of the expert pool in targeted disciplines. Unique colors represent the experts who can also portray other disciplines/specializations.

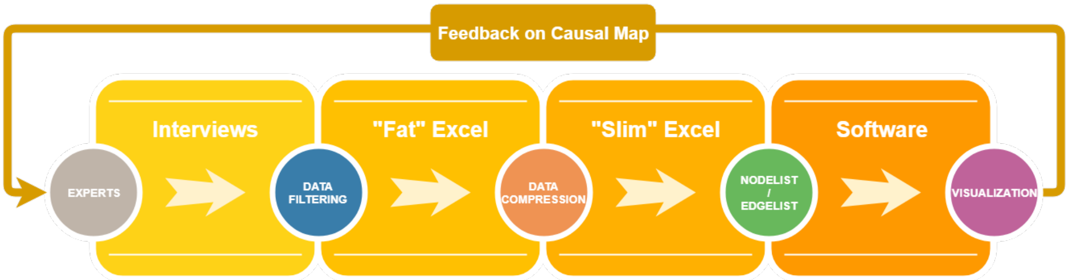


FIGURE 3.3: Data work flow regarding conducting and processing Expert Interviews.

were satisfactory similar, the output was accepted for further compression and eventually added to the database in the "slim" formatted CSV file. However, if there were unclear instances, they were discussed in the upcoming research group meeting (see figure 3.4). This process eventually led to a nodelist containing 182 nodes. 60 in the general category, 36 in onset, 44 in maintenance and 42 in relapse. Overlap and doubles were removed after combination, which led to a remainder of 90 nodes. Afterward, missing DSM-V symptoms were added to the list, thus adding another 8 nodes to the list. Lastly, a categorization was made into timescales. This led to dividing the list of 88 factors over 6 different timescales.

3.2.6 Group modelling

The iterative process of categorizing the nodes in timescales (i.e. immutable, 18 years, year, month, week, day, within day), temporal importance (i.e. onset, maintenance and relapse), spatial importance (i.e. macro, meso and micro), and external/internal (i.e. does the node influence the network while it is not influenced itself by the network, or is it also influenced by the network itself) led to a framework which promoted thinking within category networks and network relations that trickle down from the slowly changeable (or even immutable) levels to the faster levels. This was done by members of the research group on a semi-weekly basis.

These timescales required that a node had to be categorized on the fastest time scale in which it could change meaningfully. During this process the exact timespans were not defined, nodes would only be categorized relative to one another as

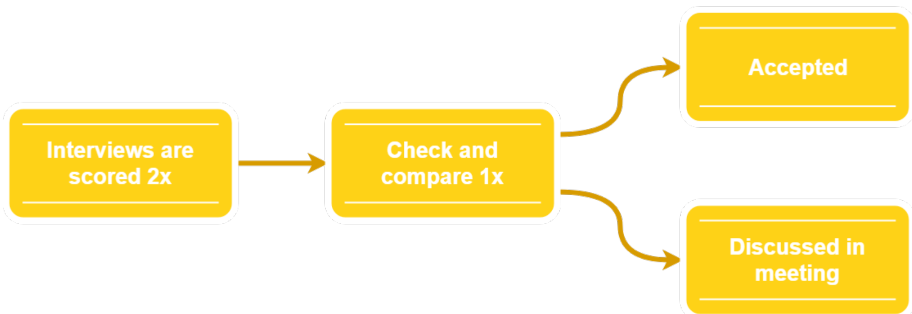


FIGURE 3.4: Data filtering: Each interview is scored by two individuals who's output are compared by a third person that can either accept or reject their output by a measure of how identical the output is. Upon rejection, the case will be brought up for discussion with the group.

a certain node could be faster or slower than its counterparts. From this process, a distinction emerged naturally and these time-scales followed instinctively. Namely, *immutable* nodes are nodes that cannot change over time, nodes that at most change over a period of around *18 years, one year, Month to a week, day and within a single day*. Furthermore, nodes were divided by their spatial scale (micro (i.e. within a person), meso (i.e. a person's direct surroundings) and macro (i.e. at a higher level like government regulation)), temporal importance (i.e. which part of the clinical picture is the node important?) as discussed earlier, and whether nodes receive feedback from the system itself (internal) or the node influences the system from outside (external). Also, a distinction was made for external nodes that are present from early childhood (initial external).

After the categorization of the nodes was done, our research group firstly continued to draw edges from the slowest time-scale to each adjacent faster time-scale using the method described in section 2.3 and subsequently the edges within a time-scale itself was drawn out.

A choice was made to only go from the slower time-scale to the faster time-scale as the number of possible connections was otherwise top big. This way there were only 5 inter-time-scale levels and 6 within time-scale levels to consider, otherwise, there would be 21 possible connections between time-scales.

3.3 Results

3.3.1 Definitions

Out of a total of 20 experts, 14 experts gave a clear description of their definition and its relation to the DSM-5 categorization of MDD. 6 experts were fine with the definition given by the DSM albeit with some downsides. All experts agreed that the DSM-5 was mostly a communicational tool, i.e. a framework in which to navigate whilst communicating about this phenomenon called Major Depressive Disorder. Also, all agreed that the true clinical picture was not properly represented through the DSM-5, but was only an approximation. Through the large heterogeneity in both the development and maintenance of MDD, one singular definition would almost always be an oversimplification. But, as is with the DSM, this would still be a useful tool to stimulate fruitful discussion and facilitates research. Critical notes were given regarding the duration of symptoms before they are eligible to count as a symptom for diagnosis. Furthermore, some experts suggested that the notion of 5 symptoms could be too restrictive or too free in individual cases. The DSM-5 is therefore seen as a catch-all solution, that could break down in individual cases. Also, the DSM is founded to increase between-psychologists diagnosis reliability and it does a fairly good job with that, however, in terms of validity there is still a lot of work to be done. In terms of subtypes of MDD there, however, was some controversy regarding whether there are or aren't any subtypes. One expert opted for a threefold of natural phenotypes:

1. **Sickness behavior:** Adaptive phenotype which might not even get a diagnosis in the DSM, might be a sickness.

2. **Starving depression:** Loss of interest in sex, humor, social companionship.

3. **Melancholic subtype:** Characterized by anhedonia, sleeplessness, eating less.

TABLE 3.2: Number of nodes on each timescale.

	Immutable	~18 Years	~Year	~Month/ week	~Day	Within a Day
# of nodes	5	20	4	24	7	26

In terms of personal definitions given by our experts, they were all rather similar. MDD is mostly characterized, not by an illness entity, but by a cluster of symptoms that as a whole impair quality of life, functioning in daily life and mostly come with comorbidity of other problems. It can be described as an absence of vitality, interest, and energy. One expert described mood, anhedonia and physical phenomena as the holy trinity of MDD. Another expert reported that he did not use algorithmic approaches for defining depression (e.g. the DSM-5) but rather looked at a combination of information concerning symptom severity, number of symptoms, type of symptoms and functioning of the individual.

3.3.2 Node list

Appendix B exhibits a node list distilled from interviews with our experts. This is a joint list consisting of 90 nodes gotten from questions regarding depression in general and the onset, maintenance, and relapse of depression. Figure F.6 displays the categorization of this node list distilled from a combination of our obtained expert knowledge and own group model building sessions. Here the main categorization consists of a division in timescales. For quick reference, see table **??** for timescales and the corresponding number of nodes within said timescale.

Figure 4.3 depicts the resulting network of the categorization process and is made using a slightly modified version of the Pymnet Python library which is based on a paper by Kivelä et al., 2014. The figure contains 1415 edges over these 90 nodes. All edges are unidirectional and between-timescale edges flow from slowest to fastest level (e.g. the immutable level has no incoming edges, but only outgoing towards the ~18 years level). As figure 4.3 is somewhat chaotic, depicting the intricacy and complexity of the network is difficult so a broken-down version is visualized in figure 3.7 where combinations of adjacent temporal levels are given. The choice to handle the drawing of edges in a slow-to-fast fashion was based on the research of Lunansky, Borkulo, and Borsboom, 2019 as they have already made progress in simulating models from multiple time-scales. An interesting finding is that the timescales with the largest amount of nodes and highest density are the scales that fluctuate within a single day, in about a week to a month, and those that change over approximately two decades. Furthermore, all nodes that operate within the span of a single day seem to be influencing every node with which it is sharing its temporal level meaning that at this level it resembles an almost complete graph. To see a subset of figure 4.3 based on the categorizations, please refer to appendix D.

Immutable

Gender / XY chromosome (Mi, O, IE)
Genetic polymorphism (Mi, O|M|R, IE)
Medical Factors/Birth Defects (Mi, O|M|R, IE|E|I)
Origins/Ancestry (Mi, O|M|R, IE)
Psychopathology Parents (Mi, O|M|R, IE)

~18 Years

Attachement (Mi, O|M|R, IE)
Attributional Style (Mi, O|M|R, IE|I)
Child Abuse (Me, O, E)
Chronic Medical Factors (Mi, O|M|R, E|I)
Coping/Self regulation (Mi|Me, M, I)
Cultural Factors (Ma, M, IE|E)
Early Life Trauma/Early Life Experience (Mi, O|M|R, IE)
Epigenetics (Mi, O|M|R, I)
Expectation Patterns (Mi, O, I)
Gender (Mi, O, IE|I)
Intelligence (Mi, O|M|R, I)
Internalising (Mi, R, I)
Neural Pathways (Mi, R, I)
Neuroticism (Mi, M, IE)
Other mental health conditions (Mi, IE|E|I)
Schema (Mi, O|M|R, I)
SES (Me|Ma, M, IE|E|I)
Social Domain (Ma, O|M, E|I)
Structural Racism/Injustice/Prejudice (Ma, O|M|R, EI|E)
Temperament (Mi, O|M|R, IE)

~Year

Level of Education (Mi|Me|Ma, O|M, E)
Poor housing (Me, O|M, IE|E|I)
Poverty (Mi|Me|Ma, IE|E|I)
Previous Episodes (Mi|Me, M|R, I)

SPATIAL SCALE: Mi - Micro | Me - Meso | Ma - Macro
TEMPORAL IMPORTANCE: O - Onset | M - Maintenance | R - Relapse
INTERNAL/EXTERNAL IE - Initial External | I - Internal | E - External

~Month/week

(Stopping) Anti-depressants (Mi|Me, R, E)
Advocacy/Agency/Autonomy/Mastery (Mi, O|M|R, E|I)
Bad Therapist/Inadequate Treatment (Me, M, E)
Chronic Stress (Mi|Me, O|M|R, I)
Complicated Problems (Mi|Me|Ma, O, E|I)
Diet (Mi|Me, O|M|R, E|I)
Dynamic medical factors (Mi, O|R, E|I)
Endocrine System (Mi, M, I)
Imminent Threat (Mi|Me|Ma, O, E)
Inflammation (Mi, M|R, I)
Lack of Treatment (Mi, M, E)
Neurotransmitter System (Mi, M, I)
Other mental health conditions (Mi, IE|E|I)
Own Perspective on Disease (Mi, O, I)
Positive Mental Health (Mi, M|R, I)
Problem solving (Mi, O|M|R, I)
Relationship (Mi|Me, M, I)
Role (Mi|Me, O|M|R, I)
Role Transition (Mi|Me, O|M|R, E)
Social Isolation (Me, O|M|R, E|I)
Social Support (Me, O|M|R, I)
Substance Use (Mi|Me, O|M|R, IE|I)
Suicide of other (Mi, O|M|R, E)
Trauma/Life Experience (Mi|Me|Ma, O|M, E)
Treatment Response (Mi|Me, M, I)
Weight gain/loss (DSM SYMPTOM)

~Day

Compliance/Adherence to Therapy (Mi|Me, M|R, I)
Daily Structure (Mi|Me, M, I)
Sleep (Mi, O|R, I)
Sleep behaviour (Mi, O|R, I)
Sleep disturbance (DSM SYMPTOM)
Social Interaction (Me, O|M|R, I)
Stress (Mi|Me, O|M|R, I)

Within a Day

Abuse (Me, O|R, E)
Agitation (Mi, O|M|R, I)
Agitation/Retardation/psychomotor (DSM SYMPTOM)
Avoidance of potential rewards (Mi, M, I)
Behavioural Avoidance (Mi|Me, O|M|R, I)
Biological Stress (Mi, O|M|R, I)
Cognition (Mi, O|M|R, I)
Concentration (DSM SYMPTOM)
Coping/Self regulation (Mi, M, I)
Demoralisation (Mi|Me, M|R, I)
Depressed mood (DSM SYMPTOM)
Eating/Apetite (DSM SYMPTOM)
Feeling of worthlessness (DSM SYMPTOM)
Help Seeking (Mi|Me, M, I)
Hope (Mi, M, I)
Inactivity/Passive Withdrawal (Mi|Me, M, I)
Loss of Interest (DSM SYMPTOM)
Non-physical/creative Activities (Mi|Me, O|M|R, I)
Perceived Stress (Mi|Me, O|M|R, I)
Physical Activities (Mi|Me, O|M|R, I)
Reduced/low self esteem (Mi, O|R, I)
Retardation (Mi, O|M|R, I)
Rumination (Mi, M, I)
Self Blaming (Mi, M, I)
Sense of Control (Mi, R, I)
Social Referencing/Modelling (Mi, O|M|R, I)
Suicidal Ideation (DSM SYMPTOM)
Withdrawal (Mi|Me, O|M|R, I)

FIGURE 3.5: Nodes were divided by temporal scale, their spatial scale (micro (i.e. within a person), meso (i.e. a persons direct surroundings) and macro (i.e. at a higher level like government regulation)), temporal importance (i.e. which part of the clinical picture is the node important?) which is discussed earlier, and whether nodes receive feedback from the system itself (internal) or the node influences the system from outside (external). Also a distinction was made for external nodes that are present from early childhood (initial external).

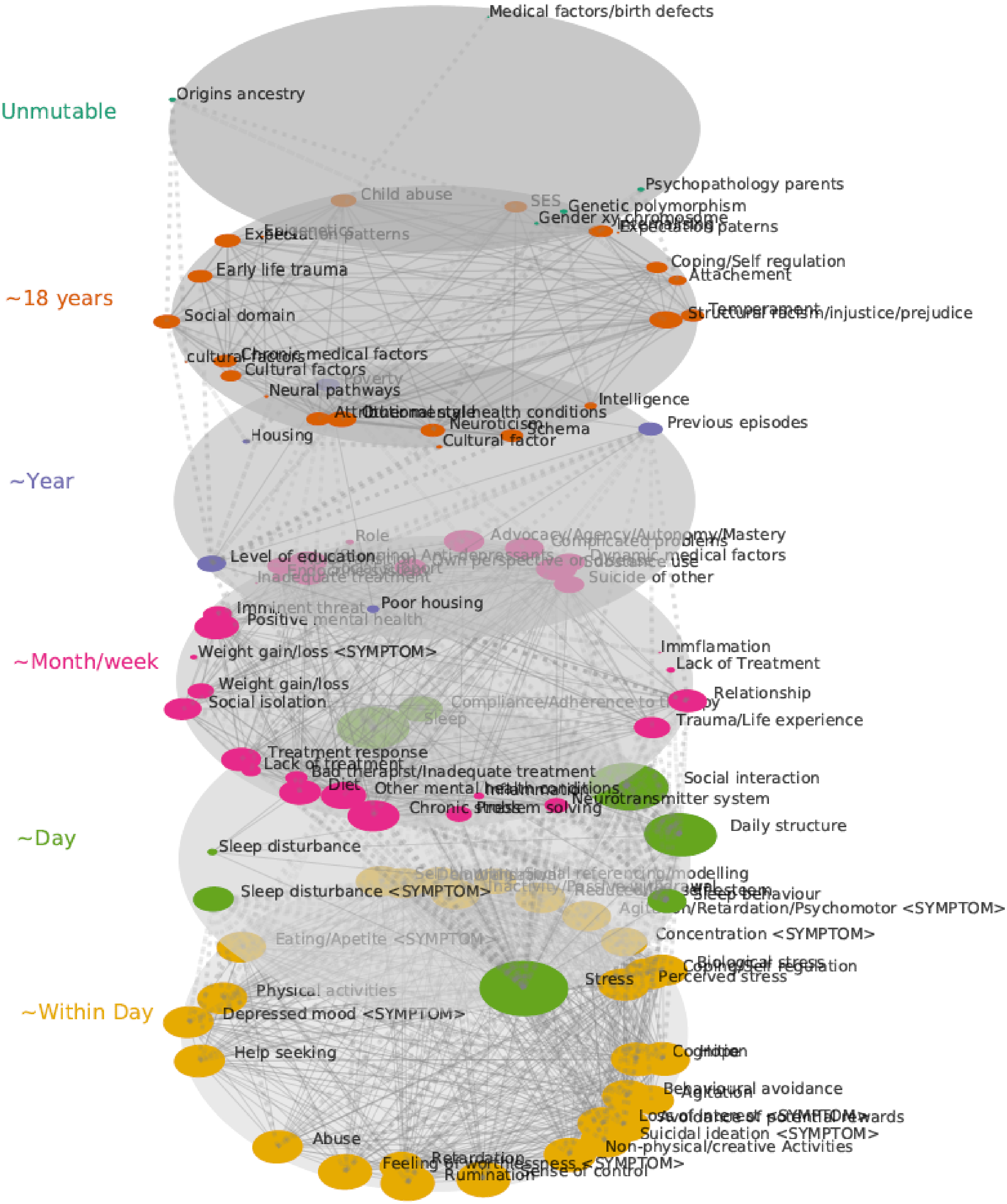


FIGURE 3.6: Depiction of the network over all time-scales including all edges within and between time-scales.

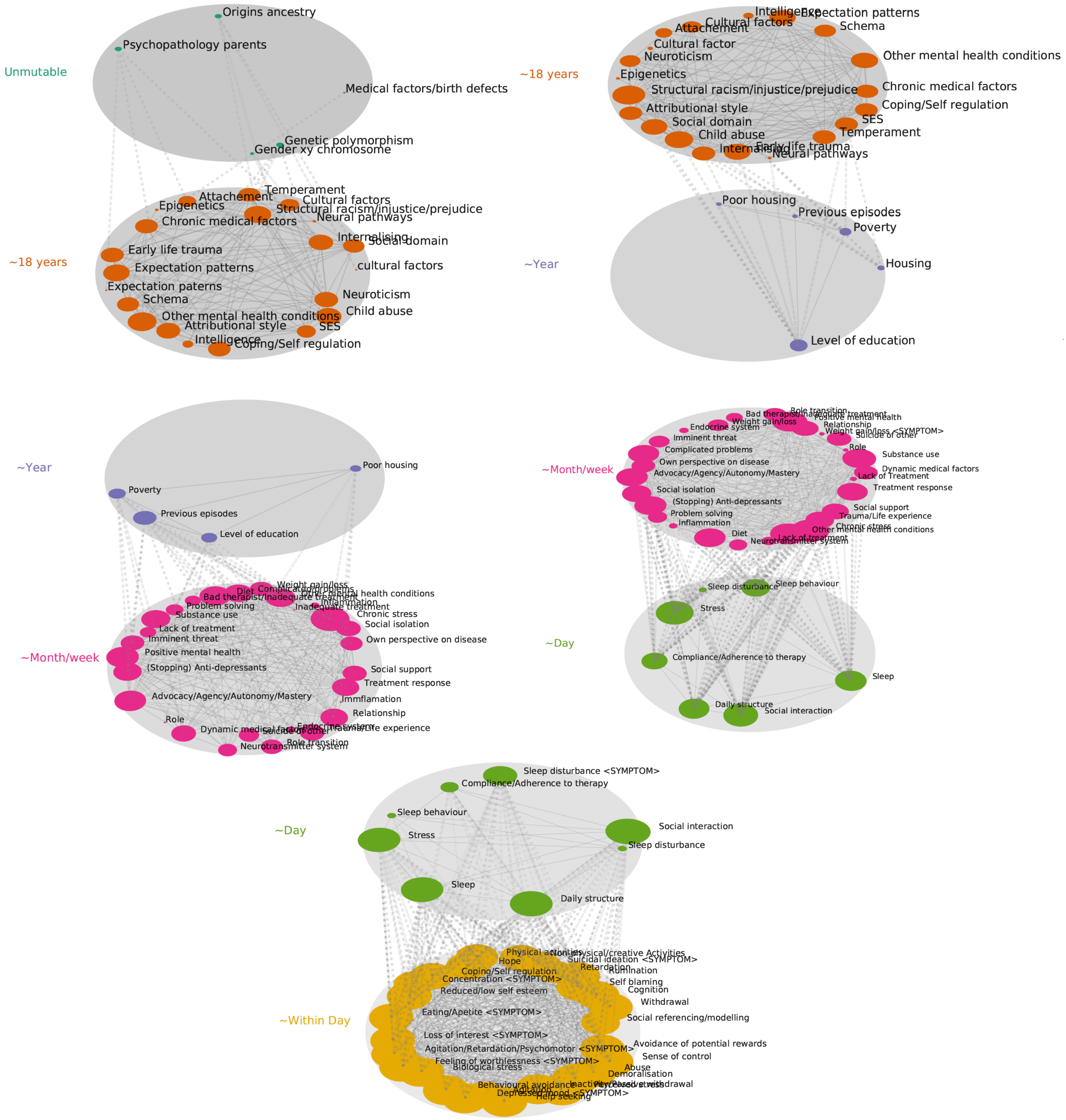


FIGURE 3.7: Each combination of adjacent time-scale are visualized separately in order to get a clearer view of within and inter-level edges.

3.4 Discussion

Our interviews have shown that there is great heterogeneity within Major Depressive Disorder, both in how it is defined and how it comes about. Furthermore, a causal network has been created in which it has become clear that the nodes that operate within a day together with the nodes that fluctuate with around a month or week and two decades are numerous and highly connected.

In this chapter, we aimed to provide a starting point for further research. Future steps should involve a feedback round where experts are asked to provide their opinion with regards to the network and categorizations given in this chapter. The causal mapping visualizations presented in the results section could be used as a tool to guide systems thinking and enhance further collaboration. There are, however, some limitations to the methods used that need to be highlighted. Even though the quality of our expert pool was superb, planning and keeping the interview focused was difficult via Skype as a communication medium. These issues will be further highlighted in section 5.1 as a possible solution will be presented there. Also, a lot of the work was performed by our research group with a single hour worth of interview per expert as input. This was enough to get a thorough overview of possible causal nodes, however, there was not enough time to get the information needed to form an extensive causal node list. This edgelist was therefore created by our own vision. So even though the edgelist was created by rigorous group model sessions, they are, in fact, not necessarily a reflection of the opinions of the interviewed experts. Lastly, as interviews were performed one expert at a time, there was no way of communication between experts. A group modeling session could have been more fruitful as there would then be a way to consult the experts directly in collaboration efforts and immediately assess the result. Yet, this method poses the difficulty of requiring experts to be in the same room which would be impossible for our diverse and international expert pool. Although there were some difficulties with our current method, this chapter could be an excellent stepping stone for future research. After feedback has been obtained, dynamics could be implemented using system dynamics techniques posed in chapter 2 using actual patient data. And the effectiveness of the topology of the resulting network could then be analyzed by the methods posed in the following chapter.

Chapter 4

Quantifying reliability of system dynamics topology

4.1 Introduction

Investigation of the behavior of certain phenomena as a system is gaining increasing popularity in a great diversity of disciplines. Disciplines such as biology (Wang and Michoel, 2019), Economics and Management (Long and Chen, 2019, Yang et al., 2019), public health (Luke and Stamatakis, 2012), psychology (such as with depression (Wittenborn et al., 2016)), geography education (Cox, Steegen, and Elen, 2018) are increasingly approaching their academic issues in terms of systems thinking. Through this holistic and inclusive approach, an understanding is gained of the relationships and patterns that arise from complex problems (Haraldsson, 2004). These efforts (normally brought about from a collaboration between modelers, literature and expert knowledge) result in a causal mapping or diagram of the phenomenon of interest.

For example, Kenzie et al., 2018 noticed an urgent need for innovative approaches to merge knowledge gained from literature and interviews to improve clinical outcomes for concussion treatment. They had combined rounds of literature and individual interviews with 26 experts in the field which resulted in an extensive causal diagram. Ford and Sterman, 1998 has developed a method for eliciting, articulation and description of expert knowledge for specifying parameters and relationships for formal system dynamics modeling. But where Kenzie et al., 2018 stops at the formation of a CLD and Ford and Sterman, 1998 seeks for ways of eliciting information for the implementation of dynamics, some questions spring to mind. How will the model be affected if there are structural mistakes within the topology of the network that are unknown to the experts?

To validate these types of models, usually, data is used to fit dynamics and parameters within the system. One way to go about implementing dynamics in a causal diagram would be to translate the network into a system of ordinary differential equations (ODEs) as is discussed in chapter 2. But what if the resulting causal diagram from these interviews contains structural mistakes or what if it is missing parts entirely? The whole premise of this complex system approach is that there are behaviors that transcend their individual parts in one way or another. But is it needed to represent the whole network correctly to get a good approximation of the real system? The aim of this chapter focuses on the effect of increasingly faulty systems and their descriptive power concerning the real, original system that is to be examined. Babtie, Kirk, and Stumpf, 2014 has found that even minor changes in a model can alter the conclusions that we draw. They used a method called Topological Sensitivity Analysis (TSA) to consider the sensitivity of model dynamics concerning both parameter and structural variation. They have not, however, mapped the decline of prediction concerning several types of mistakes. In light of this, this chapter proposes a similar approach using a genetic algorithm (GA) to assess underlying parameter values and compare the outcome of the fitting procedure to a set of Erdos-Renyi random graphs. This is done regarding a combination of two types of mistakes. One type is the reversal of relation weight polarity whilst the other type consists of rewiring edges altogether (i.e. A $\rightarrow$ B changes randomly in D $\rightarrow$ Z). This way a measurement can be made regarding the effect of changes within the true topology of a network as the network becomes more and more random. The hypothesis is that there is a gradual decline of predictability due to the introduction of increasing (combinations of) mistakes. Furthermore, an assessment is made on the effect of knowing the parameter values or not. Not knowing parameter values would indicate the need for parameter fitting (i.e. a genetic optimization algorithm) which could find a different set of parameters that facilitate a good fit to data even if there are structural problems within the network. Lastly, the addition and removal of edges in the perfect topology will lead to an understanding of whether it is advisable to provide a network with either too few or too many edges. By adding too many edges in a system a fit to data can be overly tweaked which would always lead to a better fit (i.e. it would essentially create room for overfitting) which complicates inferring causality in the network.

4.2 Methods

4.2.1 System of ODEs

As it is not needed for our purposes to adhere to systems of ODEs that are known in the literature, a random network generator was written to find systems with N nodes and E edges with interesting dynamics. To determine whether a graph was "interesting" some constraints were put into place. These constraints entailed a minimal variance over each time step of more than the arbitrarily chosen number of 50 to assure that variables within the system did not follow the exact same path. Also, for each variable on its own, a minimal variance of 10 was required over the span of the time series to assure that lines did not flatline. Furthermore, a minimal and maximal value of -1500 to 1500 was given to prevent the system from going towards infinity in either a positive or negative direction. Within these constraints, certain graphs were produced with oscillatory qualities that were used for the experiments discussed in this chapter.

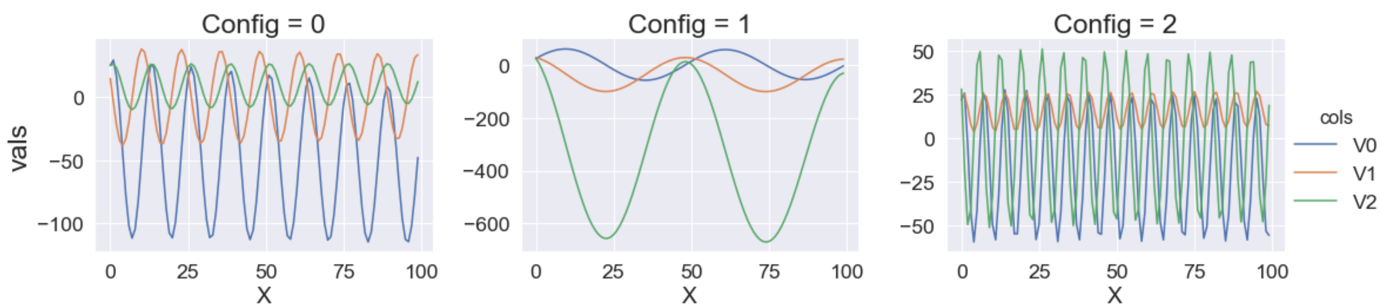


FIGURE 4.1: 3 distinct networks consisting of 3 nodes and 5 edges. All three network topologies differ with their own set of system parameters. For underlying ODEs and initial values please refer to appendix E.

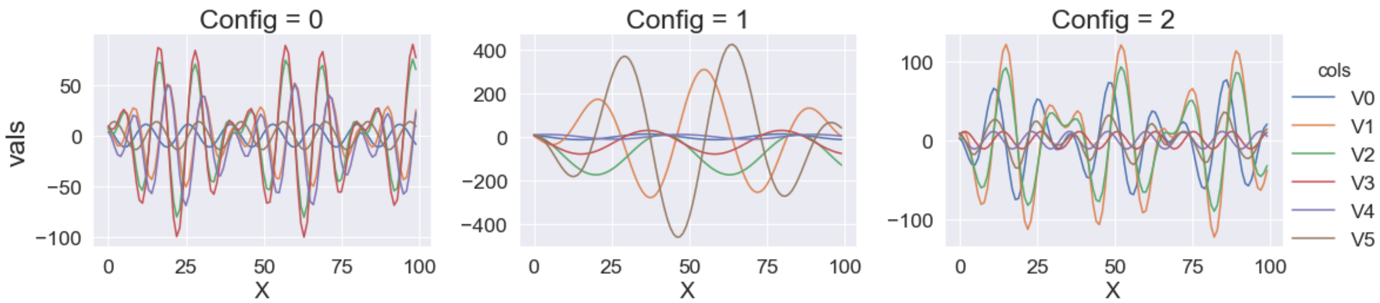


FIGURE 4.2: 3 distinct networks consisting of 6 nodes and 10 edges. All three network topologies differ with their own set of system parameters. For underlying ODEs and initial values please refer to appendix E.

The systems of ODEs were solved by using the `scipy.integrate.odeint` function and taking the dot product of the values of the system with the system parameters ($v_{t+1} = v_t \cdot P$ where v_t are the values of the nodes at time t and P is an adjacency matrix containing the parameter set). Initial values and parameter values regarding each edge were generated randomly.

Using the ODE generator, 2 different system sizes were chosen, each with 3 distinct systems. This was to infer the effect of both system size and assure robustness of results within the same system size. 3 systems had sizes $N = 3$ nodes and $E = 5$ edges which can be viewed in figure 4.1. Depicted in figure 4.2 are 3 networks all with $N = 6$ nodes and $E = 10$ edges. For underlying ODEs and initial values please refer to appendix E.

4.2.2 Parameter fitting - genetic algorithm

A Genetic Algorithm (GA) is based on ideas inspired by biological evolution. Our implementation consisted of a population of 50 possible solutions to the problem. This population then evolved over 120 generations towards an acceptable solution set. Biological processes such as reproduction, recombination, mutation, immigration and selection that are prevalent within the theories of biological evolution are also to be found with the GA paradigm. In pseudo-code, such an algorithm would look as shown in algorithm 2. The algorithm used various techniques of which some are designed to stimulate convergence towards a solution, while other techniques are designed to prevent the population from converging too rapidly into a local minimum. The following techniques are employed in this papers' GA:

- **Selection:** The selection scheme chosen in this GA was kept as simple as possible. Out of each generation, the top 10% scoring individuals proceeded unaltered towards the next generation and were also chosen as parents for the next generation. This was done to ensure a relatively quick convergence towards an acceptable answer. The rest of the parent pool was selected randomly to keep diversity in the gene pool.

- **Crossover:** Out of the parent pool sets of parents were paired up randomly and a range in the domain of parameters is chosen at random which are then swapped around between parent and thus creating "offspring".

- **Mutation:** Mutations can be seen as point-mutations on a continuous scale, a parameter is randomly chosen and perturbed slightly by multiplying a value sampled

Algorithm 2 Example of Genetic Algorithm pseudocode

Result: Convergence towards (possibly locally) optimum solution
START
  Generate initial population
  Compute fitness of individuals within the population

  while *stopping criterion has not been met* **do**
    Select best scoring individuals
      Crossover
      Mutation
      Immigration
      Compute the fitness of individuals within the new population
end
STOP

from a uniform distribution between 0.95 and 1.05. Each parameter had a 5% chance to undergo mutation.

- **Immigration:** To keep the algorithm from converging too fast towards a local minimum, immigration is put into place. Here new randomly generated solutions are introduced into the already existing population keeping the gene-pool divers.

The operators used are also depicted in table 4.1. To assess a "goodness of fit" a score function has to be used when optimizing for parameter values in a system. The score function in this research consisted of the Mean Absolute Difference (MAD) between the true data and simulation dynamics. As this score function is used as a minimization problem, the GA seeks to minimize this MAD-score (i.e. a lower score is a better fit). Algorithms were written mostly using Numpy and parallelized over 16 nodes using the python multiprocessing library on the DAS-4 supercomputer. The stopping condition was a fixed amount of generations to keep each run through the algorithm as similar as possible. The runtime of a single algorithm with a population of 50 and 120 generations was ± 2 minutes. Each network configuration was computed through the algorithm 200 times, for each up to 9 mistakes for

TABLE 4.1: Generalist genetic algorithm

Representation	Vector of N floats and an $E \times E$ adjacency matrix with corresponding weights
Recombination	Random domain crossover
Recombination probability	100%
Mutation	Probability driven single point mutation
Mutation probability	5%
Parent selection	Elitism & Random
Survival selection	Elitism
Number of elitists	10%
Number of immigrants	10%
Population size	50
Number of offspring	10%
Initialisation	Random
Termination condition	120 generations

the larger network and 4 mistakes for the smaller networks. This is done for both mistake configurations (rewiring, reversed polarity) and their respective combinations. Rewiring happened by changing the edges within the system while reversing polarity concerned flipping the weight polarity of the edge (+/-).

4.2.3 Erdos-Renyi

To compare the expert network including its introduced mistakes a set of 1000 Erdos-Renyi random graphs was generated for both system sizes. Erdos-Renyi graphs are random graphs in which the number of nodes and edges within the system are constant but the edges themselves are wired randomly.

4.2.4 Small networks vs. Large networks

First, an evaluation was performed to determine if there are differences between small networks and larger networks. For analysis of the system the mean (μ) of the resulting score distributions was considered to be a measure of accuracy (i.e. a low score would mean a small mean absolute difference indicating a high accuracy), while the variance (σ) is considered to be a measure of precision (i.e. a low variance indicates that the GA repeatedly ends up in the same score which means a high precision).

Second, a normalization towards the Erdos-Renyi set was performed to determine the "randomness" of the system (i.e. efficacy of the true or "expert" topology compared to a randomly generated graph). A value of 1 indicates that the MAD-score of the network has reached the same score as the set of randomly generated graphs. An overshoot of this mean would mean that the MAD-score of the algorithm performs more poorly than a random graph would. While a score closer to zero would indicate better performance compared to the random Erdos-Renyi graphs. The normalization was performed using equation 4.1 where EA_{score} is the GA MAD-score performance measure while ER_{score} corresponds to the Erdos-Renyi MAD-score.

$$\frac{(\frac{EA_{score}}{ER_{score}})}{max(EA_{score})} \tag{4.1}$$

4.2.5 The effect of known and unknown parameter values

After the two system sizes were compared an experiment was run to examine the effects of mistakes in system topology where the system parameter values were known. The system dynamics model was run using two methods, for the first method the parameter values were unknown and, as such, have to be fitted using the GA. In the second method, the whole network was known, topology *and* parameter values, but mistakes were also introduced by simply perturbing the edge list or reversing polarity on a parameter. This was done to examine the effect of mistakes within the network in the hypothetical case that the expert has a perfect notion of the system but is still lacking some knowledge. To obtain these results, again, 200 EAs were performed for each combination of mistakes in network topology starting from zero mistakes and, as such, a perfect topology without the need for parameter fitting.

4.2.6 The effect of under/overpopulating the edge list

As a last experiment, the effect of under- or- overpopulating the edge-list was examined. This entails adding or removing random edges in/from the weighted adjacency matrix and applying the GA on those topologies. Again this was done 200 times for each addition/removal of n edges.

4.2.7 Used software/hardware

All simulations were written using the following software and hardware:

Software: Python 3.5 $\rightarrow$ `Multiprocessing` (McKerns et al., 2012), `Numpy` (Oliphant, 2006), `Pandas` (McKinney, 2010), `SciPy` (Jones, Oliphant, and Peterson, 2001), `Seaborn` (Waskom et al., 2017).

Hardware: Distributed ASCI Supercomputer 4 (DAS-4) (Bal et al., 2000).

4.3 Results

4.3.1 Genetic algorithm reliability

Mean Absolute Difference scores

Upon viewing the score distributions of each network configuration (figures 4.3 and 4.4) an assessment can be made of the accuracy and precision of the EA. In this sense, a lower score would mean a higher accuracy while a low variance in the distribution would map to high precision. Each configuration is assessed using the true topology of the network using the mean MAD-score (μ), variance (σ) and minimal score (min). For the larger system with 6 nodes and 10 edge parameters this yielded the following for configuration 0: $\mu = 21.67$, $\sigma = .86$, $min = 19.65$, configuration 1: $\mu = 69.51$, $\sigma = 58.61$, $min = 53.12$, and configuration 2: $\mu = 23.50$, $\sigma = 3.37$, $min = 17.74$. For the smaller system of 3 nodes and 5 edge parameters we see the following for configuration 0: $\mu = 27.68$, $\sigma = .53$, $min = 25.82$, configuration 1: $\mu = 109.49$, $\sigma = 2.95$, $min = 106.09$, and configuration 2: $\mu = 24.13$, $\sigma = 4.98$, $min = 8.35$.

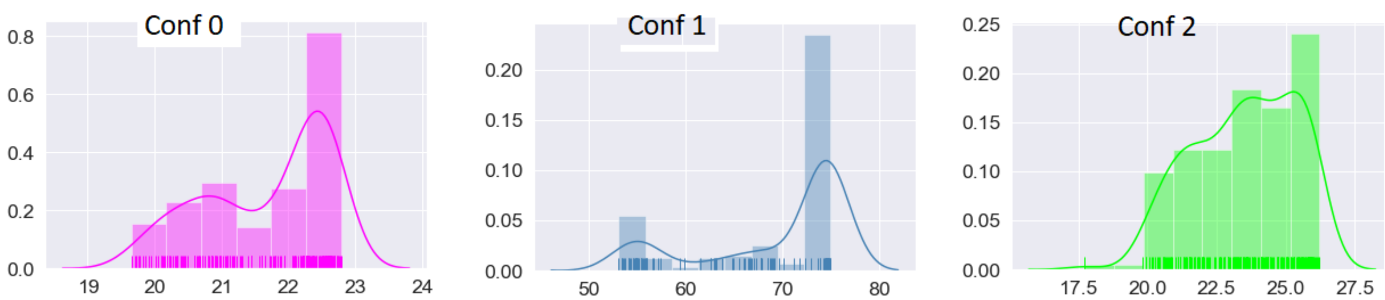


FIGURE 4.3: Score distributions after running each configuration 200 times for a **10 edge parameter, 6 nodes system**. Conf 0: $\mu = 21.67, \sigma = .86$, Conf 1: $\mu = 69.51, \sigma = 58.61$, Conf 2: $\mu = 23.50, \sigma = 3.37$

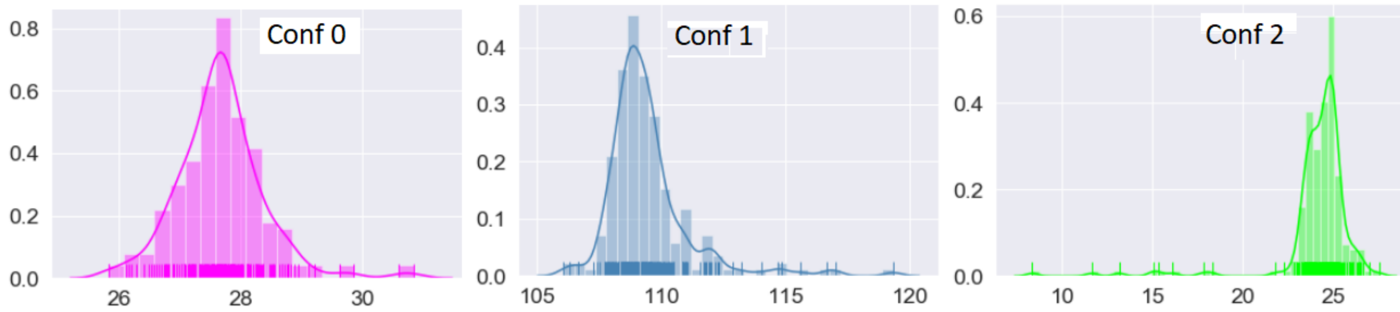


FIGURE 4.4: Score distributions after running each configuration 200 times for a **5 edge parameter, 3 node system**. Conf 0: $\mu = 27.68, \sigma = .53, min = 25.82$, Conf 1: $\mu = 109.49, \sigma = 2.95, min = 106.09$, Conf 2: $\mu = 24.13, \sigma = 4.98, min = 8.35$

Efficacy of the GA in finding parameter values in small and large system sizes

In appendix F the distribution of found edge parameter values are depicted. Here it can be seen that the error in finding the actual parameter value differs amongst parameters. Most parameters seem to be close to the top of the kernel density estimation (KDE), i.e. the theoretically most occurring value of the distribution, and/or the best scoring solution. There are, however, a lot of cases in which both the KDE top values and best-scoring solutions find very different values than that of the real parameter value. The MAD-scores of the score distributions for each configuration of the top KDE value and best scoring solution compared to the real parameter values as depicted in table 4.2. For the small system, this score is lowest with 1.07 on average for the best performing solutions. Furthermore, the average KDE and best solutions score were similar for larger systems with a MAD-score of 1.51 and 1.52. Ending last is the KDE parameter estimator with an average score of 1.85 in the small system.

4.3.2 Effect of mistakes within a network topology in small networks vs. larger networks

Figure 4.5 shows both the mean and variance of each combination of mistakes for all configurations of the large network while figure 4.6 depicts this for the small network. Noticeable is all three configurations seem to be fitted around the same order of magnitude for both system sizes. This could be indicating that the GA performs similarly over multiple networks regardless of system size. Furthermore, in terms of the accuracy of parameter prediction, each cumulative mistake adds another margin of error to the fit. Especially the combination of both reversing parameter polarity

TABLE 4.2: MAD scores for differences between actual parameter values and top KDE values or best scoring solution

Configuration	Small System KDE	Best	Large System KDE	Best
0	1.77	2.10	1.06	1.20
1	1.75	0.66	1.91	1.39
2	2.02	0.47	1.55	1.98
Mean	1.85	1.07	1.51	1.52

and edge topology seems to be synergistic in decreasing both accuracy and precision. Furthermore, it is notable that the transition towards higher scores is not necessarily a smooth shift as there are cells with a lot of mistakes and still perform well. For example, if we look at configuration 1 of the larger network there is a cell that contains 9 mistakes of polarity and three mistakes of rewiring and still has a score of $2e^{03}$ compared to its direct neighbors of $3e^{06}$, $6e^{11}$, $6e^{09}$.

Figure 4.7 depicts the mean scores normalized towards the mean of the Erdos-Renyi set. The overshoot is visualized by the steel-blue boxes surrounding the cell (the cells that perform worse than the Erdos-Renyi random set). Here we can see the same trends that are spotted in figures 4.5 and 4.6. The networks all grow towards randomness, growing several orders of magnitude per added mistake. Some even performing worse than the mean of the fully random graphs.

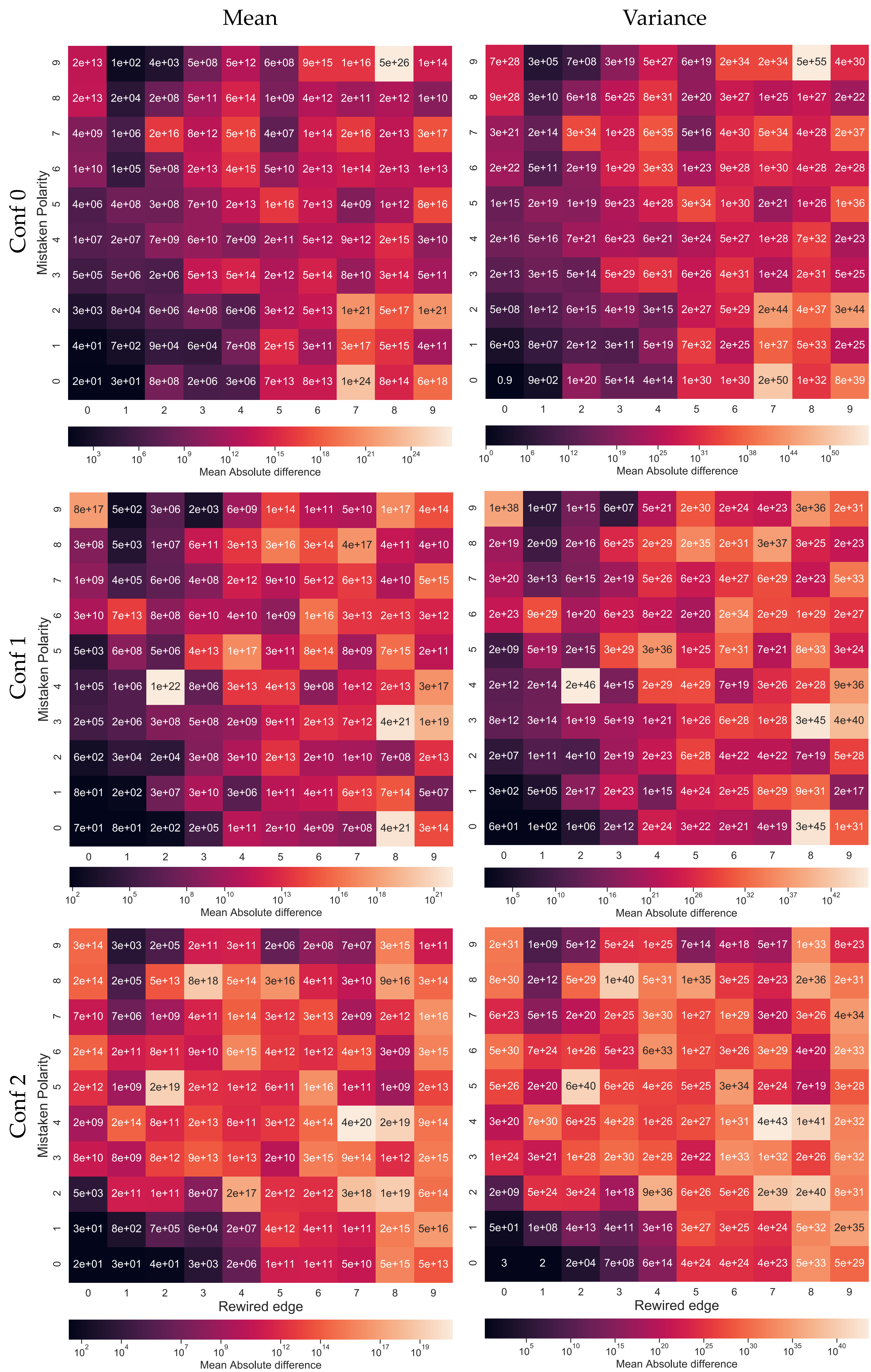


FIGURE 4.5: Mean and variance taken from the MAD-score distribution of 200 GA runs for each configuration in all combinations of mistakes for the larger network.

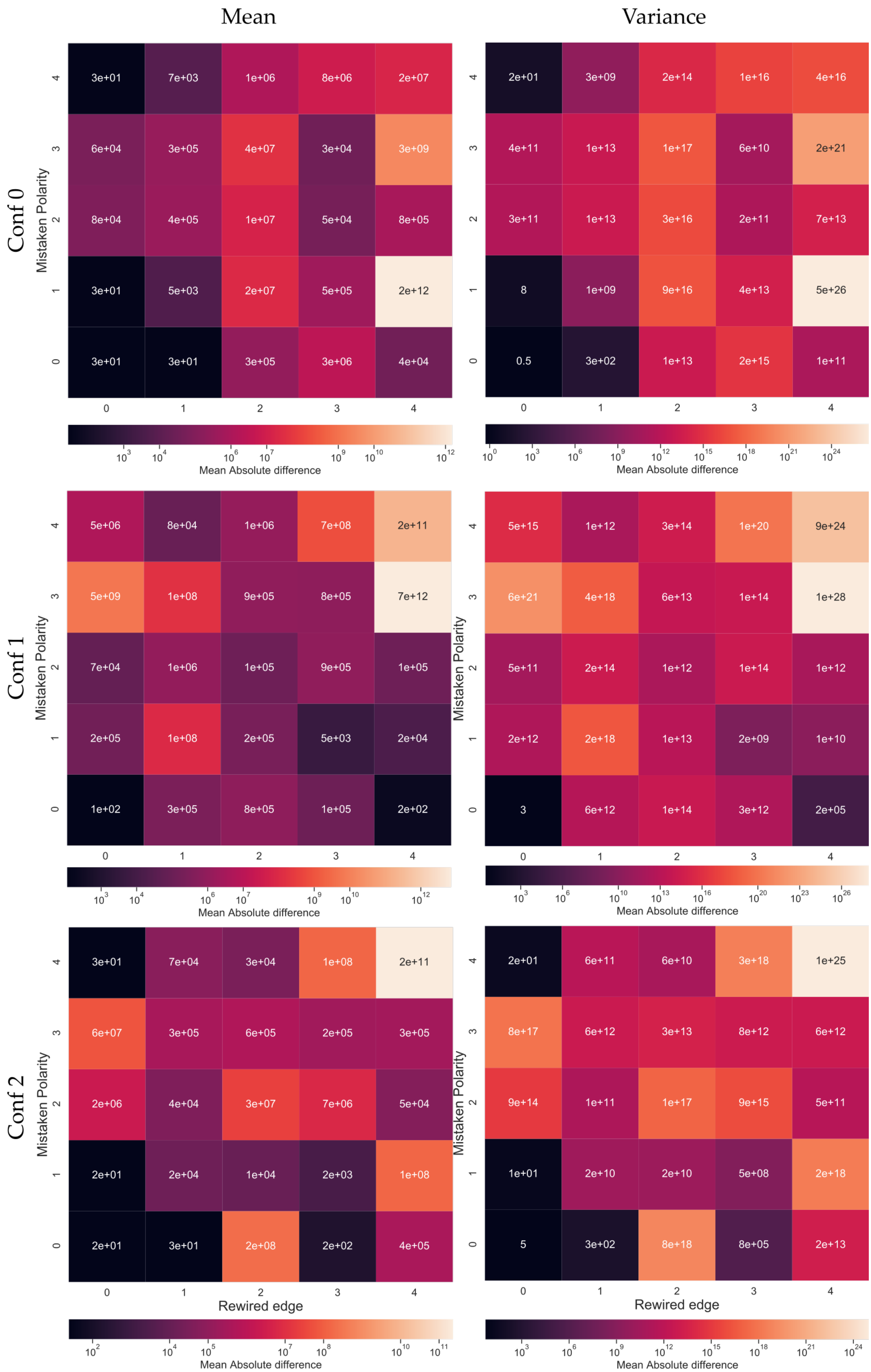


FIGURE 4.6: Mean and variance taken from the MAD-score distribution of 200 GA runs for each configuration in all combinations of mistakes for the smaller network.

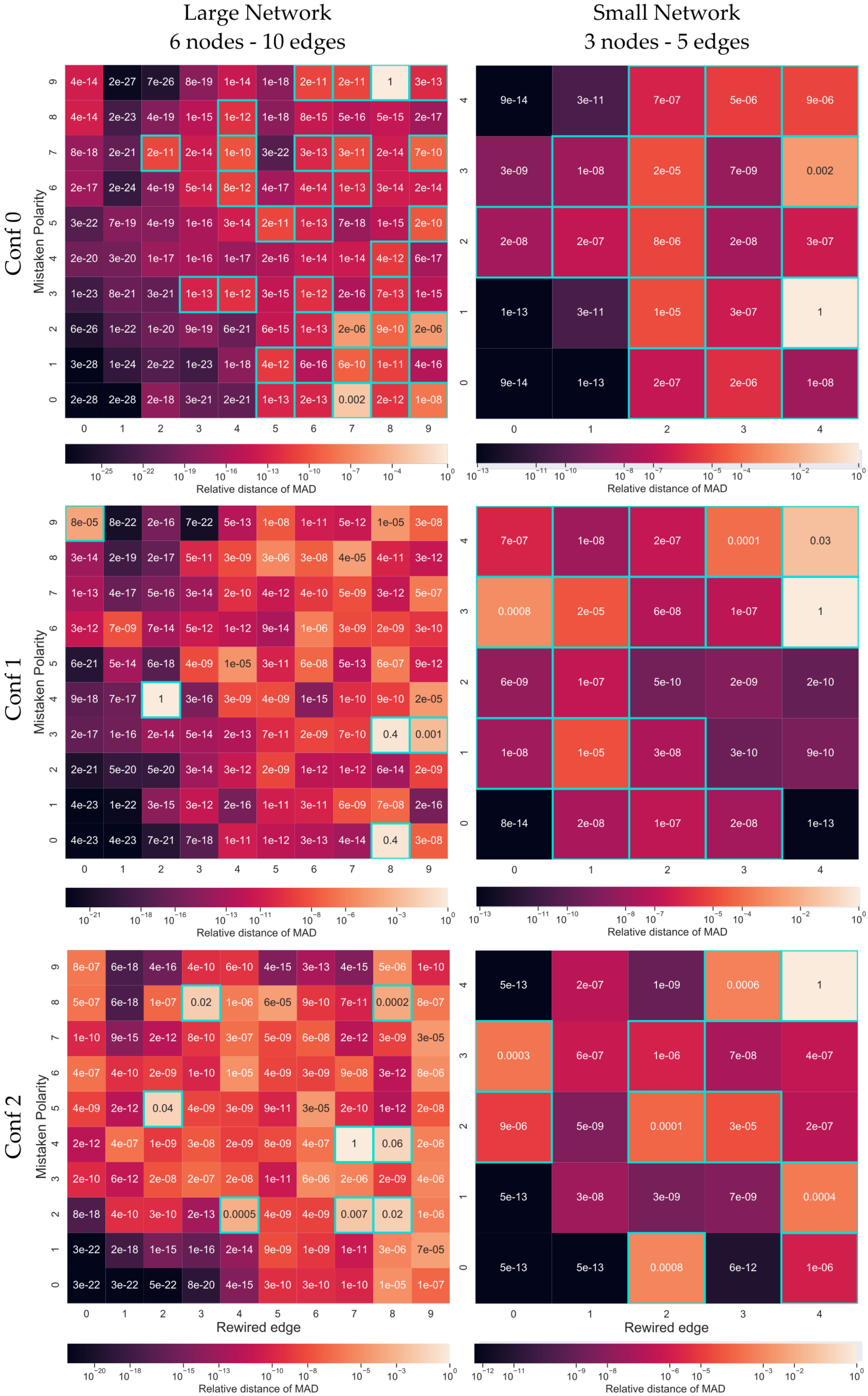


FIGURE 4.7: Normalized differences between Erdos Renyi graph scores and GA scores taken from the MAD-score distribution of 200 GA runs for each configuration in all combinations of mistakes. Here the larger networks are depicted on the left and the smaller networks on the right.

4.3.3 Mistakes in a perfect topology

As there seems to be no real difference between the smaller network and the bigger network, all results will henceforth concern the larger networks.

Here we have applied the same method of introducing mistakes as in section 4.3.2, but to the perfect network topology and all parameters are known. As there is no need for parameter fitting in this system, all simulations were run without the GA. In figure 4.9 the same trends as in section 4.3.2 are found when using a genetic algorithm to fit the parameters except that the order of magnitude is much higher. Also, rewiring edges yield a much larger error than reversing edge polarity. However, the combination of rewiring edges and reversing the polarity, again, seem synergistic in the sense that their combination yield much higher errors than errors within a single axis.

4.3.4 Under-/overpopulation of the edge-list

By adding a random edge or removing one, we can examine the effect of being too hesitant or reluctant to add edges in the system or being too overzealous. Figure 4.8 is a depiction of the MAD-score for all larger system configurations. Here we see that by removing edges the error stays low while adding several edges more than that of the actual system, the error can increase by several orders of magnitude.

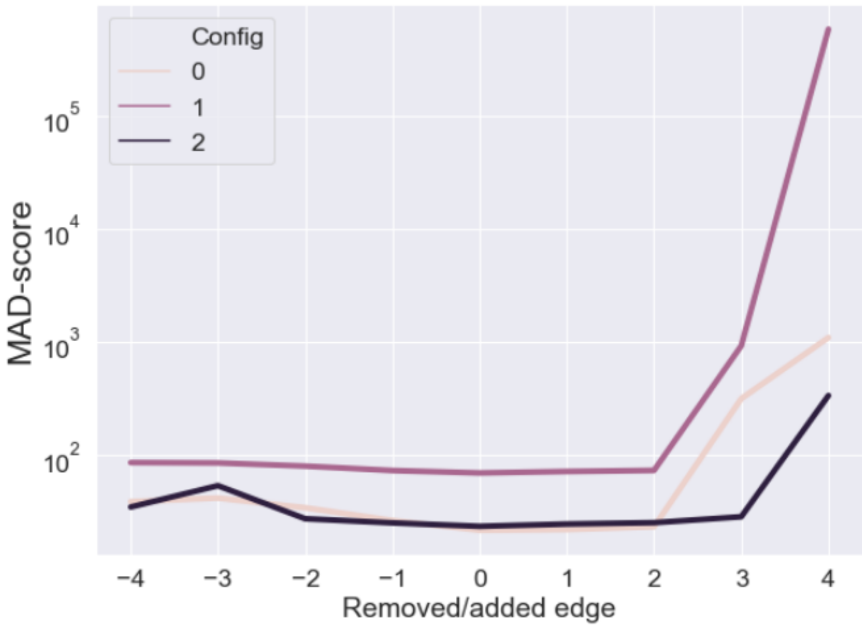


FIGURE 4.8: The removal or addition of edges in the perfect network topology with respect to the MAD-score produced by 200 GA runs.

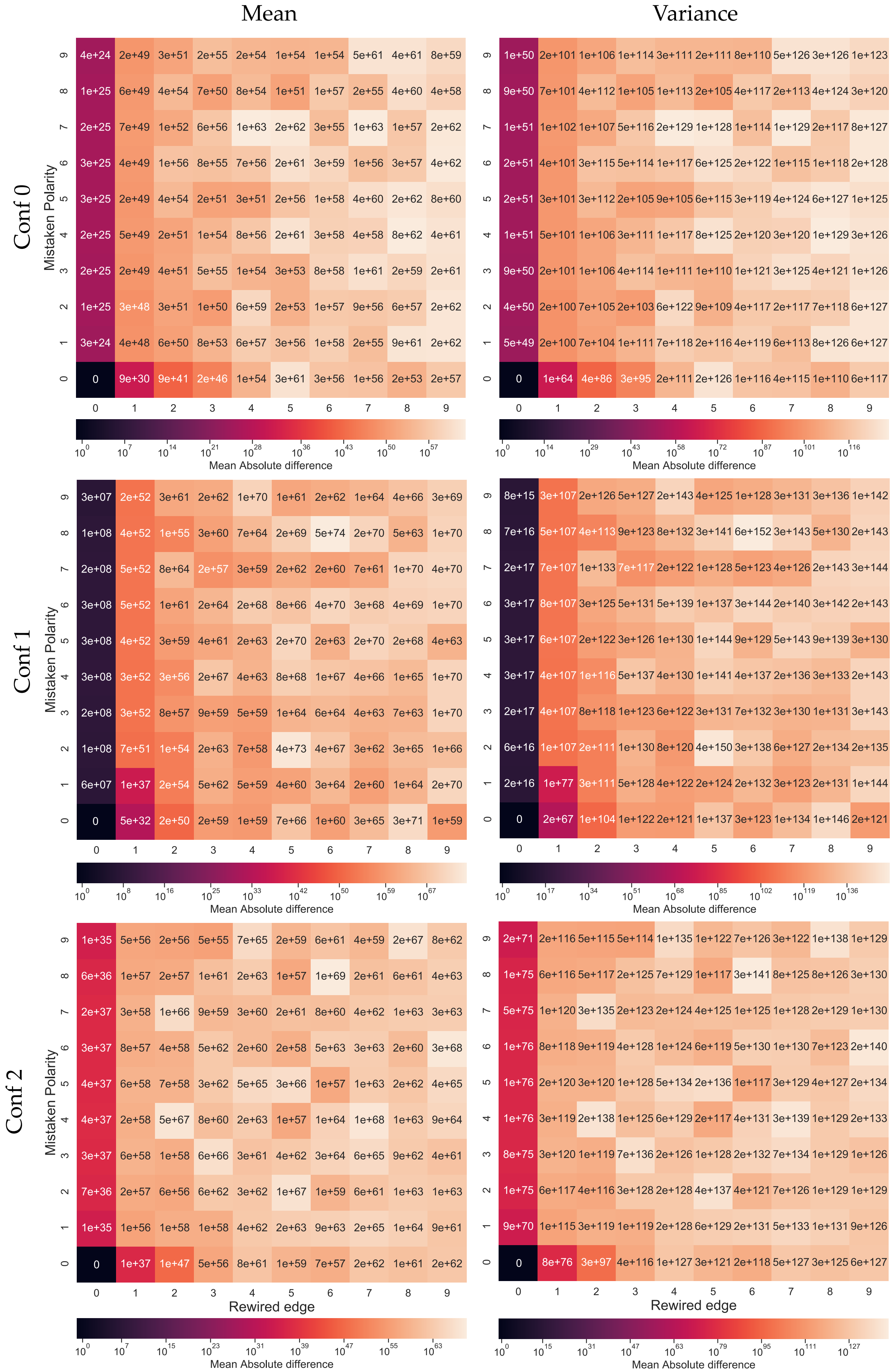


FIGURE 4.9: Mean and variance taken from the MAD-score distribution of 200 perfect topology and parameter runs for each configuration and all combinations of mistakes.

4.4 Discussion

Concerning the efficacy of the GA in finding parameter values, we saw that certain parameters were very hard to estimate as opposed to others. This might be a result of insensitive edges that are insensitive to parameter values. It can also be that the set of parameters as a whole can be set in such a way that the resulting dynamics account for the dynamics seen in the data. This would mean that various parameter sets could approximate the dynamics "well", but not perfectly. This could account for the lack of precision in some parameters as a certain "closeness" to the real parameter set would not necessarily mean a better score, but the combinations within the parameter set could.

As the purpose of this study is to find the effects of faulty network topology and not GA optimization, the GA itself is not optimized as much as it could have been leaving the accuracy somewhat lacking. This, however, is not very problematic as we are interested in how certain changes in topology affect GA outcomes relative to one another. As all configurations are simulated and optimized by the same EA, their resulting trends in MAD-scores can, therefore, be compared as opposed to the efficacy of finding the parameters themselves.

As we compare the two system sizes to one another it seems that the difference between both system sizes was significant between configurations within each system size. A direct comparison to assess whether or not there is a real difference between system size is therefore difficult. However, as can be seen in figures 4.5, 4.6, and 4.7 the trends they show are very similar. Each mistake within the wiring or the polarity in the system seems to introduce several orders of magnitude of error in both mean MAD-score and variance and converges rapidly towards scoring the same as a random Erdos-Renyi graph for both system sizes. These error bounds also seem to be similar when comparing the number of introduced mistakes in both the smaller and the larger system indicating the overall size of the system would matter less than the number of mistakes introduced. This can be viewed as a cautionary tale in which network topology can severely affect its approximation of the real underlying system. These findings underline the need to check expert knowledge-driven network topology performance by comparing their network to random generated Erdos-Renyi graphs using the same parameter estimation methods.

This has raised the question of what the effect of these types of mistakes is without the need for parameter fitting. What would be the effect if the experts even knew each parameter value perfectly but incorporated some topological mistakes? By using the perfect topology and parameter set we have eliminated the need for the GA. By employing this method the effect of these purely topological changes is exposed. As we have seen in figure 4.9 the gradient that the combinations of mistakes bring about is highly similar to that of the comparisons between the larger and smaller systems. The premiere difference being that the order of magnitude of the mean and variance is much bigger. This suggests that there is indeed a mitigating effect within the genetic algorithm, possibly finding parameter sets that differ from the true parameter set but have a better fit to the data under the altered network topology.

Finally, we have aimed to investigate the effect of under populating or overpopulating the edge-list. By adding or removing a certain number of randomly selected edges from the list we have seen that decreasing the number of edges increases the errors more slowly than adding additional edges. This can, however, also be due to the fact that removing edges can render a node to be completely flatlining over time as their incoming edges could be removed. This flatlining keeps the system from growing towards large positive or negative numbers. So by rendering the system incapable of change, or at least decreasing the flow of information within the system, it becomes less reactive and, thus, easier to fit for the GA. By adding more edges in the system we also add to the parameter space that needs to be fitted. This could be an explanation of why adding extra edges results in a worse fit in addition to the numbers being more easily able to grow towards higher numbers (opposite of removing edges). A solution to this could be by binding the ODEs between values, but that was outside the scope of this research and can be used in future endeavors.

This study has used a genetic algorithm to fit the parameter values of the networks. The optimization of the GA itself, however, was manually done by a combination of trial and error and literature. Several methods aim to optimize GA parameters (e.g. population size, number of generations, immigration flow, etc.). The parameters themselves could, for example, be made self-adaptive to control convergent and divergent forces in the population (Pellerin, Pigeon, and Delisle, 2004). Furthermore, as fitness values are recalculated on every iteration and the diversity of the population decreases as it converges to a solution, a strategy for speeding up the GA can be to save fitness values into a hash table. By doing this Povinelli and Feng, 1999 has found a performance improvement of up to 50% for complex real-world problems. This speedup can cause a decrease in runtime, creating an opportunity for increasing the number of generations that can be computed or simply increasing the number of GAs run for a better estimation of the score distributions. Another way to optimize the algorithm could be by following the Taguchi method or a variant of it Yildiz, 2013. This method is used to optimize the parameters of the GA itself. Also, as this study is more of a preliminary overview of topological effects it is needed to employ more statistically rigorous methods for future efforts. Where this study presents trends, future study needs to present statistically sound conclusions.

Only simple sets of ODEs were used in this study. Future research can provide insight into the effects of non-linear functional forms of edges in terms of accuracy and precision of several types of mistakes. Furthermore, this method has only been applied to random graphs that generate dynamics within certain constraints. It would be important to test the hypotheses posed in this chapter to real-world-systems and validate on real data. With this in mind, it would be interesting to take the causal diagram data from chapter 3 and turn it into a system dynamics model using the methodology from chapter 2 to test our hypotheses on the resulting network with real-world data.

This chapter has aimed to investigate the importance of topological truth concerning these two types of mistakes. But moreover, a method was developed to gain an indication of how expert knowledge-driven networks can be compared to random graphs and their relation to data. This type of topological sensitivity analyses can be applied by a modeler after a model has been formulated using methods from chapters 2 and 3.

Chapter 5

General discussion

This thesis has attempted to handle most of the aspects regarding depression and, more generally, system dynamics modeling. Chapter 2 has provided a step by step guide to move from a factor list towards a fully quantifiable system dynamics model. Recommendations were made to facilitate creating a fully labeled CLD and transforming that into a simulation in its own right. Where research tends to stop when creating a CLD, this chapter has provided tools for going beyond that. By implementing dynamics it is easier to look at causal relations within a network as the possibility then arises for simulated interventions. This chapter has made it easier to start from a question and end up in the more interesting phases more quickly, viewing causal relations and investigating dynamics within a system.

Chapter 3 provides a database that was created using expert knowledge which can henceforth be used as a reference for simulating depression. It must be said, however, that this database is not without limitations as is discussed in chapter 3. Yet it can serve as a framework for future research. This chapter can function as a template by adapting its timescales, categorizations, nodes, and edges. Each categorization serves as a level of granularity within a model and thus facilitates easy selection from the database enabling moving to modeling using the steps discussed in chapter 2.

Chapter 4 has found that making mistakes can be quite costly in terms of both precision and accuracy of a simple model regardless of system size. It also provided a way to compare the network topology generated from the intersection of the literature and expert knowledge to randomly generated Erdos-Renyi graphs. If there is no difference in precision and accuracy upon perturbing the network topology, one can ponder about the effectiveness of the network itself as it already performs closely to a random network. Furthermore, the mitigating effect on the error of using a GA to estimate parameter values was also found to be substantial as perturbing a perfect topology lead to errors several orders of magnitude higher. Finally, the effect of overpopulating versus under populating the edge-list was examined. The results seem to suggest that under populating the edge list would be better, however, definitive advice regarding this has to wait for further research.

Through three different chapters, this paper has sought to be as complete as possible regarding the early formation of complex network analysis and has made recommendations in its quantification. Due to time constraints, however, it has not been able to cover the gathering of feedback after the first round of interviews. In the next section 5.1 a proposition is made for how that feedback can be acquired.

5.1 Future research

It has become apparent that scheduling interviews and feedback moments pose some difficulty regarding time management. This is due to the fact that our Expert pool was scattered across the globe and all had a very busy schedule. To accommodate these schedules and, without compromise, get the best set of experts for purposes of the study an internet-based platform is proposed to create and deliberate causal mappings and automatically generate system dynamics models. By making this communication platform available on the internet experts can develop a sense of ownership with the model, as they can directly influence it. In this way, a model can be made with little effort for any single individual consisting of the shared ideas of many experts which can, in itself, lead to more understanding. This section will elaborate on an internet-based template prototype that could suit for each study aiming to create a causal diagram and/or system dynamics model. By keeping its specifics readily changeable for the administrator account the template is flexible for any topic seeking to use this same method. An example of this website can be found at www.baschatel.nl [1]. To view contents outside the scope of the introduction page an account needs to be created.

5.1.1 Software proposition

To facilitate collaboration between experts and modelers, we propose a general website template that enables communication and speeds up the process by taking the modeler by the hand and guiding them through the process. This alters the workflow given in figure 3.3 where the modeler would need to ask directly for feedback and would need to schedule a new round of interviews for feedback. The new workflow (5.1) would consist of sending experts a link to the website where they can "take

[1] The code can be viewed at https://github.com/popoiopo/psychoSystems.

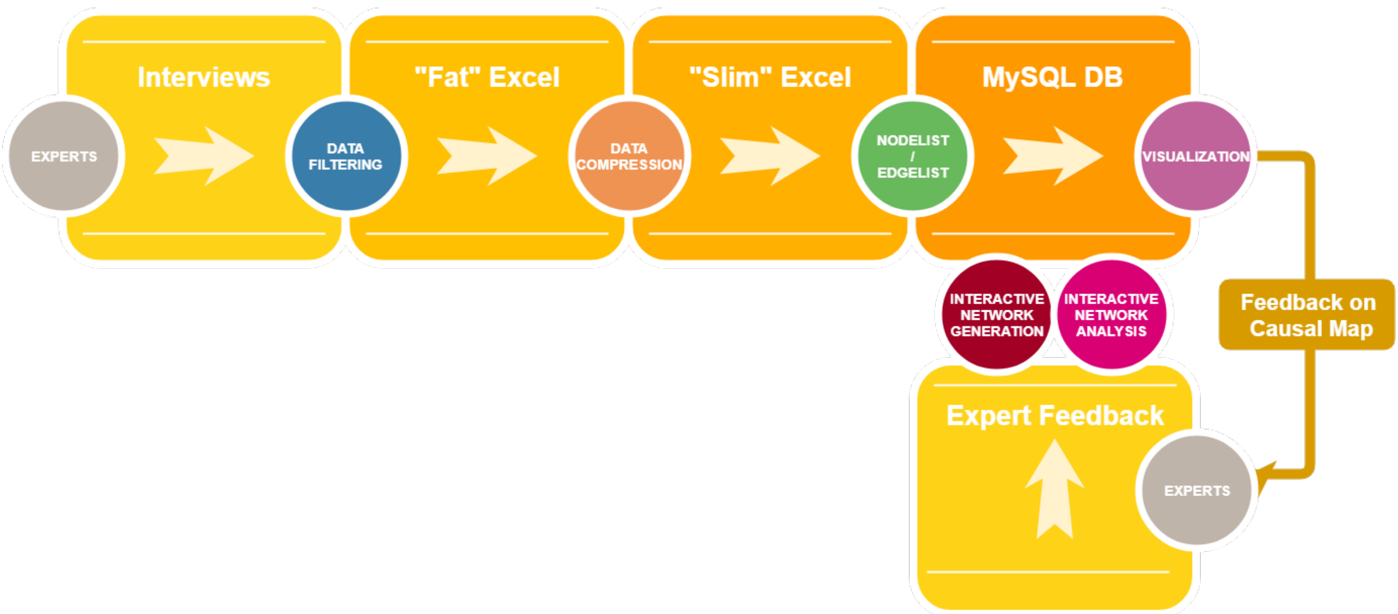


FIGURE 5.1: Altered data workflow regarding conducting and processing Expert Interviews. The process ends in a collaboration platform where experts can directly vote for what is in their mind the optimal network, and interactively test their hypothesis. This provides the modeler with tools to complete and elevate their data without being too intrusive to experts.

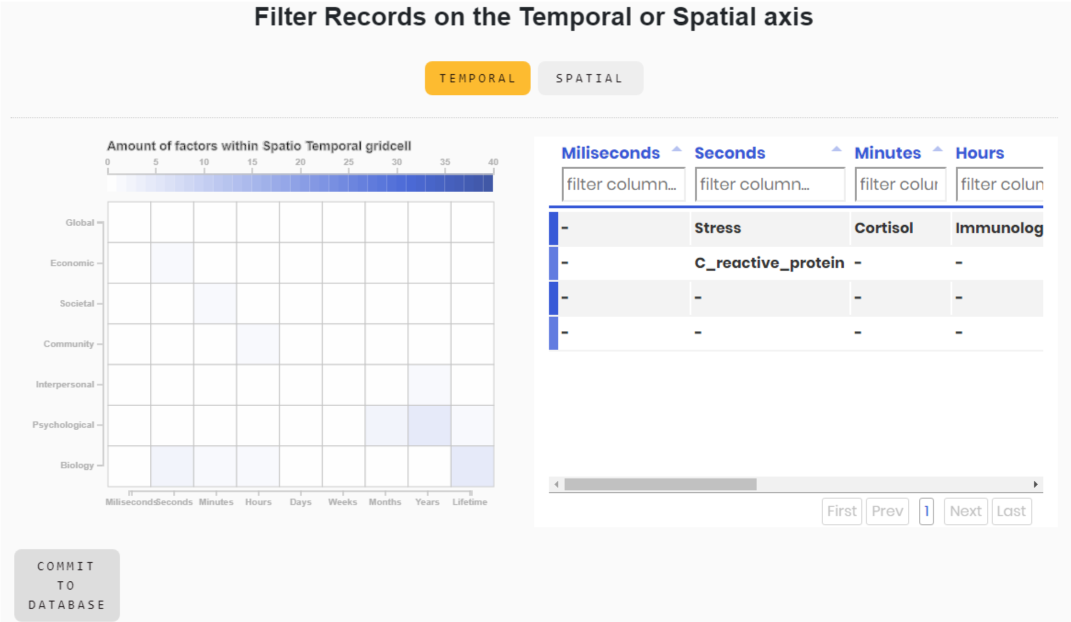


FIGURE 5.2: Gridplot with two ordinal axes representing the spatial and temporal components of MD factors. Upon clicking on a grid cell, the table next to it gets upgraded showing the individual factors inside said cell based on the spatial or temporal criterion. The amount of factors within a cell changes its color.

on the role" of the modeler, activating the experts' engagement resulting in a cooperation platform where experts can vote for their optimal network.

CHATEL-package

Hardware and software The application was programmed using the Python Flask environment in combination with SQLalchemy for database queries. Furthermore, jquery was used to handle communication between Python and JavaScript, visualizations were either made with D3.js or vis.js. The server is a 2GB RAM, 20GB memory virtual private server (VPS) using the Ubuntu16.3 OS. A full tutorial on how to create a VPS and deploy the website on it is given in appendix G.

To optimize the feedback process a visual representation was created for each phase of a three-part process. The first phase (Fig 5.2) entails a grid plot with at its y-axis the spatial element of depression (e.g. Biology, Interpersonal relations, Community effects, Societal, Economy, etc.) while the x-axis contains the temporal component of the factors (i.e. the timescale at which a factor operates (seconds, hours, lifetime, etc.)). With this visualization, users can interactively add a node based on where they would think the node operates in time and space. These axes can be changed by the admin account into, for example, the micro, meso and macro subdivisions earlier posed in this thesis.

The second phase involves building the edgelist. Here one can simply create an edge between nodes by choosing two factors in the "from" and "to" sections with some additional information regarding the speed and strength of the edge followed by clicking on the "Create Relation" button. After this, the simple network visualization will dynamically be updated visualizing the flow of information. Figure 5.3 depicts this phase. Feedback loops where an upstream node affects a downstream node is visualized in a red flowing edge.

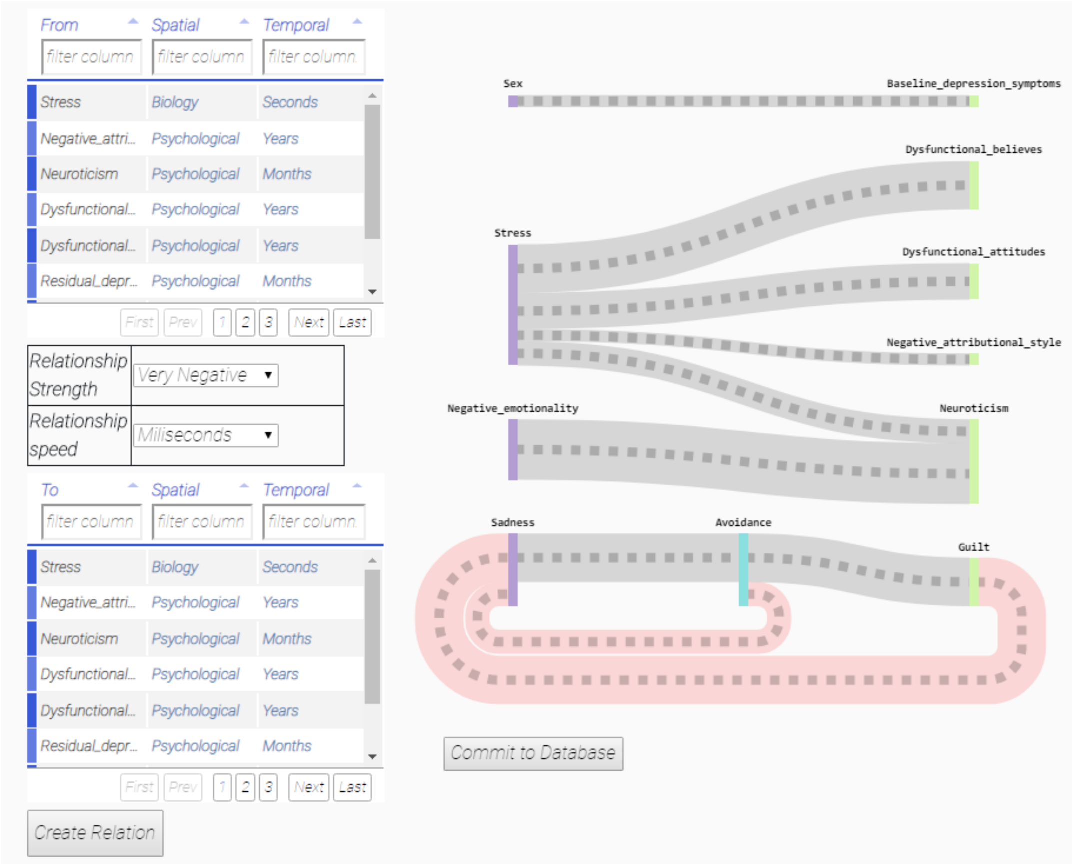


FIGURE 5.3: Edge generator: One can simply create an edge between nodes by choosing two factors in the "from" and "to" sections with some additional information regarding the speed and strength of the edge followed by clicking on the "Create Relation" button. After the simple network visualization will dynamically be updated.

The third phase form contains an interactive mapping of the causal diagram itself. Here all nodes are depicted with their respective edges to each other. New nodes and edges could then be added/removed at the users own accord and thus, create the causal mapping as the users sees fit (this freedom can be switched on or

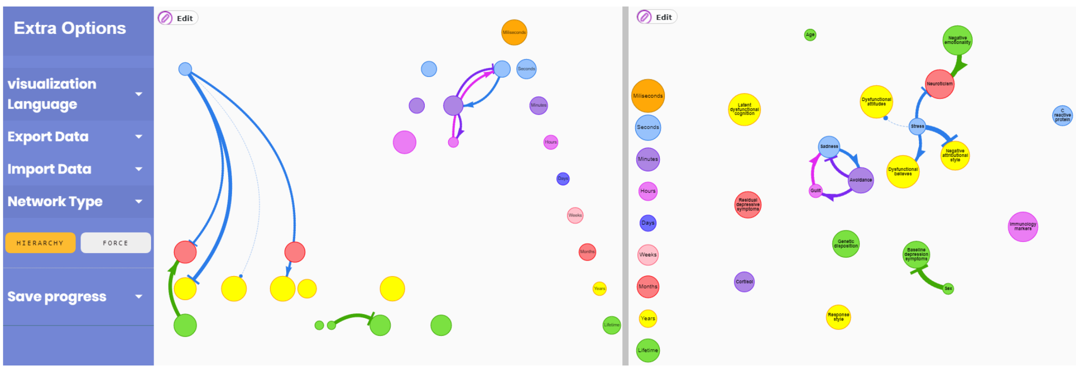


FIGURE 5.4: Network Exploration and additional information: Here experts can see how the network looks like in two different modes. One mode is hierarchy based on the speed (this is, however, variable and can change) while the other mode is force-directed. Also, additional information for nodes and edges can be provided like literature, sensitivity, temporal importance (e.g. onset, maintenance, relapse), some general notes, etc.

off by the admin as it can be crucial for the process to not be able to or maybe penalize being able to add edges and nodes at certain times). By adding another node or edge, questions arise about the nature of the new entry like these spatial and temporal aspects but also notes, supporting literature, possible weights or thresholds, type (i.e. is it a stock, flow, variable, etc.). Also previously added nodes and edges can be updated. The application will then automatically position the information as can be seen in Fig 5.4.

There are still some problems within the code of this project, and as this was outside the scope of this paper, not everything could be automized. The deployment, however, takes about one hour by following appendix G. Future work will also need to verify if this way of collaboration is indeed better than performing interviews or partake in group modeling sessions. There are still some ways to go until this template is a viable application that is user-friendly to non-programmers, but this way of presenting information and creating room for dialog that is expert-based without being intrusive can help to create extensive holistic models in the future.

5.2 Conclusion

As it is apparent that multivariate system simulations become increasingly indispensable, this paper has sought to provide a new way of approaching MDD. A causal network database was presented based on the collective information gathered from 20 interviews with experts in the field of MDD. And a methodology for creating a causal network diagram and turning that into a dynamic model was formalized. Proving an overview of key aspects in the process of system dynamics modeling. Furthermore, a topological sensitivity analysis was proposed to assess the dynamics created from these dynamical models. This, together with a proposition of an internet-based collaboration platform, this paper has provided insight into how Major Depression Disorder can be approached. As the multivariate dynamical system that MDD is, and by approaching it as such, we might just be able to find new treatment targets and eventually get treatment efficacy above this 50% threshold.

Appendix A

Expert list

- Dr. Andrews (Psychology/Evolutionary psychology, McMaster University)
- Dr. Bosch (Clinical Psychology, University of Amsterdam)
- Prof. Cuijpers (Psychology/Clinical/Neuro and Developmental psychology, VU University)
- Prof. Derubeis (Social Science, University of Pennsylvania)
- Prof. First (diagnostics of mental health conditions, Colombia University)
- Prof. Harmer (Cognitive Neuroscience/Psychopharmacology and Emotional Research Lab, University of Oxford)
- Emily Haroz, PhD (Psychiatry/Psychiatric epidemiologic methods/Advanced measurement models, John Hopkins bloomberg school of public health)
- Henrie Henselmans (Member of the board of directors (healthcare; GGZ))
- Prof. Hollon (Psychology/Intervention/Etiology, Vanderbilt University)
- Prof. Kendler (Biology/Psychiatry/(Molecular) Geneticist, Virginia Institute for Psychiatric and Behavioral Genetics)
- Prof. Margraf (Sociology/Psychology/Clinical psychology and psychotherapy, Ruhr-universität Bochum)
- Jeroen Muller (CEO at Arkin)
- Prof. Oldehinkel (Epidemiologist, University Medical Center Groningen)
- Prof. Ormel (Sociologist/Epidemiologist, University Medical Center Groningen)
- Prof. Patel (Psychiatrist, Global mental health, Harvard University)
- Prof. Schnabel (Sociology/Policy making/Director Sociaal en Cultureel Planbureau, Utrecht University)
- Dr. Greg J. Siegle (Psychiatry/Cognitive Affective Neuroscience, University of Pittsburgh)
- Prof. Stronks (Program leader Health Behaviours & Chronic Diseases, University of Amsterdam)
- André Tomlin (Managing Director of Minervation, Founder of The Mental Elf website)
- Prof. Zachar (Psychology, Auburn University Montgomery)

Appendix B

Factor list

1. (Stopping) Anti-depressants
2. Abuse
3. Advocacy/Agency/Autonomy/Mastery
4. Agitation
5. Agitation/Retardation/Psychomotor <DSM SYMPTOM>
6. Attachement
7. Attributional style
8. Avoidance of potential rewards
9. Bad therapist/Inadequate treatment
10. Behavioural avoidance
11. Biological stress
12. Child abuse
13. Chronic medical factors
14. Chronic stress
15. Cognition
16. Compliance/Adherence to therapy
17. Complicated problems
18. Concentration <DSM SYMPTOM>
19. Coping/Self regulation
20. Daily structure
21. Demoralisation
22. Depressed mood <DSM SYMPTOM>
23. Diet
24. Dynamic medical factors
25. Early life trauma

26. Eating/Apetite <DSM SYMPTOM>
27. Endocrine system
28. Epigenetics
29. Expectation patterns
30. Feeling of worthlessness <DSM SYMPTOM>
31. Gender xy chromosome
32. Genetic polymorphism
33. Help seeking
34. Hope
35. Housing
36. Imminent threat
37. Inactivity/Passive withdrawal
38. Inadequate treatment
39. Inflamation
40. Inflammation
41. Intelligence
42. Internalising
43. Lack of treatment
44. Level of education
45. Loss of interest <DSM SYMPTOM>
46. Medical factors/birth defects
47. Neural pathways
48. Neuroticism
49. Neurotransmitter system
50. Non-physical/creative Activities
51. Origins ancestry
52. Other mental health conditions
53. Own perspective on disease
54. Perceived stress
55. Physical activities
56. Poor housing

57. Positive mental health
58. Poverty
59. Previous episodes
60. Problem solving
61. Psychopathology parents
62. Reduced/low self esteem
63. Relationship
64. Retardation
65. Role
66. Role transition
67. Rumination
68. SES
69. Schema
70. Self blaming
71. Sense of control
72. Sleep
73. Sleep behaviour
74. Sleep disturbance <DSM SYMPTOM>
75. Social domain
76. Social interaction
77. Social isolation
78. Social referencing/modelling
79. Social support
80. Stress
81. Structural racism/injustice/prejudice
82. Substance use
83. Suicidal ideation <DSM SYMPTOM>
84. Suicide of other
85. Temperament
86. Trauma/Life experience
87. Treatment response

88. Weight gain/loss <DSM SYMPTOM>
89. Withdrawal
90. cultural factors

Appendix C

Interview script

C.1 Introduction

First, we would like to thank you for taking the time to participate in this interview. We really appreciate it, as your input is of vital importance for our research project. Let me start by introducing ourselves, the goal of this research project, and the aim of this interview

We are «insert name» and «insert name» and we work together with «insert names of other team members» at the Institute for Advanced Study of the University of Amsterdam. The Institute for Advanced Study is an interdisciplinary institute where researchers from different backgrounds collaborate on scientific topics. This gives us an opportunity to discuss our research with experts outside our own fields. Our own team is also multi-disciplinary: with a team consisting of psychiatrists, psychologists, methodologists, and computer scientists. We aim to create a model of major depression using complexity modeling. When constructing a computational model in which factors interact, it is important to design the model in such a way that it can explain or accommodate the most important phenomena relating to depression. For instance, the model should be able to explain the fact that some individuals recover from depressive episodes, and that relapse is prevalent. We hope that such a model will lead to a better understanding of major depression and possibly may allow us to identify new targets for treatment.

We use expert interviews to determine which factors this model should contain and how it should work. That is where you come in. Previously, we performed several meta-analyses on the evidence of leading biopsychosocial models. With the interviews, we try to create an additional dataset giving us access to information beyond the scope of literature. The goal of the interviews we are doing is to create a causal map, which we will use to create the final model. I will send you an example of a causal map of the environmental impact of passenger cars now to show what we want to create. It is just meant as an example, so please do not pay too much attention to the contents.

<< Give time to look at causal map >>

We value all input, so feel welcome to interrupt whenever you want. The interview will take around one hour and before we officially start we would like to ask you if we can record the interview? «wait for confirmation of the participant» Do you have any questions before we start?

C.2 General questions

1. What is your definition of major depression?

   (a) How does your definition relate to the DMS-V diagnosis of major depression?

2. Which are the 5 most important factors you would like to include in a causal map of major depression?

    (a) Which of these factors interact causally?

    (b) When these factors arise, does that reflect a gradual process, like growth, or the occurrence of a specific event, like a lightning strike? Is once enough or is repeated exposure necessary?

    (c) Are there any factors that are often forgotten about?

    (d) How do individual differences arise with respect to those factors?

So if I understood correctly you consider << summarize general factors >> to be the most important factors regarding major depression? We will now move on to more specific questions firstly regarding onset, secondly maintenance and finally relapse.

3. Are the factors you identified so far as the most important factors also the most important for the onset of major depression?

    (a) If not, which factors are the most important for the onset?
        i. Which of these factors interact causally?
        ii. Is it a gradual process or the occurrence of a specific event? Is once enough or is repeated exposure necessary?
        iii. Why do some people develop depression while others do not?

    (b) Are there any factors regarding the onset that are often forgotten about?

4. Are the factors you identified so far as the most important factors also the most important for the maintenance of major depression?

    (a) If not, which factors are the most important for the maintenance?
        i. Which of these factors interact causally?
        ii. Is it a gradual process or the occurrence of a specific event? Is once enough or is repeated exposure necessary?

    (b) Are there any factors regarding the maintenance that are often forgotten about?

5. Are the factors you identified so far as the most important factors also the most important for the relapse into major depression?

    (a) If not, which factors are the most important for relapse?
        i. Which of these factors interact causally?
        ii. Is it a gradual process or the occurrence of a specific event? Is once enough or is repeated exposure necessary?

    (b) Are there any factors regarding relapse that are often forgotten about?

6. Which factors could be considered the most important factors to treat?

    (a) If this factor was treated, how would it affect the causal map?

    (b) <<If any more interesting factors were named, ask about them>>

Thank you for these factors, they will help in making a model to better understand major depression. When we construct a computational model in which these factors interact, it is important for us to design the model in such a way that it can explain or accommodate the most important phenomena relating to depression. For instance, the model should be able to explain the fact that some individuals recover from depressive episodes, and that relapse is prevalent.

7. What aspects of depression do you think are vital for our model to be able to explain?

8. What is the most important article in your field regarding major depression?

   (a) Any other authors who could be important?

C.3 Personal questions

If there is time left.

C.4 Closing

Is there anything you would like to add before we end the interview? <<wait for confirmation of the participant>> Thank you very much for all your information! We can now start making our causal map based on the literature, your and a number of other interviews. When we have constructed this causal map, it is important for us to obtain feedback to make sure that we have incorporated the most important factors of depression. Could we come back to you at a later time for this very short feedback moment to ask you for your opinion on the preliminary causal map? << wait for response >> Lastly, would you like to stay informed about the progress of the study? << wait for response >> With that we have come to an end. Once again, we would like to thank you for your time. << Turn off skype >>

Appendix D

Split categorized expert networks

The figues in this appendix are presented in the following order of selection scales: Micro, Meso, Macro, Onset, Maintenance, Relapse.

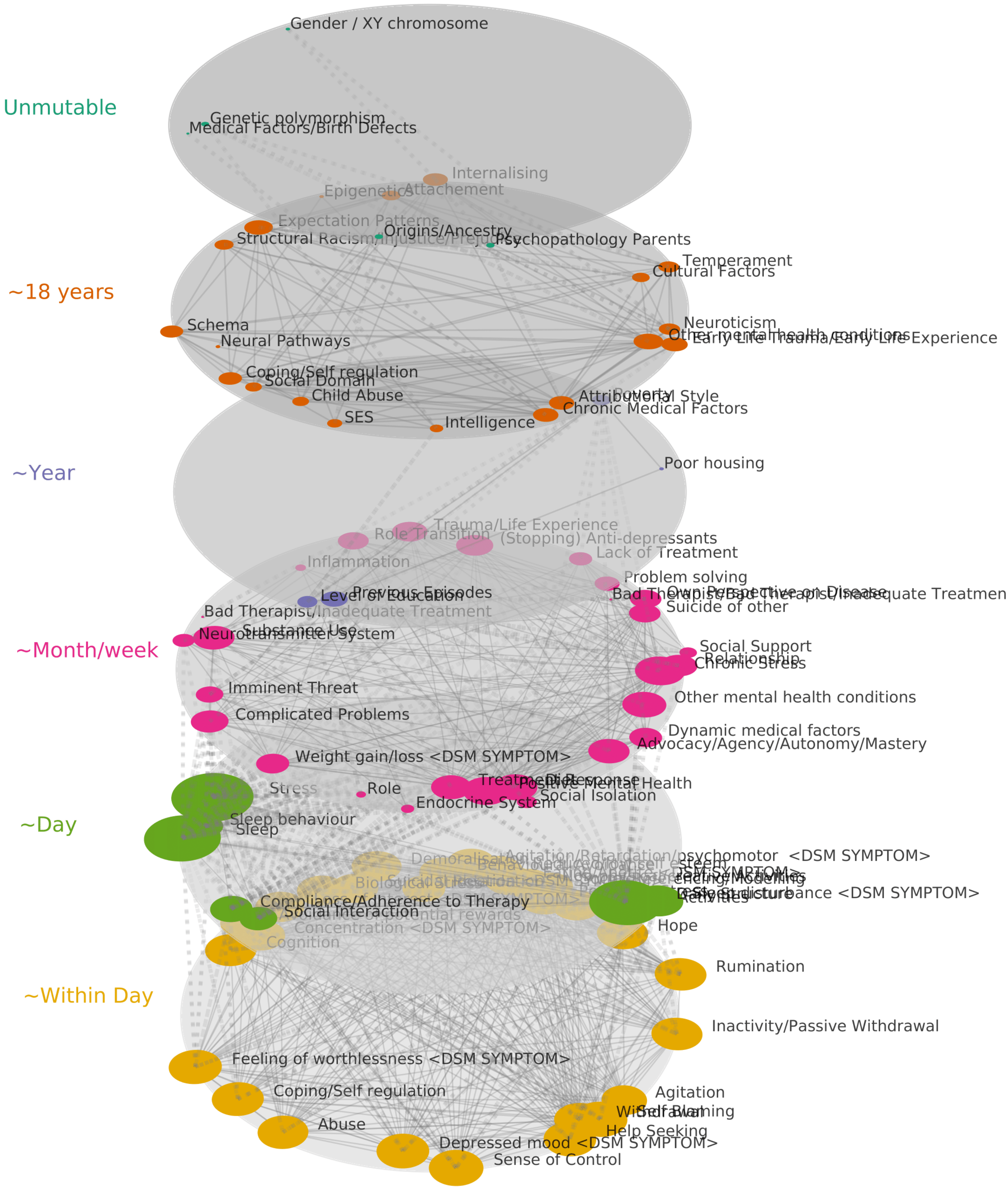


FIGURE D.1: Network with only nodes and edges that work on the spatial **micro** level.

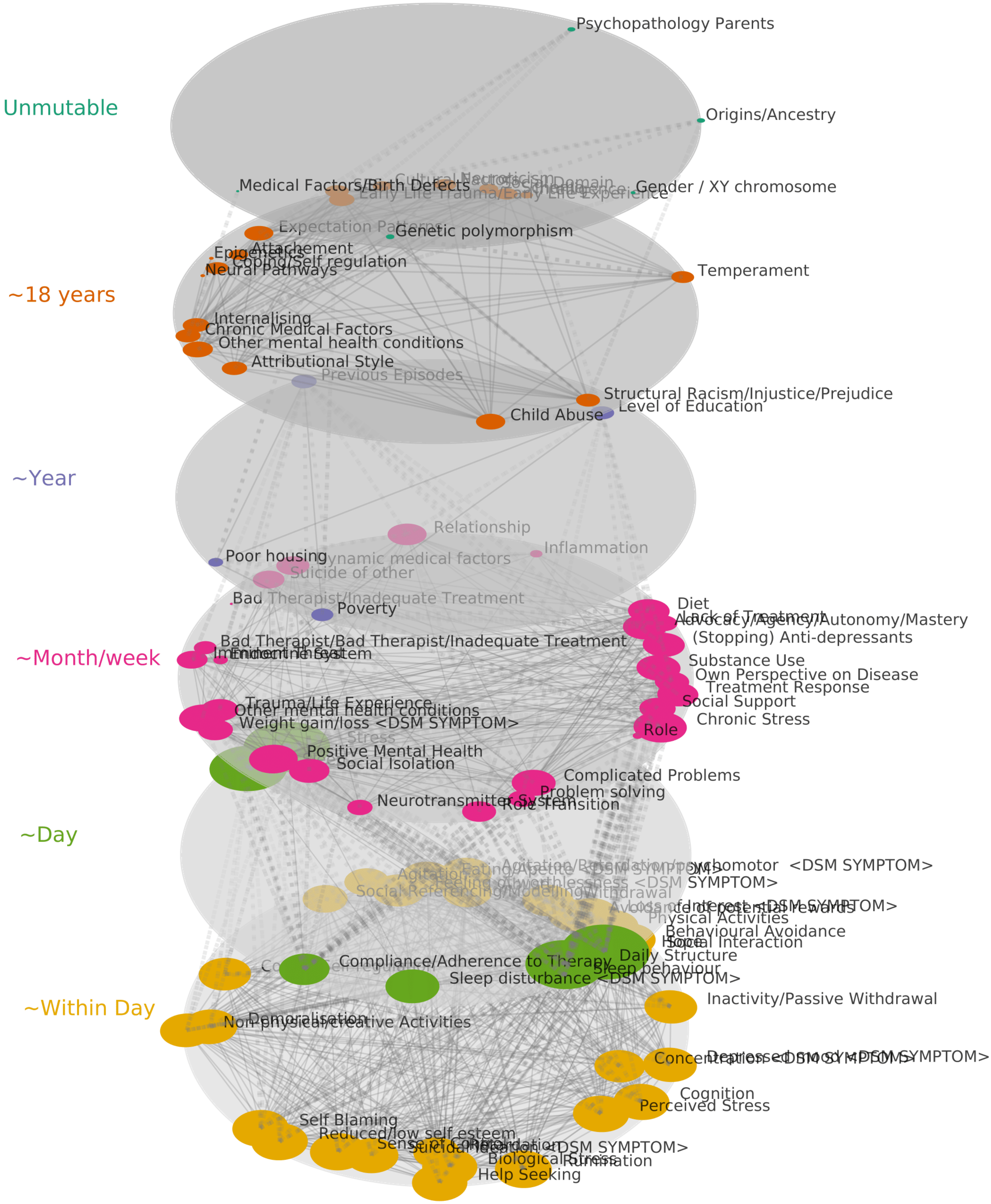


FIGURE D.2: Network with only nodes and edges that work on the spatial **meso** level.

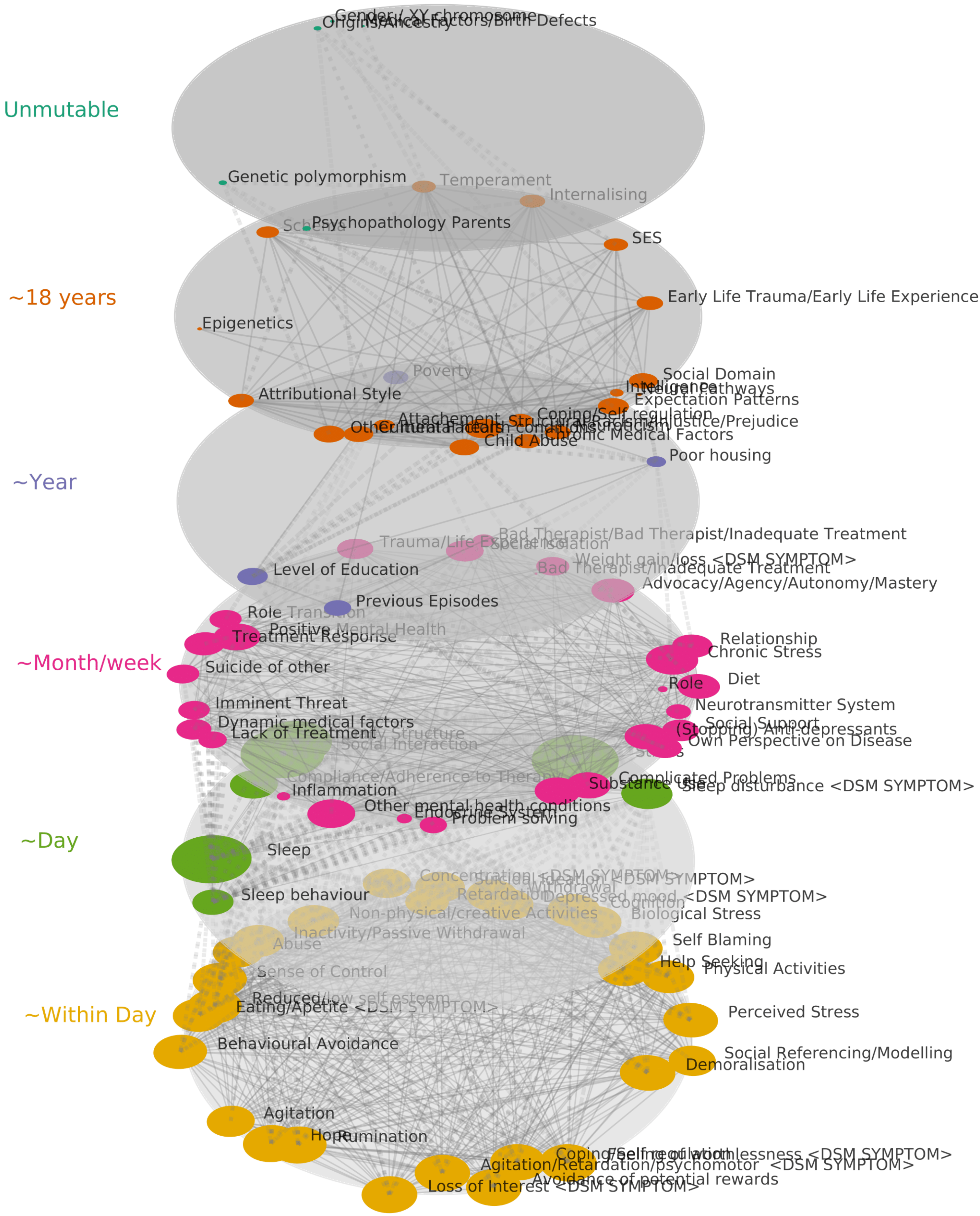


FIGURE D.3: Network with only nodes and edges that work on the spatial **macro** level.

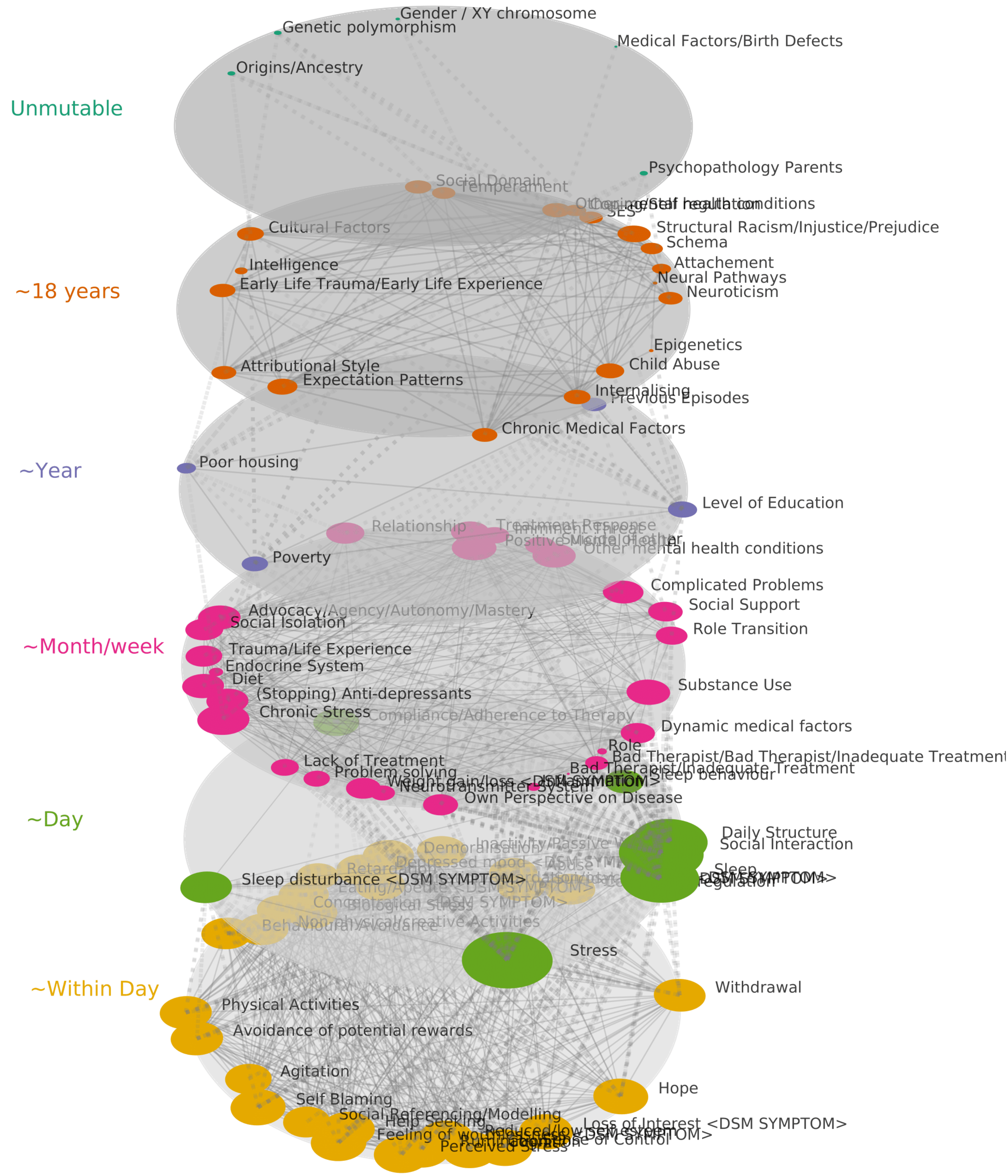


FIGURE D.4: Network with only nodes and edges that work on the temporal level of MDD **onset**.

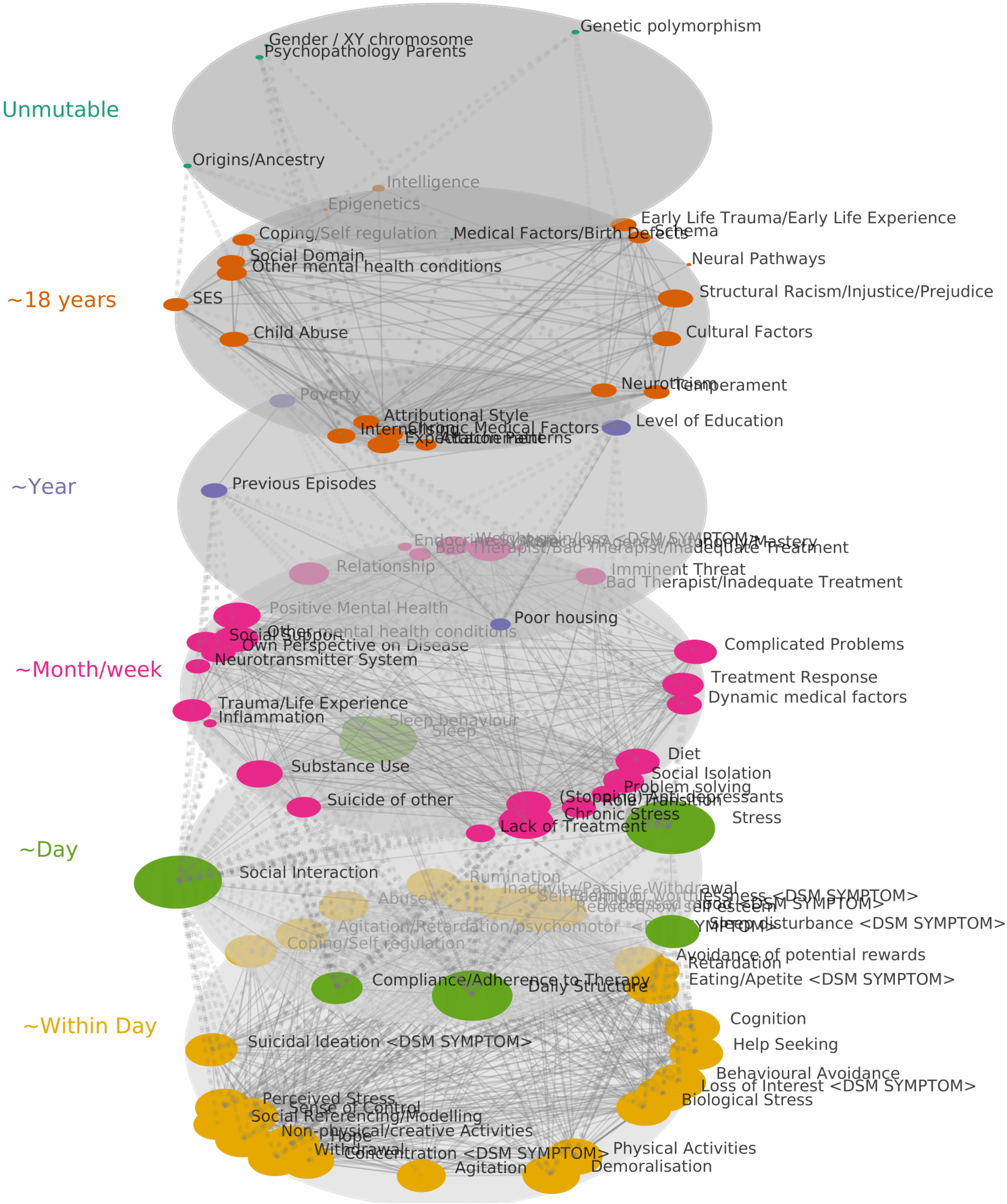


FIGURE D.5: Network with only nodes and edges that work on the temporal level of MDD **maintenance**.

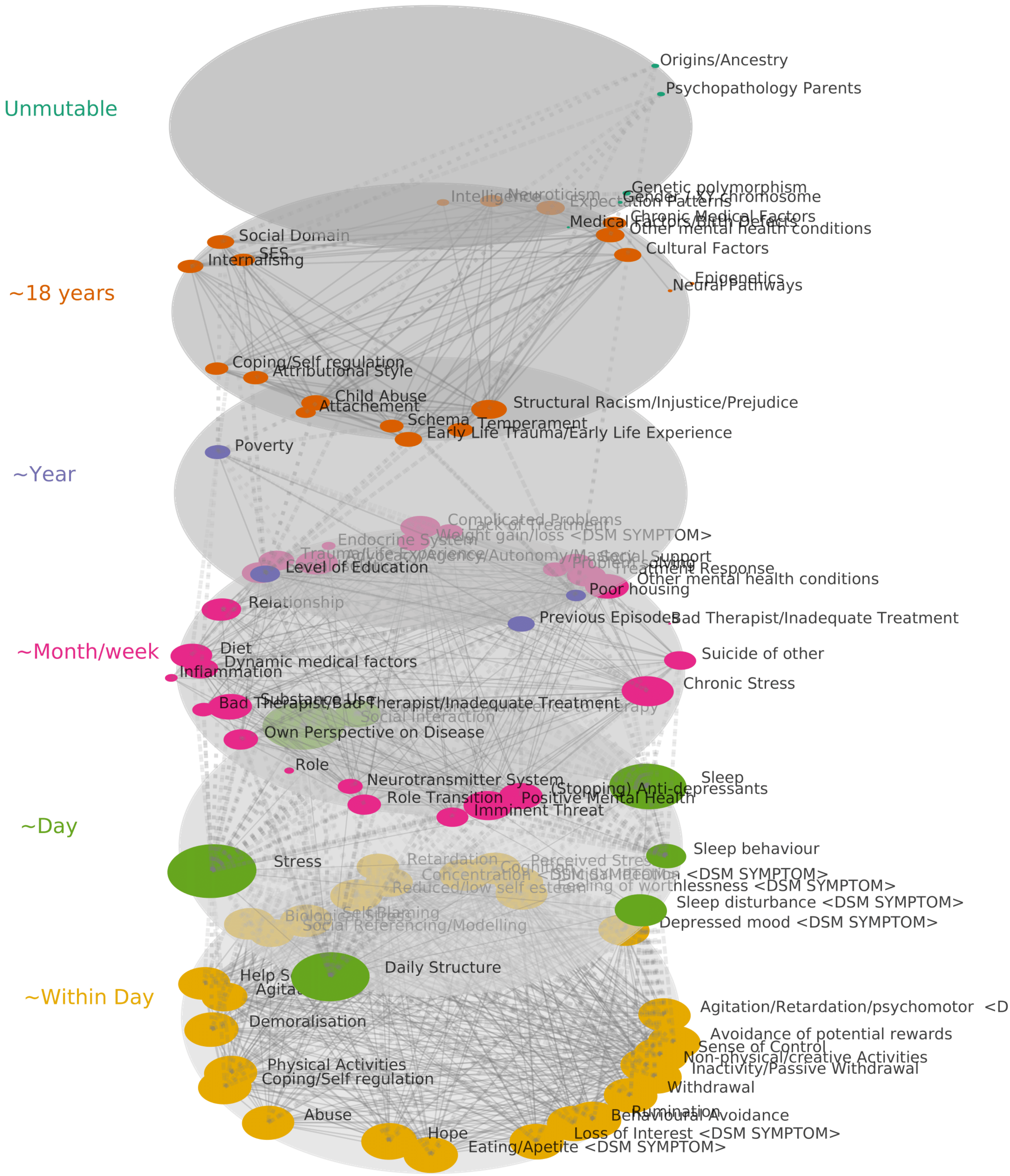


FIGURE D.6: Network with only nodes and edges that work on the temporal level of MDD **relapse**.

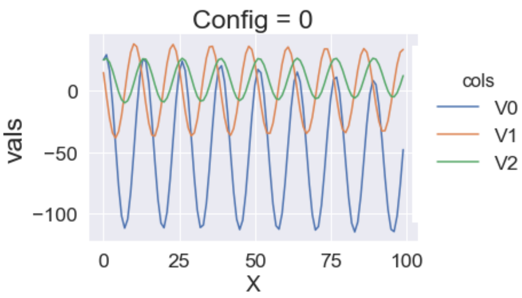
Appendix E

Ordinary Differential Equations

Given in this Appendix are the ordinary differential equations and initial values for each configuration used in chapter 4.

Small networks - 3 variables and 5 relations

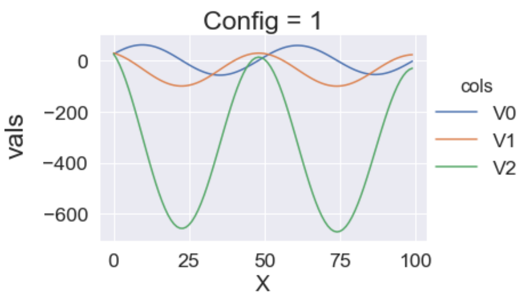
$$\frac{\delta V_1}{dt} = .90V_2 - .03V_3 \tag{E.1}$$
$$\frac{\delta V_2}{dt} = -.11V_1 - .59V_3 \tag{E.2}$$
$$\frac{\delta V_3}{dt} = .23V_2 \tag{E.3}$$



Starting Values: [24.96, 14.69, 24.90]

FIGURE E.1: Network configuration 0

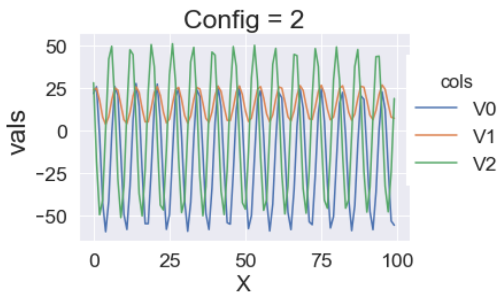
$$\frac{\delta V_1}{dt} = .27V_2 - .03V_3 \tag{E.4}$$
$$\frac{\delta V_2}{dt} = -.13V_1 \tag{E.5}$$
$$\frac{\delta V_3}{dt} = .01V_2 - .71V_1 \tag{E.6}$$



Starting Values: [23.86, 28.71, 25.35]

FIGURE E.2: Network configuration 1

$$\frac{\delta V_1}{dt} = .78V_3 - .01V_2 \tag{E.7}$$
$$\frac{\delta V_2}{dt} = .20V_3 \tag{E.8}$$
$$\frac{\delta V_3}{dt} = -.85V_1 - .92V_3 \tag{E.9}$$

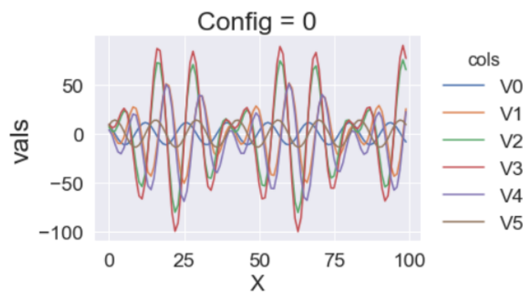


Starting Values: [21.65, 23.71, 27.97]

FIGURE E.3: Network configuration 2

Larger networks - 6 variables and 10 relations

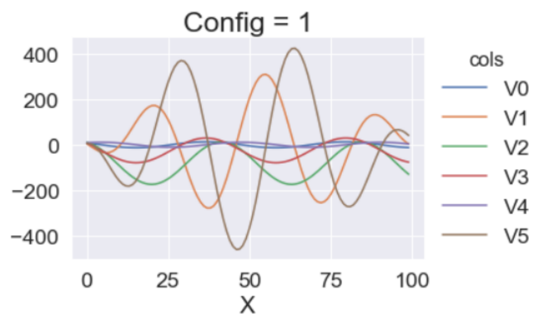
$$\frac{\delta V_1}{dt} = -.37V_6 \tag{E.10}$$
$$\frac{\delta V_2}{dt} = .46V_3 - .59V_6 \tag{E.11}$$
$$\frac{\delta V_3}{dt} = -.82V_2 \tag{E.12}$$
$$\frac{\delta V_4}{dt} = .24V_1 - .93V_2 \tag{E.13}$$
$$\frac{\delta V_5}{dt} = .14V_1 + -.55V_3 - .78V_6 \tag{E.14}$$
$$\frac{\delta V_6}{dt} = .56V_1 \tag{E.15}$$



Starting Values: [8.76, 7.58, 6.81, 9.70, 3.87, 8.74]

FIGURE E.4: Network configuration 0 (6 variables, 10 relations)

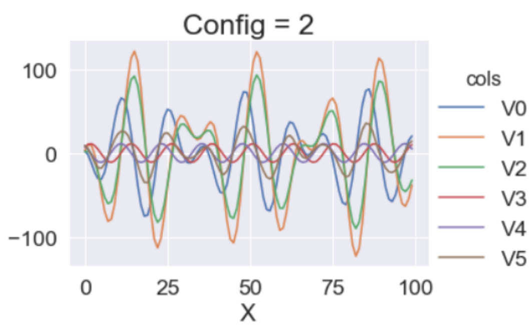
$$\frac{\delta V_1}{dt} = -.17V_5 \tag{E.16}$$
$$\frac{\delta V_2}{dt} = -.92V_5 - .14V_6 \tag{E.17}$$
$$\frac{\delta V_3}{dt} = .67V_1 - .95V_5 \tag{E.18}$$
$$\frac{\delta V_4}{dt} = -.72V_5 \tag{E.19}$$
$$\frac{\delta V_5}{dt} = .12V_1 \tag{E.20}$$
$$\frac{\delta V_6}{dt} = .88V_1 + .31V_2 - .87V_5 \tag{E.21}$$



Starting Values: [7.73, 4.39, 9.89, 7.83, 8.77, 6.27]

FIGURE E.5: Network configuration 1 (6 variables, 10 relations)

$$\frac{\delta V_1}{dt} = .5V_5 - .17V_2 - .78V_4 \tag{E.22}$$
$$\frac{\delta V_2}{dt} = .29V_1 + .84V_6 \tag{E.23}$$
$$\frac{\delta V_3}{dt} = .49V_1 \tag{E.24}$$
$$\frac{\delta V_4}{dt} = .5V_5 \tag{E.25}$$
$$\frac{\delta V_5}{dt} = -.51V_4 \tag{E.26}$$
$$\frac{\delta V_6}{dt} = .96V_3 - .84V_2 \tag{E.27}$$



Starting Values: [2.70, 3.12, 1.80, 7.09, 8.68, 9.09]

FIGURE E.6: Network configuration 2 (6 variables, 10 relations)

Appendix F

Parameter fitting

Given in this Appendix are the parameter values found by the genetic algorithms and the distributions kernel density function (KDE) for each configuration used in chapter 4. Here the best solutions parameters are represented by the red vertical line, the steelblue vertical line represents the top of the KDE (i.e. the value that is theoretically most prevalent) and the real parameter value is represented by the purple vertical line.

Small networks - 3 variables and 5 relations

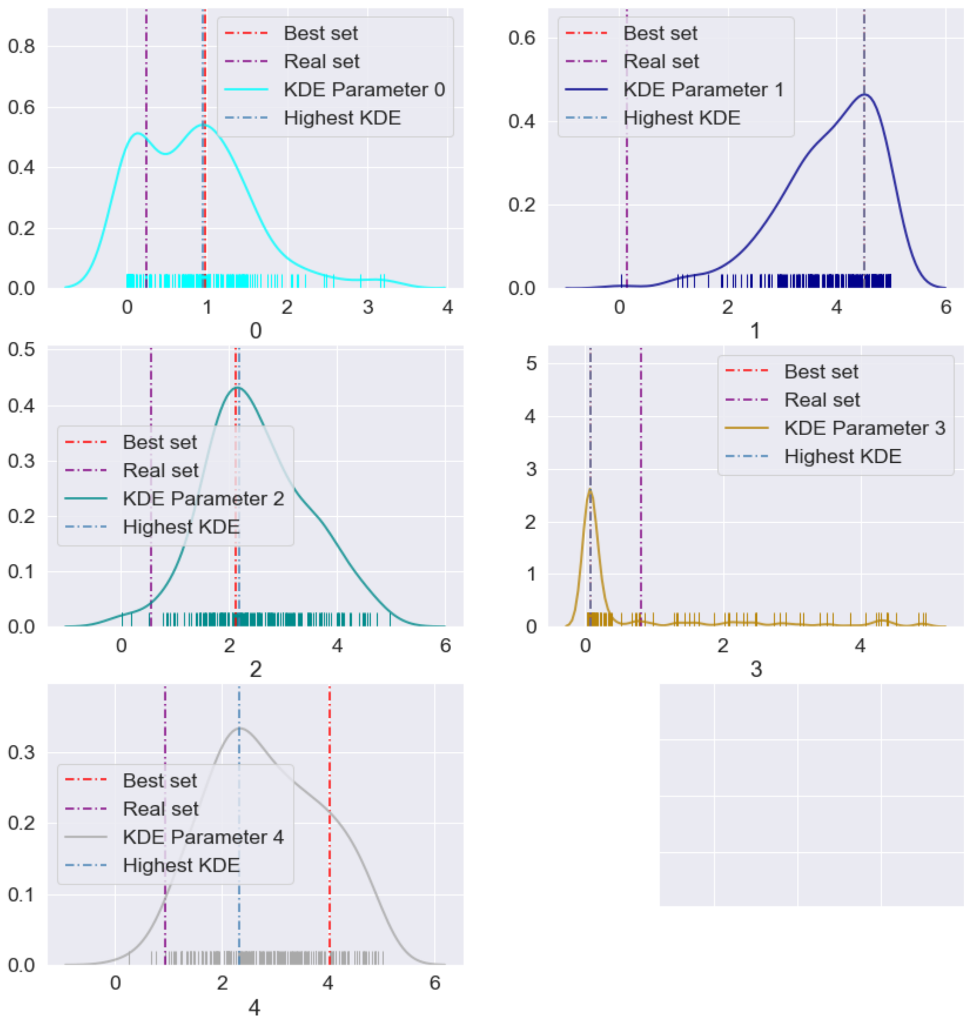


FIGURE F.1: Distributions from the best performing parameters over 200 evolutionary algorithm runs for configuration 0.

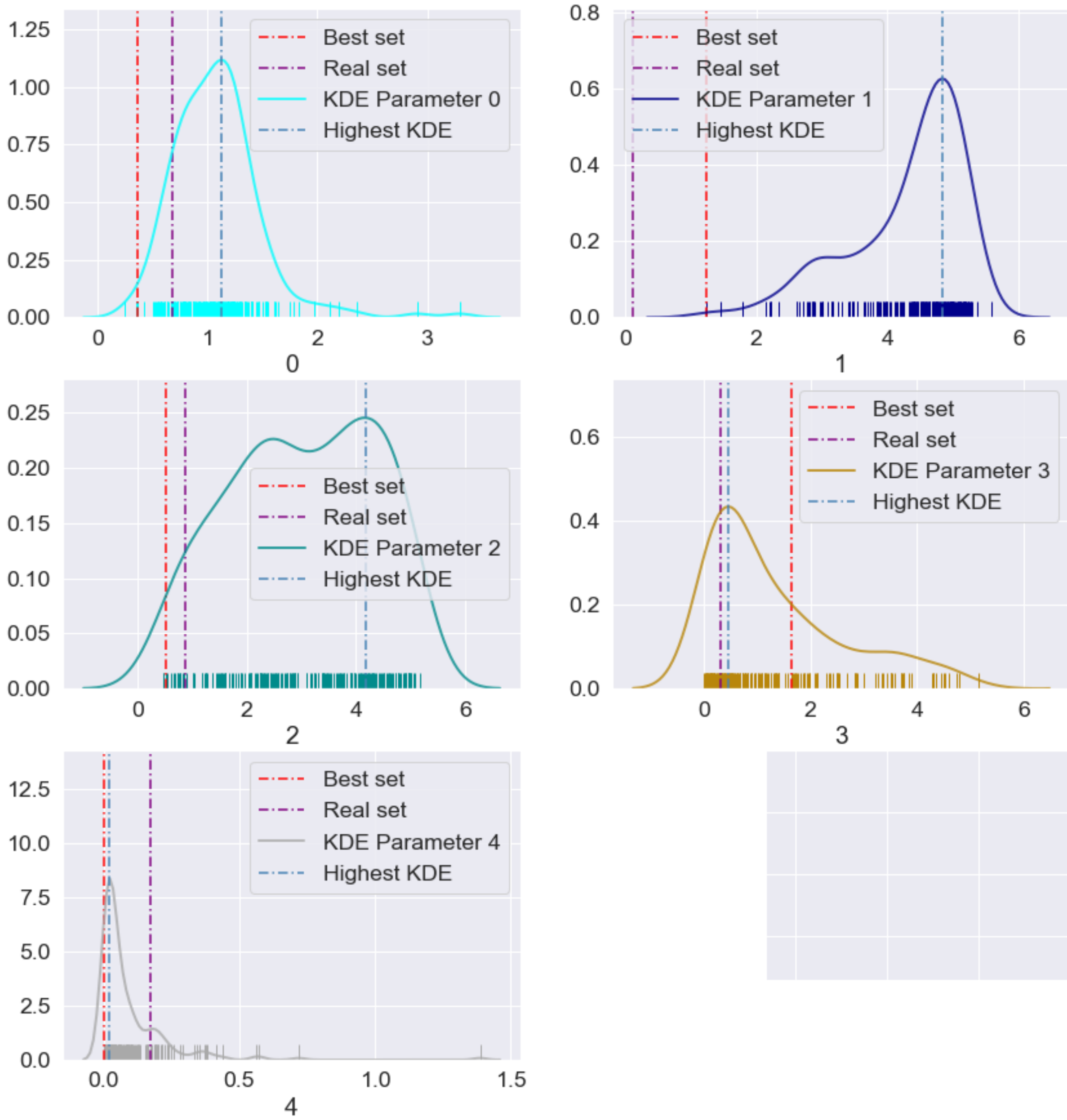


FIGURE F.2: Distributions from the best performing parameters over 200 evolutionary algorithm runs for configuration 1.

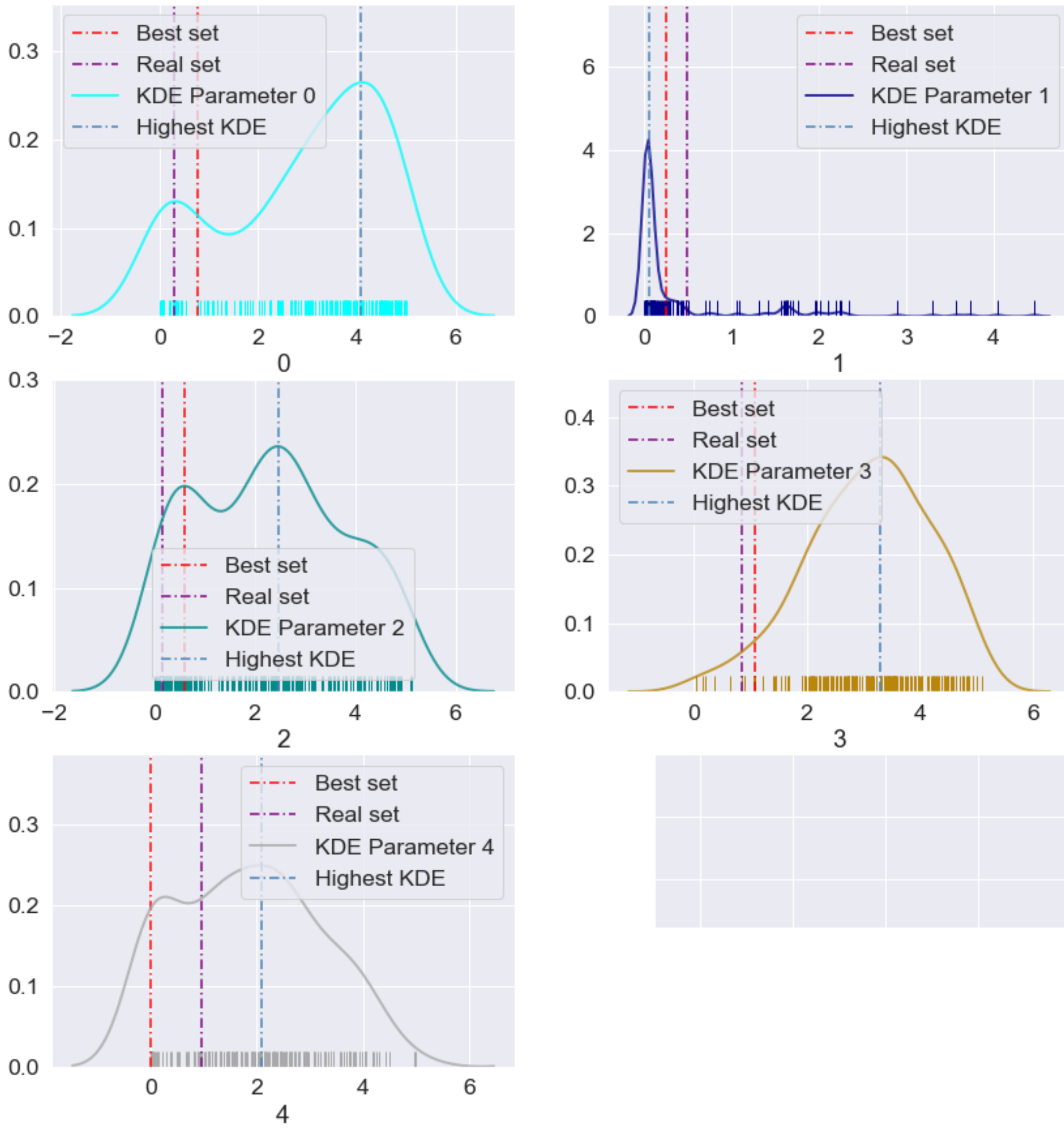


FIGURE F.3: Distributions from the best performing parameters over 200 evolutionary algorithm runs for configuration 2.

Larger networks - 6 variables and 10 relations

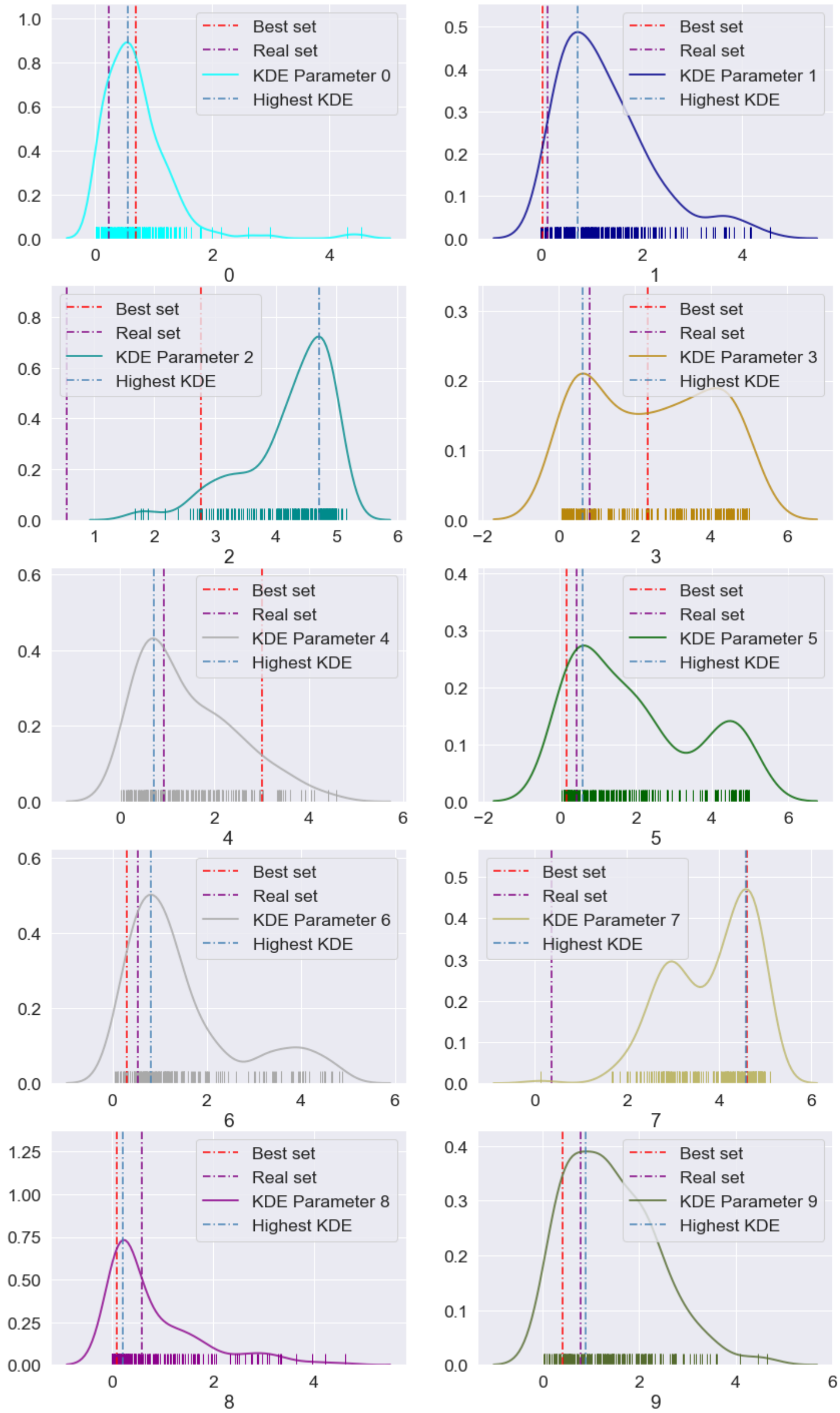


FIGURE F.4: Distributions from the best performing parameters over 200 evolutionary algorithm runs for configuration 0.

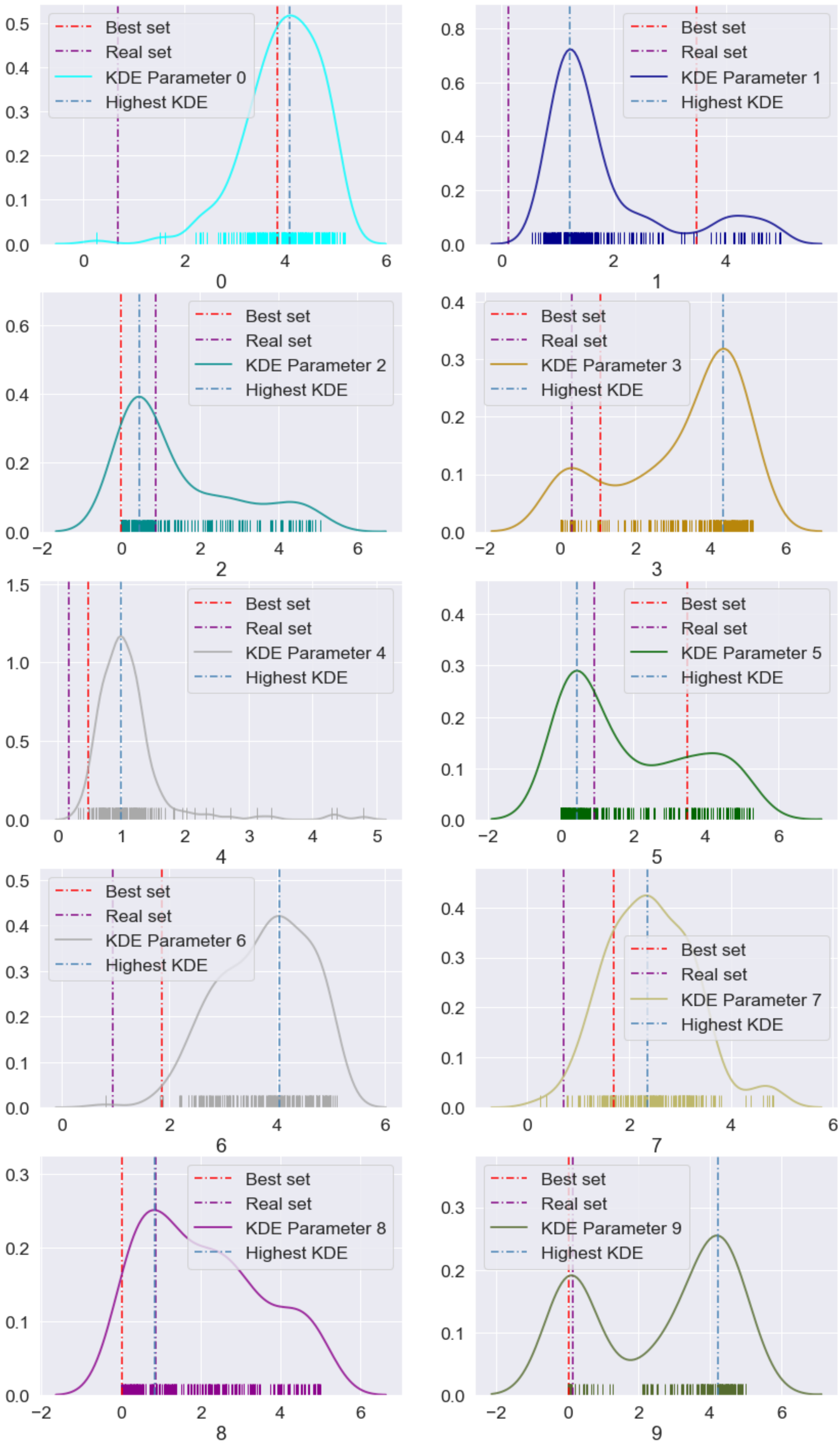


FIGURE F.5: Distributions from the best performing parameters over 200 evolutionary algorithm runs for configuration 1.

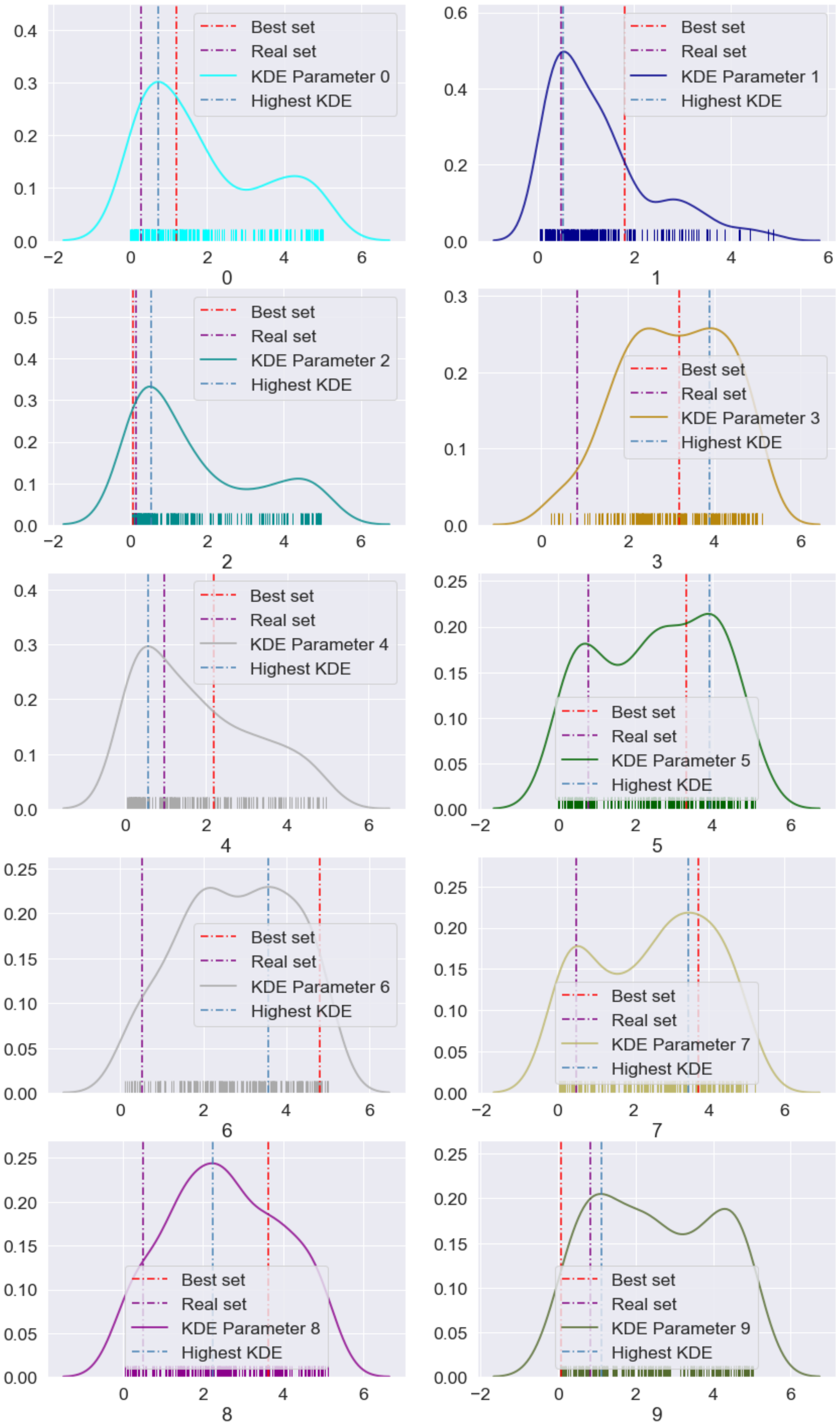


FIGURE F.6: Distributions from the best performing parameters over 200 evolutionary algorithm runs for configuration 2.

Appendix G

CHATEL

G.1 Complexity HTTP Automated Teammate Experimentation Lab

A high level python based website to facilitate conducting qualitative expert interviews regarding a Complex System problem This project is meant for making it easier to construct qualitative models from qualitative knowledge through structured interviews and collaborative work. The website guides a researcher in a coherent and incremental process of defining a complex problem by providing a multiphase structure. This structure first lets the researcher concisely define the problem and work out several variable values concerning the problem statement. These steps can be found in section Step- by- step conceptualization of your project. Eventually, the website will require the researcher to answer most of these questions in the dashboard page to configure the rest of the application as these are used throughout the application.

This README will explain how to configure a VPS (Virtual Private Server) to act as a web server, serve the website and configure its settings (see section Configuring a Virtual Private Server (VPS)). As setting up the VPS itself will take about an hour and the costs are 3 euros a month for hosting, a researcher investigating a complex system problem can worry over substantive questions regarding the problem statement instead of dealing with issues of visualization, ways of data gathering and structuring and overall structuring of the experimental layout. This Application is, as the name suggests, a web based tool designed for experimenting with qualitative views of different experts regarding a complex systems problem..

G.2 Step by step conceptualization of your project

1. **Define the Problem**

   - Prior to all research it is off course necessary to get a grasp of what the current scope of knowledge is surrounding the problem statement. To achieve this perhaps some kind of **Meta-analysis** can help with these questions:
   - What don't we know?
   - And maybe more importantly, what can be found in literature (i.e. what do we know)?
   - By providing a web-structure, we force ourselves to structure problems as we write it for a broad audience. Even as its intended target group might not need a step by step introduction of the problem and its importance, it does help to

think about what we are actually trying to accomplish. Upon writing the content of the introduction page we define what it is that we don't know, and what it is that we do know.

2. **Defining parameters**

- What type of factors are we dealing with?
- Variable, Stock, Constant, etc.
- At what spatial or temporal scale do these factors exist?
- Spatial: Biology, Psychological, Interpersonal .. Global etc
- Temporal: Milliseconds, Seconds, Minutes .. Years, Lifetime etc
- At what scale do their interactions take place?

3. **Interactions**

- What factors interact with each other?
- What is the nature of the interaction?
- If one increases does the other also increase or decrease?
- How strong are these interactions?
- How sensitive are the nodes to change?
- At what point in the process of the problem are the factors and their interactions of importance?
- Onset, Maintenance, relapse in a disease process
- Do the interactions need operators?
- A leads to B but only IF C is also active
- Operators like: IF, IFF, OR, XOR, AND

4. **Experts**

- In order to obtain a data set beyond the scope of literature by merging quantitative knowledge found in literature and qualitative knowledge obtained from experts, we need to have expert knowledge representing all fields relating to the problem set. It is important to think about the different perspectives needed to form a complete overview of the problem statement and identify key figures that are linked to the field scientifically, experience based, policy wise etc.
- For this we need some information of each expert:
- What is their affiliation with the problem statement? (Experience Expert, Researcher, Policy maker, etc.)
- In discipline are their working? (Biology, public health, physics, etc.)
- What is their specialization? (epidemiology, genetics, methodology, social media, etc)
- Also ask the experts whether they would be open for further feedback moments in unclear situations and whether they want to be named as a participant. As it might affect transparency and reliability of the qualitative data, a default term of conditions would be to state that any participant needs to agree with their name being associated with the expert list.

G.3 Configuring a Virtual Private Server (VPS)

Though this might seem daunting, and perhaps for some technologically challenging, we provide a step by step tutorial of how to configure and publish the website within the hour and for a cost of about 3 euros per month hosting costs (if Hetzner is chosen as VPS host). A VPS was chosen in order to cut budget costs and retain a lot of freedom to personalize, play around and alter anything you want. This tutorial consists of several parts guiding you from start (reading this README!) to finish (a fully functional website).

G.3.1 Choosing and configuring VPS service

As our web application will attract only few users (mostly being your expert pool) we don't need a lot of RAM, Disk space, CPU etc. For our purposes a 2 GB RAM with 20 GB Disk space, 1 CPU and 20 TB traffic is more than enough. Greenhost, DigitalOcean or Hetzner for example provide fine VPS services that let you do whatever you want with them for a small price. Out of these three, Hetzner is the cheapest with good reviews, so throughout this tutorial we will use this VPS provider. Throughout this tutorial three different notation styles will be used in order to clarify what needs to be typed in where. Upon talking about buttons that need to be clicked and text outside the Linux terminal it will be presented between "quotations marks". Within the Linux terminal we will continuously use the following box:

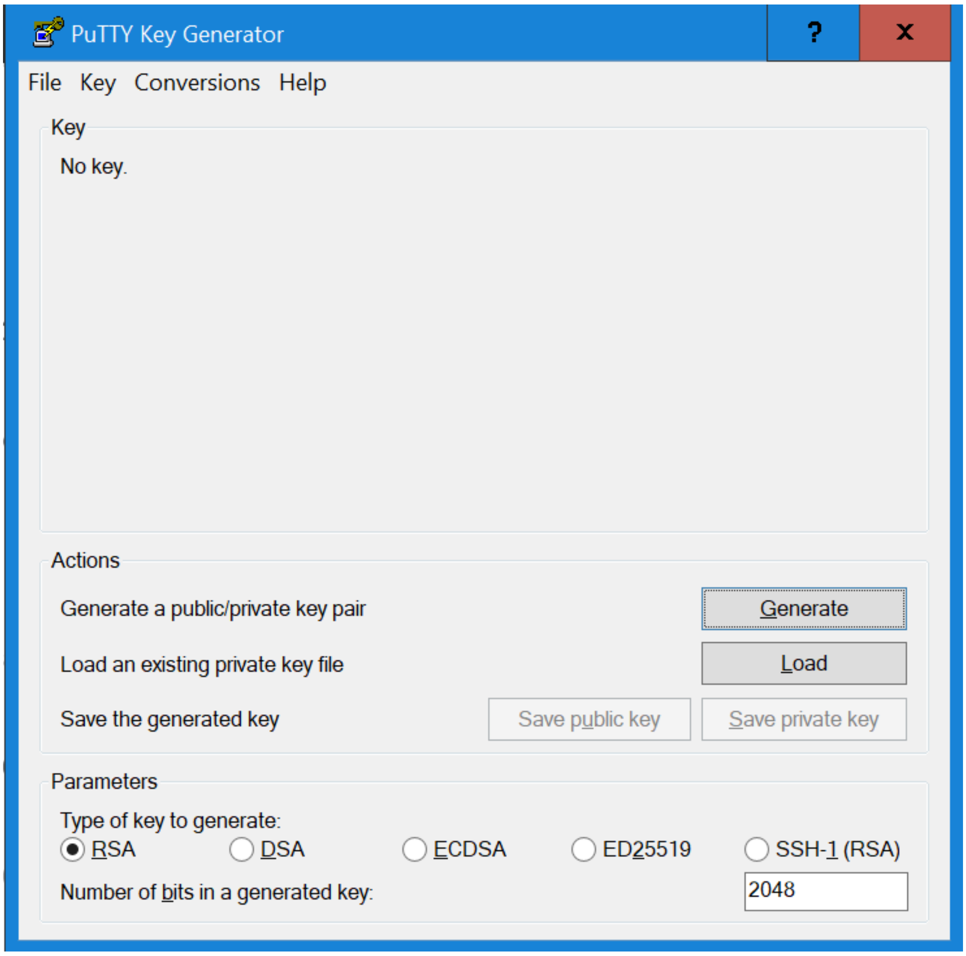
```
Code that is in here is meant as a terminal command!
```

Creating a VPS

1. At the Hetzner website make an account, don't forget to verify via your email.

2. Now that we have an account we first need to create an SSH key in order to be able to safely log in to the server we're about to create.

   1. Upon using a Windows' computer I would recommend using PUTTY. After installing this press the windows button on your keyboard, type **"puttygen"** and press enter.
   2. When the PUTTY key generator has been opened, press the **"generate"** button and move your mouse around within the PUTTY window. The program uses your mouse movements to generate a random key (as it is very hard to figure out what movements someone made during the creation of this key it is very safe).



3. Press the **"Safe private key"** button and safe this somewhere on your computer where you can find it (you'll need this later to log into your VPS!).

4. Copy the text within the box underneath "Public key for pasting into OpenSSH authorized_keys file:", the text should look a bit like this:

```
ssh-rsa
AAAAB3NzaC1yc2EAAAABJQAAAQEAiFSzr0GXlpxdhDinznQyha5
IZJ1DjtKGVu/woBnYw+OK0EzSj3pmmld233hetCekqCCGPpLG4X
/rdy1e2shPXlBjvAzb8seMoF+YU+tVWtVfjm/kUQyu5WFRBNf5x
NReFFfmko0640avJZmXARFTMQkZRqUn0hrkhK+brubModRyQn6g
vpewDLf0RcgyPCC4SNmFCyI9c3Dvfvs1rcdqy673C0c81lwpIgC
k41EByZLJ0Cx20e5gCerm/KoM0M8opeKmd9mnRbArXaTmSe4zE7
5AoDtvZiVtVyXeXlnnsKbthTQdN0ET03Q2hiJJU/Pxr2FQn1FmA
qQGZ919bu8f0w
== rsa-key-20190330
```

3. Now that we have our SSH key and a Hetzner account has been created, go to the Hetzner website, log in, and click on the servers tab. Here click **"ADD SERVER"**.

4. Here we see some choices given to us regarding the configuration of the server. Pick a location, Ubuntu 16.04 and the 3.01/mo payment plan with 1vCPU, 2 GB RAM, 20 GB SSD and 20 TB Traffic. Leave the rest blank, except point 6 where we click on **"Add SSH key" and paste our public SSH key** and 7 in which you can give your server a name. This can be any name, as long as you can remember it.

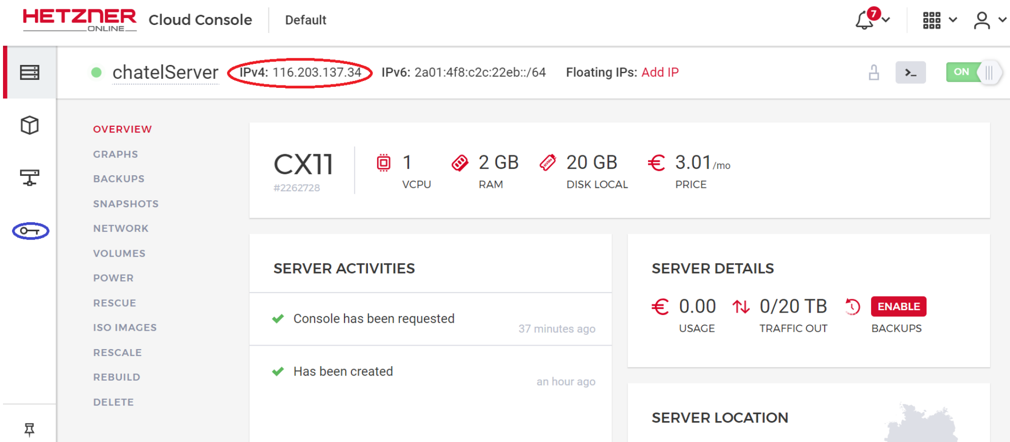
5. Click create and buy, and you'll have your very own VPS in a matter of seconds!

6. Now that we have our VPS set up, we still need to access and configure it! To do this, we'll need to access the VPS for which we can also use PUTTY, but I personally think it is a bit easier to use winSCP for this as it also provides a folder structure instead of just a terminal. You can download winSCP here.

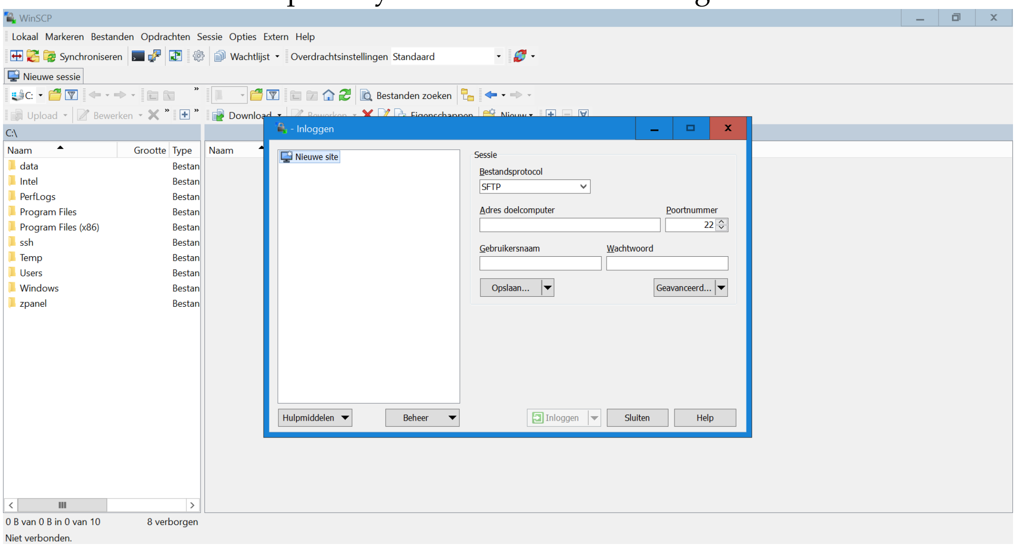
Accessing the VPS

7. Once winSCP is installed, we can access the new VPS.

   1. Go to your Hetzner server and copy the IP address (if you look at the overview page, you should be able to see your IP address next to IPv4 on top of the page). The IP address looks like 116.203.137.34 but with some different numbers. From now on if I talk about IP address or your_ip_address, this is what I'm talking about.

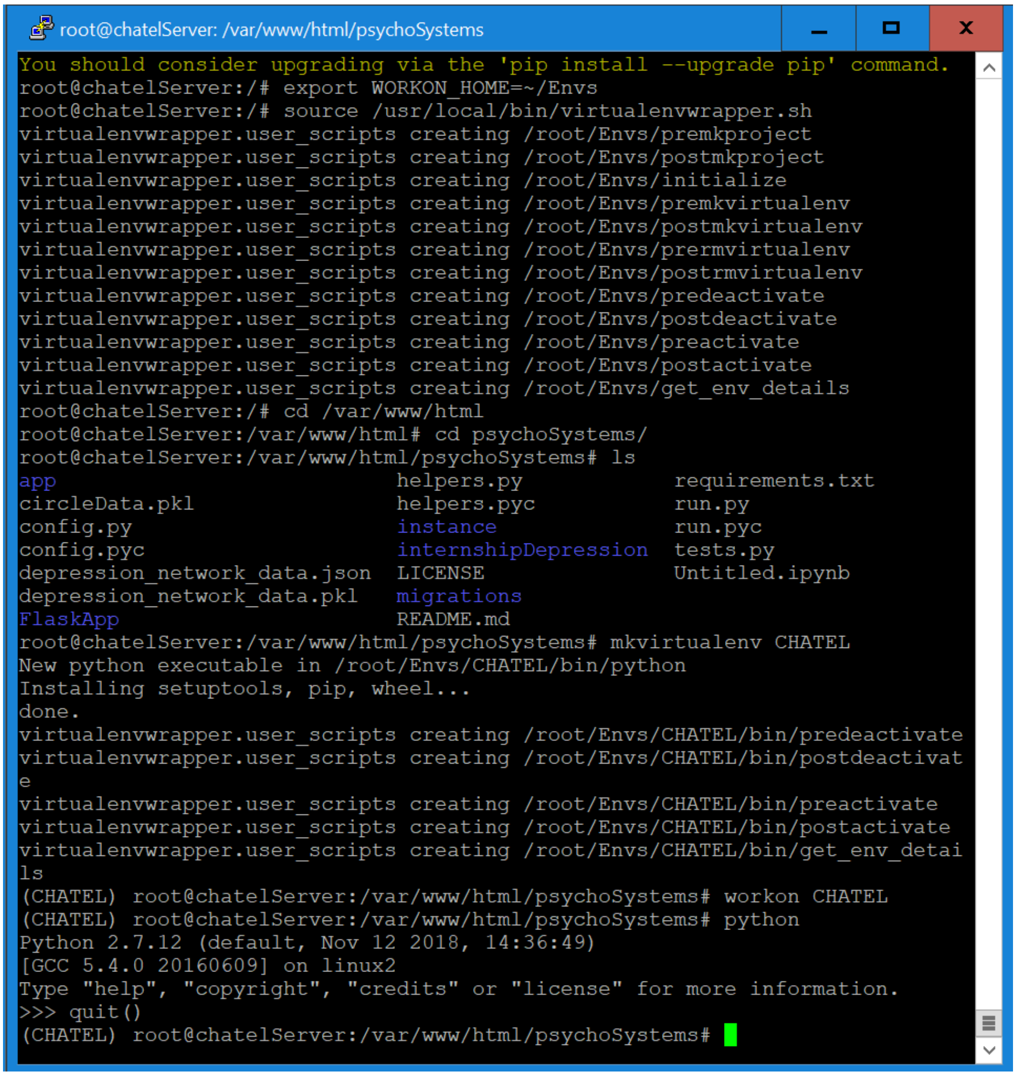


   2. Press the Windows button and type **"winSCP"** and press enter.
   3. Now that winSCP is opened you should see something like:



   4. Keep "File protocol" (1) on SFTP, paste your IP address in "Address targetcomputer" (2) , as "Username" (3) type root.
   5. Click on **"Advanced"** (4).
   6. To the left under the SSH tab, you can see **"Authentication"** (5), press here and then you'll be able to browse (click on the button with three dots (6)) to your private key that we saved in step 2.3.

7. Press **"OK"** (7), then **"Save"** (8) where you can enter a name. I recommend it being called the same as your server name (step 2.4).
8. Finally, press **"Log in"** (9) and we can access our VPS!
9. As you can see, there are lots of buttons visible and a file structure as you are used to seeing using windows. We will first just use the terminal to do some configurations. Press the **"Open Session in PUTTY"** (10), it's that small button looking like two computers with a lightning strike in between, just to the right of "synchronize".
10. If all went well, you should be able to see a terminal opening looking like this:



11. As of this moment, we will work a lot in the terminal which might seem a bit intimidating if you're not used to working with this, but I'll guide you every step of the way so it'll be fine!

Configuring your VPS as a webserver To be able to host your own web server, we need to install some things to be able to make your VPS visible to the internet, and therefore your experts and other people! We're about to install LAMP, which is an abbreviation for Linux operating system, the Apache HTTP Server, the MySQL relational database management system (RDBMS), and the PHP programming language. For this part I will follow another tutorial which you can find here but for our purposes it will be tweaked here and there as we don't need WordPress.

8. First we need to make sure that our server is up-to-date by running the following commands:

```
apt-get update && apt-get upgrade -y
```

9. Now that our server is up-to-date, we'll move on by installing Apache2 which is one of the most popular and widely used web servers as it is both fast and secure. Run the following command:

```
apt-get install apache2 -y
systemctl start apache2
systemctl enable apache2
```

To check whether this has worked, you can go open your web browser and enter your server IP address (same one as in step 7.1!). If Apache has successfully been installed, you should see the default Apache welcome page.

10. After Apache, we come to the next letter in the LAMP abbreviation, the M that stands for MySQL Database server. Here we will eventually store all information regarding the website, our factors, relations, expert profiles and even the contents of the website itself! To install MySQL type the following command:

```
apt-get install mysql-server -y
```

During the installation you will be prompted to enter a password for the MySQL root user (which is you!), so make sure it's a strong password nobody else could guess, press enter and re-enter the same password.

11. For security reasons lets configure the MySQL database by typing the following:

```
mysql_secure_installation
```

Now the program will ask your password again, following some other questions.

1. It asks to install a password validation plugin so users are required to enter strong passwords. Let's answer `y`!
2. You can choose how picky the validation system is going to be, lets pick the safest so `2`.
3. It'll prompt you if you want to change your root password by giving its security score. If you are happy with this score type: `n`
4. Remove anonymous users? `y`
5. Disallow root login remotely? As we later want to use phpMyAdmin, this is an `n`.
6. Remove test database and access to it? As we don't need it, `y`.
7. Reload privilege tables now? `y`, and that's it!

12. Now that we answered all the questions, we will enable the database to start automatically upon startup of the server by giving the following command:

```
systemctl start mysql
systemctl enable mysql
```

13. Now that MySQL has been installed, we can go ahead with installing PHP by typing the following command:

```
apt-get install php7.0 libapache2-mod-php7.0 php7.0-mysql php7.0-curl
php7.0-mbstring php7.0-gd php7.0-xml php7.0-xmlrpc php7.0-intl
php7.0-soap php7.0-zip -y
```

To test if PHP has been installed successfully we can create a file with the nano text editor:

```
nano /var/www/html/info.php
```

Enter the following content inside the file:

```
<?php
phpinfo();
?>
```

and save it by pressing "ctrl+x" then y and lastly press "enter" and restart the server:

```
systemctl restart apache2
```

If all went well, you can again go to your IP address in the browser as in step 9, but this time followed by "/info.php". Like so, **http://your_ip_address/info.php** which will show the page as shown in figure G.1.

As this file is of no further use to our website, we can safely remove it by typing the following:

```
rm /var/www/html/info.php
```

If you now revisit **http://your_ip_address/info.php**, you'll see that it will return a Not Found error.

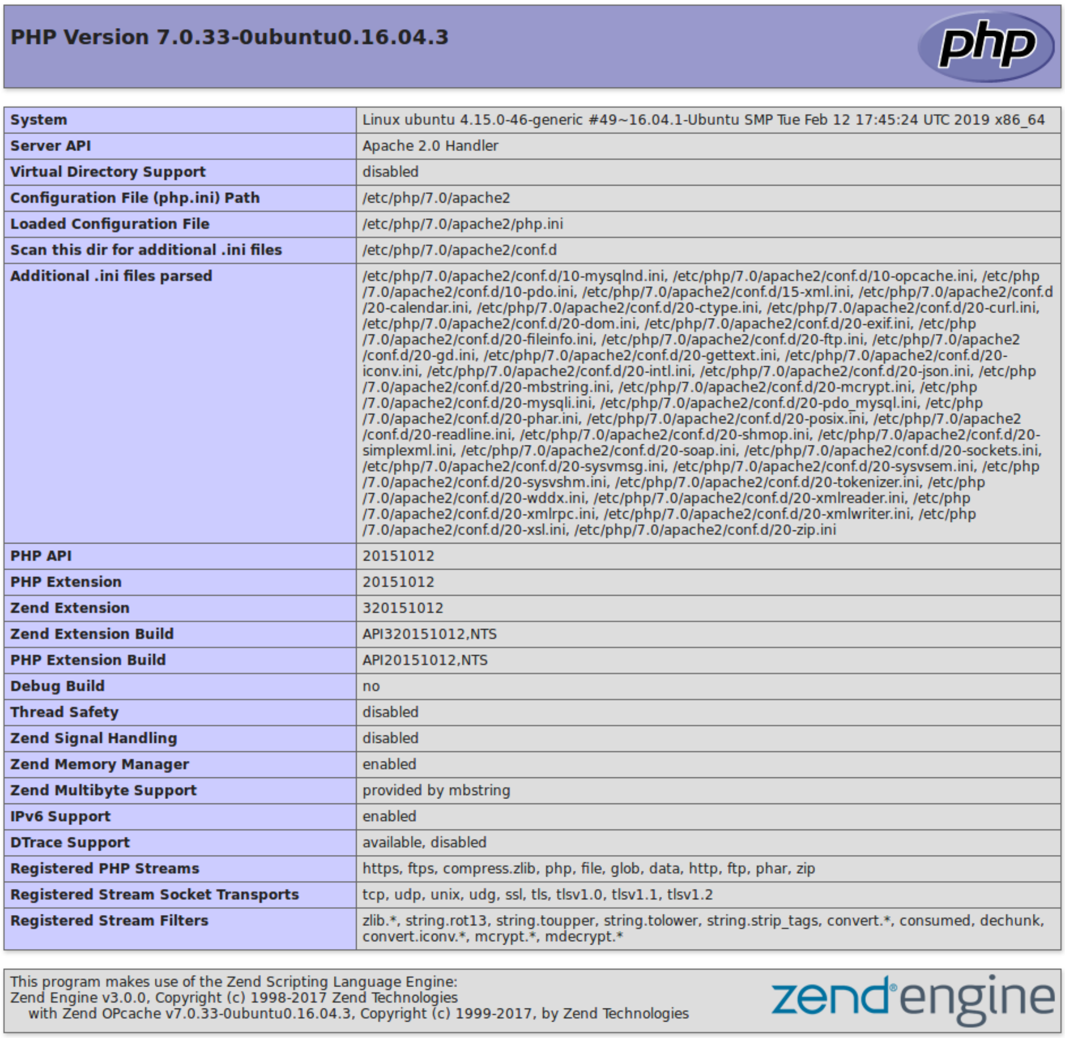
14. Now that we've installed the LAMP software, we would like to have a user interface to eventually work with regarding the database. This will make loading some defaults and debugging if a problem occurs easier. For this we will use phpMyAdmin by entering the following command:

    ```
    apt-get install phpmyadmin -y
    ```

    By doing this you will be prompted (fig G.2 to select a web server to configure (like in the figure underneath), hit the **"space bar"** to select Apache2 and Enter to confirm and continue (don't forget the space bar to actually select the apache2 option or else it won't work). On the next screen, select YES to configure a database for phpMyAdmin with dbconfig-common. And finally set your password for phpMyAdmin.

15. To check whether installing phpMyAdmin has worked, go to
    http://your_ip_address/phpmyadmin
    and you should see the following page depicted in figure G.3.

Your Username is "root", and the password is the password you've set in the previous step.



PHP Version 7.0.33-0ubuntu0.16.04.3

System	Linux ubuntu 4.15.0-46-generic #49~16.04.1-Ubuntu SMP Tue Feb 12 17:45:24 UTC 2019 x86_64
Server API	Apache 2.0 Handler
Virtual Directory Support	disabled
Configuration File (php.ini) Path	/etc/php/7.0/apache2
Loaded Configuration File	/etc/php/7.0/apache2/php.ini
Scan this dir for additional .ini files	/etc/php/7.0/apache2/conf.d
Additional .ini files parsed	/etc/php/7.0/apache2/conf.d/10-mysqlnd.ini, /etc/php/7.0/apache2/conf.d/10-opcache.ini, /etc/php/7.0/apache2/conf.d/10-pdo.ini, /etc/php/7.0/apache2/conf.d/15-xml.ini, /etc/php/7.0/apache2/conf.d/20-calendar.ini, /etc/php/7.0/apache2/conf.d/20-ctype.ini, /etc/php/7.0/apache2/conf.d/20-curl.ini, /etc/php/7.0/apache2/conf.d/20-dom.ini, /etc/php/7.0/apache2/conf.d/20-exif.ini, /etc/php/7.0/apache2/conf.d/20-fileinfo.ini, /etc/php/7.0/apache2/conf.d/20-ftp.ini, /etc/php/7.0/apache2/conf.d/20-gd.ini, /etc/php/7.0/apache2/conf.d/20-gettext.ini, /etc/php/7.0/apache2/conf.d/20-iconv.ini, /etc/php/7.0/apache2/conf.d/20-intl.ini, /etc/php/7.0/apache2/conf.d/20-json.ini, /etc/php/7.0/apache2/conf.d/20-mbstring.ini, /etc/php/7.0/apache2/conf.d/20-mcrypt.ini, /etc/php/7.0/apache2/conf.d/20-mysqli.ini, /etc/php/7.0/apache2/conf.d/20-pdo_mysql.ini, /etc/php/7.0/apache2/conf.d/20-phar.ini, /etc/php/7.0/apache2/conf.d/20-posix.ini, /etc/php/7.0/apache2/conf.d/20-readline.ini, /etc/php/7.0/apache2/conf.d/20-shmop.ini, /etc/php/7.0/apache2/conf.d/20-simplexml.ini, /etc/php/7.0/apache2/conf.d/20-soap.ini, /etc/php/7.0/apache2/conf.d/20-sockets.ini, /etc/php/7.0/apache2/conf.d/20-sysvmsg.ini, /etc/php/7.0/apache2/conf.d/20-sysvsem.ini, /etc/php/7.0/apache2/conf.d/20-sysvshm.ini, /etc/php/7.0/apache2/conf.d/20-tokenizer.ini, /etc/php/7.0/apache2/conf.d/20-wddx.ini, /etc/php/7.0/apache2/conf.d/20-xmlreader.ini, /etc/php/7.0/apache2/conf.d/20-xmlrpc.ini, /etc/php/7.0/apache2/conf.d/20-xmlwriter.ini, /etc/php/7.0/apache2/conf.d/20-xsl.ini, /etc/php/7.0/apache2/conf.d/20-zip.ini
PHP API	20151012
PHP Extension	20151012
Zend Extension	320151012
Zend Extension Build	API320151012,NTS
PHP Extension Build	API20151012,NTS
Debug Build	no
Thread Safety	disabled
Zend Signal Handling	disabled
Zend Memory Manager	enabled
Zend Multibyte Support	provided by mbstring
IPv6 Support	enabled
DTrace Support	available, disabled
Registered PHP Streams	https, ftps, compress.zlib, php, file, glob, data, http, ftp, phar, zip
Registered Stream Socket Transports	tcp, udp, unix, udg, ssl, tls, tlsv1.0, tlsv1.1, tlsv1.2
Registered Stream Filters	zlib.*, string.rot13, string.toupper, string.tolower, string.strip_tags, convert.*, consumed, dechunk, convert.iconv.*, mcrypt.*, mdecrypt.*

This program makes use of the Zend Scripting Language Engine:
Zend Engine v3.0.0, Copyright (c) 1998-2017 Zend Technologies
with Zend OPcache v7.0.33-0ubuntu0.16.04.3, Copyright (c) 1999-2017, by Zend Technologies

FIGURE G.1: Successfull PHP

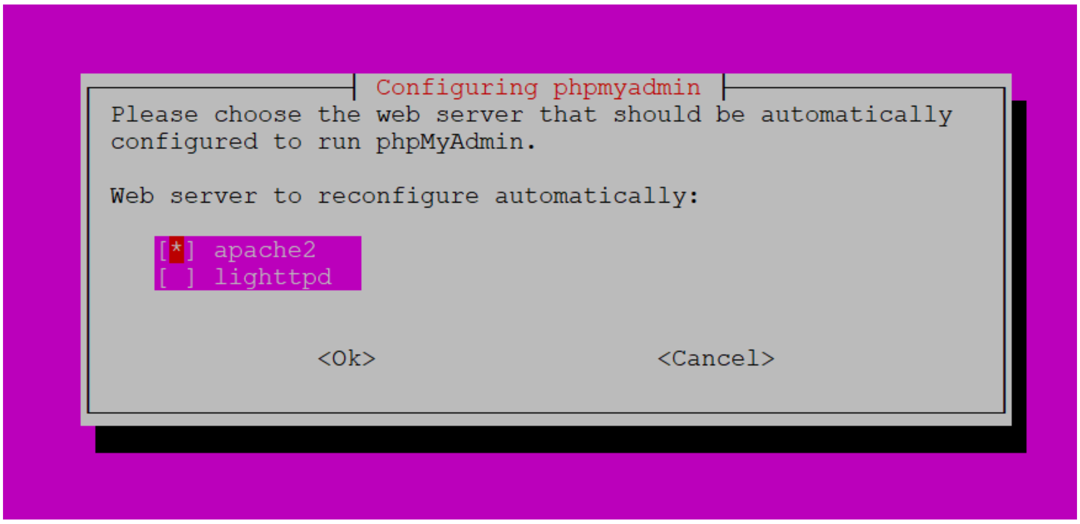


FIGURE G.2: phpMyAdmin Prompt

FIGURE G.3: Example PHPMyAdmin

Bringing in the website!

16. Now that we've downloaded pretty much all the essentials for running a web server, it's time to download the actual website onto the server. We are going to do this by cloning the CHATEL git repository to the file that Apache2 looks to for its website. This can be done by entering the following into the terminal:

```
cd /var/www/html
git clone https://github.com/popoiopo/psychoSystems.git
```

If you now type in "ls" than you should see an "index.html" and the psychoSystems folder.

17. Next we will download Python onto the server as the backbone of the website is written with the Python Flask library. For this we first need to install all required dependencies:

```
sudo apt-get install build-essential checkinstall -y
sudo apt-get install libreadline-gplv2-dev libncursesw5-dev libssl-dev
libsqlite3-dev tk-dev libgdbm-dev libc6-dev libbz2-dev -y
```

18. Then we will download Python itself:

```
cd /usr/src
sudo wget https://www.python.org/ftp/python/2.7.12/Python-2.7.12.tgz
```

And extract and compile the downloaded package:

```
sudo tar xzf Python-2.7.12.tgz
cd Python-2.7.12
sudo ./configure --enable-optimizations
sudo make altinstall
```

19. Now that we have Python up and running, we'll want the pip package manager installation software by typing:

```
apt install python-pip -y
```

So far so good! Now that we can install python stuff, lets make a virtual environment in which our website will function (virtual environments are used, so we can kind of have all our code run neatly within a single box instead of it being installed globally).

```
pip install setuptools
pip install virtualenv
pip install virtualenvwrapper
```

20. Now that we have the necessary libraries to make a virtual environment paste these to lines at the bottom of the basrc file. We can open this file with:

```
sudo nano ~/.bashrc
```

And then scroll down using your down arrow to the bottom of the file and paste in the next two commands:

```
export WORKON_HOME=~/Envs
source /usr/local/bin/virtualenvwrapper.sh
```

Then press “ctrl-x”, type “y” and enter to safe this file. After the next two commands, every time you open a terminal you can return to the CHATEL project by typing “workon CHATEL”.

```
mkvirtualenv CHATEL
deactivate
```

21. You should now see that (CHATEL) is displayed on the left side of your terminal as we are now working in the virtual environment called CHATEL. Now any dependencies within the website we will install inside this virtual environment. Whenever you want to work on the project from within the VPS, you can type **workon CHATEL** and you can access the environment again. So now that we have our virtual environment going we will perform a small magic trick and download all website requirements at once!

```
cd /var/www/html/psychoSystems
sudo apt-get install libmysqlclient-dev -y
sudo apt-get install python-dev -y
pip install wheel
pip install -r requirements.txt
```

22. And we have all our dependencies up and running! Now we can start configuring the flask backbone of the website:

```
export FLASK_CONFIG=production
export FLASK_APP=run.py
```

And set up the database by opening MySQL via the terminal by typing:

```
mysql -u root -p
```

After entering your MySQL password (given during step 10), you now type “show databases;” which gives a list showing all databases that is inside your MySQL database. You can also check this by logging into **http://your_ip_address/phpmyadmin**.

23. Now we will move creating a database and a user that receives all rights:

```
CREATE DATABASE psychoSystems;
CREATE USER 'psychoSystems_admin'@'localhost'
IDENTIFIED BY  'Psycho1!';
GRANT ALL PRIVILEGES
ON psychoSystems . *
TO
'psychoSystems_admin'@'localhost';
quit
```

Then link our python environment with the database by the following:

```
export SQLALCHEMY_DATABASE_URI=
'mysql://psychoSystems_admin:Psycho1!@localhost/psychoSystems'
```

24. Then setup the website database and migrate it to MySQL by:

```
rm -r migrations
flask db init
flask db migrate
flask db upgrade
```

You can check if this worked by logging into **http://your_ip_address/phpmyadmin** where you can see a database called psychoSystems with multiple tables in there.

25. Now that we have our database installed like we need it, we can continue to paste in the default. These defaults can later be changed within the website itself on the dashboard page when logged in as administrator. We do this by clicking on the psychoSystems database (1), then clicking on the SQL tab (2) and pasting the content of the following file: https://raw.githubusercontent.com/popoiopo/psychoSystems/master/instance/db_standards.sql Now we can run the code by pressing "Go" (3). Afterwards, you can check if it worked by looking in the tables if you'd like.

26. As the database is now filled, and the website is ready to go (yay!) we need to make it visible to the world. In order to accomplish this, we need to tell Apache that it needs to search its root file (beginning of the website) elsewhere (i.e. instead of the index.html we need it to look at our run.py file). This can be accomplished by putting a configuration file as follows:

```
cd /etc/apache2/sites-available
sudo nano CHATEL.conf
```

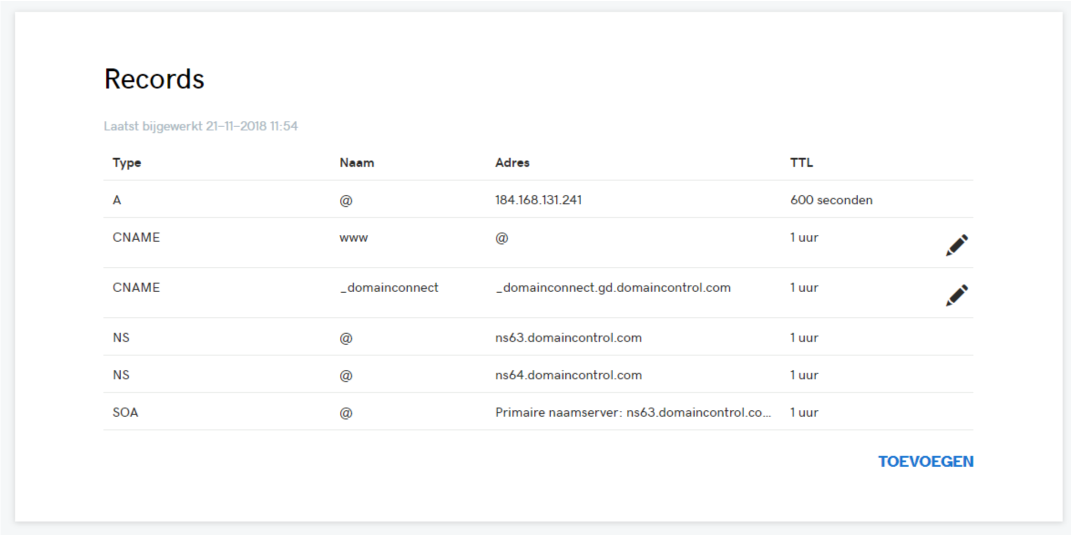
And paste the following code within this file:

```
<VirtualHost *:80>
    ServerName 116.203.137.34
    ServerAdmin root@116.203.137.34

    WSGIDaemonProcess CHATEL python-home=/root/Envs/CHATEL
    WSGIProcessGroup CHATEL

    WSGIScriptAlias / /var/www/html/psychoSystems/CHATEL.wsgi

    <Directory /var/www/html/psychoSystems/app/>
        Order allow,deny
        Allow from all
    </Directory>
    Alias /static /var/www/html/psychoSystems/app/static
    <Directory /var/www/html/psychoSystems/app/static/>
```



Records

Laatst bijgewerkt 21-11-2018 11:54

Type	Naam	Adres	TTL
A	@	184.168.131.241	600 seconden
CNAME	www	@	1 uur
CNAME	_domainconnect	_domainconnect.gd.domaincontrol.com	1 uur
NS	@	ns63.domaincontrol.com	1 uur
NS	@	ns64.domaincontrol.com	1 uur
SOA	@	Primaire naamserver: ns63.domaincontrol.co...	1 uur

TOEVOEGEN

FIGURE G.4: Domain Settings

```
            Order allow,deny
            Allow from all
        </Directory>
        ErrorLog ${APACHE_LOG_DIR}/error.log
        LogLevel warn
        CustomLog ${APACHE_LOG_DIR}/access.log combined
    </VirtualHost>
```

You can save the file by pressing "ctrl-x" followed by y and an "enter". Now that apache is looking at the right files, we still need one last dependency download and a restart:

```
apt-get install libapache2-mod-wsgi
systemctl restart apache2
```

Finally we need to set permissions for the datafiles to be downloaded by entering the following in the terminal:

```
cd /var/www/html/psychSystems/app/static/data
chmod -R 777 .
```

And we are open for business! Please open the website at **"http://your_ip_address"** and enjoy the website! If you'd like a more professional look and have the URL be like a normal website instead of an IP address, you could go to godaddy.com and buy a domain name. If you have one and are logged in, you can click on your domain name followed by "Manage DNS" and configure it as follows:

Bibliography

Association, American Psychiatric et al. (2013). *Diagnostic and statistical manual of mental disorders (DSM-5®)*. American Psychiatric Pub.

Babtie, Ann C, Paul Kirk, and Michael PH Stumpf (2014). "Topological sensitivity analysis for systems biology". In: *Proceedings of the National Academy of Sciences* 111.52, pp. 18507–18512.

Bal, Henri et al. (2000). "The distributed ASCI supercomputer project". In: *ACM SIGOPS Operating Systems Review* 34.4, pp. 76–96.

Batty, Michael and Paul M. Torrens (2019). "Modelling complexity : The limits to prediction". In: *journals.openedition.org*. DOI: 10.4000/cybergeo.1035.

Berrios, German E (1999). "Classifications in psychiatry: a conceptual history". In: *Australian and New Zealand Journal of Psychiatry* 33.2, pp. 145–160.

Binder, Thomas et al. (2004). "Developing system dynamics models from causal loop diagrams". In: *Proceedings of the 22nd International Conference of the System Dynamic Society*. Citeseer, pp. 1–21.

Bockting, Claudi L et al. (2015). "A lifetime approach to major depressive disorder: The contributions of psychological interventions in preventing relapse and recurrence". In: *Clinical psychology review* 41, pp. 16–26.

Borsboom, Denny and Angélique OJ Cramer (2013). "Network analysis: an integrative approach to the structure of psychopathology". In: *Annual review of clinical psychology* 9, pp. 91–121.

Brouwer M., Williams A. Kennis M. Fu Z. Cuijpers P. Bockting C. (2019). "Psychological theories of depressive relapse and recurrence: A systematic review and meta-analysis of prospective studies". In:

Burns, James R and Philip Musa (2001). "Structural validation of causal loop diagrams". In: *Proceedings of the Atlanta SD Conference: July 2001; Atlanta*.

Chatfield, Chris (1995). "Model uncertainty, data mining and statistical inference". In: *Journal of the Royal Statistical Society: Series A (Statistics in Society)* 158.3, pp. 419–444.

Chib, Siddhartha and Edward Greenberg (1995). "Understanding the metropolis-hastings algorithm". In: *The american statistician* 49.4, pp. 327–335.

Chukhraev, N et al. (2017). "Combined physiotherapy of anxiety and depression disorders in dorsopathy patients". In: *Journal of Physical Education and Sport* 17.1, p. 414.

Churchman, Charles West (1968). *Challenge to reason*. McGraw-Hill New York.

Coello, Carlos A Coello, Gary B Lamont, David A Van Veldhuizen, et al. (2007). *Evolutionary algorithms for solving multi-objective problems*. Vol. 5. Springer.

Cox, Marjolein, An Steegen, and Jan Elen (2018). "Using causal diagrams to foster systems thinking in geography education". In: *International Journal of Designs for Learning* 9.1, pp. 34–48.

Coyle, Geoff (2000). "Qualitative and quantitative modelling in system dynamics: some research questions". In: *System Dynamics Review: The Journal of the System Dynamics Society* 16.3, pp. 225–244.

Cramer, Angélique OJ et al. (2016). “Major depression as a complex dynamic system”. In: *PloS one* 11.12, e0167490.
Day, World Health and Said W H O Director-general (2018). “Media centre " Depression : let ′ s talk " says WHO , as depression tops list of causes of ill health”. In: 38.March, pp. 2017–2019.
Dias, Amit et al. (2019). “Effect of a lay counselor intervention on prevention of major depression in older adults living in low-and middle-income countries: a randomized clinical trial”. In: *JAMA psychiatry* 76.1, pp. 13–20.
Diewert, W Erwin (2008). *Cost functions*. Springer.
Dobson, Keith S. et al. (2008). “Randomized Trial of Behavioral Activation, Cognitive Therapy, and Antidepressant Medication in the Prevention of Relapse and Recurrence in Major Depression”. In: *Journal of Consulting and Clinical Psychology* 76.3, pp. 468–477. ISSN: 0022006X. DOI: 10.1037/0022-006X.76.3.468.
Ferrari, Alize J et al. (2013). “Burden of depressive disorders by country, sex, age, and year: findings from the global burden of disease study 2010”. In: *PLoS medicine* 10.11, e1001547.
Ford, David N and John D Sterman (1998). “Expert knowledge elicitation to improve formal and mental models”. In: *System Dynamics Review: The Journal of the System Dynamics Society* 14.4, pp. 309–340.
Freeman, Marlene P (2019). “Perinatal Depression: Recommendations for Prevention and the Challenges of Implementation”. In: *Jama* 321.6, pp. 550–552.
Frodl, Thomas (2016). *Systems neuroscience in depression*. Academic Press.
Fu Z., Williams A. Kennis M. Brouwer M. Cuipers P. Bockting C. (2019). “Psychological theories of depression etiology: A systematic review and meta-analysis of prospective studies of depression onset.” In:
Greenberg, Paul E et al. (2015). “The economic burden of adults with major depressive disorder in the United States (2005 and 2010)”. In: *The Journal of clinical psychiatry* 76.2, pp. 155–162.
Haman, Kirsten L. et al. (2005). “Prevention of Relapse Following Cognitive Therapy vs Medications in Moderate to Severe Depression”. In: *Archives of General Psychiatry* 62.4, p. 417. ISSN: 0003-990X. DOI: 10.1001/archpsyc.62.4.417.
Haraldsson, Hörđur V (2004). *Introduction to system thinking and causal loop diagrams*. Department of Chemical Engineering, Lund University.
Hassan, Junaid (2008). “Capturing the Dynamics of Panic Disorder. A System Dynamics Translation of the Contemporary Biological and Psychological Conceptualization of Panic Disorder”. MA thesis. The University of Bergen.
Hernando, Leticia, Alexander Mendiburu, and Jose A Lozano (2018). “Hill-Climbing Algorithm: Let’s Go for a Walk Before Finding the Optimum”. In: *2018 IEEE Congress on Evolutionary Computation (CEC)*. IEEE, pp. 1–7.
Herrera, Daniel and IL Bleijenbergh (2016). “Cutting the Loops of Depression: a System Dynamics Representation of the Feedback Mechanisms Involved in Depression Development”. In:
Hetlevik, Øystein et al. (2019). “Patient-reported depression treatment and future treatment preferences: an observational study in general practice”. In: *Family Practice*.
Jones, Eric, Travis Oliphant, Pearu Peterson, et al. (2001). *SciPy: Open source scientific tools for Python*. [Online; accessed]. URL: http://www.scipy.org/.
Kassebaum, N. J. et al. (2016). “Global, regional, and national disability-adjusted life-years (DALYs) for 315 diseases and injuries and healthy life expectancy (HALE), 1990–2015: a systematic analysis for the Global Burden of Disease Study 2015”.

In: *The Lancet* 388.10053, pp. 1603–1658. ISSN: 1474547X. DOI: 10.1016/S0140-6736(16)31460-X.

Kendler, Kenneth S, Charles O Gardner, and Carol A Prescott (2002). “Toward a comprehensive developmental model for major depression in women”. In: *American Journal of Psychiatry* 159.7, pp. 1133–1145.

— (2006). “Toward a comprehensive developmental model for major depression in men”. In: *American Journal of Psychiatry* 163.1, pp. 115–124.

Kennis M., Gerritsen L. van Dalen M. Williams A. Cuijpers P. Bockting. (2019). “Lack of evidence for predictive biomarkers of Major Depressive Disorder: A Systematic Review and Meta-Analysis of Prospective Studies”. In:

Kenzie, Erin S et al. (2018). “The dynamics of concussion: mapping pathophysiology, persistence, and recovery with causal-loop diagramming”. In: *Frontiers in neurology* 9, p. 203.

Kirkpatrick, Scott, C Daniel Gelatt, and Mario P Vecchi (1983). “Optimization by simulated annealing”. In: *science* 220.4598, pp. 671–680.

Kivelä, Mikko et al. (2014). “Multilayer networks”. In: *Journal of complex networks* 2.3, pp. 203–271.

Kupfer, David J. and Ellen Frank (2001). “The interaction of drug- and psychotherapy in the long-term treatment of depression”. In: *Journal of Affective Disorders* 62.1-2, pp. 131–137. ISSN: 01650327. DOI: 10.1016/S0165-0327(00)00357-8.

Lai, Danny and Rebekah Wahba (2001). “System dynamics model correctness checklist”. In: *MIT D-Memo D-4851* 1, p. 15.

Levine, Ralph L (2000). “System dynamics applied to psychological and social problems”. In: *Proceedings of the 18th International Conference of the System Dynamics Society*.

Libuda, Lars et al. (2019). “Vitamin D and the Risk of Depression: A Causal Relationship? Findings from a Mendelian Randomization Study”. In: *Nutrients* 11.5, p. 1085.

Long, JiaoYue and Chuan Chen (2019). “A Causal feedback relationship model for critical success factors of PPP projects based on system dynamics”. In: *International Academic Conference on Frontiers in Social Sciences and Management Innovation (IAFSM 2018)*. Atlantis Press.

Luke, Douglas A and Katherine A Stamatakis (2012). “Systems science methods in public health: dynamics, networks, and agents”. In: *Annual review of public health* 33, pp. 357–376.

Lunansky, Gabriela, Claudia D van Borkulo, and Denny Borsboom (2019). “Personality, resilience, and psychopathology: A model for the interaction between slow and fast network processes in the context of mental health”. In:

Mauss, Iris B and Michael D Robinson (2009). “Measures of emotion: A review”. In: *Cognition and emotion* 23.2, pp. 209–237.

McKerns, Michael M et al. (2012). “Building a framework for predictive science”. In: *arXiv preprint arXiv:1202.1056*.

McKinney, Wes et al. (2010). “Data structures for statistical computing in python”. In: *Proceedings of the 9th Python in Science Conference*. Vol. 445. Austin, TX, pp. 51–56.

Oliphant, Travis E (2006). *A guide to NumPy*. Vol. 1. Trelgol Publishing USA.

Pellerin, Eric, Luc Pigeon, and Sylvain Delisle (2004). “Self-adaptive parameters in genetic algorithms”. In: *Data Mining and Knowledge Discovery: Theory, Tools, and Technology VI*. Vol. 5433. International Society for Optics and Photonics, pp. 53–64.

Povinelli, Richard J and Xin Feng (1999). "Improving genetic algorithms performance by hashing fitness values". In: *Artificial Neural Networks in Engineering*, pp. 399–404.

Schmelzer, Martin, Richard P Dwight, and Paola Cinnella (2019). "Machine Learning of Algebraic Stress Models using Deterministic Symbolic Regression". In: *arXiv preprint arXiv:1905.07510*.

Siegle G. J., Cramer A. O. J. van Eck N. J. Spinhoven P. Hollon S. D. Ormel J. Bockting C. L. H. (2019). "Where are the leaks from theory to clinical practice (and back) in addressing the depression epidemic? An empirical graph-theoretic approach". In:

Smith, Caroline A et al. (2018). "Acupuncture for depression". In: *Cochrane database of systematic reviews* 3.

Sterman, John D (2000). *Business dynamics: systems thinking and modeling for a complex world*. HD30. 2 S7835 2000.

Susta, Marek and Gustav Bizik (2011). "Human stress response from the system dynamics point of view". In: *The 29th International Conference of the System Dynamics Society*.

Tucker Lima, Joanna M et al. (2017). "Does deforestation promote or inhibit malaria transmission in the Amazon? A systematic literature review and critical appraisal of current evidence". In: *Philosophical Transactions of the Royal Society B: Biological Sciences* 372.1722, p. 20160125.

Tuomilehto, Henri PI et al. (2009). "Lifestyle intervention with weight reduction: first-line treatment in mild obstructive sleep apnea". In: *American journal of respiratory and critical care medicine* 179.4, pp. 320–327.

Valles-Colomer, Mireia et al. (2019). "The neuroactive potential of the human gut microbiota in quality of life and depression". In: *Nature microbiology* 4.4, p. 623.

Vins, Holly et al. (2015). "The mental health outcomes of drought: a systematic review and causal process diagram". In: *International journal of environmental research and public health* 12.10, pp. 13251–13275.

Wang, Lingfei and Tom Michoel (2019). "Whole-transcriptome causal network inference with genomic and transcriptomic data". In: *Gene Regulatory Networks*. Springer, pp. 95–109.

Waskom, Michael et al. (2017). *mwaskom/seaborn: v0.8.1 (September 2017)*. DOI: 10.5281/zenodo.883859. URL: https://doi.org/10.5281/zenodo.883859.

Wittenborn, AK et al. (2016). "Depression as a systemic syndrome: mapping the feedback loops of major depressive disorder". In: *Psychological medicine* 46.3, pp. 551–562.

Yang, Miles M et al. (2019). "Analysing the dynamics of mental models using causal loop diagrams". In: *Australian Journal of Management*, p. 0312896218823831.

Yildiz, Ali R (2013). "Hybrid Taguchi-differential evolution algorithm for optimization of multi-pass turning operations". In: *Applied Soft Computing* 13.3, pp. 1433–1439.

Zimmermann, Martha and Anthony Papa (2019). "Causal explanations of depression and treatment credibility in adults with untreated depression: Examining attribution theory". In: *Psychology and Psychotherapy: Theory, Research and Practice*.