UNIVERSITY OF PANNONIA

DOCTORAL THESIS

Multidimensional network-based analysis of complex systems

Gergely Marcell Honti
Author

Prof. Dr. János Abonyi
Supervisor

4th August 2021

DOI:10.18136/PE.2021.804

Multidimensional network-based analysis of complex systems

Thesis for obtaining a PhD degree in the Doctoral School of Chemical Engineering and
Material Sciences of the University of Pannonia

Written by Gergely Marcell Honti

Supervisor:
propose acceptance yes / no . . . .
Prof. Dr. János Abonyi

As a reviewer, I propose acceptance of the thesis:

Name of the Reviewer:

. . . . yes / no . . . .
(reviewer)

Name of the Reviewer:

. . . . yes / no . . . .
(reviewer)

The PhD-candidate has achieved ....% at the public discussion.

Veszprém . . . . , ...
(Chair of Committee)

The grade of the PhD Diploma .... (...... %).

Veszprém . . . . , ...
(Chair of UDHC)

Abstract

Multidimensional network-based analysis of complex systems

The focus of this research is in the area of multidimensional networks. A multidimensional network is a network in which among the nodes there may be multiple different qualitative and quantitative relations. Such a study is essential to understand real-world data, and it also provides data heterogeneity, which enables complex algorithms and knowledge extraction. The research approach in this dissertation is to transform real-world data to the multidimensional network and extract domain-specific knowledge, which otherwise would still be hidden or non-trivial, difficult to prove. The transformation is key to achieve homogeneity in data. The dissertation shows three non-comparable nor related datasets production system, systems dynamics models and linked data transformed differently then treated homogeneously, with ease. The transformation steps will start from bi- and multipartite networks, raw model data, and to query results, and will end in different types of multidimensional networks depending on the application. The examples ultimately the applications provided enable the analysis thru a multidimensional network of discrete production environments; abstract systems such as world sustainability or linked data. The main conclusions drawn from this study are that real-world data, can be transformed into multidimensional networks; the transformation steps are non-trivial and have include domain-specific knowledge. Once the transformation is successful, the transformed data acts the same way as previously, no truncation will occur. The new structure of the data reveals the under-laying system with essential structural and dynamical knowledge which can be ultimately used by decision-makers. This dissertation recommends the usage of data enrichment by standardized ontologies and taxonomies, to ease the transformation steps and that transformations from data to a multidimensional network could occur automatically.

Keywords: multidimensional network, Linked Data, Industry 4.0.

Kivonat

Többdimenziós hálózat alapú komplex rendszer elemzés

A kutatás központjában a többdimenziós hálózatok állnak. A többdimenziós hálózatok olyan hálózatok, melyekben a csomópont között több kvalitatív és kvantitatív reláció is leírható. A tanulmány sarokköve a komplex algoritmusokon alapuló tudásbányászat. Valós életből származó adatok többdimenziós hálózatokra lettek átalakítva, melyek egyfajta köztes, ugyanakkor homogén leírásmódra adtak lehetőséget, lehetővé téve szakterület specifikus tudás kinyerését, melyek máskülönben rejtettek vagy nem triviálisak, vagy egyszerűen csak nehezen bizonyíthatók. A kulcs mindenkor a transzformációban rejlett, melyek az adatokat homogénné tették. A disszertáció három teljesen különböző adathalmazt mutat be, melyek témaköre rendre a termelési rendszerek, a rendszerdinamikai modellek valamint az általános láncolt adatok (linked data). Az transzformációs lépések rendre bi- és multipartite hálózatokból, nyers modell adatokból, valamint lekérdezési eredményekből indulnak ki majd alkalmazástól függően különböző típusú többdimenziós hálózattá alakulnak. Végeredményt illetően a dolgozat bemutatja, hogy többdimenziós hálózatokkal hogyan lehet diszkrét termelői rendszert, vagy teljesen absztrakt témaköröket mint a fenntarthatóságot vagy a láncolt adatokat elemzeni. A tanulmány fő eredménye, hogy bemutatja hogyan lehet valós életből származó adatokból többdimenziós hálózatokat létrehozni. Rámutat ezen lépés összetettségére, hogy az átalakításhoz szakterület specifikus tudás szükséges, viszont, ha ezen lépés sikerek akkor az adott adathalmaz csonkítás mentesen, homogén módon kezelhető. Az új adatstruktúra a hálózattudomány eszköztárával elemezhetővé teszi az alrendszerek struktúráit és dinamikáit. A dolgozat továbbá rámutat arra, hogy az adatok bővítése szabványosított ontológiákkal és taxonómiákkal, segíteni tudja a transzformációs lépéseket valamint hogy ezen lépések automatizálhatóvá is válhatnak.

Kulcsszavak: többdimenziós hálózat, Linked Data, Ipar 4.0.

Auszug

Komplex System Analysis mit mehrdimensionaler Netzwerke

Der Schwerpunkt dieser Forschung liegt im Bereich mehrdimensionaler Netzwerke. Ein mehrdimensionales Netzwerk ist ein Netzwerk, in dem zwischen den Knoten mehrere verschiedene qualitative und quantitative Beziehungen bestehen können. Eine solche Studie ist unerlässlich, um reale Daten zu verstehen, und sie bietet auch Datenheterogenität, die komplexe Algorithmen und Wissensextraktion ermöglicht. Der Forschungsansatz in dieser Dissertation besteht darin, Daten aus der realen Welt in das mehrdimensionale Netzwerk zu transformieren und Domänen-spezifisches Wissen zu extrahieren, was sonst noch verborgen oder nicht trivial wäre, und daher schwer zu beweisen wäre. Die Transformation ist der Schlüssel zum Erreichen der Homogenität der Daten. Die Dissertation zeigt drei nicht vergleichbare oder verwandte Datensätze: Produktionssystem, System-Dynamik-Modelle und verknüpfte Daten, die mit Leichtigkeit unterschiedlich transformiert und dann homogen behandelt werden. Die Transformationsschritte beginnen mit zwei- und mehrteiligen Netzwerken, Rohmodelldaten und Abfrageergebnissen und enden - je nach Anwendung - in unterschiedlichen Arten von mehrdimensionalen Netzwerken. Die bereitgestellten Beispiele ermöglichen letztendlich die Analyse durch ein mehrdimensionales Netzwerk diskreter Produktionsumgebungen, abstrakter Systeme wie World Sustainability oder Linked Data.

Die wichtigsten Schlussfolgerungen aus dieser Studie sind, dass Daten aus der realen Welt in mehrdimensionale Netzwerke umgewandelt werden können. Die Transformationsschritte sind nicht trivial und beinhalten Domänen-spezifisches Wissen. Sobald die Transformation erfolgreich ist, verhalten sich die transformierten Daten wie zuvor - es erfolgt keine Kürzung. Die neue Struktur der Daten offenbart das darunter liegende System mit wesentlichen strukturellen und dynamischen Kenntnissen, die letztendlich von Entscheidungsträgern genutzt werden können. Diese Dissertation empfiehlt die Nutzung der Datenanreicherung durch standardisierte Ontologien und Taxonomien, um die Transformationsschritte zu erleichtern und Transformationen von Daten in ein mehrdimensionales Netzwerk automatisch ablaufen zu lassen.

Schlüsselwörter: Mehrdimensionales Netzwerk, Linked Data, Industrie 4.0

List of appended papers

1^{st} thesis: I have created a multidimensional network model to model the broad spectrum of production systems. The method supports the production flow analysis and even pinpoints the development potentials. With the clustering of similar components and machines the method achieves similar manufacturing optimization as the most advanced manufacturing cell algorithms.

Author	Title	Publication date	Publisher	IF
Tamás Ruppert, Gergely Honti, János Abonyi	Multilayer network-based production flow analysis	2018	Complexity	2.591

2$^{\text{nd}}$ thesis: I have created a method, that can automatically transform a system dynamics model to a network, enabling the network science methods to capture key elements, similar elements and modules, and it could also capture the dynamics between modules. The transformed network supported view of stock and flow diagram to map the network, the state space view and the views were separately analyze to support both the modeler and the communication of the model. The method enabled comparation between system dynamics models.

Author	Title	Publication date	Publisher	IF
Gergely Honti, Gyula Dörgő, János Abonyi	Review and structural analysis of system dynamics models in sustainability science	2019	Journal of Cleaner Production	7.246

3$^{\text{rd}}$ thesis: I have created a method to analyze Linked Data by network science methods. I have introduced a new multidimensional network notation to support this method, by introducing the dimensions of the nodes. I demonstrated that frequent pattern mining can be applied to reveal statistically significant correlations between layers.

Author	Title	Publication date	Publisher	IF
Gergely Honti, János Abonyi	Frequent itemset mining and multi-layer network-based analysis of linked data	2021	Mathematics	1.747

Other published papers and books, which were published during the PhD period, but not included in the dissertation.

Author	Title	Publication date	Publisher	IF
Tímea Czvetkó, Gergely Honti, János Abonyi	Regional development potentials of Industry 4.0: Open data indicators of the Industry 4.0+ model	2021	PlosOne	3.2
Tímea Czvetkó, Gergely Honti, Sebestyén Viktor, János Abonyi	The intertwining of world news with Sustainable Development Goals: an effective monitoring tool	2021	HELIYON	1.8
Gergely Honti, Tímea Czvetkó, János Abonyi	Data describing the regional Industry 4.0 readiness index	2020	Data in Brief	0.1
János Abonyi, Tímea Czvetkó, Gergely Marcell Honti	Are Regions Prepared for Industry 4.0?	2020	SpringerBriefs in Entrepreneurship and Innovation	-
Gergely Marcell Honti, Gyula Dörgő, János Abonyi	Network analysis dataset of System Dynamics models	2019	Data in Brief	0.1
Gergely Marcell Honti, János Abonyi	A Review of Semantic Sensor Technologies in Internet of Things Architectures	2019	Complexity	2.59
Gyula Dörgő, Gergely Marcell Honti, János Abonyi	Automated analysis of the interactions between sustainable development goals extracted from models and texts of sustainability science	2018	Chemical Engineering Transactions	0.76

Contents

List of Figures

List of Tables

Chapter 1

Introduction

Network science is an interdisciplinary endeavour, with methods and applications drawn from across the natural, social, and information sciences. Most of the networks studied in the literature are monodimensional: there can be only one link between two nodes. Therefore in this context, network analytics has focused on the characterization and measurement of local and global properties of such graphs, such as diameter, degree distribution, centrality, connectivity—up to more sophisticated discoveries based on graph mining, aimed at finding frequent subgraph patterns and analyzing the temporal evolution of a network.

However, in the real world, networks are often multidimensional, i.e there might be multiple connections between any pair of nodes. Therefore, multidimensional analysis is needed to distinguish among different kinds of interactions, or equivalently to look at interactions from different perspectives. This is analogue to multidimensional analysis in OLAP systems and data warehouses, where data are aggregated along various dimensions. In analogy, the important part is the different interactions between two entities as dimensions. Dimensions in-network data can be either explicit or implicit. In the first case, the dimensions directly reflect the various interactions in reality; in the second case, the dimensions are defined by the analyst to reflect different interesting qualities of the interactions, that can be inferred from the available data. Therefore the use of multidimensional networks is advised. There are different types of multidimensional networks, and in addition, the thesis will also use several of them. The most popular multidimensional network is the multilayer network. The multilayer network can be extended and unfolded into different networks[6] like multiplex network[7], temporal networks[8], edge labeled multigraphs[9], interacting networks[10], interdependent networks[11], multilevel networks[12] and hypernetworks[13]. The different types of multidimensional networks allow different types of investigations.

Data is typically a jumble of raw facts, and users need to sift through it to properly interpret and organize the data. Only then does the data become useful. This dissertation introduces a viewpoint, where the raw data is interpreted as a multidimensional network, and further knowledge can be extracted. Figure 1.1 represents the general idea of all the thesis points, the idea of the paradigm of data-to-network-to-knowledge (D2N2K). The paradigm covers the transformation into a relatively structured heterogeneous multidimensional network, then mine and apply different tools on the structure-rich heterogeneous network to generate useful knowledge.

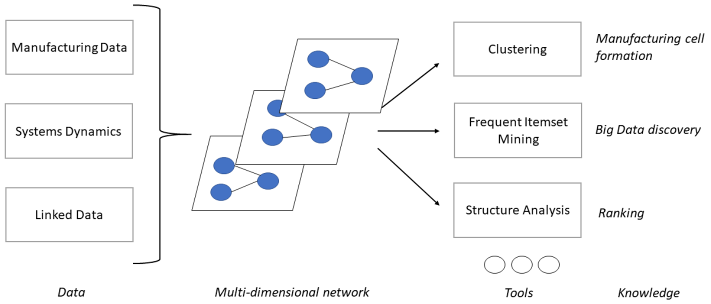


Figure 1.1. Representation of the thesis points and the general idea of "data-to-network-to-knowledge".

The current work will show an insight into D2N2K on the following topics:

1. Production-flow can be modelled as a multilayered network.
2. Systems Dynamics models effectively capture the system as a whole, and the conversion of the models to multidimensional network enables efficient structural analysis of the system.
3. Linked Data is effectively analyzable by multidimensional networks.

Production-flow data can be transformed into a multilayered network enabling further analysis with the tool-set of network science. For example the detection of mesoscopic structures, or as in-network science called, communities or cohesive groups, or more explanatory the groups of nodes that are more tightly connected than they are to the rest of the network are revealing manufacturing cells. System dynamical models can also be transformed into a network, and with structure analysis, especially the triad search reveals the sub-dynamics of the complex systems. Since its creation, the digital world has been evolving at exponential rate, presenting both its developers and users with the constant challenge of updating their skills and knowledge in order to support its development, and take advantage of its potential. Alongside with the more popular World Wide Web and Web 2.0, another version of the web has been developing quietly, compared to the spectacular growth of its relatives: the Semantic Web, also known as Web 3.0 or Web of Data. If the expression 'Semantic Web' reflects the more general concept, 'linked data' (LD) can be defined as the key tool to realize the idea. The main and innovative concept that underlies LD is building relationships between raw data, in a way, that is understandable by computers to improve the discoverablity and interpretability of the raw data[14]. Although the potential behind LD is widely perceived, there is still confusion and reserve on how to benefit from this tool. The path towards its implementation has revealed several challenges, e.g., bottlenecks in communication, raw data handling, mapping of different ontologies. However, LD promises to allow linkage to other services, improve data recovery, enable interoperability without affecting data source models and improve the credibility of the end user resources annotations[15]. Linked data is the ultimate tool for representing and sharing contextual data, with the proper transformation it can be transformed into multidimensional network, enabling the previously mentioned tool-set to analyze the linked dataset.

Chapter 2 showcases the transformation of raw production system data to the network and the information gain from this network, on a real-life example of a mostly manual, discrete production environment of wire-harness production.

Chapter 3 showcases the transformation of system dynamics model, which are capable to implicitly capture the whole system behaviour, to network adding valuable insights into dynamics of the models through community discovery and the discovery of dynamical substructures through triads. This is presented through a very important aspect of both Industry 4.0 and future trends, namely sustainability.

Chapter 4 showcases the transformation of linked data to a multidimensional network, also introducing a novel notation of the multidimensional networks, where nodes will also be scaled to dimensions according to their ontological role.

Each chapter contains an introduction to the problem, a roadmap to the chapter, literature review and an in-depth described methodology, including the mathematical background and also a conclusion.

Chapter 2

Production Systems as Networks

This chapter shows how a production system can be represented by a set of bipartite networks and transformed into a special multi-dimension network, a multilayered network. This bi- and multipartite representation is beneficial in production flow analysis (PFA) that is used to identify improvement opportunities by grouping similar groups of products, components, and machines. It is demonstrated that the goal-oriented mapping and modularity-based clustering of multilayer networks can serve as a readily applicable and interpretable decision support tool for PFA, and the analysis of the degrees and correlations of a node can identify critically important skills and resources. The applicability of the proposed methodology is demonstrated by a well-documented benchmark problem of a wire-harness production process. The results confirm that the proposed multilayer network can support the standardized integration of production-relevant data and exploratory analysis of strongly interconnected production systems.

2.1 Introduction

Industry 4.0 is a strategic approach to design optimal production flows by integrating flexible and agile manufacturing systems with Industrial Internet of Things (IIoT) technology[16] enabling communication between people, products and complex systems[17, 18, 19].

The integration of manufacturing and information systems are, however, a challenging task[20]. Horizontal- and inter-company integration should connect the elements of the supply chain[21], while vertical integration should connect information related to the entire product life cycle[22].

According to this new concept, improvement and optimization of production technologies based on Cyber-Physical-Systems (CPS) are realized by the simultaneous utilization of information related to production systems[23], products, models[24], simulators and process data[25, 26].

CPS- and Industry 4.0 type solutions also enable the compositions of smaller cells providing more flexibility with regard to production[27]. This idea leads to decentralized manufacturing[28] and emerging Next Generation Machine Systems[29]. This trend highlights the importance of the relationship between flexibility and complexity[30].

The complexity of production systems can be divided into the physical and functional domains[31]. To analyze this aspect our focus is on the production flow analysis of production systems as production analysis has multiple perspectives according to the hierarchical decomposition of the production system: 1.) Production Flow Analysis studies the activities needed to make each part and machines to be used to simplify the material flow, 2.) Company Flow Analysis studies the flow of materials between different factories to develop an efficient system in which each facility completes all the parts it makes, 3.) Factory Flow Analysis plans the division of the factory into groups or departments each of which manufactures all the parts it makes and plans a simple unidirectional flow system by joining these departments, 4.) Group analysis divides each department into groups, each of which completes all the parts it makes- groups which complete parts with no backflow, crossflow (between groups) no need to buy any additional equipment, 5.) Line analysis analyses the flow of materials between the machines in each group to identify shortcuts in the plant layout and 6.) Tooling analysis tries to minimize setup time by finding sequences that minimize the required additional tooling for the following job[32].

Production flow analysis (PFA) is a technique to identify both groups and their associated “families” by analyzing the information in component process routes which show the activities (often referred as operations) needed to make each part and the machines to be used for each activity[33, 34]. Every production flow analysis begins with data gathering during which non-value adding activity should be optimized[35]. When dealing with large quantities of manufacturing data, a representational schema that can efficiently represent structurally diverse and dynamical system have to be taken into consideration. Standards like ISO 18629, 10303 (STEP), 15531 (MANDATE) support information flow by standardizing the description of production processes[36]. Based on these standards and web semantics, a manufacturing system engineering (MSE) knowledge representation scheme, called an MSE ontology model was developed as a modeling tool for production[37]. The MSE ontology model by its very nature can be interpreted as a labeled network.

A simple multidimensional representation is proposed that can unfold the complex relationships of production systems. Network models are ideal to represent connections between objects and properties[38]. However, as a multidimensional problem that requires flexibility due to the continuously growing amount of information is in question, a new multidimensional approach in the form of a multilayer network[39] is presented.

For the analysis of the resultant ontology-driven labeled multilayer network, techniques to facilitate cell formation and competency assignment for operators were developed.

Manufacturing cell-formation aims to create manufacturing cells from a given number of machines and products by partitioning similar machines which produce similar products. Standard cell-formation problems handle products and machines while their connections are represented by two-layered bipartite graphs or machines-products incidence matrices. Classical algorithms are based on clustering and seriation of the incidence matrices. Recently various alternative algorithms have been developed, for example, self-organizing maps[40] of fuzzy clustering-based methods[41]. What is common in most of these approaches is that they only take two variables, machines and parts, into account[42]. However, complex manufacturing processes should be characterized by numerous properties, like the type of products, resources and the required skills of operators should be also taken into account at successful line balancing since the skills of the operators are influencing the speed of the conveyor belt[43]. Dynamic job rotation[44] also requires efficient allocation of the assembly tasks whilst taking into account the constraints related to the available skills of the operators.

To handle these elements of the production line, the traditional cell-formation problem was extended into a multidimensional one. The main idea is to represent these problems by multilayered graphs and apply modularity analysis to identify the groups of items that could be handled together to improve the production process.

An entirely reproducible benchmark problem was designed to demonstrate our methodology. As an example, the problem of process flow analysis of wire-harness production was selected as this product is complex, and varies significantly[45] as the geometries and components of the harness vary depending on the final products[46]. Since there are challenges in the selection of the cost-effective design[47] and the demand for flexibility and a short delivery time urge the definition of product families produced from the submodules[48], the problem requires the advanced integration of process- and product-relevant information.

The remaining part of the chapter is structured as follows. In Section 2.2 a multilayer network model is formalized that was developed to represent production systems. In Section 2.3 how production flow analysis problems can be interpreted as network analysis tasks is discussed. Section 2.3.1 describes the applicability of network science in PFA. Section 2.3.2 formalizes the projection of the multilayer networks and studies how conditional connections can be defined, while Section 2.3.4 applies this projection to calculate the node similarities. The group formation task is described in Section 2.3.5, where the results of this approach on benchmark examples are also presented. The detailed case study starts in Section 2.4 with the definition of the wire-harness production use case. The details of the problem are given in the Appendix. Section 2.4.1 demonstrates the applicability of similarity and modularity analysis. The workload analysis is given in Section 2.4.2, while interesting applications related to the evaluation of the flexibility of operator-task assignment problems are discussed in Section 2.4.3. Finally, conclusions are drawn in Section 2.5.

2.2 Multilayer-network representation of production systems

Essential information about the products to be assembled, parts to be manufactured, materials to be used, methods and techniques to convert the material to the required finished components and manpower to operate the plant is usually available to a company, but rarely in an appropriate form for ease of digestion by the manager[49]. In this section, we propose a network-based model to study the relationship between these elements.

As can be seen in Figure 2.1, the proposed network consists of a set of bipartite graphs representing connections between the sets of products: $\mathbf{p} = \{p_1, \ldots, p_{N_p}\}$, machines/workstations: $\mathbf{w} = \{w_1, \ldots, w_{N_w}\}$, parts/components: $\mathbf{c} = \{c_1, \ldots, c_{N_c}\}$, activities (operations): $\mathbf{a} = \{a_1, \ldots, a_{N_a}\}$, and their categorical properties (referred as activity types): $\mathbf{t} = \{t_1, \ldots, t_{N_t}\}$), and skills of the operators needed to perform the given activity $\mathbf{s} = \{s_1, \ldots, s_{N_s}\}$.

The relationships among these sets are defined by bipartite graphs $G_{i,j} = (O_i, O_j, E_{i,j})$ represented by $\mathbf{A}[O_i, O_j]$ biadjacency matrices, where O_i and O_j are used as a general representation of a sets of objects, as $O_i, O_j \in \left\{\mathbf{p}, \mathbf{w}, \mathbf{c}, \mathbf{a}, \mathbf{a}^{'}, \mathbf{t}, \mathbf{s}, \mathbf{o}, \mathbf{m}\right\}$.

The edges of these bipartite networks can represent material, energy or information flows, structural relationships, assignments, attributes as well as preferences, and the edge weights can be proportional to the number of shared components/resources, or time/cost needed to produce a given product (see Table 2.1).

The proposed model can be considered as an *interacting or interconnected network*[39], where the family of bipartite networks defines crossed layers. Since different types of connections between the nodes can be defined, the model can also be handled as a *multidimensional network*. Both of these models are the special cases of multilayer networks, which representation is beneficial, since the layers represent the direct connections defined by the bipartite graphs, while the interlayer connections help in term of the visualization of the complex system by arranging the corresponding nodes at the same place within the layers (as it is illustrated in Figure 2.2).

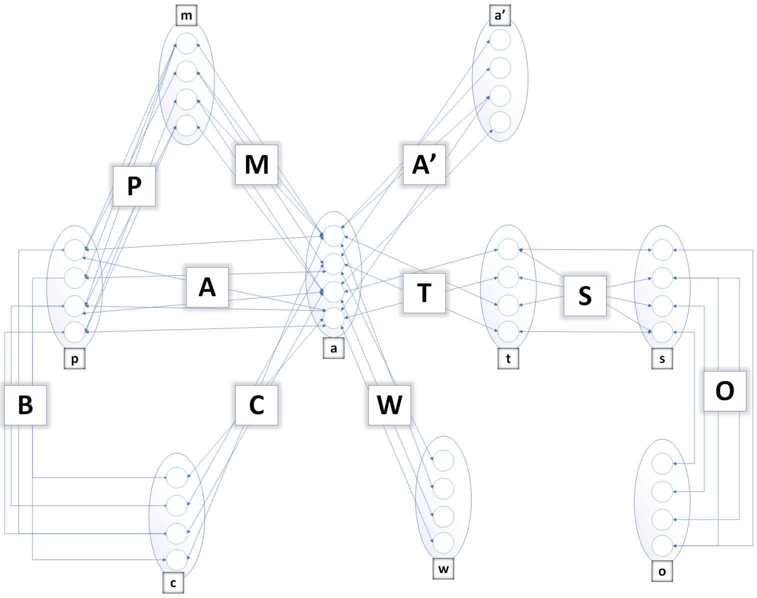


Figure 2.1. Illustrative network representation of a production system. The definitions of the symbols are given in Table 1.

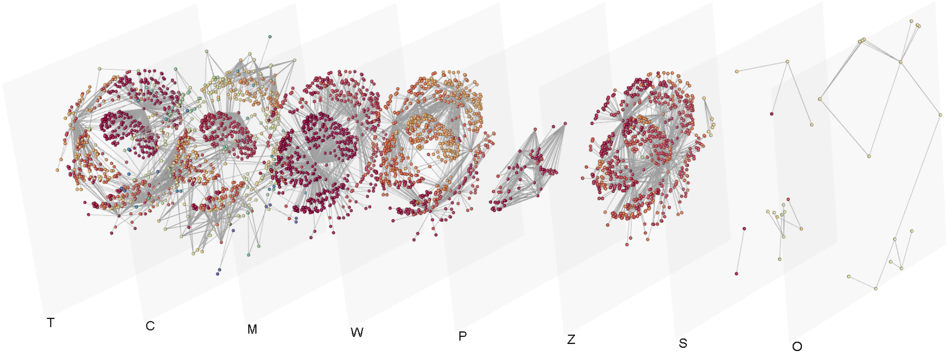


Figure 2.2. Visualization of the illustrative network as a multilayer/multiplex network highlights how the complex production system can be grouped into modules based on the 'viewpoints' of the layers.

The previously presented example serves only as an illustration. For real-life applications, the model should be extended and standardized. Manufacturing systems and their information can be organized by following the 5Ms and 5Cs concepts. The 5Ms stand for

Materials (properties and functions), Machines (precision and capabilities), Methods (efficiency and productivity), Measurements (sensing and improvement), and Modeling (prediction, optimization, and prevention). The 5Cs stand for Connection (sensors and networks), Cloud (data on demand and at anytime), Content (correlation and purpose), Community (sharing and social) and Customization (personalization and value)[23]. Based on the characteristic elements and connections of production systems the type of nodes and edges of their network[50] can be defined, the relevant information is summarized in Tables 2.2, 2.3 and 2.4. Although these concepts are already useful in structuring information, as a standardized solution the application of the ADACOR Predicates that established relationships among the essential concepts of production management are recommended[4] (see Table 2.5).

Thanks to the recent standardization and integration of enterprise resource planning (ERP), manufacturing execution systems (MES), shop floor control (SFC) and product lifecycle management (PLM), it is straightforward to identify the connections of the standardized variables of production management and transform them into a multidimensional network model. The model is capable of representing information at different levels, so it

Table 2.1. Definition of the biadjacency matrices of the bipartite networks used to illustrate how a production system can be represented by a multidimensional network.

notation	nodes	description	size
$\mathbf{A}$	product ($\mathbf{p}$) - activity ($\mathbf{a}$)	activity required to produce a product	$N_p \times N_a$
$\mathbf{W}$	activity ($\mathbf{a}$) - workstation/machine ($\mathbf{w}$)	workstation assigned for the activity	$N_a \times N_w$
$\mathbf{A}'$	activity ($\mathbf{a}$)- activity ($\mathbf{a}'$)	precedence constraint between activities	$N_a \times N_a$
$\mathbf{B}$	product ($\mathbf{p}$) - component/part ($\mathbf{c}$)	component/part required to produce a product	$N_p \times N_c$
$\mathbf{P}$	product($\mathbf{p}$) - module ($\mathbf{m}$)	module/part family required to produce a product	$N_p \times N_p$
$\mathbf{C}$	activity ($\mathbf{a}$) - component ($\mathbf{c}$)	component/part built in or processed in an activity	$N_a \times N_c$
$\mathbf{M}$	activity ($\mathbf{a}$) - module ($\mathbf{m}$)	activity required to produce a module	$N_a \times N_m$
$\mathbf{T}$	activity ($\mathbf{a}$) - activity type ($\mathbf{t}$)	category of the activity	$N_a \times N_t$
$\mathbf{S}$	activity type ($\mathbf{t}$) - skill ($\mathbf{s}$)	skill/education required for an activity category	$N_t \times N_s$
$\mathbf{O}$	skill ($\mathbf{s}$) - operator ($\mathbf{o}$)	skills of the operators	$N_s \times N_o$

can support factory flow analysis, departmental flow analysis, or, according to the concept of Industry 4.0, it can also integrate inter-organizational supply chains. The development of organizational models is also supported, for this purpose solutions following the standard of UN/EDIFACT (the United Nations rules for Electronic Data Interchange for Administration, Commerce and Transport) could be used.

The extracted models lend themselves to be handled in the databases of graphs[51, 52] or RDF based ontologies[53]. In our work, the related technical details of building and storing graph-based decision systems is not the focus; rather, how information from this model can be extracted to support production flow analysis is of concern. In the next section, such techniques are presented.

Table 2.2. The characteristics of the edges of the proposed multilayer network.

	Flow type	Attribute type
Definition	Material, energy or information flow between the nodes	Representation of the property of the node
Edge weight	Physical attributes of the flow, like quantity, or during discrete events the frequency of the flow, like the number of hours between events	Similarity measure, meaning the quantity of equal attributes or the similarity of an attribute based on a scale
Self-Loop	Inner activities	Not interpreted, as self-similarities are trivial
Parallel edges	Multiple flows - can be represented by multilayer/multidimensional networks.	Multi-aspect similarities can be converted in to edge weights
Serial connections	Paths of the flow of different entities	Interpreted in terms of the time-varying case; shows spreading of a property
Modularity	Highly cooperative nodes	Highly similar nodes

Table 2.3. The characteristics of the node types of the proposed network.

	Event type	Resource type	Competency type
Fundamental properties	Occurrence probability, failure rate, cycle time, etc.	Physical properties, quality parameters (capacity, idle state, etc.)	Not generalizable, concept-dependent quantity and quality parameters
Node degree	Event frequency	Resource usage metric	Spreading competency
Modularity	Example: event sequence	Example: resources with the same usage parameters	Example: Competencies possessed by the same resources/operators

Table 2.4. The characteristics of node and edge matchings in the proposed network.

	Flow type **(edges)**	Attribute type **(edges)**
Event type **(nodes)**	Process steps **(nodes)** and their input-output connections **(edges)**	Independent variables **(nodes)** and their settings **(edges)**
Resource type **(nodes)**	Information exchange **(edges)** between information systems **(nodes)**	Colleges working **(nodes)** on the same workstations **(edges)**
Competency type **(nodes)**	Commitment reporting between **(edges)** and jobs **(nodes)**	Same competency demanding **(edges)** jobs **(nodes)**

Table 2.5. The ADACOR predicates can be directly applied to define layers of the network[4]. (Please note that we use the term activity to refer to operations)

Predicates	Description
ComponentOf(x,y)	Product x is a component of product y
Allocated(x,y,t)	Operation x is allocated to resource y at time t
Available(x,y,t)	Resource x is available at time t for operation y
RequiresTool(x,y)	Execution of operation x requires tool y
HasTool(x,y,t)	Resource x has tool y available in its magazine at t
HasSkill(x,y)	Resource x has property (skill) y
HasFailure(x,y,t)	A disturbance x occurred in resource y at time t
Precedence(x,y)	Operation x requires previous execution of y
UsesRawMaterial(x,y)	Production order x uses raw material y
RequestSetup(x,y)	Operation x needs the execution of setup y
HasProcessPlan(x,y)	Production of x requires process plan y
OrderExecution(u,x,w,y)	Operation u is listed in process plan w (describing production of y) for production order x
HasRequirement(x,y)	Operation x requires property y
HasGripper(x,y,t)	Resource x has gripper y in its magazine at time t
ExecutesOperation(x,y)	Work order x includes operation y

2.3 Production flow analysis relevant operations on networks

2.3.1 From problems of production analysis to tools of network science

The main benefit of the multidimensional network model is that it provides a transparent and easily interpretable integration of process- and product-relevant information and as well as facilitating the tools of network science for production flow analysis.

The aim of production flow analysis (PFA) is to identify bottlenecks and groups in products, components, and machines to highlight possible improvements by redesigning the layout, forming manufacturing cells, scheduling the activities, or identifying line families of products based on clustering the sequences of machine usage.

Modules/part families are sets of machines and parts that are highly likely to work together in one group or be processed in a similar order. Since this definition is similar to the concept of modules in networks, it is assumed that finding modules in (multidimensional) networks can be considered as a useful heuristical approach of PFA.

The application of heuristics in PFA is a well-accepted approach since in most cases the economic benefits are complicated and time-consuming to calculate, and the resultant complex optimization problems are not easy to solve with classical optimization algorithms/operation research tools. In this chapter we suggest that the following network analysis tools should serve as a good heuristic solutions for specific PFA problems:

- Calculation of the loads, usage frequencies - identification of the bottlenecks
  - Calculation of unknown dependencies
  - Analysis of node and edge centralities
- Group formation - clustering nodes, identifying communities
  - Rank-order based clustering
  - Similarity-based clustering
    - Calculation of node similarities of (projected) networks
    - Clustering nodes and edges based on the calculated similarities
    - Joining of clusters of different objects to form modules
  - Finding modules in the (multilayer) network

- Line formation - ordering modules to minimize sequential transfers
  - Ordering based on the ratio of in/out degrees - Hollier's method[54]
  - Application of graph layout techniques

2.3.2 Projections of the multilayer network and calculation of transitive connections

As Figure 2.3 illustrates, when relationships among the O_i and O_j sets are not directly defined, it is possible to evaluate the relationship between its $o_{i,k}$ and $o_{j,l}$ elements as the number of possible paths or the length of the shortest path between these nodes.

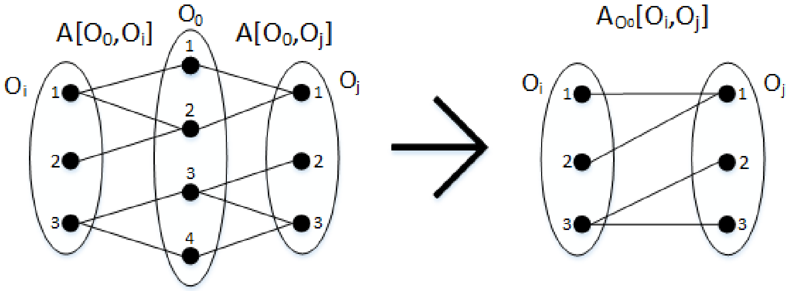


Figure 2.3. Projection of a property connection.

In the case of connected unweighted multipartite graphs the number of paths intersecting the O_0 set can be easily calculated based on the connected pairs of bipartite graphs as:

$$\mathbf{A}_{O_0}[O_i, O_j] = \mathbf{A}[O_0, O_i]^T \times \mathbf{A}[O_0, O_j] \,. \tag{2.1}$$

2.3.3 Conditional connection

Conditional connections could also provide useful information in terms of PFA. To demonstrate the problem, let's have a look at Figure 2.4 shows the network defined in Eq. (2.2). In this example, although operators o_1 and o_3 do not share any machines, the fact that machines m_1 and m_2 produce identical products results in the $\mathbf{A}[O_2|_{O_1}(O_0, O_0)]$ projection operators defining a connection between these operators.

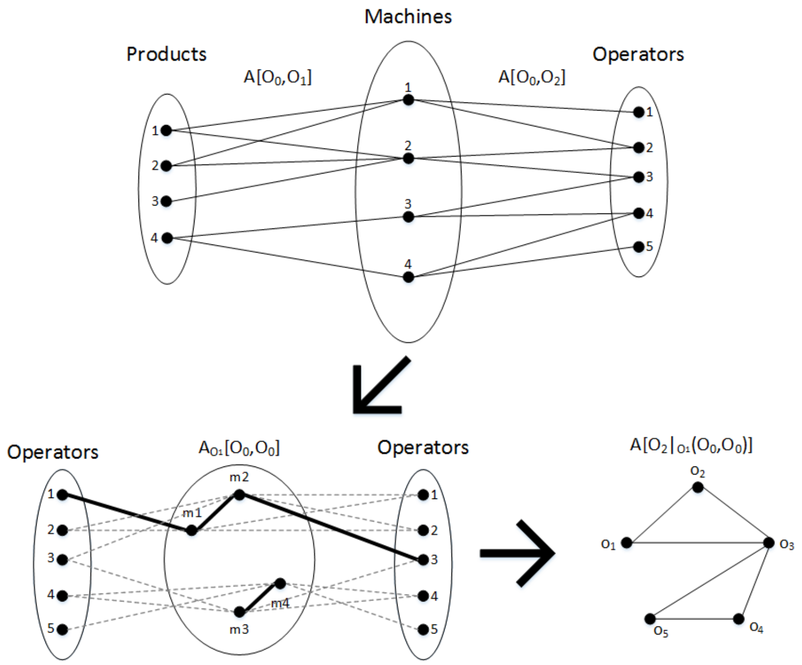


Figure 2.4. The advantage of complex conditional analysis using inner-network

$$\mathbf{A}[O_0, O_1] = \begin{bmatrix} 1 & 1 & 0 & 0 \\ 1 & 1 & 1 & 0 \\ 0 & 0 & 0 & 1 \\ 0 & 0 & 0 & 1 \end{bmatrix}, \mathbf{A}[O_0, O_2] = \begin{bmatrix} 1 & 1 & 0 & 0 & 0 \\ 0 & 1 & 1 & 0 & 0 \\ 0 & 0 & 1 & 1 & 0 \\ 0 & 0 & 0 & 1 & 1 \end{bmatrix},$$

$$\mathbf{A}[O_2|_{O_1}(O_0, O_0)] = \begin{bmatrix} 2 & 4 & 2 & 0 & 0 \\ 4 & 9 & 5 & 0 & 0 \\ 2 & 5 & 4 & 2 & 1 \\ 0 & 0 & 2 & 4 & 2 \\ 0 & 0 & 1 & 2 & 1 \end{bmatrix}. \tag{2.2}$$

Formally, in some cases the $\mathbf{A}[O_i|_{O_k}(O_j, O_j)]$ conditional projections might be of interest defined by:

$$\mathbf{A}[O_i|_{O_k}(O_j, O_j)] = \mathbf{A}[(O_j, O_i)]^T \times (\mathbf{A}[(O_j, O_k)] \times \mathbf{A}[(O_j, O_k)]^T) \times \mathbf{A}[(O_j, O_i)] \tag{2.3}$$

where the resultant $\mathbf{A}[O_i|_{O_k}(O_j, O_j)]$ network states that the ith property set is analyzed based on the $\mathbf{A}_{O_k}[(O_j, O_j)]$ inner network defined by the inner projection of the objects to the jth set.

The projections are not applicable for all types of edges (e.g. the projection with precedence constraints does not result in interpretable networks). Generally, the projections calculate the number of paths between the nodes which number is directly interpretable (e.g. it can reflect the number of assignable operators for a given workstation).

To support these calculations its is beneficial to utilise the adjacency matrix of the whole multiplex network obtained by *flattening or matricization* :

$$\mathbf{A}_M = \begin{bmatrix} \mathbf{0}_1 & \mathbf{A}_{1,2} & \dots & \mathbf{A}_{1,N} \\ \mathbf{A}_{2,1} & \mathbf{0}_1 & \dots & \mathbf{A}_{2,N} \\ \vdots & \vdots & \dots & \vdots \\ \mathbf{A}_{N,1} & \mathbf{A}_{N,2} & \dots & \mathbf{0}_N \end{bmatrix} \tag{2.4}$$

where $\mathbf{A}_{i,j}$ is used to represent the $\mathbf{A}[O_i, O_j]$ biadjacency matrices of the $G_{i,j}$ bipartite graphs.

2.3.4 Calculation of node similarities

Node similarities can reveal useful information with regard to PFA, for example, if the similarities of the machines need to be defined based on how many common parts they are processing. When the machines are denoted as k and j, and S_k and S_j as the sets of parts that are connected to these machines, the similarities of the machines can be evaluated according to the Jaccard similarity index[55]:

$$sim(k, j) = \frac{|S_k \cap S_j|}{|S_k| + |S_j| - |S_k \cap S_j|} \tag{2.5}$$

The proposed network based representation is also beneficial in similarity analysis. When $O_0 = \mathbf{w}$ represents the set of machines/workstations and $O_i = \mathbf{c}$ represents the set of components, the $a_{j,i} = 1$ edge weight stored at the intersection of the j-th row and i-th column of the the $\mathbf{A}[O_0, O_i]$ biadjacency matrix represents that the i-th type of component is built in at the j-th workstation and the degree of the j-th node, $k_j = \sum\limits_i a_{j,i}$ is identical to the cardinality of the $|S_j|$ set, which means how number of component types are built in at the j-th workstation.

We can generate two projections for each bipartite network. The first projection connects two O_o-nodes (in our case two workstations) by a link if they are linked to the same O_i-node (same components). As Figure 2.5 illustrates, the $|S_k \cap S_j|$ cardinality is identical to the $j-k$ edge weight of the projected network which represents how many identical components are built in at the k-th and j-th workstation:

$$\mathbf{A}_{O_0}[O_0, O_i] = \mathbf{A}[O_0, O_i]^T \times \mathbf{A}[O_0, O_i] \tag{2.6}$$

The second projection connects the O_i-nodes (in our case two components/parts) by a link if they connect to the same O_o-node (workstations), which projection represents how parts are connected by the machines:

$$\mathbf{A}_{O_0}[O_0, O_i] = \mathbf{A}[O_0, O_i] \times \mathbf{A}[O_0, O_i]^T \tag{2.7}$$

When the similarities of more layers are taken into account, multiple projections on the same machines can be defined by the weighted sum of their projections:

$$\mathbf{A}[O_0, O_0] = \sum_i w_i \mathbf{A}[O_0, O_i] \times \mathbf{A}[O_0, O_i]^T \tag{2.8}$$

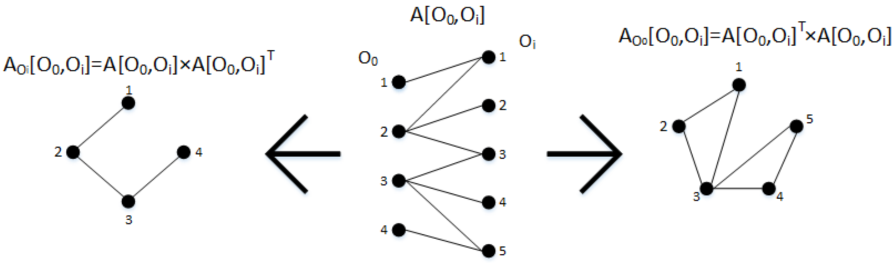


Figure 2.5. Two different projections can measure how the neighboring node set generates connections among the objects.

2.3.5 Identifying modules for group formation

Communities are locally dense connected subgraphs in a network, so nodes that belong to a community have a higher probability to link to the other members of that community than to nodes that do not belong to the same community. Our key idea is that finding communities in (multilayer) networks of the proposed models can be used to solve group/cell formation problems of PFA. To formalize the cell formation problem we utilized the modularity measure introduced by Newman[56] and improved for bipartite graphs by Barber[57].

A module of the network consists of a subgraph whose vertices are more likely to be connected to one another than to the vertices outside the subgraph. Modularity reflects the extent, relative to a random configuration network, to which edges are formed within modules instead of between modules. The modularity can be determined for each community of a network (in PFA this means the modularity of each production cell can be calculated) . For a network with n_c communities, the following modularity value is used to determine the modularity value of community Q_c in terms of each C_c community with N_c nodes connected by L_c links, $c = 1, \dots n_c$:

$$Q_c = \frac{1}{L} \sum_{(i,j) \in C_c} (a_{i,j} - \frac{k_i k_j}{L}) = \frac{L_c}{L} - \frac{k_i k_j}{L^2} \tag{2.9}$$

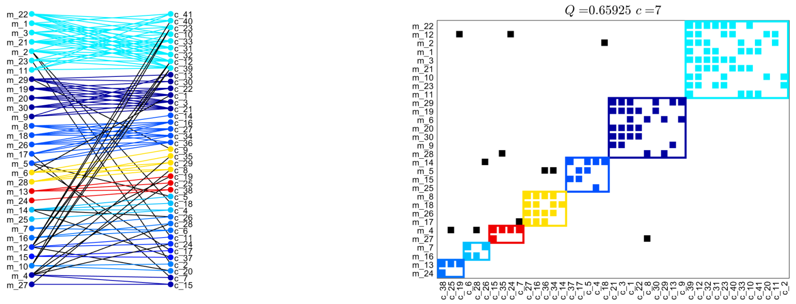
If the Q_c modularity value of a cluster is a positive value, then the subgraph C_c tends to be a community. The modularity of the full network can be evaluated by summing Q_c over all n_c communities, $Q = \sum_c Q_c$.

As can be seen, the definition of modularity perfectly fits the problem of manufacturing cell formation. Therefore, we propose a graph modularity maximization based approach for this purpose. In this study we adapt the Newman[56], LP-BRIM[58] and Adaptive BRIM[57] algorithms available in the BiMAT MATLAB toolbox[59].

To illustrate the applicability of this approach, Figure 2.6 visualizes a cell formation problem and how the extracted modules can be assigned as manufacturing cells.

The efficiency of the formation of the cell can be evaluated based on $\mathbf{e}$, the total number of activities, $\mathbf{e_0}$, the number of exceptional elements that are excluded from the cells, and $\mathbf{e_v}$, the number of zeros in the cells[60]:

$$\Gamma = \frac{e - e_0}{e + e_v} \tag{2.10}$$



(a) The rows and columns of the biadjacency matrix of the bipartite graph can be reordered to visualize the similarities of the modular graph layout.

(b) After reordering/serialization of the biadjacency matrix the modular structure of the problem is revealed.

Figure 2.6. Modularity analysis of the 30x41 machine-part benchmark example.

Table 2.6 compares the efficiencies of cell formation achieved by the proposed clustering and the modularity-based algorithms of cell formation with recently developed advanced goal-oriented optimization results in several benchmark problems of Ref.[60]. As can be seen, modularity-based algorithms perform surprisingly well, the Γ values (given as rounded parentages) are near to the optimized performances, and most importantly, the number of machine-part matchings outside of the modules (e_0 values) and the number of modules are much smaller in almost all cases than the optimized reference solutions.

The BRIM algorithm assigns nodes into modules successively to maximize the per-node contribution of modularity given prior assignments. In that way, each set of nodes recursively induces the other set of nodes. BRIM assigns nodes of each type to modules until a local maximum is reached[57]. LP is simple iterative method, where initially, every node is assigned with a unique label, which represents the community it belongs to. At every step, each node updates its label to a new one which is the most of its neighbors have. If a node has two or more maximal labels, it picks one. In this iterative process, densely connected group of nodes can reach a consensus on a unique label and form a community quickly. LP-BRIM is a combination of LP and BRIM. BRIM refines the result. Adaptive BRIM introduces an other heuristic to BRIM, by randomly selecting initial groups[61].

Based on this success, several modularity optimization algorithms were applied. As will be demonstrated in the following section, the approach is also applicable when searching for modules in multiple layers by the multilayer InfoMap algorithm[62, 63].

Table 2.6. Cell-formation efficiency of bipartite-modularity optimization algorithms. (The Γ values are given as rounded parentages.)

Problem size	Optimization[60]			Newman			LP-BRIM			Adaptive BRIM		
	#c	Γ[%]	e_0	#c	Γ[%]	e_0	#c	Γ[%]	e_0	#c	Γ[%]	e_0
14 x 24	7	72	10	4	67	2	4	67	2	8	62	19
20 x 20	5	43	50	4	41	48	4	40	48	4	41	50
24 x 40	11	53	50	7	41	51	7	40	48	8	43	50
28 x 46	10	45	60	4	37	58	3	33	49	5	39	63
30 x 41	10	59	40	6	45	11	7	51	11	8	52	12
30 x 50	12	60	75	9	44	59	10	47	66	9	44	63
37 x 53	3	59	337	4	49	391	3	53	338	2	53	301

2.4 Application to the analysis of wire-harness production

To provide a detailed and reproducible case study for production flow analysis, an open-source benchmark model of modular wire-harness production was developed. The details of the model are given in the supplementary material of the paper (see the Appendix). The multilayer network model of the production flow analysis problem is formed and analyzed in the MuxViz framework developed for the interactive visualization and exploration of multilayer networks[64]. The established network is depicted in Figure 2.2.

2.4.1 Similarity and modularity analysis

Analysis of the reducibility of a multilayer network provides useful information about the similarities of the layers[65, 66]. To demonstrate the applicability of this metric the **C**, **Z**, **S**, **O** and **T** layers were analyzed (see Figure 2.7).

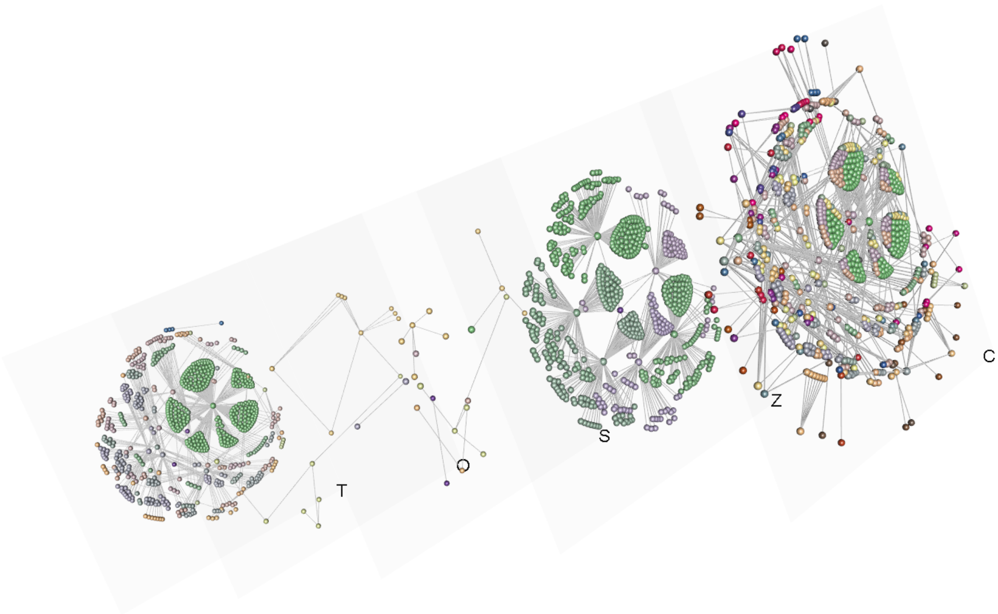


Figure 2.7. Multilayer network representing the details of the work of the operators (built in components: **C**, zones of the activities: **Z**, skills: **S**, assignment of the operators to the workstations: **O** and activity types: **T**. (see Table 1 for the detailed definition of the layers)

As can be seen in Figure 2.8, based on the reducibility of the network two clusters were formed. The first cluster is related to product-process (Z-T-C) layers, while the second collects the operator-skills (O-S)-relevant information. The importance of the definition of the activity types (layer T) is also highlighted.

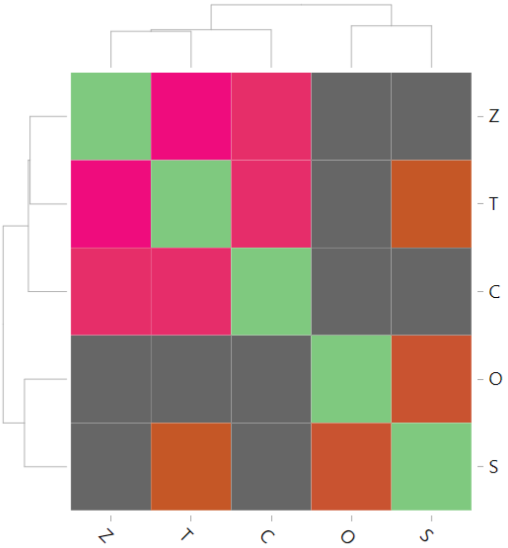


Figure 2.8. Analysis of the reducibility of the model provides useful information about the similarities of the layers. In our case the two clusters related to product-process (Z-T-C) and operator-skills (O-S) were revealed. The importance of the definition of the activity types (layer T) is also highlighted.

Although our network defines part families indirectly in layer **M** and also groups of these activities (in layer **T**), it is interesting to observe how the multilayer network is structured and how the analysis of the modularity of the network can form part and activity groups. For this purpose, a multilayer InfoMap algorithm was applied[62, 63].

The analysis yielded useful and informative results. 26 modules were identified. Although layer **M** which represents how the activities are grouped according to different products, this analysis was able to detect the modules of the products $(m_1, \ldots, m_7)$ in terms of the types of the activities $(t_1, \ldots, t_{16})$. This result confirms that the analysis of the modularity of the proposed multilayer network model is useful in fine-tuning the existing part families based on multiple aspects representing the layers of the model.

To demonstrate how such information is useful in the early process-design phase to define technical modules, layer T of the $C - Z - S - O - T$ multilayer network is shown in Figure 2.9. As can be seen, the most significant module is separated into six smaller groups by following the structure of layer Z that defines in which zone the activities occur. The central role of the most frequent and widely distributed t_{10} type of activity (wire-terminal attachment) is also highlighted.

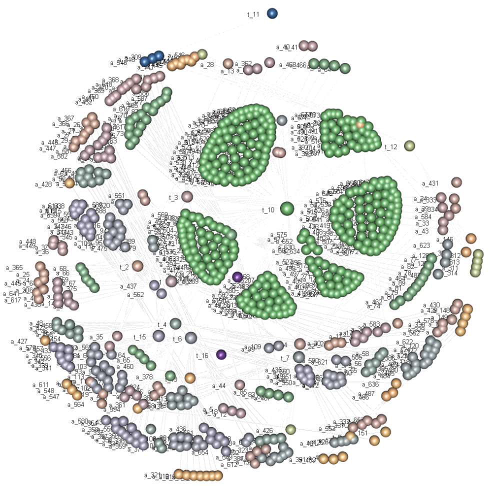


Figure 2.9. Layer $\mathbf{T}$ of the network defines the types of activities. The six clusters formed in this layer reflect the effects of how the activities are distributed among the zones (defined by layer $\mathbf{Z}$), which illustrates the benefit of the multidimensional network-based visual exploration of the production data.

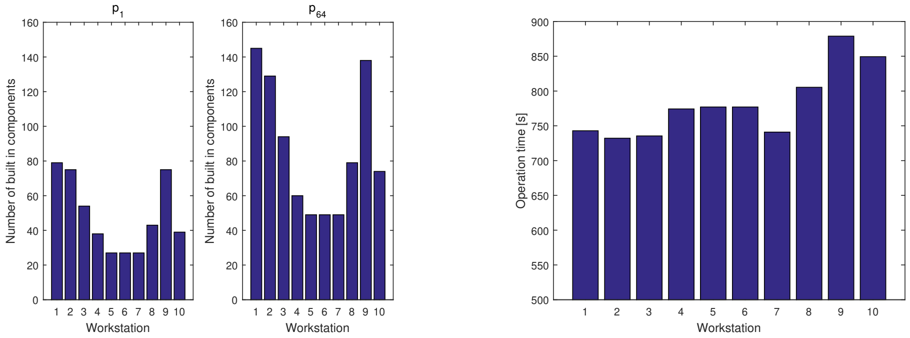
2.4.2 Workload analysis

The balancing of modular production is challenging due to the great diversity of products[67]. Besides group formation, the analysis of the workloads is also an important task of in production flow analysis. The proposed bipartite network-based model can be directly applied for this purpose as the biadjacency matrices of the layers result in simple calculations. To illustrate this applicability let us consider the analysis of how well the production line is balanced. The equation $\mathbf{L}_a = \mathbf{M}\mathbf{P}'_p$ represents the activities of the production of the p-th product (where $\mathbf{P}_p$ represents the p-th column of the P product-module matrix). As these activities are assigned to the workstations as $\mathbf{L}_w = diag(\mathbf{L}_a)\mathbf{W}$ and $\mathbf{T}'\mathbf{L}_w$ represents the number of activities grouped by activity types and $\mathbf{T}'\mathbf{C}\mathbf{C}'\mathbf{L}_w$ is the number of built-in components at the workstations, the total activity time at the workstations can be calculated by the following equation, where θ_t represents the elementary activity times given in the appendix:

$$\mathbf{l}_{time} = \left[\mathbf{T}'\mathbf{L}_w\,,\mathbf{T}'\mathbf{C}\mathbf{C}'\mathbf{L}_w\right]\theta_t \tag{2.11}$$

As Figure 2.10 illustrates, the calculations above can be used to check how the process is balanced and how the complexity of the product influences the workloads of the workstations.

Although the presented workload analysis is not unique to the proposed model, we believe that the results demonstrated the rich information content and broad applicability of multilayer networks which can also be interpreted as a linear algebraic approach model of the system.



(a) Number of built-in components at a given workstation. The figure shows how the workload differs during the production of the base module (p_1) and the most complex product (p_{64}).

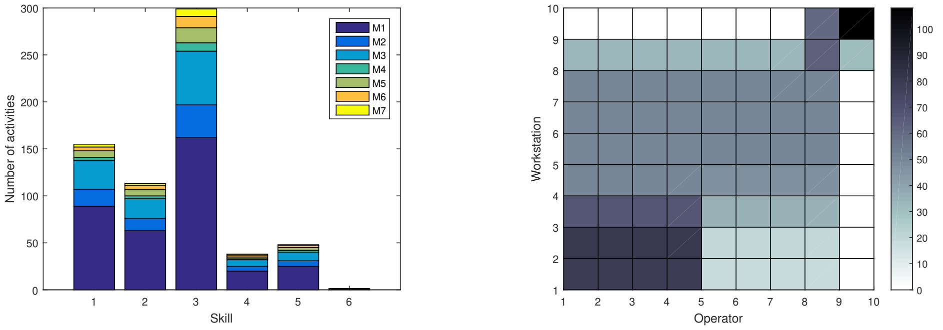
(b) Total station times during the production of the $16th$ product.

Figure 2.10. The workloads (number of activities, built-in components and total activity times) can be easily calculated based on the biadjacency matrices of the proposed model, which supports the balancing of the conveyor belt.

2.4.3 Analysis of the flexibility of operator assignment

In early 80's[68] suggested that organisational research should incorporate network perspective. In early 90's six themes (turnover/absenteeism, power, work attitudes, job design, leadership, motivation) dominated the research of micro-organisational behaviour[69]. Recently multilayer networks are becoming widely used in the analysis of social networks where people interact with each other in multiple ways like via mobile phone and emails[70, 71, 72, 73, 74]. In this chapter, we make the first attempt to integrate such analysis to the modelling and optimisation of production process.

For successful line balancing of wire-harness production, the skills of the operators influencing the speed of the conveyor belt should also be studied[43] and handled[75]. Dynamic job rotation[44] requires efficient allocation of the assembly tasks whilst taking into account the constraints related to the available skills of the operators. Figure 2.11 shows the distribution of the required skills as a function of different product modules, $\mathbf{M'TS}$. As can be seen, the most in demand is the s_3 terminal-attaching skill, while s_6 is the visual testing skill which is required only once during production. The abilities of the operators can also be calculated, e.g. $\mathbf{W'TSO'}$ yields how many activities can be performed at a given operator-workstation assignment (see the left side of Figure 2.11).



(a) Distribution of the required skills as a function of the modules of the product.

(b) The number of activities that can be performed during a given operator-workstation assignment. The skills of the operators constrain the flexibility of line balancing.

Figure 2.11. Analysis of the demand of skills and the flexibility of the operator-workstation assignment.

The presented analysis can be useful in designing the sessions of the operators by determining the components of critical skills and knowledge. Figure 2.12 shows the layers $\mathbf{S}$ and $\mathbf{O}$ of the network. Five groups of activity, skill and operator nodes were identified with the help of multilayer modularity analysis. The smallest module contains the t_{15} clip installation activity type which requires specialist skills.

As can be seen, the skill s_4 can be considered a key piece of knowledge, because it is related to five types of activities. Operators o_9 and o_{10} possess specialist knowledge, while s_3 consists of group-wise knowledge because it is the most related to the operators.

The presented analysis demonstrated that the analysis of the node degrees can identify the critically essential skills and resources. Skills that have small degrees in the $\mathbf{O}$ layer can be considered as the knowledge of specialist, while skills with large degrees are quantified as group-wise knowledge. Skills that have no links at the $\mathbf{S}$ layer are useless, while skills

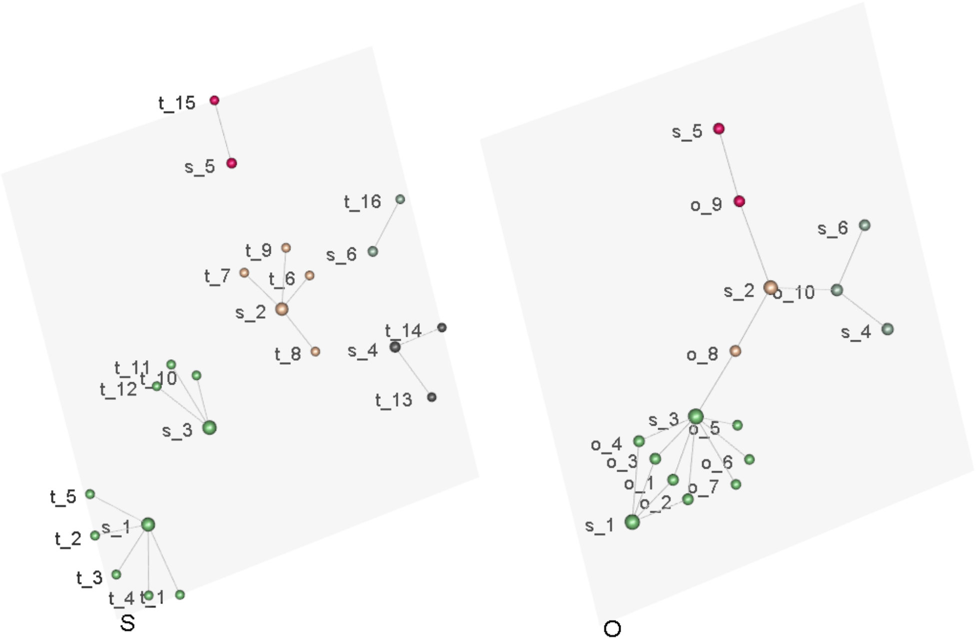


Figure 2.12. Skill (**S**) and operator (**O**) layers define the network that can be used to determine elements of critical knowledge which is useful in terms of the design of training programs for the operators.

that have a small degree at the **O** layer and high degree at the **S** layer are critical, as this reflects that a small number of operators can be assigned to a large number of tasks which requires this knowledge.

2.5 Conclusions

A multilayer network model was developed for production flow analysis to represent the physical and functional domains of production systems by taking into account the aspects of the structure of the system, the variety of machines, products, components and operators, and their interdependencies.

Most of the layers of the model are represented by a bipartite graph, where edges represent material, energy or information flows, and attributes of the objects represented by the nodes of the graph. It was highlighted that the nodes and connections could be easily defined based on standards of process management. As the layers of the network represent different aspects of the production system, the proposed model is flexible and easily extendable.

Following the introduction of the new modeling concept, it was demonstrated how the tools of network science should be used to support production flow analysis. Firstly, it was shown that the analysis of the paths in the network provides useful information about hidden, previously undefined connections. It was recognized that modularity analysis of the network is a promising tool for forming groups in PFA, and the performances of advanced (bipartite and multilayer) network modularity algorithms (like InfoMap) are comparable to the most advanced optimization algorithm tailored to the problem of cell formation.

A detailed benchmark problem was developed to make the research of multivariable algorithms of production flow analysis reproducible. With the help of the studied wire-harness process, the benefits of the modularity analysis of problem-specific sets of layers were demonstrated. The results confirm that the detected groups of activities are useful in terms of fine-tuning of modules (part families). Workload and capability-related network measures were developed. Along with analysis of the node degrees and their correlations, individual-, key- and group-wise skills could be identified. The biadjacency matrices of the network lead to the calculation of workloads and the investigation of how the production line is balanced. Besides the numerical analysis, visualizations were presented to demonstrate how multilayer networks provide insights into the critical factors of interconnected production systems, and the results of which confirm that multilayer networks can support the integration of production-relevant data and decision-making related to complex production systems.

Since the handling of the time-varying behaviour of process systems is becoming ever more critical in the field of cyber-physical systems, the future of this work will focus on the integration of historical process data to define networks of sequential procedures and temporal connections.

Chapter 3

System dynamics models

System Dynamics models are being published rapidly, their automated structural analysis is still absent. This chapter provides a methodology to structure and visualise the information content of these models. The novelty of the present approach is the development of a network analysis-based tool for modellers to measure the importance of variables, identify structural modules in the models and measure the complexity of the created model, and thus enabling the comparison of different models. The typical topics and complexity of SD models highlight the need for tools that support the automated structural analysis of sustainability problems. For practising engineers and analysts, nine models from the field of sustainability science, including the World3 model, are studied in details. The results highlight that with the help of the developed method the experts can highlight the most critical variables of sustainability problems (like arable land in the Word3 model) and can determine how these variables are clustered and interconnected (e.g. the population and fertility are key drivers of global processes).

3.1 Introduction

Given the globally interlinked and constantly changing economic, environmental and social aspects of today, the understanding of ongoing processes is of crucial importance in meeting the needs of current and future generations and, therefore, in establishing and maintaining sustainable development. However, scientific knowledge plays a key role in the determination of the aspects of unsustainability, moreover, with regard to forming the policies of sustainable development, the development of such a knowledge base for the promotion of sustainable development faces significant difficulties[76]. One of the most urgent problems is that although several policy documents are being published[77], the demand for in-depth knowledge is not sufficiently met, thus, the formation of policies is lagging behind the currently emerging challenges[78] and this is a core problem in such dynamically improving fields as the green supply chain management as well[79]. The design and monitoring of sustainability policies should rely on models that can capture the complex dynamics of interconnected variables and sustainability-related subsystems. Systems thinking and system dynamics (SD) explicitly aim to facilitate the understanding of such complex systems and the construction of models that describe their characteristics. Realising that more and more system dynamics models are being published in the field of sustainability and the potential in their structural analysis, the present study aims the introduction of a methodology for their automated structural analysis.

Several questions have arisen given the constant changes and improvements of sustainability related issues in addition to the close connection between policy-making and model building. Policymaking and models of system dynamics in the field of sustainability should go hand in hand, meaning that policymakers should validate their assumptions with regard to dynamic models, and the developed models should incorporate the latest aspects of sustainability. Therefore, it is reasonable to question how these close interactions can be handled manually? How can the ever more widespread models, that often consists of hundreds of variables and even more interactions, be tested for all the modelled variables and interactions? How should the policymakers process long documents and compare them with the results of dynamic models? In the present chapter, just such a methodology to support the automatic analysis of issues concerning sustainability is proposed.

Our contribution to the analysis of SD models is a novel methodology as are the related algorithms for the extraction of interactions from dynamic models and their network-based representation and analysis. This network-based representation can help to unfold the missing aspects of policies and reveal the oversimplified or neglected subsystems of dynamic models. The representation also makes the toolbox of network analysis applicable to the characterisation of the nodes and their connections[80]. The model indicates the complexity of the real system and helps the different models to be compared since complexity measures also help to standardise the comparison of models[81]. Our goal is to give feedback to the modeller about how complex the created model is and provide further insights into each element of the model, e.g. different sensitivity and importance measures, as well as overall structural metrics, namely modularity, density and connectivity. Besides the network metrics, the models are also presented as cognitive maps and hierarchical CLDs to further simplify the understanding of the models.

The applicability of the developed methodology and the related tool is presented in terms of the analysis of the most widespread system dynamics model of sustainability according to the citations, namely the World3 model. Furthermore, nine models have been selected with different complexities according to our variable and edge number-based subjective judgment to prove the efficiency of the developed methodology as well as present the applicability of the relationship between different metrics.

3.1.1 Motivation & Contributions

As the complexity of the sustainability-related system dynamics models increases, it is getting more and more difficult to understand the interactions between the different model elements and obtain a clear vision of the undergoing processes. However, the activity of model building needs to be validated, elaborated on and supported to maintain the preciseness of its meaning.

Therefore, during the construction and validation of these models, several research questions arise. What are the key elements of the analysed model? How interconnected are the elements or are there any well-separated subsystems in the model? What is the purpose of these subsystems? Which are the most important variables of the model that are suitable for controlling or monitoring the undergoing processes? These are just some exemplary research questions that can be hardly answered in the case of complex system dynamics models with numerous variables.

In the present study, just such a methodology is provided to answer these questions and facilitate model building and analysis. The main steps of the presented methodology are depicted in Figure 3.1. The building of models usually starts with the definition of the modelled aspects and variables. Then, the interactions of the defined variables are theorized in the form of a cognitive map, which serves as the outline of the constructed system dynamics model. After the validation of the model with measured datasets, the analysis of the results can be carried out. However, complex models with numerous variables are conveniently analysed in network-based representations.

Finally, the network related measures give a brief and fast view of the modelled process and to the interactions in the model. Moreover, having multiple models of the same phenomena, the comparison of the models can be carried out using the generated metrics.

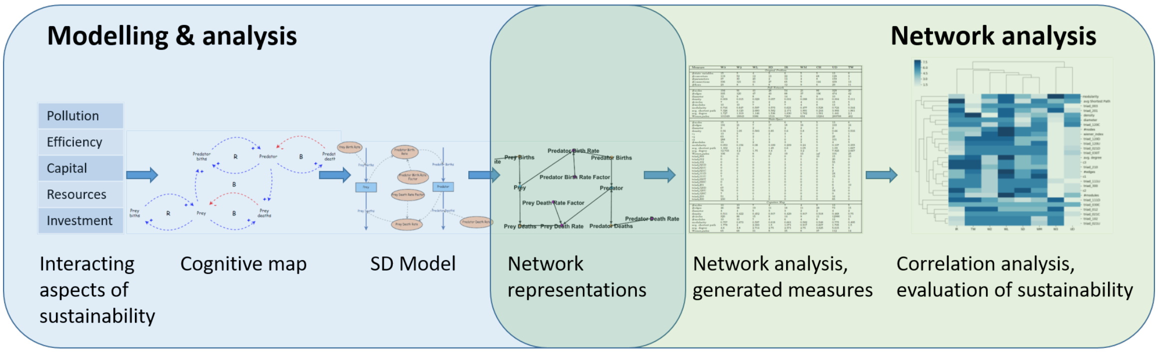


Figure 3.1. Workflow of the proposed methodology

In accordance with the previously presented workflow, the contributions of the present chapter are multiple-fold:

- A complete and categorized bibliography is provided for the reader in Section 3.2.
- The different network-based representations of system dynamics models are provided and are ready to use by the reader for analysis purposes (Section 3.3.1).
- The metrics suitable for the characterisation of system dynamics models can be applied to compare the different models (Section 3.3.2).
- A software is developed in the Python Programming Language following step by step the introduced methodology for the automated analysis of SD models. The software tool is openly and fully available on the website of the authors[1].
- An exemplary application of our algorithm is presented through the analysis of nine models with varying degrees of complexity providing a benchmark application example for practising engineers and decision makers.

Based on the previously listed contributions, the present study, the related software tool, the provided models and their benchmark analysis results provide a ready-to-use solution

---

[1] www.abonyilab.com

for practising engineers for the automated structural analysis of system dynamics models. Even though the tool was explicitly designed for the analysis of system dynamics models in the field of sustainability, the system dynamics models of any field can be analysed analogously to steps presented in the methodology.

The roadmap of the chapter is as follows: The Methodology section (Section 3.3) starts with the description of the state-space model-based representation of SD models with the description of how the SD models are adapted to labelled networks, then to the state-space model (SS), and finally to the cognitive map-based representation (CM) according to the proposed methodology. In the following, the possibility of the application of the toolbox of network analysis using the resultant network-based representation of SD models is provided. The methodological part concludes with the description of the developed software tool. After the in-depth description of the proposed methodology, its applicability is presented in terms of the analysis of the World3 model and other selected SD models of sustainability (Section 3.4).

3.2 System dynamics models of sustainability

In the present section, first, a brief introduction to the origins of SD modelling in the field of sustainability is provided in Section 3.2.1. Then, the toolbox of SD together with the overview of its application in sustainability-related studies from the past few years is described in Section 3.2.2. This is followed by the description of the analysis tools of system dynamics in Section 3.2.3.

3.2.1 The origins of SD modelling in the field of sustainability

The origin of systems thinking in terms of sustainability dates back to Forrester's revolutionary work in both system dynamics and sustainability, World Dynamics in 1971[82], who was the first to interpret the trends in sustainability as a result of dynamic processes. One of the most well-known works concerning the topic, "The Limits to Growth", was published by[83], which not only focuses on the simulation of the issues of sustainability worldwide but on the deep conceptual understanding of the modelled socio-economic and environmental systems as well. The World3 model published in this book is a system dynamics model discussing the issues of sustainability in the view of the interactions between population, industrial growth, food production and the limits of the ecosystem of the earth. The model has undergone several improvements by scientists all over the world, e.g. the WorldWater model was introduced by[84] and the calibration of the original model was updated by[85]. The model was also improved by the authors themselves, first revising several scenarios of growth pointing out that the global industrial system has already overshot some of the Earth's ecological limits[86] and then they published the 30-year update of the original book[87].

3.2.2 The toolbox of system dynamics and the overview of its application in sustainability-related studies

The modelling of sustainability-related complex systems can be considered to be almost impossible due to practically infeasible experiments. Therefore, the lack of accurate measurements, the immaturity of theories, the susceptibility of the system to unpredictable external perturbations, and nonlinear system behaviour are pressing issues[88]. By supporting the understanding of forces that interact in complex systems[89], system dynamics is a powerful tool for the analysis of these systems and their underlying dynamics, providing a useful tool for decision support[90].

System Dynamics covers a set of qualitative tools for the analysis of dynamic processes, e.g. Causal Loop Diagrams (CLD), Stock and Flow Diagrams (SFD) simulation and optimisation software[91]. These tools may facilitate a wide range of computational experiments to examine the various possible outcomes, suggest hypotheses to explain puzzling data, discover significant connections, classes or thresholds among the ensemble of plausible models. In addition, they support reasoning based upon the results of risk analysis, opportunities or scenarios[88]. Even though these dynamic simulations of sustainability-related issues can assist the work of policymakers and engineers, the principle that all models are as good as their underlying assumptions and simplifications must not be ignored.

The application of the SD toolbox has received criticism as well. Formerly[91] discussed the bad practices concerning the application of SD for policymaking. According to critics, SD models are often biased due to the combination of strategic resources and capabilities identified by the modeller in advance[92], and as a result of the problems with regard to understanding and managing reinforcing and balancing feedback loops, which is often achieved by testing[93]. Moreover, scientists must also face the fact that the models and theories that are used to generate their results relate to a number of their own assumptions, beliefs and values[94], thus all models are wrong and systems thinking requires humility about the limitations of our own knowledge[95]. Here, without the aim for completeness, we need to shortly mention that not only these SD models are available for the analysis of sustainability issues, but other data-based approaches, like fuzzy logic reasoning and sensitivity analysis[96], Monte Carlo-based sensitivity analysis of models[97] or Markov chain-based modelling[98].

SD models in the area of sustainability have been published in numerous journals, and their areas of applicability areas also cover a wide range of topics. In order to help the researchers and experts of the field of modelling in sustainability science, in the following an overview of the recently published articles is provided. The database of the Web of Science contains 874 articles by simultaneously searching for the terms "sustainability" and "system dynamics". Most of these articles are in the field of environmental sciences (192), green sustainable science technology (103), engineering environmental (81) and environmental studies (79).

A more insightful categorization is based on the analysis of the abstracts of the 731 articles selected from the database of Scopus by the simultaneous search for the terms "sustainability" and "system dynamics". The results of text mining highlight the frequently co-occurring word pairs in the analysed abstracts as depicted in Figure 3.2. The nodes represent the keywords of each abstract and two nodes are connected by a link if they frequently co-occur in abstracts. The communities of the resultant network (nodes of the original network that are more connected to each other than to the rest of the network) were determined and the detected modules labelled by differently coloured nodes. As a result of this analysis, the colours of the nodes also represent the large categories and therefore, the topics of the related articles. The module presented in green collects the articles from the applications, e.g. climate change or the topic of water - these articles form the majority of the publications in the field of sustainability. The module presented in pink indicates the articles that deal with indicators, policymaking and management. The light blue nodes represent the articles dealing with the social aspects of sustainability. Finally, the nodes presented in blue denote the publications that deal with the modelling of sustainability-related issues. The present article belongs to this category, therefore, our aim is not to investigate a concrete problem but analyse the models and provide a tool for engineers and sustainability experts for the structural analysis of system dynamics models in the sustainability domain.

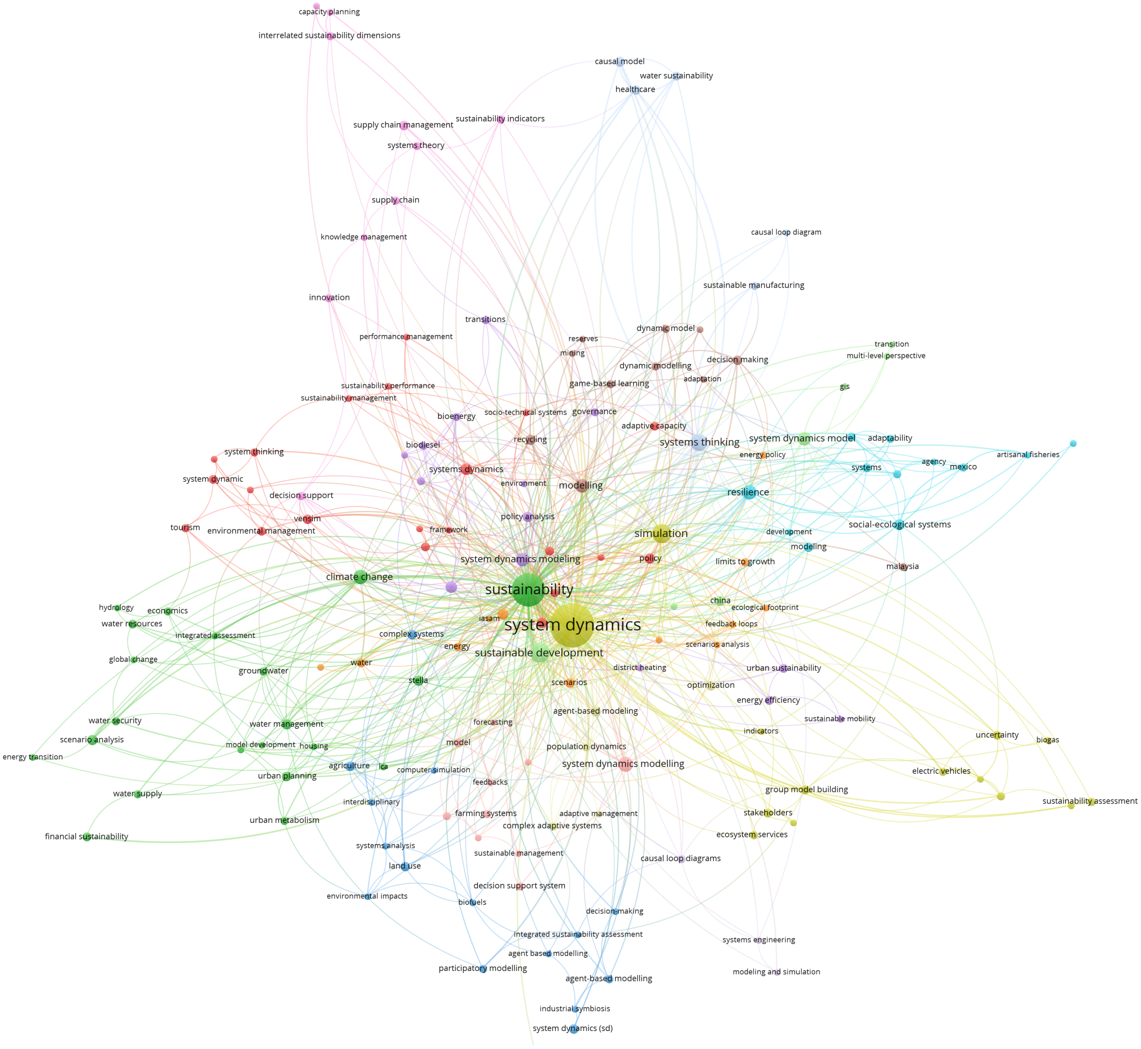


Figure 3.2. Network-based representation of the co-occurring words mined from the abstracts of the articles queried from the database of Scopus by the simultaneous search for the terms of "sustainability" and "system dynamics".

The core concept of counting the state variables was based on the idea that modelling is a cognitive process and it is possible to simultaneously think of 7 ± 2 things[99], which strongly supports the critique of SD models. The distribution of the number of state variables in the 130 models collected is presented in Figure 3.3. The overview is attached in the appendix Chapter 5.2. The presented distribution confirms that the primary aim of modelling is often didactic or demonstrative to support knowledge transfer and, therefore, the number of models containing only one or two state variables is very high. The number of models containing two or three submodules is also outlined by the peaks of 13 and 23 variables, respectively. Here, it should be mentioned that the models contain three submodules in most of the cases, that discuss the environmental, financial and social aspects of sustainability. The literature review confirms the demand for tools that can effectively handle highly complex models comprised of multiple hierarchical submodules. A summary of existing tools is provided in the next section.

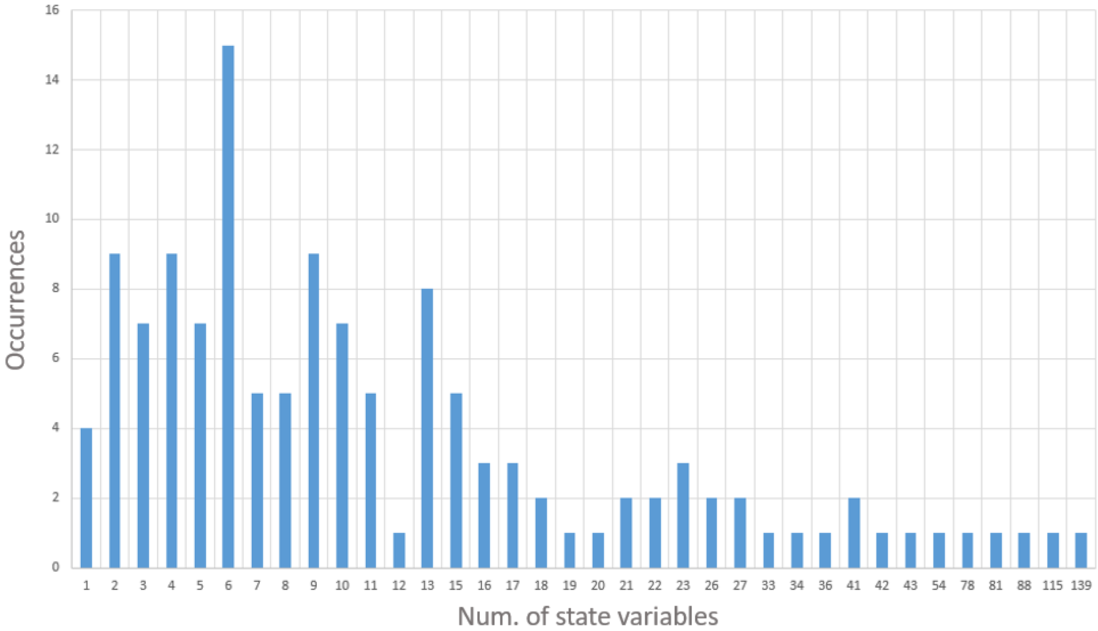


Figure 3.3. The distribution of state variables over the 130 sustainability-related system dynamics models collected from the past five years (2013 - early 2019).

The high proportion of models with numerous state variables highlights the complexity of the analysed models and draws attention to the urgent need for a methodology that supports the construction of sustainability-related system dynamics models. Our aim is to support the construction and structural validation of models and highlight the character of the internal structure of these models.

3.2.3 Analysis tools of system dynamics

A wide range of software supports the construction and analysis of SD models. Table 3.1 is a compilation of the most frequently applied programs for the construction and simulation of SD models. The table summarises the software according to licensing, the analytics of the supported model and the computational technique of dealing with uncertainty, e.g. Monte Carlo simulation, Latin Hypercube sampling, direct manipulation, optimisation possibilities or external data collection. Besides the software listed in Table 3.1, two important applications must be mentioned which only analyse SD models, namely PySD[2] and SDEverywhere[3], by automated documentation and information extraction. The main features of the available software programs together with their web page address in footnotes are provided in Table 3.1.

Table 3.1. Comparison of SD simulation softwares
MA - Model analysis, MCS - Monte Carlo Simulation, DM - Direct Manipulation, OPT - optimization; DATA - external data collection

Software	License	MA	MCS	DM	OPT	DATA
Analytica [4]	subscription based	modules	✓	✗	✗	✓
Vensim [5]	commercial	consistency check	✓	✓	✓	✓
Stella/iThink [6]	commercial	modules	✓	✓	✓	✗
InsightMaker [7]	free	-	✗	✗	✓	✗
AnyLogic [8]	commercial	modules	✓	✓	✓	✓
Powersim [9]	commercial	-	✓	✓	✓	✓

In the present chapter, the construction and analysis of sustainability-related SD models are supported by making them more understandable according to a simple, step-by-step methodology for their analysis. The motivation of our methodology is based on the assumption that the building of SD models and policymaking should go hand in hand, meaning that policymakers should test their assumptions using the tools of SD, while the key variables and subsystems of the models should reflect the latest aspects of sustainability. The method presented here also supports the advice of[100] that the models should be more open - in other words, more interpretable in the present context. This chapter is an extension of our previous conference publication[101] with a more detailed elaboration on the

---

[2] https://pysd.readthedocs.io/en/master/

[3] http://sdeverywhere.org/

[4] https://lumina.com/

[5] https://vensim.com/

[6] https://www.iseesystems.com/

[7] https://insightmaker.com/

[8] https://www.anylogic.com

[9] http://www.powersim.com/

analysis of the network-based representation of SD models and the compilation of a review on sustainability-related SD models from the past five years (2013-early 2019).

3.3 Methodology

In the present section, first, the state-space model-based representation of SD models is discussed together with the introduction of the methodology for the construction of the network-based representation of stock and flow diagrams (SFDs) and the determination of modules inside the networks (Section 3.3.1). This is followed by the description of how the resultant networks should be analysed together with the introduction of the network-related metrics in Section 3.3.2. The section concludes with the discussion of the developed software tool in Section 3.3.3.

3.3.1 Hierarchical Network Representation of System Dynamics Models

The tools of system dynamics modelling are attractive as they provide an easily understandable representation of state-space models. The state-space representation of a system is based on a set of first-order differential equations that describe how the $\mathbf{x}(t) = [x_1(t), \ldots, x_n(t)]^T$ state variables evolve over time as a function of the $\mathbf{u}(t) = [u_1(t), \ldots, u_{n_u}(t)]^T$ inputs and $\mathbf{p}(t) = [p_1(t), \ldots, p_{n_p}(t)]^T$ parameters:

$$\frac{d\mathbf{x}(t)}{dt} = f\left(\mathbf{x}(t), \mathbf{u}(t), \mathbf{p}(t)\right) \tag{3.1}$$

The output equation of the state-space model describes how the measured output variables depend on the values of the state variables:

$$\mathbf{y}(t) = g\left(\mathbf{x}(t)\right) \tag{3.2}$$

The beauty of the models of system dynamics is how state-space models are represented in the form of the easily interpretable stock and flow diagrams. Figure 3.4 is an example of the model to calculate the number of individuals in an imaginary population. As it can be seen in Figure 3.4, stocks (denoted by blue squares) represent the state variables according to the following differential equation that describes the behaviour of the x_i state variable which accumulates or is depleted over time, while flows (denoted by the orange inflow and outflow of the x_i square) affect the rate of change in the stocks represented in terms of the first-order differential equations. Moreover, the models often include parameters (represented by the grey circles) and converters that define the functional relationship between different state variables (which are not presented in Figure 3.4).

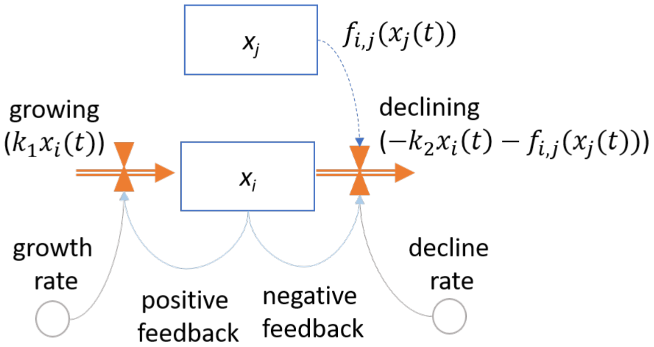


Figure 3.4. The key elements of stock and flow diagrams.

Figure 3.4 can be denoted with the following equation:

$$\frac{dx_i(t)}{dt} = k_1 x_i(t) - k_2 x_i(t) - f_{i,j}\left(x_j(t)\right) , \tag{3.3}$$

where $x_i(t)$ can be considered as a stock variable (e.g. the number of individuals in a population), and $k_1 x_i(t)$ denotes a flow that affects the rate of change in the state variable (e.g. the number of births). Flows are often affected by parameters, e.g. k_1 and k_2 represent the birth and mortality rates, respectively. The state variables can influence each other. Such influence is handled by converters in SFDs, e.g. the $f_{i,j}\left(x_j(t)\right)$ converter, which holds information about what affects the rate of the flows or the value of other converters. The value of these converters in the present chapter is referred to as *variable* and this nomenclature will be used henceforth. In the following, a didactic description of how the state-space models and SFDs are converted into networks that represent different details of the modelled phenomena is given.

Step 1: Transformation of SFDs into labelled networks

As the causal loop diagram-based representation of the SFD formerly presented in Figure 3.4 is depicted in Figure 3.5, our methodology starts with the transformation of system dynamics models into networks, which is by no means a trivial task. SFDs are transformed into labelled networks, where the parameters, converters and stocks are represented by nodes, while the connections between these nodes are represented by the edges.

Although the SFD enables a direct network interpretation, dynamical systems possess internal logic which should also be taken into account. A significant contribution of our work is that attention is drawn to the influence of the outlet flow on the stock itself and

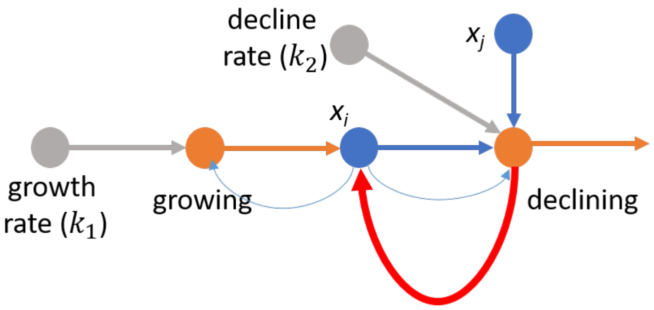


Figure 3.5. Causal loop diagram representation of the converted stock and flow diagram.

the network is extended with the additional edge highlighted in red in Figure 3.5. In the following, the resultant labelled network will be referred to as *Full Network (FN)*.

An example of this representation is given based on the World2 model[82] depicted in Figure 3.6.

Step 2: Extraction of the network of state variables

After the generation of the full network, the network of state variables (stocks) was extracted by network reduction, excluding the nodes that are not state variables. The $\frac{dx}{dt} = \mathbf{A}(t)x(t)$ general state-space model representation of the model easily lends itself to the application of this methodology and the network-based visualisation of state-space models (Figure 3.6(**c**)). The nodes of the resultant matrix are the state variables (x), while the edges are derived from the structure of state-transition matrix $\mathbf{A}$ of the state-space model, as if the element $\mathbf{A}_{i,j} \neq 0$ (j^{th} column of the i^{th} row), then an edge is present from x_j to x_i. This matrix representation is beneficial as it supports the analysis of the dynamical order (complexity), observability and controllability of the system. The characterisation of a dynamical system by a state-transition matrix requires the linearisation of the system around the operating point analysed. Even though the values of the matrix $\mathbf{A}$ can often change in the case of different operating points, the structure of the analysed system, and therefore, the core driving force of dynamics usually remain constant, thus, structural analysis can be conducted using a single linearised state-transition matrix. An important advantage of the applied state-transition matrix is that the adjacency matrix $\mathbf{A}_{adj}$ of the system can easily be formed from it by changing every nonzero element to one and transposing this binomial matrix. Therefore, the element $a_{i,j}$ of the adjacency matrix of a system is one if an edge points from x_i to x_j.

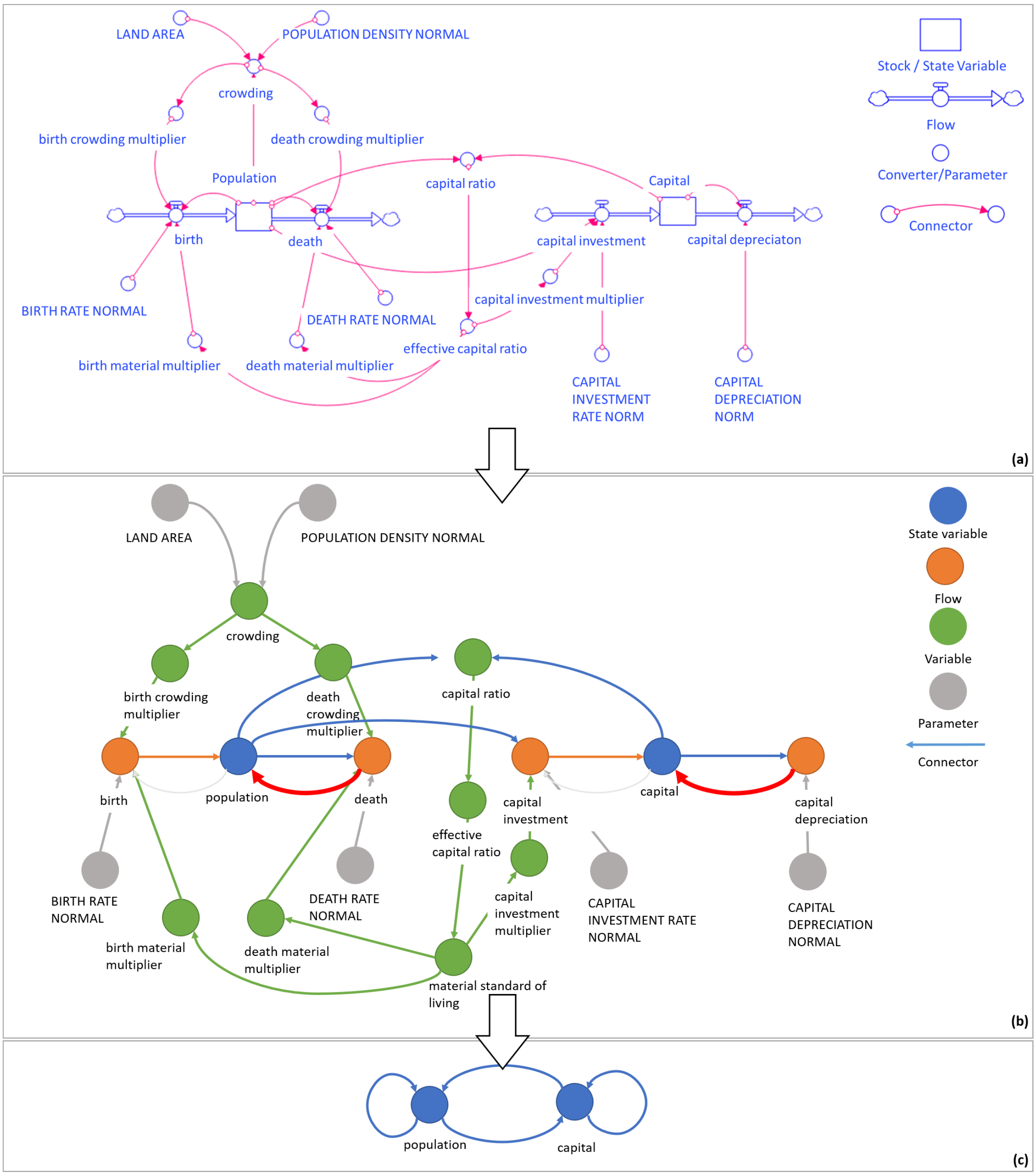


Figure 3.6. (a) Part of the Stock and Flow Diagram (SFD) of the World2 model, (b) full Network-based representation of the Stock and Flow Diagram, (c) reduced state-space representation, where only the effects between state variables are represented.

Step 3: Extraction of network modules to automatically define hierarchical models With the aid of community detection algorithms in networks that identify groups of nodes with more connections between them than with the nodes of the rest of the graph[80], communities in the network are detected by the Girvan-Newman algorithm[102] and it is assumed that the extracted communities are corresponding to the topics of the (network) model which serves as a purely automatized structure-based topic-extraction method. After the extraction of the network, the modules are entitled based on the name of the node that possesses the highest PageRank centrality within the community. An in-depth description of the applied metrics of network theory is provided in the next section.

The detection of submodules inside the original network structure is conducted by the use of the modularity metric. *Modularity* is the fraction of edges that fall within the given groups (modules) minus the fraction of edges that were distributed[103]. These more connected modules are determined by the application of a benefit function Q, which yields high values when the network is divided well into communities:

$$Q = \frac{1}{L}\sum_{i,j}[a_{i,j} - \frac{k_i^{out}k_j^{in}}{L}]\delta(c_i, c_j) \tag{3.4}$$

where the Kronecker delta function (δ) is equal to one when nodes i and j are classified as being in the same module (i.e. they possess the same label value) or zero otherwise and c_i stands for the label of the community to which node i is assigned. k_i^{out} and k_j^{in} denote the in- and out-degrees of a node, respectively, and L represents the number of edges in the network (described in more depth in the next section which is explicitly dedicated to network analysis). If the edges were randomly distributed in the network, $\frac{k_i^{out}k_j^{in}}{L}$ would be the expected number of links between nodes i and j.

The visualization of the network of the communities can provide an easily interpretable overview of the system. Based on an inspection of how the modules are formed and connected as well as what the nodes that connect them are, the more general structure of the model can be revised and the interfaces between modules determined.

Step 4: Extraction of structural motifs as building elements of complex dynamical behaviour

In SD models, it is crucial to investigate both feedback and balancing loops. In our study, it is shown that the complexity of some SD models extends the limits of the classical technique that investigates these loops, thus, they can only be investigated by simulations. The search for non-trivial loops and their quantification is facilitated by network representation. The smallest structures that possess the true character of a "society" are the triads (any triple of nodes). If three nodes and the edges between them are selected in the directed network, one of the 16 combinations is always obtained as illustrated in Figure 3.7. As these triads are the simplest structures in which the emergence of a hierarchy can be seen, the dynamical behaviour of a network is explicitly defined by the number and type of triads. A significant additional function of the proposed methodology is the determination of these triads and their common motifs, which has hardly been interpreted manually, however, a correlation between the motifs and complexity of a model exists.

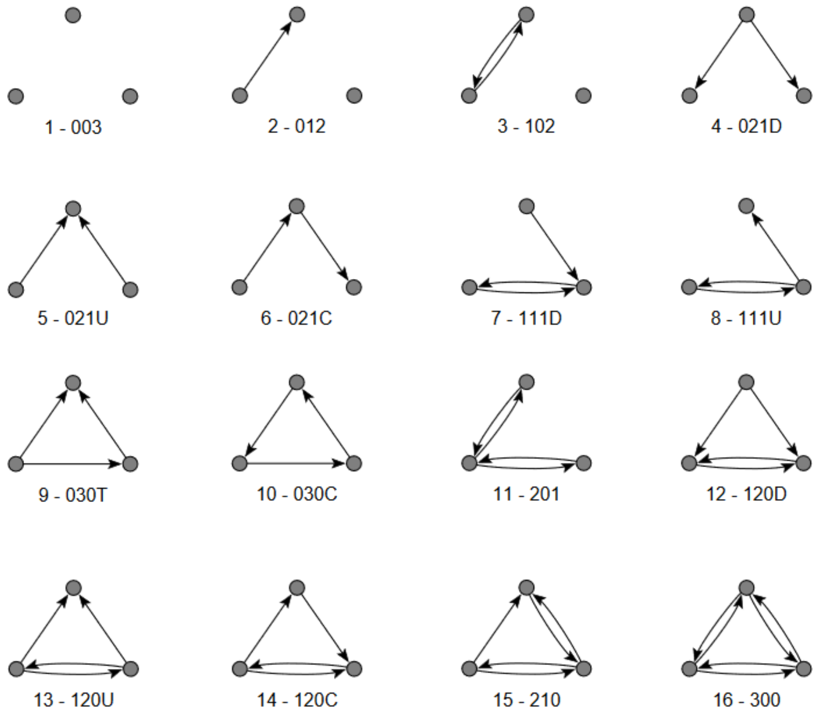


Figure 3.7. Types of triads (Figure adopted from[1])

3.3.2 Network analysis

The network representation of the model facilitates the efficient structural analysis of SD models as it supports the measurement of the importance of the variables, detection of the loops and evaluation of the complexity of the interconnectedness. The main indicators that were utilized to evaluate the networks as a representation of the original system are summarized in Table 3.2.

Table 3.2. The meaning of structural measures

Notation	**Measure**	**Description**
N	# of nodes	# of building blocks (e.g. number of state variables)
L	# of edges	# of interconnections in the network
$\bar{k}$	Average degree	Average # of edges per node
c_1	# of self loops	# edges that point back to the node of origin
c_2	# of circles with 2 nodes	# of back-and-forth relationships
c_3	# of circles with 3 nodes	# of 3-node-long loops
d	Density	Ratio of actual to all possible edges
W	Wiener Index[104]	The sum of the shortest-path lengths between each pair of reachable nodes
M	Modularity	Representing the hierarchical structure of the model

From the viewpoint of the analysis of the dynamics of the related models, the loops play a very important role since they represent the existing cross-effects between variables. c_1 loops can be considered as self-loops that point back to their node of origin, while c_2 loops represent the direct effects of variables on each other. c_3 loops can be considered as the smallest real circles of the model and the building blocks of dynamical systems. The three-element relationships are explicitly analysed in the form of triads as described in the previous section.

The introduction of these basic metrics provides the possibility for the definition of further metrics from the toolbox of network theory. Maybe the most evident metrics for the characterisation of a node in a network are the *in-degree* (k_i^{in}) and *out-degree* (k_i^{out}) metrics, namely the number of incoming and outgoing edges of the related node, respectively, and the sum of these numbers, which provides the overall *degree* of the node (k_i^{tot}):

$$k_i^{in} = \sum_j a_{j,i} \tag{3.5}$$

$$k_i^{out} = \sum_j a_{i,j} \tag{3.6}$$

$$k_i^{tot} = k_i^{in} + k_i^{out} \tag{3.7}$$

For a more in-depth description of the related metrics refer to the work of[105]. By applying in- and out-degrees, the number of neighbours of a given sector was determined and those which have the ability to influence (maximum out-degree) or observe (maximum in-degree) many other sectors were established.

Another important metric is the *betweenness centrality*[106], which measures the number of shortest paths between pairs of nodes that pass through a given node. This measure provides a good insight into how important the presence of the related node with regard to the connection of parts of the network is, serving as an interface between communities. Moreover, nodes with high betweenness centralities are sensitive to changes in other parts of the network.

The *PageRank* metric is recursive: a node that is connected to many other nodes by a high PageRank metric receives a high rank itself as well[107].

3.3.3 The developed program

The results are calculated using our developed tool, which will be fully available on the website of the authors[10]. The software is developed in the Python Programming Language for the automated analysis of SD models as it follows step by step the introduced methodology. The program reads pySD parseable models (currently Vensim and Stella exports) as well as models exported from InsightMaker and it also supports the manual input of the models. Besides the structural analysis it provides, it is also a well-designed infrastructure to build upon. The software enables the manipulation of parameters of all elements of the model due to the flexible classes applied and Python Data Analysis Library (pandas)-based serialisations. For the convenient analysis of the network-represented models, the toolbox of NetworkX is ready to use during all stages of the methodology (FN, SS, CM).

---

[10] www.abonyilab.com

3.4 Results and discussion

Using the developed algorithms, first, the applicability of the proposed methodology through the in-depth analysis of the World3 model is presented (Section 3.4.1), which is followed by the discussion of the network theory-based metrics together with other sustainability-related SD models in Section 3.4.2.

To stimulate further research, the Python code of the developed software tool is publicly available on the website of the authors (www.abonyilab.com).

3.4.1 Detailed analysis of the World3 model

The World3 model is a system dynamics model originally documented and applied in the book 'The Limits to Growth'[83] to obtain a better understanding of the sustainability-related questions and simulate trends with regard to the global population, industrial growth, food production and limits in the ecosystem of the Earth. The main systems in the original model were that of food (dealing with agriculture and food production), industry, population, non-renewable resources and pollution.

First, the stock and flow diagram of the World3 model is transformed into a full network-based representation as presented in Figure 3.8. By detecting the modules in the network of nodes of the full network, the key areas of the modelled processes can be determined according to Figure 3.8. The colours represent the modules of the model, while the size of the nodes and labels is proportional to the PageRank of the nodes. The thoroughness of the modellers is well reflected in this representation since these modules reflect the key areas of the Club of Rome project published in the book, The Limits to Growth (1972), which served as the documentation of the model[83].

The reduced network, excluding every node that is not a state variable is shown in Figure 3.9. This network represents the variables of the state-space model and their structural relationship providing a clearly visible picture of the cross-effects and causal relationships. The size of the nodes is proportional to their PageRank and similarly, the thickness of the edge is proportional to the PageRank of the node of origin. The high PageRank value of the "Arable Land" state variable reflects its central role in the model.

The state variables, i.e. the stocks of the original stock-and-flow diagram, can be characterized based on their proximity to each other, which is an important measure of the ease of controlling or observing these nodes. Therefore, the number of edges along with the

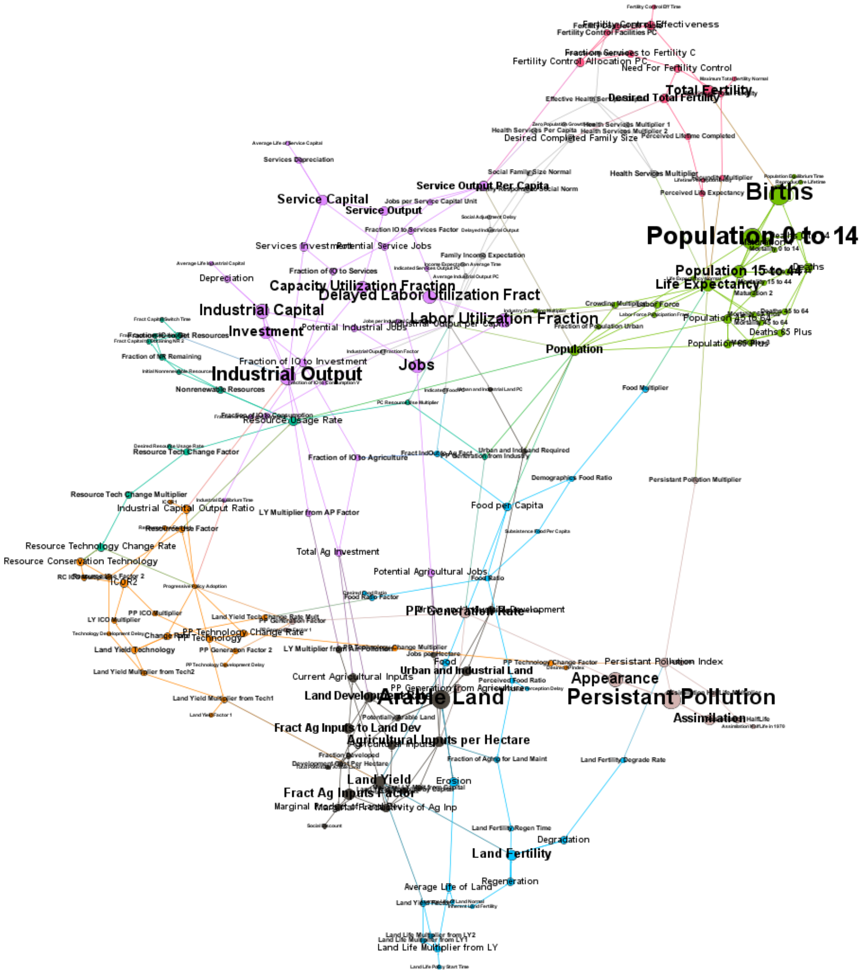


Figure 3.8. The network extracted from the World3 model.

shortest path between two nodes (state variables) is investigated in Figure 3.9. The heat-map in Figure 3.10 shows this metric which is closely related to the influence of the nodes on each other. The causal node is presented in the rows of the matrix, while the columns indicate the effected nodes. The self-loops contain only one edge starting and ending at the same node, therefore, their metric is equal to one. The closely connected nature of the sustainability-related variables of the world is well reflected as most of the variables are separated by only one edge, therefore, they have a direct effect on each other.

Next, the modules of the original full network of Figure 3.8 were investigated yielding the cognitive map of the World3 model presented in Figure 3.11. If the detected module

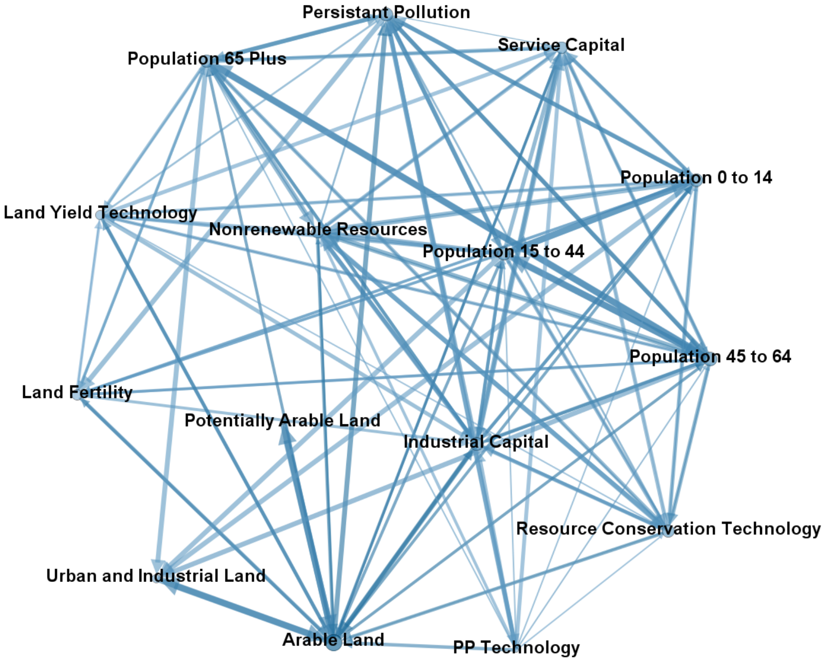


Figure 3.9. The network representing the state-space model of the World3 model. The network contains only the state variables represented by the nodes, whose size is proportional to their PageRank value.

contains state variables, then the name of the module is the name of the state variable with the highest PageRank value. If the module contains no state variables, then the name of the module is the name of the variable with the highest PageRank value. The size of the modules is proportional to the number of incorporated nodes inside them. The figure provides an overview of how the main topics of the World3 model are connected and how they effect each other. While the node with the highest PageRank value possesses a central node in the related module, the nodes with the highest betweenness centrality are the interfaces between the related parts. The role of these nodes is to connect the different modules.

To obtain a view of the nodes with central and interface roles inside the detected modules, the nodes with the highest PageRank and betweenness centrality are presented in Table 3.3. As the underlying assumption of the PageRank metric is that more important nodes are likely to receive more edges from other nodes, the variables with high PageRanks play a central role in the structure of the related community. The betweenness centrality

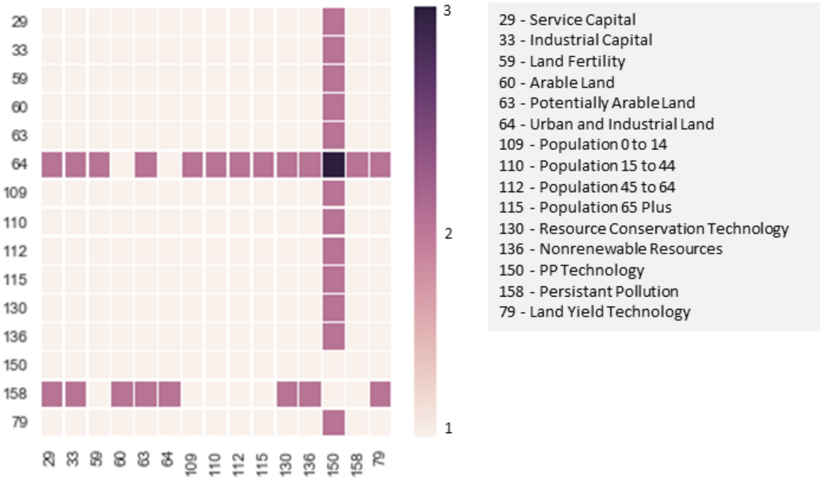


Figure 3.10. The distance between state variables. The numbers denote the IDs of the state variables presented on the right, the rows show the causal variables, while the columns indicate the effected ones.

shows the number of shortest paths between pairs of nodes that pass-through a given node and, therefore, highlights well the variables that connect different modules. As a result, for example, the module ”Industrial Capital” is mainly connected to other modules through the ”Industrial Output” and the central role in this community is played by the ”Investment” flow, which is clearly not surprising given the nature of industrial trends.

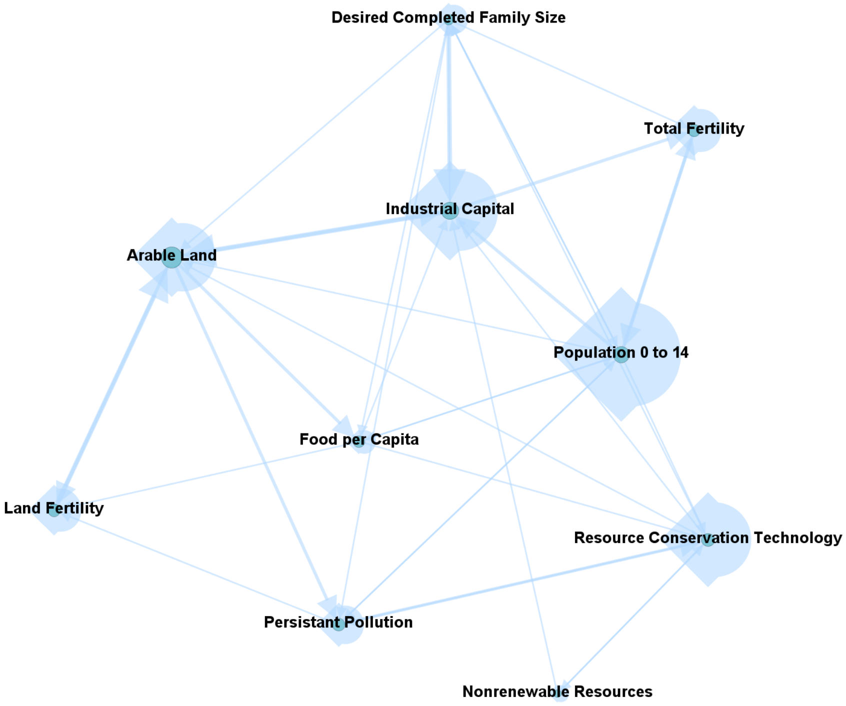


Figure 3.11. Based on the detected communities, the cognitive map of the World3 model can be produced.

Table 3.3. Detailed network metrics of the World3 model, ranked according to PageRank (PR) and Betweenness Centrality (BC) in the left- and right-hand columns, respectively, and grouped by the modules.
(Abbreviations: sv-state variable; v-variable; f-flow)

Name	**Type**	**BC**	**PR**	**Name**	**Type**	**BC**	**PR**
Module 0: Industrial Capital							
PageRank order				*Betweenness order*			
Investment	f	0.066	0.023	Industrial Output	v	0.285	0.019
Industrial Output	v	0.285	0.019	Labor Utilization Fraction	v	0.112	0.016
Labor Utilization Fraction	v	0.112	0.016	Delayed Labor Utilization Fract	v	0.111	0.014
Module 1: Land Fertility							
PageRank order				*Betweenness order*			
Erosion	f	0.030	0.014	Land Fertilitystate	v	0.038	0.008
Regeneration	f	0.010	0.009	Perceived Food Ratio	v	0.038	0.003
Degradation	f	0.025	0.009	Fraction of agricultural inputs for land maintenance	v	0.037	0.003
Module 2: Arable Land							
PageRank order				*Betweenness order*			
Arable Land	v	0.172	0.023	Arable Land	sv	0.172	0.023
Fraction of Agricultural Inputs Factor	v	0.042	0.011	Land Yield	v	0.12	0.011
Agricultural Inputs per Hectare	v	0.066	0.011	Land Development Rate	f	0.09	0.01
Module 3: Population 0 to 14							
PageRank order				*Betweenness order*			
Births	f	0.118	0.044	Population	v	0.23	0.009
Population 0 to 14	sv	0.138	0.023	Life Expectancy	v	0.187	0.013
Life Expectancy	v	0.187	0.013	Population 0 to 14	sv	0.138	0.023
Module 4: Resource Conservation Technology							
PageRank order				*Betweenness order*			
Resource Technology Change Rate	f	0.038	0.011	Resource Usage Rate	f	0.081	0.01
Resource Usage Rate	f	0.081	0.01	Industrial Capital Output Ratio	v	0.079	0.007
Persistent Pollution Technology Change Rate	f	0.038	0.008	Industrial Capital Output Ratio2	v	0.073	0.006
Module 5: Persistant Pollution							
PageRank order				*Betweenness order*			
Appearance	f	0.108	0.026	Persistant Pollution	sv	0.124	0.018
Persistant Pollution	sv	0.124	0.018	Persistant Pollution Index	v	0.121	0.007
Assimilation	f	0.013	0.015	Persistent Pollution Generation Rate	v	0.109	0.011
Module 6: Nonrenewable Resources							
PageRank order				*Betweenness order*			
Fraction of Industrial Capital to Get Resources	v	0.043	0.005	Fraction Industrial Capital to Get Resources	v	0.043	0.005
Non-renewable Resources	sv	0.039	0.004	Fraction of Non-renewable Resources Remaining	v	0.042	0.003
Fraction of Non-renewable Resources Remaining	v	0.042	0.003	Non-renewable Resources	sv	0.039	0.004
Module 7: Total Fertility							
PageRank order				*Betweenness order*			
Total Fertility	v	0.104	0.013	Total Fertility	v	0.104	0.013
Desired Total Fertility	v	0.056	0.009	Service Output Per Capita	v	0.075	0.009
Service Output Per Capita	v	0.075	0.009	Desired Total Fertility	v	0.056	0.009
Module 8: Desired Completed Family Size							
PageRank order				*Betweenness order*			
Desired Completed Family Size	v	0.043	0.006	Industrial Output per Capita	v	0.203	0.005
Industrial Output per Capita	v	0.203	0.005	Desired Completed Family Size	v	0.043	0.006
Fraction of Industrial Output to Consumption	v	0.017	0.004	Family Income Expectation	v	0.021	0.003
Module 9: Food per Capita							
PageRank order				*Betweenness order*			
Food per Capita	v	0.169	0.007	Food per Capita	v	0.169	0.007
Food	v	0.12	0.006	Food	v	0.12	0.006
Fraction of Industrial Output to Agricultural Fact	v	0.038	0.004	Food Ratio	v	0.075	0.003

3.4.2 Illustration of the automated analysis of SD models

To prove the efficiency of the developed tool for the automated analysis of models, illustrate the applicability of the different metrics and reveal the relationships between them, we have selected ten models with various degrees of complexity according to our subjective judgment based on the number of variables and connections. In the following, the analysed sustainability-related models are briefly introduced, the models used henceforth are denoted by bold letters.

The World3 model (**W3**), introduced and analysed in-depth in the previous section, serves as a reference point to interpret the results of the models analysed in this section more easily.

The World2 model (**W2**) is the previous generation of the World3 model. The network-based measures also indicate the superiority of W3 in terms of complexity[108].

In connection with sustainability, another well-known model of demographic, economic and environmental interactions is the Wonderland Model (**WL**)[109].

The next model describes sustainable development from a market-oriented point of view, taking into consideration the connection between economic capital, non-renewable resources, renewable resources, waste and life-supporting systems[110]. This model is denoted by **SD**.

Testing of policy is another area of application of the models of system dynamics. To incorporate models of that test policy in our analysis, an irrigation model (**IR**), which evaluates the impact of water management policies, was analysed to identify critical feedback in the system[111].

Another type of model in the literature aims to demonstrate some complicated options that aid the understanding of decision-makers as well as educational and presentation purposes. The water management options in Las Vegas, Nevada is an example of this type of model showcasing the availability of sustainable drinking water to the city (**WM**)[112].

A recent model of sustainable urbanization performance is a case study of China that discusses the most common sustainability indicators (**CH**)[113].

Another model of urban dynamics (**UD**) is based on the original urban dynamics model by[114].

Lastly, a small model as another reference point, which is an easy-to-understand computer recycling model in Taiwan (**TW**), for the testing of electronic waste recycling policy was created by[115].

First, the full network-based representation of the models is obtained. The influence of state variables is investigated in the network of state variables which is generated as the next step by network reduction excluding the nodes that do not represent state variables. This representation provides visual feedback and an immediate opportunity to validate how state variables influence each other. Then, the cognitive map of the SD model is generated. The cognitive map also holds visual feedback about the predefined structure of the SD model. For validation purposes, it can be used to validate the predefined structural importance and significance of the designated groups. Ideally, each model contains at least one state variable, which also implies the correct configuration of variables and converters in the models.

By analysing the nine models listed previously using the developed software tool, the network theory-based measures were determined. These metrics are presented in Table 3.4. The results are grouped by the analysed representation type and the sign "#" stands for the "number of" expression.

The results presented in Table 3.4 suggest several things. First, based on the results, it is assumed that a good complexity measure of SD models is the density of the state-space (SS) representation. This assumption can be considered straightforward as the density measures the ratio of actual to all possible edges between the nodes, therefore, can be considered as a normalised measure of how closely connected the analysed network is. This definition can be completed with the nature of the SS, in other words, it approximates the dynamics of the model. Since it shows the direct influence between the state variables, an insight can be gained into why the density measure of a network describes well the complexity of the dynamics of the original model.

By observing the metrics of the SS representation of the **CH** model, hardly any interpretable results are seen. First of all, no edges are present between the state variables, therefore, the SS representation is a disjoint network. As a result, the modeller only focuses on the analysed indicators, e.g. resource index, environment index, social index, etc., individually and the only role of other flows, parameters and variables is to calculate these state variables. The analysis shows that such changes are independent of each other and solely dependent on time, which is not in line with a classical SD-based modelling philosophy as the cross-effects are not modelled. On the other hand, the **CH** model is a good example of the visualisation of different indicators and their interconnections aiding a better comprehension. This is also supported by the cognitive map (CM) representation in Figure 3.12,

Table 3.4. Measures of the selected models

Measure	**W3**	**W2**	**WL**	**SD**	**IR**	**WM**	**CH**	**UD**	**TW**
Original Problem									
#state variables	15	5	4	5	6	5	5	13	6
#converters	119	52	12	13	22	3	68	129	3
#parameters	37	30	20	2	14	4	8	159	1
#connections	336	121	41	27	65	9	144	439	13
#flows	23	8	5	8	12	9	6	29	11
Full Network									
#nodes	194	91	41	28	54	21	86	329	20
#edges	335	120	47	43	88	37	136	474	42
diameter	12	9	8	8	14	6	9	10	4
density	0.009	0.015	0.029	0.057	0.031	0.088	0.019	0.004	0.111
#circles	7	0	0	2	8	4	0	15	5
#modules	10	10	7	4	7	4	8	13	5
modularity	0.716	0.647	0.597	0.573	0.621	0.477	0.528	0.724	0.344
avg. shortest path	7.326	3.120	2.680	5.582	2.171	1.881	0.244	3.993	1.861
avg. degree	1.727	1.319	1.146	1.536	1.630	1.762	1.581	1.441	2.1
Wiener_index	101349	18843	3386	1519	7263	634	13264	289788	462
State-Space									
#nodes	15	5	4	5	6	5	5	13	6
#edges	191	21	7	17	18	16	0	103	16
diameter	3	2	2	3	2	2	0	2	2
density	0.91	1.05	0.583	0.85	0.6	0.8	0	0.66	0.533
c_1	13	5	0	2	3	0	0	7	1
c_2	63	7	2	1	1	4	0	25	5
c_3	288	7	1	3	0	2	0	101	0
#modules	13	5	2	3	3	3	5	4	5
modularity	0.053	0.156	0.08	0.189	0.209	0.24	0	0.107	0.355
avg. shortest path	1.162	1.2	1	1.45	0.6	1.05	0	1.09	1.667
avg. degree	12.733	4.2	1.75	3.4	3	3.2	0	7.923	2.667
Wiener_index	106	11	8	12	19	13	inf	101	25
triad_003	0	0	0	0	0	0	10	15	10
triad_012	0	0	1	1	3	1	0	20	0
triad_102	0	0	0	0	1	1	0	55	0
triad_021D	6	0	0	0	3	0	0	3	0
triad_021U	0	0	0	0	2	1	0	8	0
triad_021C	0	0	0	0	3	0	0	2	0
triad_111D	0	1	2	3	0	2	0	24	0
triad_111U	7	1	0	0	0	0	0	15	0
triad_030T	11	0	0	0	5	0	0	1	0
triad_030C	0	0	0	0	0	0	0	0	0
triad_201	0	1	0	1	0	2	0	6	10
triad_120D	67	0	0	1	1	1	0	2	0
triad_120U	76	0	0	1	2	0	0	34	0
triad_120C	7	1	0	1	0	0	0	0	0
triad_210	51	2	1	2	0	2	0	16	0
triad_300	230	4	0	0	0	0	0	85	0
Cognitive Map									
#nodes	10	10	7	4	7	4	8	13	5
#edges	46	38	19	11	18	11	29	73	15
diameter	4	4	5	3	3	2	2	4	2
density	0.511	0.422	0.452	0.917	0.429	0.917	0.518	0.468	0.75
#circles	320	86	15	8	14	9	11	12990	11
#modules	10	10	7	4	6	4	8	13	5
modularity	0.757	0.673	0.597	0.618	0.661	0.582	0.548	0.775	0.495
avg. shortest path	1.778	2	2.333	1.5	1.071	0.917	0.607	1.705	1.5
avg. degree	4.6	3.8	2.714	2.75	2.571	2.75	3.625	5.615	3
Wiener_index	65	68	31	8	35	8	37	112	14

which shows clusters in the absence of state variables highlighting the indirect nature of the interconnectedness of the indices. The infinite value of the Wiener Index confirms that the model is very weakly connected or rather disconnected. The modeller should address this issue during its validation and also highlight it in the documentation of the model.

The results of the detection of triads can also be easily applied by the modellers. The code of the detected triads shows the type of connectivity according to Figure 3.7. In possession of a type of connection and the listed triads, the modeller can investigate the

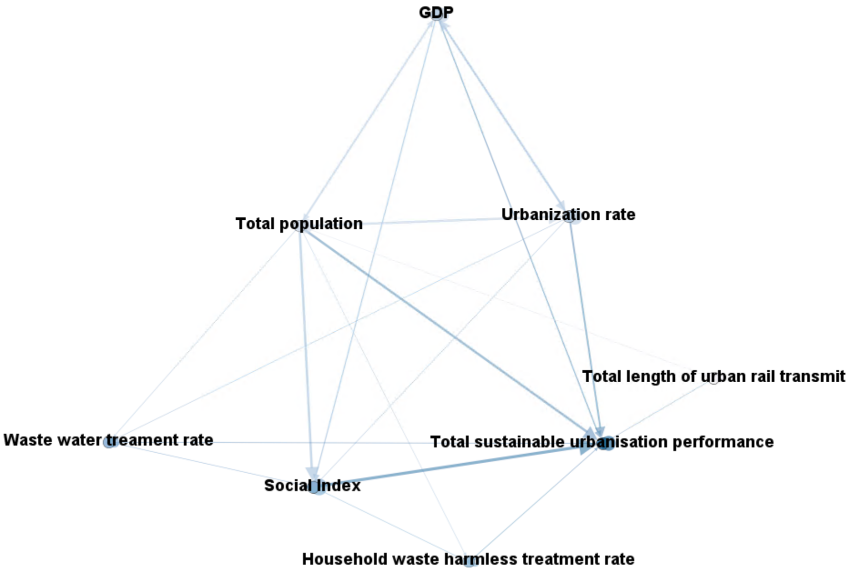


Figure 3.12. Cognitive map of the **CH** model

existence of the original assumption or search for contradictions. For example, in the triad '300' of the World3 model, 'Land Fertility' affects both 'Population 0 to 14' and 'Population 15 to 44', so they both affect 'Land Fertility' and each other. An example of '120C' triads is that both 'Persistent Pollution' and 'PP Technology' affect each other, while 'PP Technology' affects the 'Service Capital' which in turn eventually affects 'Persistent Pollution'.

How the modules of the full network show the topics of the model was presented earlier (Figure 3.11). However, after the generation of the SS model that only contains the state variables and also separates this network into modules, the main driver topics can be derived. In the case of the World3 model, these are the 'Land', 'Population' and 'Capital'.

Based on the resultant metrics, the models can be conveniently compared. The cluster map of the results based on the rank correlation analysis of the metrics is shown in Figure 3.13. The diagram groups together similar models and metrics, therefore, the columns and rows of the matrix, respectively. The rows show the related metric ranked from first to last denoted by the increasingly dark shades of blue. Therefore, the higher the rank of the related model according to the analysed metric, the darker the related bracket is.

By observing the groups of metrics, proofs of some straightforward network phenomena are determined. In the top part of the figure, modularity correlates with the triad types

’003’ and ’201’, therefore, the number of unclosed triangles in the network, as well as the number of circles, are low in these cases. In the bottom part of the figure, the complete triangles are in correlation with the edges, which is also reasonable as with more edges the probability of forming a circle is higher. As these metrics were placed in the top and bottom part of the rank correlation matrix, it can be seen that they are highly disconnected, which is once more straightforward: the measures at the top describe the disjointedness, while those at the bottom characterise the connections.

By analysing similarities between the models (the correlation between the columns of Figure 3.13), a similarity between **W3** and **UD** is apparent, which shows that their structural characters are very much alike. The most significants are the well-developed feedback loops with many states.

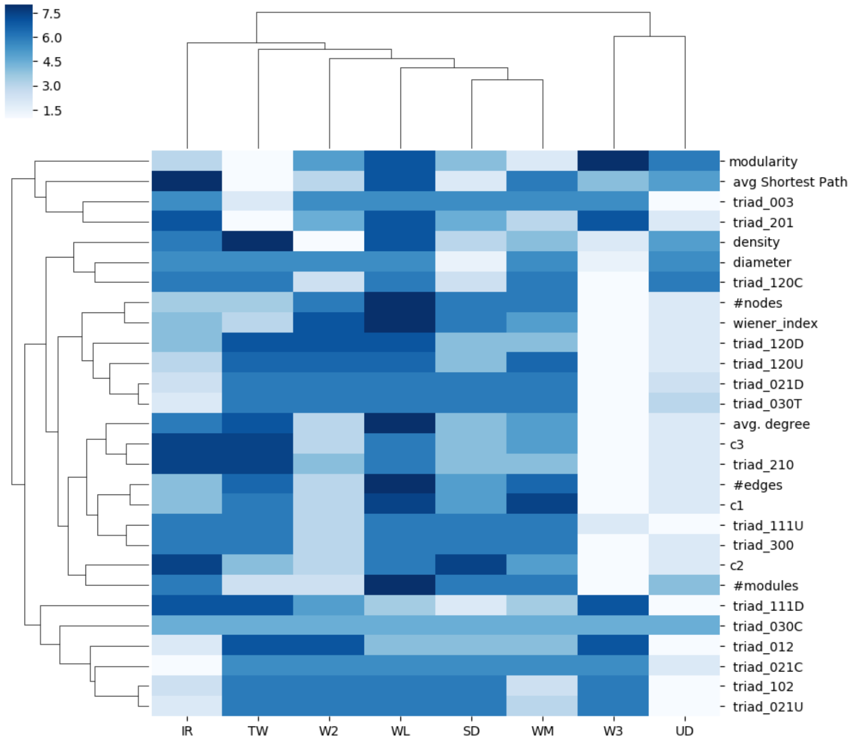


Figure 3.13. Structural comparison based on the rank correlation analysis of the metrics of the state-space representation.

The models **W3** and **W2** are very dense, which can only be achieved with numerous type ’300’ triads. This also highlights that the structure of these models is quite complex with many cross-effects and their effective analysis requires the simulation tools of system dynamics.

Providing some concluding remark to the Results and discussion section, the importance of the topic of sustainability-related model building and analysis is an intensively researched area, as it was highlighted in the Introduction section. The present chapter supports this model building-based research concept and the associated decision support process as discussed thoroughly in the work of[116]. However, the specific evaluation and interpretation of the results are always problem-dependent, the presented results provide a brief overview of the results obtainable by following the discussed methodology. In the future, the handling of the dynamical nature of models by weighted models is a challenging research aspect.

3.5 Conclusion

A methodology to validate and interpret SD models was developed. The method first transforms the model into a network. After this transformation, the variables and parameters of the model are evaluated according to their structural importance.

Regarding the results, 130 system dynamics models were overviewed from the past five years concerning the topic of sustainability. It was determined that system dynamics, by its very nature, could capture well the complex interconnections between parts of the triangle of sustainability, e.g. the environment, economy and society. The developed methodology, therefore, can be applied by policymakers and engineers to analyse SD models to achieve cleaner production and more sustainable processes.

To prove the applicability of the developed methodology and related tool, the World3 model was analysed. The model was automatically turned into a network containing 194 nodes and 325 edges, which served as the basis of further analysis. In the state-space representation 230 type '300' and 51 type '210' triads were detected, which proves the complexity of the structure and the dynamics. The high PageRank value of the "Arable Land" state variable reflects its central role in the model, while the main driver topics of the World3 model according to the SS model are the 'Land', 'Population' and 'Capital'. Furthermore, eight additional models with varying degrees of complexity were selected to prove the efficiency of the developed methodology as well as present the applicability of and relationship between different metrics.

The examples demonstrate how practising engineers and analysts can apply the method for the structural analysis and validation of complex sustainability models.

Regarding future work, the analysis of how the interconnections of the targets of sustainable development goals is a promising research objective. Researchers interested in the modeling of environmental and cleaner production systems could further develop the presented tool to support stability and resilience analysis.

Chapter 4

Linked Data

Linked data (LD) is an essential tool used to organize, store and share data with context. Knowledge extraction from a large amount of interconnected data requires effective tools and methods to address the complexity and mass of the underlying structure. The proposed method generates a compact and interpretable multi-layered network from a linked dataset or another multidimensional network, capturing the most informative context and making it analysable. Frequent itemset mining of the edge and node dimensions can extract the informative segments of a multidimensional network that are worth analysing as the layers of a multi-layer network.

4.1 Introduction

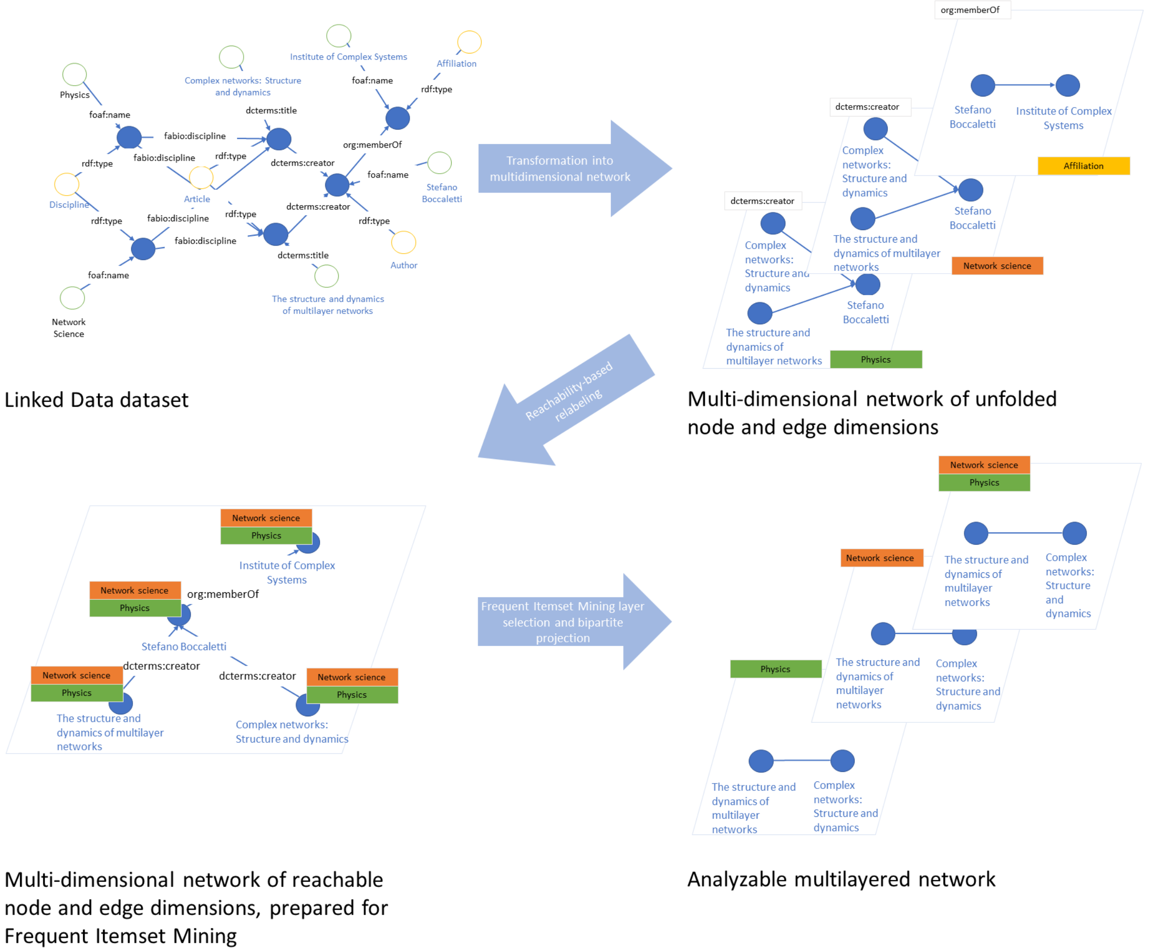
Linked data (LD) is an essential tool used to organize, store and share data with context[117]. Datasets that are published as LD are forming Semantic Web. The part of the Sematic Web which is freely accessible is called linked open data cloud (LODC). The driver of LD is the resource description framework (RDF) data model[118], which is standardized by the World Wide Web Consortium (W3C). Databases following the RDF standard are called triplestores or RDF stores, the naming is very intuitive, thus the atomic form of the RDF is an RDF triplet in the form "subject - predicate - object (s,p,o)", which states that "an object o has a relationship p with subject s"[119]. There is little work on formalization of the RDF besides the official documents of the W3C, particularly RDF Concepts and Abstract Syntax[120] and RDF Semantics[121], due its flexibility and extensibility[122]. There are formalizations towards special representations and formalization, like the bipartite graphs as intermediate model for RDF[123]. The main concepts of the RDF are self-descriptive data, data about data[124], machine readability[125] and extendibility[126]. LOD offers large quantities of freely available, interconnected, statistical (linked open statistical data - LOSD)[127], governmental[128], scientific[129, 130] and other annotated data[131]. The collection of such databases forms the Linked Open Data Cloud (LODC)[132], which consists of 2,973 datasets with 149.5 billion triplets.

In knowledge discovery and extraction, context is critical[133]. LD-based ontologies provide a facilitating toolset for knowledge sharing[134]. Our goal is to extract potentially useful knowledge, considering the flexibile nature and multi-aspect potential of LD, in an automated, easy-to-understand and validated way. We chose the toolset of network theory because it has a compelling perspective on these interconnected, information-dense, complex systems[135]. RDF supports the network-based perspective, as it can be interpreted as a directed labelled network[136]. When network analysis was first incorporated into the analysis of LD, classical measures such as degree distributions, small-world properties[137] and centrality-based rankings of entities[138] were measured. Topic-oriented analysis and LD-based social network analysis also arose[139]. With the introduction of special types of graphs on LD, such as labelled networks[118], hypergraphs[140] and tagged networks[141], deeper analysis was enabled. Most techniques focus on information discovery, such as tag-based clustering[142], hierarchical tag analysis[143], semantic distance measures[144], keyword clusters[145] and similarity-based rankings[146].

The main disadvantage of using multi-layered, multidimensional networks for knowledge extraction is that layer aggregation and cross-layer analysis are often difficult to keep track of when dealing with many layers because of the different overlaps[147, 39].

In the proposed methodology, aggregation and cross-layer analysis are performed with a logical description originating from an ontology, increasing the understandability of layer aggregation and analysis.

Figure 4.1 presents the most critical steps of network transformation in order. The first step is the discovery of the knowledge base and transformation of an LD dataset into a multidimensional network. In this step, the goal is to interpret the dataset by identifying entities and distinguishing between attributes and dimensions, which is important to keep the analysis transparent, without loosing information. The second step enumerates the reachable properties in the network. Algorithmic tools such as RDF chain search[148] and



Linked Data dataset

Multi-dimensional network of unfolded node and edge dimensions

Multi-dimensional network of reachable node and edge dimensions, prepared for Frequent Itemset Mining

Analyzable multilayered network

Figure 4.1. Workflow of the proposed network transformation steps towards an analysable multidimensional network from a linked data dataset.

querying property paths over distributed RDF datasets (QPPDs)[149] can be used in this step. In the methodology section, we describe a more efficient, network-focused method, developed specifically according to the nature of RDF datasets. Scanning and sampling an RDF dataset and performing analysis are often not easy tasks[150]. In addition to its large number of factors, such as different resources that may have different sets of properties, the properties themselves can be multi-valued (i.e., there can be triples in which the subject and predicate are the same but the objects are different); resources that may or may not have types[151] complicate the process even more, not to mention that the task is highly dependent on the platform, algorithm, dataset and underlying hardware. LD has its own toolsets for scanning and sampling tasks such as partitioning[152] and multi-indexing[153]. The more context-driven approaches to sampling and scanning are pattern recognition and statistics.

Frequent itemset mining (FIM) is also an option for statistical scanning, and it has been successfully carried out for synonymous property exploration [154], text extraction[148] and entity identification[155].

The frequent itemset mining can be performed based on the local database of the RDF triplets - which is the standard preprocessing and analysis procedure of LD [156], or the FIM can be performed in the cloud with the help of RDF Query Language (SPARQL)-based automatically generated queries [157].

Optional but important steps, layer selection and data enrichment, will be discussed in the next sections, as they are more subjective and situation dependent. They indicate the selection of layers for the multidimensional network based on sampling and the enrichment of the selected layers by other sources of information, respectively.

The final step represented in Figure 4.1—building the final, analysable, multidimensional network according to the previous sampling and decisions—is also carried out in the cloud by multiple systematic SPARQL queries. This step can be seen as a series of bipartite projections but is described in more detail in the next section.

The overall final step is ranking in the multi-layer network, as ranking can be considered a translation of highly complex phenomena into short, simple messages that can be easily digested[158]. Ranking, however, not only describes but also prescribes[159]; therefore, a very careful criteria selection method must be used. Ranking interconnections in the network has also been investigated for finding relevant relationships[160]. Network-based techniques are very understandable; according to a ranking[161] and with the inclusion of the statistically relevant layers, the relevant relationships are guaranteed. The aim of a complex knowledge exploration method in the LODC that takes into account the known hierarchies of the data (e.g., ontologies and taxonomies) as well as their interconnections is thereby achievable. Ultimately, the knowledge extraction performed in this way is a multi-criteria, multi-objective ranking system, in contrast to single aspect rankings and ranking only by analysing the structure.

To test and demonstrate the applicability of our methodology, we use the Microsoft Academic Knowledge Graph (MAKG)[162] to investigate the scientific realms of climate change and sustainability. The discovery process also includes a ranking of authors and institutes. The multi-aspect ranking also includes the layer similarities, determining the similarities among research fields and their combinations, which act as the dimensions of the network. The MAKG describes research fields hierarchically. The specialization of a layer can be determined by incrementing the number of elements in the itemset, interconnecting more disciplines or stepping downwards in the hierarchy tree. The incrementation of the specification yields a lower entity count and increased density and modularity. We inspect both layers, both types of community similarity, to reveal and explore overlaps and gain insight into the specifics of climate change and sustainability.

According to the main contributions, this thesis chapter is organized as follows:

- The RDF databases are represented as multidimensional networks as presented in Section 4.2.
- We propose a frequent itemset mining-based method to extract information from multidimensional network in Section 4.3.
- The resulted of the frequent itemsets of multidimensional networks can be represented as a multi-layer network that can be analysed by metrics presented in Section 4.4.
- We present the methodology through an example in which we uncover the scientific realms of climate change and sustainability, including an alternative co-author, co-organizational network ranking used to measure the impact of authors in multiple disciplines in Section 4.5.

4.2 Multidimensional network-based representation of RDF databases

Linked Data can be seen as multiple interconnected datasets in RDF format. The atomic form of an LD dataset or an RDF dataset is the RDF triplet; "an object o has a relationship p with subject s", can be seen as a single edge in a network that connects entities, nodes s and o, with a labelled attribute p. A good example is that Isaac Asimov (s) wrote (p) The Foundation (o).

The classic multidimensional networks are edge-labelled multi-graphs, which are described as $G = (V, E, D)$, where V represents the set of nodes, E the set of edges, and D the set of dimensions. The set of edges can be described as connections between nodes (u and v) along a dimension (d). The set can be written as $E = \{(u, v, d), u, v \in V, d \in D\}$.

The nodes in LD are often enriched by properties and descriptions. In the example, Isaac Asimov is both a person and a writer, and "The Foundation" is a fiction novel. These properties, such as "The Foundation" is "fiction", are also described by triplets. These triplets can be merged into a simple node or skipped if they contain irrelevant pieces of information. These ontological properties often act as dimensions. Therefore, to simplify the ideas, the notation and ultimately the analysis, we extend the description of a dimension with two sets, the dimension of the nodes (D_V) and the dimension of the edges (D_E). The union of the sets results in the dimension set ($D_V \cup D_U = D$). Then, the notation of the edges is described as $E = \{(u, d_u, v, d_v, d_e); \quad d_e \in D_E, d_u, d_v \in D_V\}$, where u and v are the nodes as before and d_u and d_v are their dimensions, respectively; d_e represents the dimension of the edge.

A multidimensional edge is represented as $E = \{(u, v, D_\alpha); \quad D_\alpha \subseteq D\}$, where D_α refers to the simultaneously matching dimensions of both the node and edge dimensions $D_\alpha = D_{\alpha,V} \cup D_{\alpha,E}$. This selection is a direct reference to a layer in a multiplex network $G = \{G_\alpha; \alpha \in \{1, \ldots, M\}\}$ where $G_\alpha = (V_\alpha, E_\alpha, D_\alpha)$. The network G_α is a network with α dimension selection, where the nodes V_α and edges E_α take the dimension nodes and edges of the selected dimensions D_α, respectively. M corresponds to the number of created layers.

A multiplex network is a particular multidimensional network in which every layer contains every node and the cross-layer edges are identifier edges, which refer to the same cross-layer node. We use this notation, with the addition that not every layer will include every node; therefore, an activity check in a multiplex network—checking whether a node is connected or disconnected in a layer—will effectively be an existence check. Extending the multiplex notation with simultaneous dimension selection, we build the edges as $E_\alpha = \{(u,v) \in V \times V; (u,v,D_\alpha) \in E \quad \text{and} \quad D_\alpha \quad | \quad d_u \in D_{\alpha,u}, d_v \in D_{\alpha,v}, d_e \in D_{\alpha,e}\}$, where $D_{\alpha,u}$, $D_{\alpha,v}$, and $D_{\alpha,e}$ refer to simultaneously matching node and edge dimensions, respectively. Returning to the example of Asimov, dimension selection would work for the network of books with the simultaneous matching attributes "fiction" and "robots". The expected result would be a network of books containing every book from the Elijah series, "The Caves of Steel" and "Robots and Empire" with the levels and dimensions of the important layer and non-layer constructing properties, such as the author "Isaac Asimov" and the main protagonist "R. Daneel Olivaw". This means that the created network is explainable by the writer or the protagonist and of course the layer constructing properties "fiction" and "robots".

The number of layers of a fully defined multi-layer network is large if we consider $n_e = |D_E|$ as the number of all edge dimensions and $n_v = |D_V|$ as the number of node dimensions. Then, the number of possible layers is $\sum_{k=1}^{n_k} \binom{n_e}{k} \sum_{j=1}^{n_v} \binom{n_v}{j}$. However, the selection of significant layers is the key to reducing the space of the analysis. The previously introduced variable M as the number of layers refers to the number of selected layers.

The transactions can be extended with the environment and the reachable labels and properties in the RDF. The right side of Figure 4.2, $G_\alpha^{(2)}$, takes the reachable tags into account.

The reachability states also enumerate all node attributes, and they take the neighbour attributes into account. Beyond the simple mapping of the dataset, they can also be used to extract information. In the example below, this means that Stefano inherits all attributes of the paper and his institute.

To examine the significant layer selection, we have to understand the reachability concept in the dataset.

To visualize reachability, Figure 4.2 represents the core idea. In figure 4.2, G_α represents an RDF dataset. It can be translated as stating that "The structure and dynamics of multilayer networks" (v_1), which is a review article (d_{v1}) in network science (d_{v2}), is written by (d_{e1}) Stefano Boccaletti (v_2), who is a physicist (d_{v3}). Stefano is affiliated with (d_{e2}) the Institute for Complex Systems in Florence (v_3), which is a research institution (d_{v5}).

The presented procedure creates multi-links as well as networks for a given set of attrib-

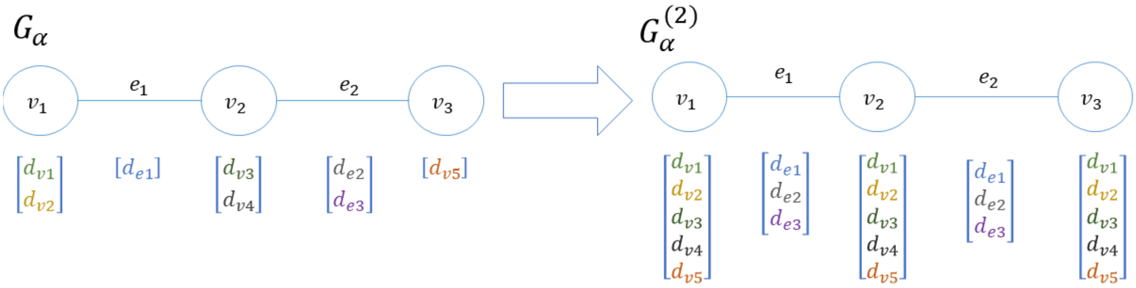


Figure 4.2. Example of frequent slicing and an application of reachability. G_α is the starting network in a non-directed format. $G_\alpha^{(2)}$ is the set of attributes reachable from G_α.

utes. The procedure is similar to multidimensional network-based analysis methods, were RDF databases were analysed in thematic dimensions[163]. The bipartite network-based analysis of RDF datasets has also been proven useful[123]. Bipartite networks are excellent to study the connections of two sets of objects. However, for multi-objective analysis, a more complex model representation is needed, which motivates the development of our method that forms sets of layers of networks where the layers represent significant subsets of the dimensions of the RDF model. According to this, the next step of the proposed method is selecting these significant sets of dimensions, which will be presented in the following section.

4.3 Frequent itemset mining in multidimensional networks

Frequent itemset mining (FIM) is a mining technique used to uncover frequent correlations in transactional datasets[164]. We can consider the itemset $I = \{I_1, \dots, I_{n_e+n_v}\}$ in the case of the FIM representing the products; in the LD, the itemset represents the dimensions $D = \{d_1, d_2, \dots, d_{n_e+n_v}\}$, or more specifically, the labels of the RDF. A transaction is defined as $\tau = (tid, X)$, where tid is the transaction identifier and X is a set of items over I $(X \subseteq I)$. The database is the set of all transactions $P = \{\tau_1, \tau_2, \dots, \tau_n\}$. The support of an itemset is equal to the count of the constellation of dimensions.

As stated before, we are interested in significant layers. We measure significance with the support of the itemsets. The set $F = [\alpha, \dots, M]$ holds the frequent itemsets. The support of an itemset is $supp(F_\alpha) = |\{\tau_i \mid F_\alpha \subseteq \tau_i, \tau_i \in P\}|$; the support of a layer (G_α) is equal to its edge count $supp(G_\alpha) = |E_\alpha|$. G_α is frequent if $supp(G_\alpha) \geq minsupp$, where

Table 4.1. Summary of the frequent itemset mining (FIM) technique notation and its multidimensional counterpart

	Frequent itemset mining	multidimensional network
Items	$I = \{I_1, \dots, I_{n_e+n_v}\}$ The products, in traditional FIM	$D = \{d_1, d_2, \dots, d_{n_e+n_v}\}$ The labels of the RDF
Transactions	$\tau_i = \{tid, X\}$ Set of items	A node with extended reachable tags
Database	$P = \{\tau_1, \tau_2, \dots, \tau_m\}$ all transactions	The enriched dataset
Support	Number of itemset occurrences	$supp(F_\alpha) = \|\{\tau_i \| \; F_\alpha \subseteq \tau_i, \tau_i \in P\}\|$
Frequent itemset	$F_\alpha \subseteq \tau_i$ if $supp(F_\alpha) \geq minsupp$	$supp(F_\alpha) \geq minsupp$

minsupp is a chosen threshold. Summarizing G_α is frequent if $supp(G_\alpha) \geq minsupp$. F_α is called a closed or frequent closed itemset if there exists no proper superset of it that cannot be extended by any dimensional data without losing support. Table 4.1 shows the technique and its multidimensional counterpart.

An effective representation of the layer selection in a multidimensional network is the multi-link. Multi-link $\overrightarrow{m}$ is an enumeration of the selected layers: $\overrightarrow{m} = [m^\alpha, m^\beta, \dots, m^M]^T$. We can now introduce the multi-adjacency matrix ($A^{\overrightarrow{m}}$), with elements $a_{ij}^{\overrightarrow{m}}$ that are equal to 1 if there is a link between node i and node j and zero otherwise[39].

$$a_{ij}^{\overrightarrow{m}} = \prod_{\alpha=1}^{M} [a_{ij}^{\alpha} m^{\alpha} + (1 - a_{ij}^{\alpha})(1 - m^{\alpha})] \tag{4.1}$$

Thus, multi-adjacency matrices satisfy the condition $\sum_{\overrightarrow{m}} a_{ij}^{\overrightarrow{m}} = 1$. The enumeration of the layers where nodes are active can serve as the input to most of the frequent itemset algorithms, as they effectively represent the itemsets. The methodology works with any FIM algorithm, including CHARM[165], FPclose[166] and FP-Growth[167]; for an exhaustive list, see the work of Chee, which also studies the scalability of FIM algorithms[168].

4.4 Analysis of the resulted multilayer network

The union, the logical aggregation of layers, can be best expressed by the overlapping edges[39]($O^{\alpha,\beta}$).

$$\begin{aligned} &G_\gamma = G_\alpha \cup G_\beta, \quad O^{\alpha,\beta} = supp(G_\gamma) \\ &O^{\alpha\beta} = \sum_{i<j} a_{ij}^{\alpha} a_{ij}^{\beta}\,, \end{aligned} \tag{4.2}$$

where G_γ is the layer formed by combining G_α and G_β and where a_{ij}^{α} expresses a simple edge in layer α connecting the nodes i and j. The count of the overlapping edges corresponds to the support of the combined layer. The aggregated layers are also frequent, as every subset of a frequent itemset is frequent[164]. Logically aggregating the layers is also an efficient technique in data discovery.

In the upcoming example of author networks, authors interact on the layer of climatology, on the layers of climatology and meteorology and in every other layer. In this case, it is difficult to keep track of all the different types of multi-links. Therefore, we can calculate the multiplicity of the overlap v_{ij} between nodes i and j, which indicates the total number of layers in which the two nodes are connected.

$$v_{ij} = \sum_{\alpha=1}^{M} a_{ij}^{\alpha} = \sum_{\alpha=1}^{M} m_{ij}^{\alpha}\,, \tag{4.3}$$

where the nodes i and j are linked by the multi-link $\vec{m} = \vec{m_{ij}}$. In weighted multidimensional networks, the weights might be correlated with the structure in a nontrivial way. To study the weights, there are two new measures, the multi-strength ($s_{i,\alpha}^{\vec{m}}$) and the inverse multi-participation ratio ($Y_{i,\alpha}^{\vec{m}}$)[169]

$$s_{i,\alpha}^{\vec{m}} = \sum_{j=1}^{N} a_{ij}^{\alpha} a_{ij}^{\vec{m}}\,, \tag{4.4}$$

$$Y_{i,\alpha}^{\vec{m}} = \sum_{j=1}^{N} \left(\frac{a_{ij}^{\alpha} a_{ij}^{\vec{m}}}{\sum_r a_{ir}^{\alpha} a_{ir}^{\vec{m}}} \right)^2 . \tag{4.5}$$

The multi-strength ($s_{i,\alpha}^{\vec{m}}$) measures the total weight of the links incident to node i in layer α that form a multi-link. The inverse multi-participation ratio ($Y_{i,\alpha}^{\vec{m}}$) is a measure of the inhomogeneity of the weights of the nodes that are incident to node i in layer α and are also part of the corresponding multi-link. Thus far, we have covered some indicators for multidimensional activities, which are very useful for dealing with many layers. The final step of knowledge extraction is ranking. Before turning to the ranking, we recall that the density, modularity and other structural measures are very different from layer to layer.

Therefore, for each node, we can write an NxM activity matrix ($\mathbf{B}$) of elements $b_{i,\alpha}$, indicating whether node n_i is present in layer α:

$$b_{i,\alpha} = n_i \in G_\alpha \,. \tag{4.6}$$

In this way, we can measure the number of layers where i is present and active[147] as

$$b_i = \sum_{\alpha=1}^{M} b_{i,\alpha} \,. \tag{4.7}$$

Additionally, the number of nodes present and active in a layer can be given by N_α:

$$N_\alpha = \sum_{i=1}^{N} b_{i,\alpha} \,. \tag{4.8}$$

The correlation between layers can be given by $Q_{\alpha,\beta}$, quantifying the fraction of nodes that are present in layer α as well as in layer β.

$$Q_{\alpha,\beta} = \frac{1}{N} \sum_{i=1}^{N} b_{i,\alpha} b_{i,\beta} \,. \tag{4.9}$$

A straightforward ranking in a network is obtained by calculating the centralities of the nodes, reflecting their importance from different viewpoints. In multidimensional networks, the most common centrality measure is to calculate the centralities of each layer and finally aggregate them according to certain weights[39]. Both the aggregation (maximum selection, minimum selection, summation, etc.) and the centrality measure used depend on the interpretation

$$\theta_i = \sum_{\alpha}^{M} w^\alpha \theta_i^\alpha \,, \tag{4.10}$$

where θ_i^α is the calculated centrality measure of node i in layer α and w^α indicates the importance of layer α. Now that the methodology has been described, the next section demonstrates the applicability of the methodology.

4.5 Results

The programs of the following case study are available at the github[1] and the raw dataset is available on the Microsoft Academic Knowledge Graph homepage[2] as well as the SPARQL endpoint[3]. The goal of this demonstration is to showcase knowledge extraction from vast linked data. Therefore, we selected the LOD catalogue for scientific publications from Microsoft, the MAKG[162]. The MAKG itself contains definitions for 209,792,741 papers and 253,641,783 authors, in RDF terms, more than eight billion triplets. The papers are categorized into 229,716 fields of studies. For the relevant results, we selected the date range 2010 - 2017; the catalogue was last updated in late 2018[162]. Our aim was to study the realms of sustainability and climate change based on the MAKG dataset on the other hand to showcase the importance of the proper focus to not get lost at scale, the applied frequent itemset mining pinpoints and keeps understandable the important areas of the data.

The first test on the dataset is reachability, to discover how to treat the dataset, which is better formulated as, what are the dimensions that we can analyse? For example, in the catalogue, the authors can be connected to universities, research organizations, and industrial laboratories. Therefore, the dataset describes the connections from *rdf:type Article* to *rdf:type FieldOfStudy* through the connection of *fabio:hasDiscipline*. The previously mentioned article connects to an *rdf:type Author* through the connection of *dcterms:creator*. Reaching the *rdf:type Affiliation*, an *org:memberOf* connection is needed.

An *rdf:type Affiliation* can be connected to an external data source, the Global Research Identifier Database (GRID), to extend the affiliations with the geo-coordinates, regions, establishment dates, etc. Therefore, regional and institutional categorization could be one aspect of the data. Another, more straightforward analysis is the analysis of the author network. The articles are sorted into multiple categories (*rdf:type FieldOfStudy*) according to the hierarchical ontology created by the MAKG. The ontology contains five levels of depth. The top level, level zero, is the major category (e.g., mathematics, medicine, engineering,

---

[1] https://github.com/abonyilab/aprioriSPARQL

[2] http://ma-graph.org/rdf-dumps/

[3] http://ma-graph.org/sparql

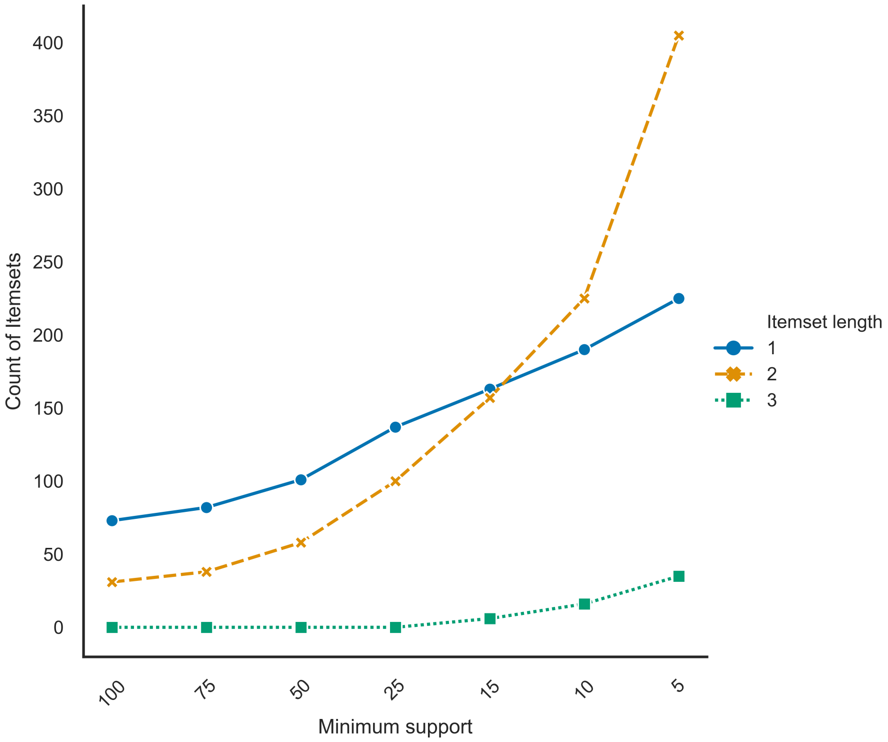


Figure 4.3. Counts of frequent itemsets by length and minimum support.

chemistry, etc.), and the next levels are their descendants, the more specialized categories (e.g., nuclear medicine, applied mathematics, etc.). We take the ontology elements and the constellation of the ontology elements as the layers or dimensions of the network. Going downwards in the ontology, increasing the specification of a layer also increases the density of the layer. Not every paper is categorized into as many matching ontology elements as possible. Therefore, the lower levels, three, four and five, are ignored, and the density does not increase. However, it is true that the more specialized a layer is, the denser it becomes, even for horizontal extensions of layers, meaning the extension of an element to another element that is on the same ontological level.

In this chapter, we focus on sustainability science and climate change. Therefore, we choose the ontological element “Climatology” as the starting point for our analysis to observe the advancements and analyse the social background of humanity’s major problem, climate change. We also restrict ourselves to analysing only the author and organization networks within the second ontological level, which is easy to understand and sophisticated enough to investigate. Now that we have a rough idea about what we want to do, we execute FIM on the dataset to sample it from multiple angles. With this technique, we want to uncover significant dimensions for the analysis and common constellations of disciplines that go hand-in-hand with the previously selected ontology element, “Climatology”. For discovery, we propose to load and execute FIM on the whole data space, as the linked data are large on average.

In this study the apriori FIM algorithm was used on the offline dump of the RDF database and SPARQL-based queries were utilised for the validation of the results.

FIM was executed on a low setting, with a minimum support of five, to probe the dataset, which means the selected timescale (2010-2018), in order to have slightly less than one article per year in the given frequent constellation of the field of studies. Figure 4.3 shows the FIM results. The results also show the optimal minimum support of 10, which is also selected for the next steps, where is a significant drop in small itemsets, but not as significant as in the longer itemsets. The lack of longer itemsets is due to the categorization of the field of study in the dataset, as an article is categorized into 1.52 fields of study, on average, in the second ontological level. The other ontological levels have much the same statistics: 1.01 on the first, top, level, 3.18 on the third, 1.75 on the fourth, 1.32 on the fifth, and 1.31 on the sixth.

A length of one for the itemset indicates that climatology can be connected with another ontology element; two indicates that it can be connected with two other elements, and three with three, while still reaching the minimum support. No more extensions than three reach the minimum support. The choice could be made here to set the minimum support to a lower position, less than five, or we could be satisfied with the choice and the count of networks, in this case, 665. This is quite a manageable size of networks, and it is also worth mentioning that the edge count of the networks is approximately 5000.

The next step is the network creation of authors and organizations on the layers. If the layers do not have enough nodes, then they significantly influence the ranking; therefore, our selection requirement for a layer is at least 40 different contributors, and that for edge formation is at least two contributions between organizations and one for authors where prescribed, based on the correlation measures between the layers (Equation 4.9). Figure 4.4 shows the similarities between the layers using the edge overlap metric.

In Figure 4.4, the darker a region is, the more similar the layers are. The dendrograms on the edge of the figure show the distances between the networks. We see that the extensions of the layers are clustered together as well as similar studies. The top left segment of the figure, including meteorology, atmospheric sciences, hydrology, oceanography, ecology and geomorphology, shows the starting points for the extensions. Those are the closest fields to climatology, which also have the most substantial support from FIM. We can also observe different views; for example, the cluster in the middle, containing agroforestry, environmental planning, economics, soil science and environmental engineering, is formed around the economic side of climatology. The cluster in the bottom right, including remote sensing combined with meteorology, artificial intelligence, and mathematical optimization, is formed around computer-based observations and modelling. A natural question about these clusters is, why are there not more extensions? Remote sensing and artificial intelligence would be a perfect match. There are such extensions, but their support is below the minimum support. The support of artificial intelligence itself is small, 482. Artificial intelligence and remote sensing together have the support of 44 papers from 2010 to 2018; however, their node count is below the selected minimum node count (10 for country networks, 40 for organization and author networks). The other combinations show the same phenomena.

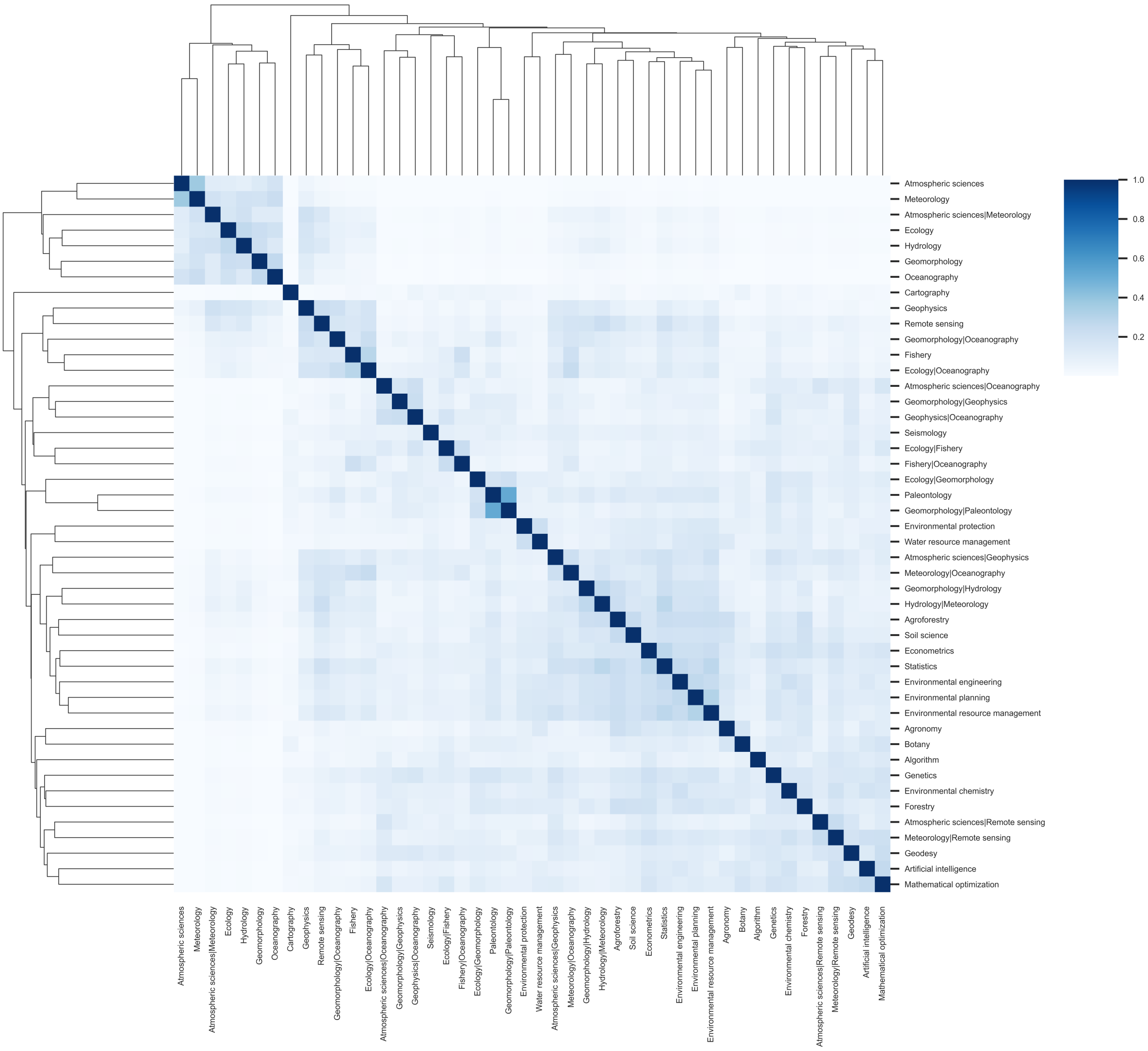


Figure 4.4. Organizational cluster map in the realm of climatology, showing how similar the disciplines are according to their networks of organizations.

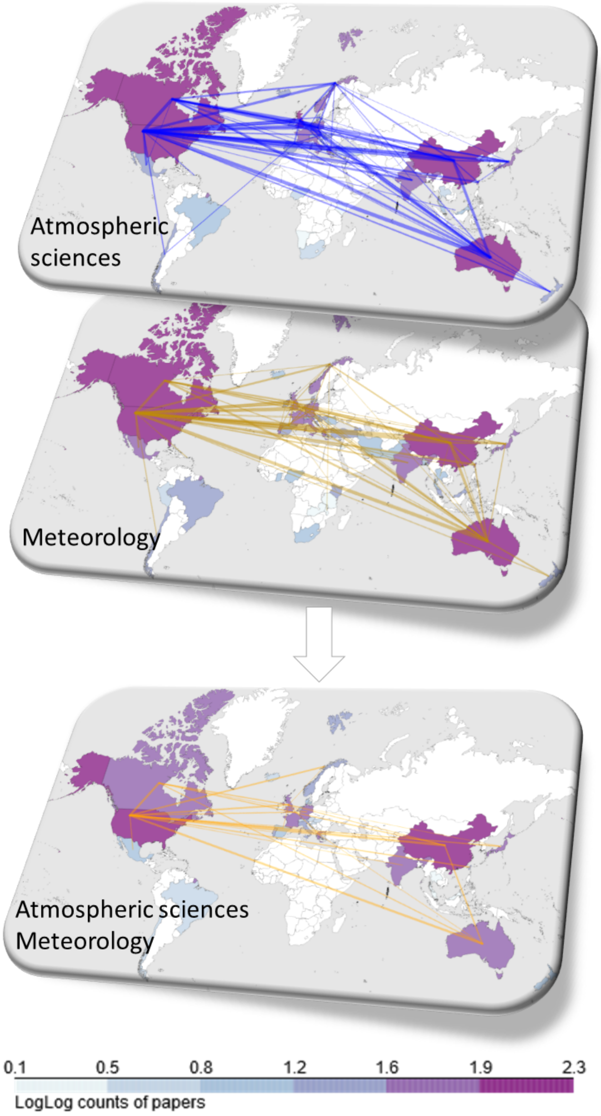


Figure 4.5. Multi-layer institutional network aggregated at the country level from the atmospheric sciences and meteorology layers and the extension of them, atmospheric sciences - meteorology

Table 4.2. Layer metrics of the institutional network

network_layer_resolver	no.nodes	no.edges	density	no.clusters	modularity	avg.clustering_coefficient
Agroforestry	208	251	0.0117	174	0.9592	0.0011
Agronomy	89	120	0.0306	70	0.9278	0.0029
Algorithm	50	118	0.0963	33	0.7231	0.0582
Artificial intelligence	55	58	0.0391	53	0.9738	0.0063
Atmospheric sciences	1285	14217	0.0172	360	0.2958	0.0011
Atmospheric sciences—Geophysics	168	213	0.0152	130	0.9431	0.0016
Atmospheric sciences—Meteorology	638	2332	0.0115	248	0.6023	0.0034
Atmospheric sciences—Oceanography	45	48	0.0485	42	0.9679	0.0000
Atmospheric sciences—Remote sensing	44	56	0.0592	35	0.8616	0.0056
Botany	67	72	0.0326	63	0.9780	0.0029
Cartography	47	52	0.0481	42	0.9608	0.0000
Ecology	999	4370	0.0088	436	0.5001	0.0020
Ecology—Fishery	58	61	0.0369	55	0.9766	0.0000
Ecology—Geomorphology	66	72	0.0336	60	0.9776	0.0000
Ecology—Oceanography	239	401	0.0141	140	0.8644	0.0064
Econometrics	126	141	0.0179	112	0.9850	0.0032
Environmental chemistry	78	92	0.0306	66	0.9698	0.0077
Environmental engineering	208	231	0.0107	186	0.9884	0.0005
Environmental planning	184	200	0.0119	169	0.9893	0.0011
Environmental protection	86	123	0.0337	77	0.8606	0.0308
Environmental resource management	201	260	0.0129	167	0.9291	0.0047
Fishery	248	385	0.0126	171	0.9085	0.0061
Fishery—Oceanography	71	96	0.0386	56	0.9015	0.0132
Forestry	99	107	0.0221	91	0.9823	0.0000
Genetics	97	117	0.0251	82	0.9654	0.0110
Geodesy	71	80	0.0322	62	0.9753	0.0000
Geomorphology	1148	5850	0.0089	368	0.4877	0.0010
Geomorphology—Geophysics	55	63	0.0424	47	0.9499	0.0000
Geomorphology—Hydrology	254	363	0.0113	186	0.9375	0.0020
Geomorphology—Oceanography	291	511	0.0121	160	0.8614	0.0080
Geomorphology—Paleontology	78	101	0.0336	60	0.9556	0.0256
Geophysics	503	1013	0.0080	286	0.7733	0.0030
Geophysics—Oceanography	42	46	0.0534	38	0.9631	0.0000
Hydrology	1128	2893	0.0046	537	0.7027	0.0008
Hydrology—Meteorology	309	411	0.0086	232	0.9548	0.0023
Mathematical optimization	60	62	0.0350	58	0.9807	0.0000
Meteorology	1904	9299	0.0051	788	0.5096	0.0005
Meteorology—Oceanography	221	320	0.0132	168	0.8769	0.0017
Meteorology—Remote sensing	75	84	0.0303	66	0.9691	0.0000
Oceanography	1164	7617	0.0113	370	0.3972	0.0022
Paleontology	145	179	0.0171	119	0.9627	0.0075
Remote sensing	448	739	0.0074	304	0.8733	0.0015
Seismology	89	101	0.0258	78	0.9756	0.0023
Soil science	128	217	0.0267	100	0.8141	0.0074
Statistics	275	314	0.0083	246	0.9813	0.0013
Water resource management	66	99	0.0462	56	0.8316	0.0210

Table 4.2 shows the important metrics of the significant institutional layers: the number of nodes, number of edges, density, modularity and average clustering coefficient. The average clustering coefficient represents the likelihood that two neighbours of a node are connected, while modularity informs us about the community structure of the network. The higher the modularity is, the more community-centred the graph. We see here that the more specific a layer is, the more community-centred, and the higher its modularity. This can be seen in atmospheric sciences by extending it with geophysics. The modularity of atmospheric sciences is 0.2989, while the extended layer modularity is 0.9436. The same phenomena can be found in all other layers and their extensions, which means that more specific layers and disciplines are owned by more interconnected communities.

Figure 4.5 shows the multi-layer visualization of (1) atmospheric sciences, (2) meteorology and (3) their interconnection, giving insight into the data. For this visualization, enrichment of the data was needed to locate the research institutions on the map. The enrichment was performed with the GRID. Not every research institution could be mapped into the GRID, and therefore, the unmapped research institutions are not counted in the country-level aggregation; however, this does not influence the overall ranking. With enrichment, we can easily observe the clusters both inside a country and across countries. For example, the USA, Canada, China and Austria are strong clusters. In layer (1), atmospheric sciences, the USA contributes 84,201 papers, with 3,400 co-contributions with Canada, 2,444 with Australia and 1,952 with China. Layer (2), meteorology, shows the same trend of the USA dominating the discipline with 43,981 papers; however, for the co-contributions, Canada is third with 582 papers, there are 688 China-USA contributions and in first place, there are 1,086 Australian-USA contributions. The greatest contribution to the interconnected layer is (3) atmospheric sciences and meteorology; the USA has 10,183 contributions, and China is second with a total of 809 contributions. The USA mostly contributes to the discipline with Canada, with 200 individual papers. The rest of the contributions are with Australia, 118, and China, 86, and there are very small amounts with Chile, Japan, Taiwan and Norway.

The following rankings are based on these insights into the data. The multi-layer representation clearly shows that the more specialized a topic is, the fewer contributors there are, but the more connected they are. The aggregated networks are denser with a higher modularity, as observed previously. The next artefact of knowledge extraction is the ranking. For the ranking, we calculate the importance of a country, institution and author with the multi-layer eigenvector centrality (Equation 4.10).

The ranking mostly depends on the layers in which an entity (country, organization or author) takes part. This is why a strong minimum support and minimum node count are needed for the analysis; otherwise, the very specialized layers will dominate, with very few nodes, which have a very high rank. Therefore, we can use weightings according to the correlations of the layers and the nodes, as described in the methodology section, or other subjective criteria to balance the sparseness of the very specific layers. The top list is represented in Table 4.3. Next to the ordering in the table, the most important layer column shows where the organization or author obtained the highest rank, and the "Agg. eigen. centr." column shows the aggregated eigenvalue centrality of the entity. With the aid of this toolset, we can observe specific connections between research areas and pinpoint

Table 4.3. Leaderboards of the top 5 institutes and authors contributing to climatology.

Organization leaderboard			
Rank	**Name**	**Most important in layer**	**Agg. eigen. centr.**
1	The United States of America (USA)	Artificial intelligence—Pattern recognition	56.0279
2	China (CHN)	Agroforestry—Hydrology	35.6429
3	Australia (AUS)	Economy	33.0746
4	Canada (CAN)	Thermodynamics	24.4631
5	United Kingdom of Great Britain (GBR)	Environmental protection	23.2615

Organization leaderboard			
Rank	**Name**	**Most important in layer**	**Agg. eigen. centr.**
1	Chinese Academy of Sciences	Geodesy	24.1449
2	National Oceanic and Atmospheric Administration	Meteorology—Remote sensing	5.0147
3	National Center for Atmospheric Research	Econometrics	3.3779
4	French National Centre for Scientific Research	Ecology—Oceanography	2.9609
5	Alfred Wegener Institute for Polar and Marine Research	Geomorphology—Oceanography	2.0770

Individual leaderboard			
Rank	**Name**	**Most important in layer**	**Agg. eigen. centr.**
1	Vijay P. Singh	Hydrology	1.4517
2	Hai Cheng	Geomorphology	1.2189
3	R. Lawrence Edwards	Geomorphology	1.0049
4	Colin Schultz	Meteorology—Oceanography	1.0005
5	Qiang Zhang	Geomorphology—Hydrology	0.8854

research constellations describing sectors. The multi-aspect ranking provides the flexibility to take significant topics into account and refine the ranking. The different metrics are the searchlights of importance and focus.

Table 4.4. Comparison between the ranks based on the publication count in sustainability science and climate change and the multi objective rank created by the multidimensional network.

		Publication count based ranks				
Organization	Multi-objective rank	Global rank	Hydrology	Ecology	Paleontology	Geophysics
Chinese Academy of Sciences	1	1	1	1	3	10
National Oceanic and Atmospheric Administration	2	76	75	51	573	39
National Center for Atmospheric Research	3	177	133	840	2558	28
French National Centre for Scientific Research	4	2	4	3	2	2
Alfred Wegener Institute for Polar and Marine Research	5	197	291	87	166	135
Russian Academy of Sciences	6	9	42	6	4	4
Potsdam Institute for Climate Impact Research	7	1323	407	1061	3521	921
California Institute of Technology	8	37	71	571	307	3
Goddard Space Flight Center	9	71	1478	2765	2637	551
Wageningen University and Research Centre	10	44	16	24	560	835
Beijing Normal University	11	227	18	241	1083	684
Lamont-Doherty Earth Observatory	12	409	396	667	310	35
Ocean University of China	13	380	345	399	617	326
United States Forest Service	14	111	33	10	718	1084
International Institute for Applied Systems Analysis	15	939	537	847	3240	3085

Table 4.4 compares the publication count-based ranks in sustainability science and climate change with the multidimensional network-based ranking. The selected topics are the

subset of the FIM-selected topics presented in Table 4.2. The Chinese Academy of Sciences has the most publications in sustainability science and climate change, also has the most publications in most of the layers, and it is also very cooperating therefore the first place for the Academy. In the comparison, we see that interestingly the National Oceanic and Atmospheric Administration is very highly ranked; however, the publication count in the shown layers predicts it otherwise. The Administration is highly embedded into the any Oceanic (e.g. Fishery) and Atmospheric sciences (e.g. Remote Sensing, Geophysics), as the name would predict. Thanks to the substantial co-operations of the institute, this organisation plays a central role in sustainability science and climate change, which would not be highlighted in classical analysis techniques.

4.6 Discussion and conclusions

Our work contributes to the knowledge extraction of linked data. It also contributes to the notation of multidimensional networks by extending the nodes with dimensions, in contrast to the formal labelled network notation. This extension is useful in high-dimensional data analysis, such as for linked open data, as the nodes are often extended with hierarchical properties and ontologies. The extraction of useful data is validated with on-demand, online, iterative SPARQL-based sampling of the dataset with frequent itemset mining.

We demonstrated the applicability of the methodology through an interesting scientometric example, co-author and co-organization rankings in sustainability and climate change. The source of the analysis was the linked open database of the Microsoft Academic Knowledge Graph. We discovered multi-disciplinary science boards using the proposed multidimensional network-based approach. We showed similarities between disciplines and the layers of the network. We also discovered that the aggregation of the layers in a multidimensional network does not always result in the loss of information, and in contrast, the aggregation of the layers results in denser, more modular information. Finally, we ranked authors and organizations with multidimensional centrality rankings and showed where sustainability and climate change are major research topics and who and which organizations are the main contributors.

The proposed methodology generates a compact and interpretable multi-layered network from a linked dataset or another multidimensional network. The methodology is applicable when there are a large number of edge and node labels, with the current reference to eight billion triplets, the dataset of the Microsoft Academic Knowledge Graph. The scalability of the methodology is not limited, however, it is more an engineering challenge, than a research objective. The most time and memory consuming operation is the Frequent Itemset Mining, where serious advancement were already made by GPU acceleration[170], Hadoop based partitioning[171], and Spark-based parallelism[172]. The endpoint capabilities limit the scalability of the FIM against the SPARQL endpoint; as it can be seen as an Application Programming Interface communication its parallelism and effective scalability have already been proven by all modern web browsers.

With the aid of the proposed methodology and toolset, we can observe, select and analyse particular connections between entities in linked data, taking ontological dimensions and specific properties into account. The multi-aspect ranking provides the flexibility to refine the ranking, while the other proposed tools act as searchlights of focus to interpret a whole set of linked data, with all its extensions and possible enrichments.

Chapter 5

Conclusion and outlook

This PhD dissertation gave an insight into multidimensional network-based knowledge extraction from different challenging information sources, summarized in Figure 5.1. The trend of data architecture and data engineering is system-independent and contextualized information, enabling abstract and more generalized applications. However, knowledge extraction from such systems is not intuitive nor trivial. The work showed a way to analyze systems through data to unfold the structure, identify key elements and components, discover component similarities. Depending on the data set used, more especially the domain of the data, gives different insights into the analyzed systems.

The rise of interconnected production, monitoring systems, as well as the rise of online publications, reviews, and social media, has created a trove of easily accessible text-, research- and raw data for researchers and data scientists to use to study different behaviours and phenomena. Due to the increased processing powers of the under-laying information systems categorization and contextual semantic enrichment of such data is currently possible. Currently, the pressing questions of the mass amount of collected data pinpoint, how to handle this magnitude of data? How to reflect, understand and optimize underlying systems, based on the insights of the data collected? Summarizing with a single question, how to gain knowledge for these data? To further press the future aspects of the Big Data era, this work can be connected to several ongoing trends such as the 4^{th} industrial revolution and modern public intelligence.

The fourth industrial revolution, Industry 4.0 (I4.0), is a paradigm shift that actively changed the way we process and produce things. The short historical overview of I4.0 begins

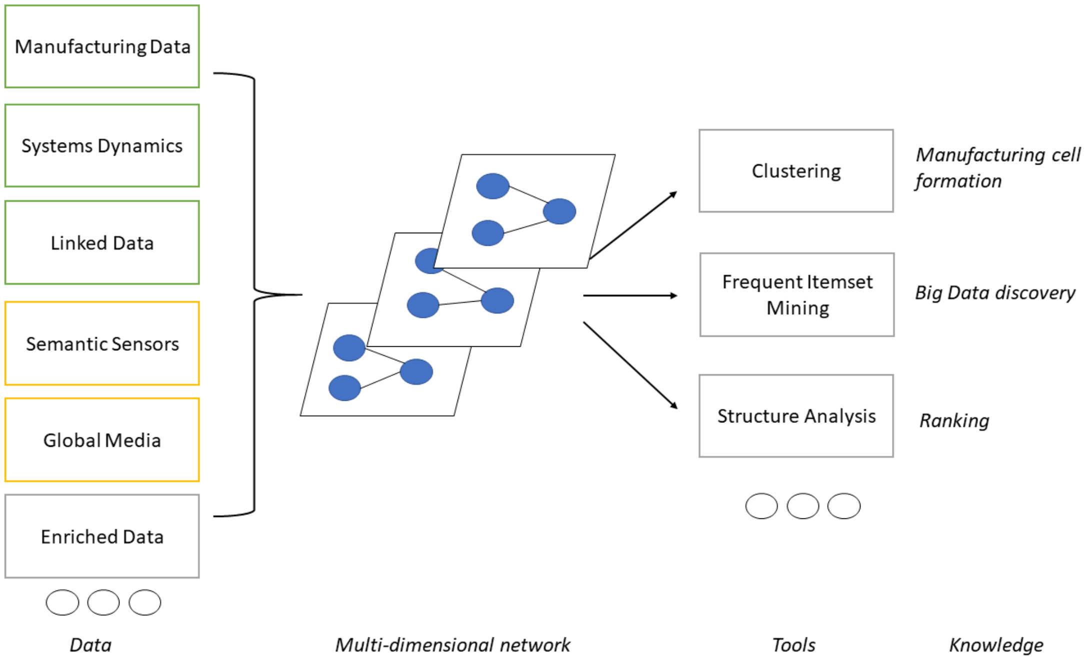


Figure 5.1. Representation of the work done, the green rectangles of the left shows the topic where this work made advancements, the yellow ones showing experiments on topic and the grey one shows possible topics to work on.

with the first industrial revolution at the end of the 18th century enabled by mechanization and steam power. The spread of electricity enabled assembly lines and mass production during the first half of the 20th century, which was the starting point of the second industrial revolution. Industry 3.0 came with the invention of the computer with an increase of automation such as programmable machines and robots. The fourth revolution is upcoming and is supposed to increase productivity and flexibility to the same extent as the previous three. The term Industry 4.0 was introduced in 2011 by the German government as a national program to boost research and development of the manufacturing industry. Many countries within the EU and also outside of the EU, has since then started similar initiatives. The non-EU, but similar initiatives are referred to as "smart factories". The aim is driven by market forces to prevent further outsourcing of production to low-cost countries by improving competitiveness with increased automation and flexibility, moving from mass production into mass customization, and reducing the 'time to market'. However, not only flexibility and the extended production capabilities were important, but also the sustainable production[173]. The implementation was slow, and many manufacturing companies had only started to computerize and were still far from being fully digitalized.

The digitalization addressed several catchy tools, like secure control systems, central assembly information systems and digital twins and broad trends such as machine learning, big data and machine-to-machine communication. The base and key enabling technology towards I4.0 was the Industrial Internet of Things (IIoT). The idea to utilize recent advances in information technologies and the Internet to interconnect machines, tools, equipment, sensors, and people into decentralized intelligent systems that can sense and adapt to the ever-changing tasks and production environment making appropriate responses, was often demanded real-time with high-level data analysis. The presented idea was still far, from capturing the whole program of which includes three dimensions: (1) horizontal integration across the entire value creation network (integration across the entire value creation network describes the cross-company and company-internal intelligent cross-linking and digitalization of value creation modules throughout the value chain of a product life cycle and between value chains of adjoining product life cycles), (2) end-to-end engineering across the entire product life cycle (intelligent cross-linking and digitalization throughout all phases of a product life cycle: from the raw material acquisition to manufacturing system, product use, and the product end of life), as well as (3) vertical integration and networked manufacturing systems (intelligent cross-linking and digitalization within the different aggregation and hierarchical levels of a value creation module from manufacturing stations via manufacturing cells, lines and factories, also integrating the associated value chain activities such as marketing and sales or technology development)[174]. The dissertation have not answered all three dimensions but gave an insight into vertical integration and also outlook, and tools to interact and analyze the other two dimensions, of horizontal integration and to the end-to-end product life cycle engineering.

Chapter 2 showcased the wire-harness production flow analysis integrating flexible and agile manufacturing systems in the name of Industry 4.0. The chapter showed that integration of enterprise resource planning, manufacturing execution systems, shop floor control and product lifecycle management, is straightforward to identify the connections of the standardized variables of production management and transform them into a multidimensional network model to support production flow analysis and departmental flow analysis according to the concept of Industry 4.0. It was recognized that modularity analysis of the network is a promising tool for forming groups in production flow analysis, and the performances of advanced (bipartite and multi-layer) network modularity algorithms (like InfoMap) are comparable to the most advanced optimization algorithm tailored to the problem of cell formation.

Systems thinking and system dynamics explicitly aims to facilitate the understanding of complex systems and the construction of models that describe their characteristics. Chapter 3 presented a tool for automated structural analysis if such models. Sustainability roughly referred to meeting the needs of ones without compromising the ability of future generations to meet their own needs. The design and monitoring of sustainability policies should rely on models that can capture the complex dynamics of interconnected variables and sustainability-related subsystems. The dissertation showed the most famous model of sustainability models in detail to showcase the automated structural analysis of system dynamics models. The in-depth analysis demonstrated how practicing engineers and analysts could apply the method for the structural analysis and validation of complex sustainability models. The chapter also raised questions about sustainability thinking and models in general. In the future, we will publish an article about the sustainable development goals appearance in the global media.

Chapter 4 showcased the Linked Data analysis. Linked Data is an approach to representing and publishing information applied the principles of the web to the challenge of integrating and using data. This was all about large scale integration and reasoning on data. One of the cornerstones of the large scale integration was the schema-free, self-descriptive data. As it was a trend, more and more domains were adopting it, with extending existing and creating new domain related ontologies. In a review article [175] we have collected the potentials and needs of applying enriched data to sensors. To further press the need of ontological enrichment of raw data is, that with the transformation of raw data to the multidimensional network could be automatic. In the chapter of Linked Data, there is a proposed method that generated a compact and interpretable multi-layered network from a linked dataset or another multidimensional network, capturing the most informative context and making it analyzable. Frequent itemset mining of the edge and node dimensions extracted the informative segments of a multidimensional network worth analyzing as the layers of a multi-layer network. The key findings of the chapter was a new notation of multidimensional networks by extending the nodes with dimensions, in contrast to the formal labelled network notation. This extension was useful in high-dimensional data analysis, such as for linked open data, as the nodes were often extended with hierarchical properties and ontologies.

5.1 Appendix to Chapter 2 - Details of the wire-harness production technology

To support the reproducible development of production flow analysis and optimization algorithms, an open source benchmark problem of a modular wire-harness production system was developed. The core of the system is a paced conveyor shown in Figure 5.2. Based on data published in [47] and [5], N_p was based on 64 products and defined N_m as a combination of 7 modules: m_1 base module, m_2 as left- or right-hand drive, m_3 normal/hybrid, m_4 halogen/LED lights, m_5 petrol/diesel engine, m_6 4 doors/5 doors and m_7 manual or automatic gearbox. N_a was defined 654 activities/tasks categorized into N_t which consisted of 16 activity types with well-modeled activity times (see Table 5.1). In these activities N_c was equal to 64 different built-in part families (component types) (among these $C_t = 180$ terminals, $C_b = 63$ bandages, $C_c = 25$ clips, and $C_w = 90$ wires). The conveyor N_w consisted of 10 workstations (tables). For every table (workstation) one operator is assigned, $N_o = 10$. The required N_s was also defined as 6 skills of the operators, namely: s_1 - laying cable, s_2 - spot-tying, s_3 - terminal attaching, s_4 - connector installing, s_5 - clip installing, and s_6 - visual testing. N_z was also defined as 6 zones for the workstations (see Figure 5.3) to study the distribution of the fixtures on the tables. The related $\mathbf{Z}$ matrix is defined based on the layout of the table and shows the relationship between the activities and zones of the workstation, which facilitates a detailed analysis of the workload in the workstations. All of this information is represented by a set of bipartite graphs defined in Table 2.1 and depicted in Figure 2.2. The related dataset is freely and fully available on the website of the research group: www.abonyilab.com

Figure 5.2. The wire-harness assembly pace conveyor [2]. The conveyor (often referred to as rotary) contains assembly tables consisting of connector and clip fixtures.

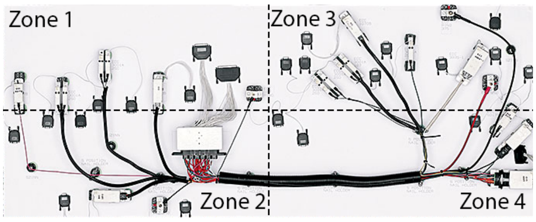


Figure 5.3. Zones were defined in the workstations to analyze the distribution of the fixtures and the related workload. The figure has been edited based on [3].

Table 5.1. Types of activities and the related activity times [5]. The activity times are calculated based on fixed and proportional values, e.g. when an operator is laying four wires over one foot, according to the t_4 model, the activity time will be $1 \times 6.9s + 4 \times 4.2 = 23.7s$

ID	Activity	Remark	Unit	Time [s]
t_1	Point-to-point wiring on chassis	Direct wiring	Number of wires	4.6
t_2	Laying in U-channel			4.4
t_3	Laying flat cable			7.7
t_4	Laying wire(s) onto harness jig	Laying flat cable	Base time Per wire	6.9 4.2
t_5	Laying cable connector (one end) onto harness jig	To the same breakout	Base time Per wire	7.4 2.3
t_6	Spot-tying onto cable and cutting it with a pair of scissors			16.6
t_7	Lacing activity		Base time Per additional stitch	1.5 3.6
t_8	Taping activity		Base time Per stitch	1.8 5.0
t_9	Inserting into tube or sleeve		Base time Per inch	3.0 2.4
t_{10}	Attachment of wire terminal	Terminal-block fastening (fork lug)		22.8
t_{11}	Screw fastening of terminal			17.1
t_{12}	Screw-and-nut fastening of terminal			24.7
t_{13}	Circular connector	Installation only		11.3
t_{14}	Rectangular connector	Latch or snap-on		24.0
t_{15}	Clip installation			8.0
t_{16}	Visual testing			120.0

.

5.2 Appendix to Chapter 3: Review of sustainability related system dynamics models

Table 5.2. Models of sustainability. The column denoted by # indicates the number of state variables in the related model.

Model	**#**	**Ref.**
2019		
Economics		
Specific dynamics were observed as oil palm cultivation increased which can be related to a reduced rate of displacement of the population, poverty and emissions of pollutants that affect air quality.	42	[176]
Modelling the presence of two categories of brand (preferred and un-preferred) and their respective processes of reverse logistics that strongly influence product shelf life, maximum allowable loss in sales and delays.	9	[177]
Water		
Investigation into the dependence on external water supplies of megacities in Shiraz, Iran by Monte Carlo simulation which suggests that the water shortage is more sensitive to the irrigation efficiency than the rate of treated wastewater and the capacity of the dam capacity.	4	[178]
2018		
Air & Climate		
Understanding how clean cooking is embedded within complex community processes to ensure its implementation, reduce household air pollution and alleviate adverse health issues and climate change.	4	[179]
Economics		
Constitutes a general framework for the representation of critical biotechonomic sectors, their land usage, production and consumption of raw material, research and development, investments and utilisation of production capacities (i.e. capital investments), usage of labour, finances (value creation, capital and operating costs as well as profit) and sustainability.	13	[180]
Highlighting the economic, environmental and social benefits that India can gain by recycling the gold contained in discarded cell phones	18	[181]
Simulation of the reaction of businesses to strategic and organisational changes with regard to performance, innovation and value creation.	11	[182]

Developement of criteria-based indicators using qualitative modelling based on mental models of decision-makers to investigate what corporate sustainability and sustainable manufacturing indicators are.	139	[183]
Investigation of the aspects of a resilient biocomposite production system, e.g. the interconnectivity including the reliability and effectiveness of the supply chain network, the resourcefulness of the production systems, adaptability, as well as adjusting the capacity of a system to be more versatile and flexible across the range of elements involved in the production system.	1	[184]
Simulation of strategies that will improve the adoption and sustainability of resource recovery systems in Placencia (Belize).	6	[185]
Model of trade scenarios to understand the impacts of import duties and non-price drivers on the relative volumes of imports for and domestic supply in China, the United States and the rest of the world.	6	[186]
Investigation of the financial impacts of maritime regulations, primarily the International Convention for the Safety of Life at Sea (SOLAS), International Convention for the Prevention of Pollution from Ships (MARPOL) and the International Convention on Standards of Training, Certification and Watchkeeping for Seafarers (STCW) by addressing ship instruments like operation, cargo, crew, environment, security and safety.	2	[187]
Identification of sustainability strategies for transportation development in emerging cities in China through the simulation of three modules, namely socio-economic, demand and supply.	10	[188]
Multilevel simulation of Coursera[1] to test hypotheses about the financial and sustainability performance of the venture over time.	13	[189]
Investigation of technological innovation systems emerging or declining in the context of various socio-technical transitions and different resourcing condition pathways.	10	[190]
A model showing the impact of the organizational behaviour of megaprojects and its performance with implications in terms of both the practical and cultural promotion of organizational citizenship behaviour.	9	[191]
Inspection of the organisational citizenship behaviour and performance at organisational megaprojects with implications in terms of both effective and cultural promotion of the behaviour.	9	[192]

---

[1] https://www.coursera.org/

Investigation of the impact of the changes in the population policy of China on social insurance by incorporating all the related factors that influence the income and expenditure of maternity insurance in Jiangsu Province. The goal is to determine a sustainable contribution rate of maternity insurance.	2	[193]
Energy		
Investigating the causal processes of biogas development to understand what endangers social, economic and environmental sustainability in Norway.	41	[194]
Investigating the changes in the bioenergy system influenced by environmental measures, economic development, and social impacts.	17	[195]
Simulation of sustainability in the cement industry using energy efficiency measures (clinker substitution, alternative fuel use and waste heat recovery).	5	[196]
A general model of technological innovation systems that overcome transformational failures as a result of different policies or policy mixes simulated in the field of electronic vehicles.	10	[190]
Mapping variables that act in an ethanol production chain in Southern Brazil to understand the interrelationships between its actors. Simulating different supply chain configurations to identify cause-and-effect interrelationships between the actors and seek solutions to the competitiveness and economic sustainability of biofuel chains by integrating associative production and family-farming.	81	[197]
Investigation of decentralised solar photovoltaic systems as a solution in unelectrified rural areas of India in terms of clean lighting and the reduced use of kerosene. The model takes into consideration the feedback between adoption, diffusion and implementation processes in resource-poor communities of low- and middle-income countries.	9	[198]
Land & Urban		
Identification of social science perspectives in terms of the food-energy-water nexus to delineate strategies that incorporate socio-ecological resilience and community capitals.	21	[199]
Exploring the main drivers of growth in the productivity of oil palms. Suggest ways to enhance the production of oil palms, albeit at the cost of unstoppable land loss.	11	[200]
Different scenarios based on the variables and their interrelationships according to the expertise and tacit knowledge of the stakeholders to determine the number of houses to be built in order to achieve a sustainable housing system in Florianopolis, Brazil.	6	[201]

Examination of the dynamics behind the formulation of policies to address issues surrounding urban rehabilitation in Hong Kong. It also presents the relationships between old buildings, rehabilitation policies and resourcing.	7	[202]
A model to identify and qualify the cross-scale interactions between each type of driver on the local dynamics of the socio-ecological system of the oasis of Comondú (Baja California Sur, Mexico).	5	[203]
Exploration of sustainable urbanisation performance and its indicators in China that account for transportation, population, land usage and GDP.	5	[113]
A model for mapping the determinant factors to identify areas of interventions in order to limit the appropriation of natural resources and support agri-food sustainability.	43	[204]
Water		
Model for formalizing the behaviours of water users and management authorities as well as the consequences of their actions on water management.	6	[111]
Highlights various variables that play a role regarding the sustainability of the water footprint management for supply chains subject to green market behaviour.	8	[205]
Characterisation of the water tariff (price)-demand-revenue system.	3	[206]
Simulation of the coordinated development of ecology and socio-economy in the Weihe River Basin, China.	15	[207]
Capturing the critical sustainability performance indicator of a poultry supply chain with a blue water footprint when significant environmental and regulatory constraints are applied. Simulations contribute to investigating the impact of two potential policy-making scenarios highlighting the water usage across the poultry supply chain.	3	[208]
2017		
Air & Climate		
Model of the proportionality factors and conversion rates six pollutants to establish quantitative connections between different types of variables, e.g. emissions and the annual average concentration of PM-2.5, to investigate atmospheric environmental capacity and atmospheric environmental carrying capacity restricted by GDP and PM-2.5.	15	[209]
Economics		
Research evaluating the role of the Latvian forest biotechonomy industry in the macroeconomic development model of the national economy.	19	[210]

Sustainability performance of construction projects is examined in terms of economic, social and environmental aspects with feedback loops based on expert opinions about causal links between the factors that affect sustainable construction and productivity.	115	[211]
Model of the sustainability, financial benefits and return on investments of Health Information Exchange as well as the correlation between Health Information Exchanges and hospital readmission reductions.	22	[212]
Modelling the dynamic interactions between students, science members, content, policy, pedagogy, community and government.	21	[213]
Implementing three scenarios to investigate sustainable development through investments in terms of environmental preservation in Beijing.	2	[214]
Evaluation of the sustainability of supply chains with regard to the use of raw materials due to the disposal fees, collection, recycling and return of some materials from desktop computers and laptops.	15	[215]
Modelling gender inequality in terms of pensions to show the differences in working hours over a lifetime due to childbearing/child-rearing and the simulation of social investment strategies and childcare policies in Norway.	3	[216]
Business simulation case study to assess product-service system-related business benefits that respond to multiple implementation scenarios and strategies.	5	[217]
Comparison between two types of ”collect to reuse” mobile phone businesses from the perspective of the potential for economic growth.	8	[218]
Simulation of several future development scenarios for the Hadaqi industrial corridor that show basic economic and social development goals which can be achieved at a reasonable rate of economic growth with an acceptable level of environmental impact with regard to the reduction of GDP per unit.	6	[219]
Energy		
Establishment and validation of the characteristics of each sustainability dimension of biofuel production. Analysis with regard to the status and sensitivity of the Colombian market.	13	[220]
Exploration of the potentials for and limitations of subsidies that might affect the future of climate-sensitive housing and night-time usage of air conditioners in Malaysia.	2	[221]
A model to analyse the long-term pattern of sustainable development with regard to renewable energy subject to feed-in tariffs and renewable portfolio standards.	6	[222]

Scenario-based analysis of the transition from conventional district heating systems to a 4th generation one to find a balance between the price of a fossil fuel and its share of thermal energy production.	15	[223]
Land & Urban		
The complexity of constructing high-rises for residential purposes is visually explained	88	[224]
Exploration of the outcomes of the rapid urbanization of land use and transportation options across multiple societal dimensions with feedback from the land, transportation, economic, equity and energy sectors.	27	[225]
Hypotheses-based modelling of small-scale agro-industrial operations and their business environments for the identification of the factors that influence factors of long-term financial sustainability.	1	[226]
Modelling four subsystems (population, economy, waste recycling and waste disposal) to test the high recycling rates that satisfy the requirements for economic growth and environmental sustainability while increasing landfill capacity for waste disposal.	8	[227]
Simulates urban–regional dynamics, including feedbacks from energy expenditure on economic growth and land scarcity in terms of the development of the Light Rail Transit between Durham and Chapel Hill.	26	[228]
Model of the Delphi transportation system that reveals key reinforcements such as policies that subsidize fossil fuels.	27	[229]
Simulation of scenarios that are indicate of the fact that appropriate governance and effective policy-making supports smallholder farming and short food supply chains in the developed world which provides promising grounds for ensuring food security and social cohesion while further promoting economic growth and environmental sustainability.	23	[230]
Model of scenarios with regard to the need to increase the profitability of farms by reducing production costs and the environmental impact of farming including the land wasted as result of the erosion. The goal is to better understand the processes as well as economic, socio-structural and biophysical variables that underpin the success of the conservation of arable cropping systems.	4	[231]
Water		
Simulation of the complex interactions between factors that govern the quantity/quality of water and their effects on social and economic conditions. The model simulates 70 years of policies and decisions that have the potential to improve conditions and prevent risks that may lead to social unrest and hinder economic development on the US-Mexican border.	6	[232]

Simulation of the management strategies of Tehran metropolitan water resources systems to reduce the deficit in water.	2	[233]
Life cycle sustainability assessment of a multi-region input-output analysis that is capable of quantitatively capturing macro-level social, environmental and economic impacts.	10	[234]
A model for understanding the possible effects of the implementations of sustainable policies in terms of environmental protection, wastewater treatment fees and the growth in the rate of the capacity of wastewater treatment for stakeholders of drainage enterprises.	4	[235]
2016		
Air & Climate		
Understandable visualization of the positive trade-offs between green practices with regard to energy efficiency and the reduction of CO2 emission.	6	[236]
Economics		
The study of policy design and management of beef cattle production systems in Malaysia to achieve sustainability	6	[237]
Dividing the capabilities of sustainable business and enterprise innovation into three aspects to analyse the influencing factors: knowledge-enabled innovation capability, product innovation capability, and marketing innovation capability.	4	[238]
Explaining the persistence of policy resistance in the context of sustainability transitions to overcome of policy resistance.	9	[239]
Analysis of the impact of different project conditions with indicators of time, cost and environmental impact.	2	[240]
Multi-level perspective on sustainability transitions to explain the emergence of functional foods as a niche in the nutrition of socio-technical systems.	22	[241]
Simulation to quantitatively analyse material and energy flow in the local eco-aquaculture industry chain. The simulation evaluates and examines the integrated effects of the ecological economics as well as their long-term evolutionary trends, moreover, and identifies defects of the system.	5	[242]
Identifying the drivers of an efficient closed-loop system in the production of mobile phones. The model proposes an eco-cycle scenario to lower pressures on resources by decreasing the demands of resources for production and increasing resource recovery.	10	[243]
A model that discovers the systemic interactions between drivers with regard to the economic, environmental and social sustainability of agricultural production.	6	[244]

A model that seeks the best coordinative strategy between the manufacturer and retailer for maximising profit in the reverse supply chain by modelling contracts in three different contracting processes.	18	[245]
Analysis of seven macro-level indicators, namely global warming potential, particulate matter formation, photochemical oxidant formation, cost of vehicle ownership, contribution to gross domestic product, employment generation and human health impacts, as well as the life-cycle cost of alternative fuel vehicles in the USA. The analysis covers their manufacture and impacts of the operation of internal combustion engine vehicles, hybrid electric vehicles, plug-in hybrid electric vehicles and battery electric vehicles.	9	[246]
Energy		
Investigation into the causal processes of biogas development to understand what endangered social, economic and environmental sustainability in Italy.	33	[247]
Identification of the optimal energy generation capacities and schedules for biomass and backup boilers, along with the optimal size of biomass orders and its storage. Investigation into the sensitivity of decisions with regard to changes in the source, types and pricing of biomass materials as well as the choice of technologies and their costs and operational performance criteria.	3	[248]
Determination of the smart self-sustainable pattern in waste-to-energy microalgae technology, including anaerobic digestion and renewable energy crops. The technology can be sustainable in terms of providing treated water, fertiliser, animal feed and energy for use, as well as reduce emissions of wastewater, resource consumption and environmental impacts in the field of waste management.	12	[249]
Land & Urban		
The study of the unpopularity of building projects within the urban fabric through a case study of a parking plaza	36	[250]
A model for identifying areas to increase the sustainability and resilience of food supply systems in city regions.	9	[251]
Several adopted models that describe a variety of problems in agriculture and natural resource management which occur globally from food supply distribution to soil erosion and irrigation.	13	[252]
Water		
The capture of the interacting factors and policy effects of the net anthropogenic nitrogen input (NANI) concept for Dianchi Lake in China.	11	[253]

Identification of leverage points used to suggest structural changes to achieve sustained benefits of water services and emphasise the importance of the enhanced robustness and reliability of water technology.	23	[254]
2015		
Air & Climate		
Scenario-based simulations to identify the most vital variable of the complex systems of firms and the impacts of climate change in order that a variety of scenarios that depict the potential evolution of the operation of a firm are tested.	9	[255]
Economics		
Company- and industry-level sustainability analysis for the investigation of decisions and policies with regard to conducting the cost-benefit scenarios for participants in closed-loop supply chains.	13	[256]
A formative scenario analysis that shows the loss of common property pastures and the resulting afforestation in Switzerland.	26	[257]
Examination of how the physical processes and information flows are interrelated in the collection of waste portable batteries and dynamic behaviours over time.	10	[258]
Understanding improvements in campus sustainability initiatives according to five program designs and two performance indicators, namely energy and monetary savings.	5	[259]
The modelling of tuna fish processing according to three limitations, the catch, sustainable yield and maximum economic yield.	20	[260]
Simulation the possibilities of a closed-loop supply chain for an electronics manufacturing company. The simulation proposes solutions to raise consumer satisfaction and the green image factor.	6	[261]
Decision support to select the best sustainable supplier based on behaviors with regard to relevant sustainability criteria.	4	[262]
A model that considers interrelated dimensions of sustainability and adopts the product life cycle with its inherent uncertainties, such as the length and pattern of the product life cycle as well as the residence index.	15	[263]
Two models to investigate the Dutch Wadden Sea region by studying the factors of tourism and the mussel fisheries. The goal is to identify critical variables such as market supply from the point of view of the fishery and experience value from that of tourism.	54	[264]

Scenarios for the dynamic control and scientific management of sustainable development of tourism in Tibet by analysing 13 tourism-related sustainability indicators, e.g. conventional tourism income, tourism resource stock, pollution stock and specific residents in the field of tourism (e.g. seasonal differences, accessibility).	6	[265]
Internal and external evaluation of the aspects of sustainable infrastructure, focusing on the unique characteristics of sustainable urban infrastructure in China to model policy settings and monitor operations.	11	[266]
The presentation of the competitive behavioural loops of contractors in specific market environments of the building and construction sector to analyse the implication of policies and sustainable developments.	4	[267]
Energy		
A conceptual system of renewable electricity accumulation to illustrate the possibility of the integration of large-scale intermittent energy sources in electrical grids.	8	[268]
The testing of biogas production policies subject to uncertainties, concerning three conflicting objectives, namely maximize production, reduce emissions and minimise costs, in the Netherlands.	4	[269]
Analysis of the long-term effects of various policy instruments that were proposed in the British Electricity Market Reform, focusing on environmental quality, security of supply and economic sustainability.	3	[270]
Land & Urban		
The identification of aspects to avoid the collapse of cocoa production systems in Malaysia	7	[271]
Study of the effects of urban transportation strategies, fuel tax, motorcycle parking management and free bus services to explore their potential in reducing the fuel consumption of vehicles and mitigating carbon dioxide emissions.	7	[272]
Water		
Basic water sustainability indicators investigated under simulated scenarios to model the impact of different policy packages	6	[273]
Capture of the dynamics of water footprints and water scarcity along agrifood supply chains to simulate water footprint management policies.	16	[274]
A hydrologic-economic model of sustainable groundwater use in drylands to model the net groundwater balance and variation in the consumers of water in terms of the population and tourism. The governing variables of the system were identified, e.g. the rise in international tourism as the main driving force that reduces emigration.	23	[275]

Model of the water demand system in China that forecasts environmental (water quality, ecosystem preservation) and socio-economic (population growth, water consumption, policy and management) factors of regional water demand as well as their nonlinear interactions with the physical elements of hydrological processes (natural runoff, groundwater recharge).	41	[276]
2014		
Economics		
Predicts the long-term trends in the hilsa fish population over several decades to assess the impacts of harvesting the juveniles and spawning adults.	7	[277]
Aiding the understanding of the interaction between the wetland environment and management policies.	16	[278]
Simulation of the behaviours and interactions in the Malaysian palm oil industry, from plantation to mill and mill-refinery to indicate how information could be linked throughout the different tiers of the supply network.	5	[279]
Exploration of feasibility, replicability and comprehensive assessment for the competitive positioning of organizations.	78	[280]
Investigation of the dynamic relationship between the supply chain, operations reference model, carbon emissions, liquid waste and employee welfare.	13	[281]
A model that combines mining, trading markets, price mechanisms, population dynamics, usage in society and waste as well as recycling into a worldwide system.	16	[282]
Model of the sustainability of the global silver supply, reserves and stocks by capturing the supply rates from mining, the reduction in ore grade and ground reserves, the peak year of reserves endowments, the general trend in the price of silver, the cumulative amount mined to date, market prices, etc.	34	[283]
A model that captures soft issues and feedback causalities with regard to the sustainability of quality improvement programs in a heavy manufacturing-engineering environment.	2	[284]
A model addressing the impact of policies and the perceptions of stakeholders with regard to the measuring of the performance of sustainability in the case of publicly funded projects.	1	[285]
A model to optimise the industrial restructuring of sustainable relationships between the environment, economy and society in the traditional industrial city of Leshan in South west China.	6	[286]
Investigation of the sustainability of construction projects in terms of technological advances and changes in people's perceptions.	1	[287]

Simulating the difficulties of improving the business, lean supply chain and environmental performance in accordance with remanufacturing by the retailer and recycling by the supplier.	17	[288]
Land & Urban		
Exploring the development of infinite city sprawl and identifying the cause-and-effect relationships in terms of the development trends of traffic congestion, high-density population aggregation in cities, land development and the disappearance of greenfield land resulting from urban sprawl in Taiwan.	9	[289]
Water		
Building operational early warning signals for the critical transitions of systems, where tipping points are suspected to exist, according to a study of an irrigation system.	2	[290]
An analysis of water conservation in Las Vegas according to water-pricing policies and smart water appliances.	6	[291]
Simulation of decisions related to variables that have an impact on the water sustainability of hospitals such as water reuse and reduction.	3	[292]
Quantitative simulation of the influence of interrelationships and feedback loops in wastewater collection network management. The model captures cost drivers and revenue sources in the system. It also includes a set of policy levers which facilitates the formulation of various financing and rehabilitation strategies.	13	[293]
2013		
Economics		
Examining the dynamics of the two dominant species in fishery (Atrina tuberculosa and Pinna rugosa).	4	[294]
Investigation into the capacity planning policies concerning the long-term profitability of a single-producer closed-loop recycling network.	13	[295]
Investigation into the long-term impacts of international outsourcing in the electrical industry from Europe to China on environmental, social and economic sustainability.	11	[296]
Energy		
Modelling the change in energy structure from carbon-based fuel (coal) to low-carbon fuel (natural gas) to reduce carbon emissions in Beijing. The model forecasts that the service sector will gradually displace the dominant status of industry in terms of energy consumption as the sector with the largest energy demand, followed by the industrial and transport sectors. The analysis suggests that a stable rate of population growth will have a far-reaching influence on energy consumption and carbon emissions.	2	[297]

Simulation that attempts to avoid a gap between electricity supply and demand, as well as sustainable, safe and cost competitive manufacturing.	8	[298]
Land & Urban		
Capturing spatio-temporal information about complex coastal systems including scenarios that simulate rises in sea level to make environmental decisions.	3	[299]
Water		
The analysis of water conservation in Las Vegas according to indoor and outdoor policies.	10	[300]
Examination of the impact of different management strategies on the water tariff regarding consistency and stability over time.	17	[301]
A model discussing the long-term impacts of various investment plans that seek to achieve water self-sufficiency according to ideas concerning underground water storage or surface water catchments.	7	[302]

Bibliography

[1] Vladimir Batagelj and Andrej Mrvar. A subquadratic triad census algorithm for large sparse networks with small maximum degree. *Social networks*, 23(3):237–243, 2001.

[2] Assembly line conveyor systems, 2015.

[3] Panduit introduces new technology for harness assembly board layout, 2015.

[4] Stefano Borgo and Paulo Leitão. Foundations for a core ontology of manufacturing. In *Ontologies*, pages 751–775. Springer, 2007.

[5] NS Ong and G Boothroyd. Assembly times for electrical connections and wire harnesses. *The International Journal of Advanced Manufacturing Technology*, 6(2):155–179, 1991.

[6] Stefano Boccaletti, Ginestra Bianconi, Regino Criado, Charo I. Del Genio, Jesús Gómez-Gardenes, Miguel Romance, Irene Sendina-Nadal, Zhen Wang, and Massimiliano Zanin. The structure and dynamics of multilayer networks. *Physics Reports*, 544(1):1–122, 2014. Publisher: Elsevier.

[7] Luis Solá, Miguel Romance, Regino Criado, Julio Flores, Alejandro García del Amo, and Stefano Boccaletti. Eigenvector centrality of nodes in multiplex networks. *Chaos: An Interdisciplinary Journal of Nonlinear Science*, 23(3):033131, 2013.

[8] Petter Holme and Jari Saramäki. Temporal networks. *Physics reports*, 519(3):97–125, 2012.

[9] Michele Coscia. *Multidimensional network analysis.* PhD thesis, Ph. D. thesis, Universitá Degli Studi Di Pisa, Dipartimento di Informatica, 2012.

[10] Jonathan F Donges, Hanna CH Schultz, Norbert Marwan, Yong Zou, and Jürgen Kurths. Investigating the topology of interacting networks. *The European Physical Journal B*, 84(4):635–651, 2011.

[11] Sergey V Buldyrev, Roni Parshani, Gerald Paul, H Eugene Stanley, and Shlomo Havlin. Catastrophic cascade of failures in interdependent networks. *Nature*, 464(7291):1025–1028, 2010.

[12] Regino Criado, Julio Flores, Alejandro García del Amo, Jesús Gómez-Gardeñes, and Miguel Romance. A mathematical model for networks with structures in the mesoscale. *International Journal of Computer Mathematics*, 89(3):291–309, 2012.

[13] Regino Criado, Miguel Romance, and María Vela-Pérez. Hyperstructures, a new approach to complex systems. *International Journal of Bifurcation and Chaos*, 20(03):877–883, 2010.

[14] Diane Pennington. Demystifying linked data: are you ready for what's next? *CILIP Update*, 2016(Jul/Au):34–36, 2016.

[15] María Hallo, Sergio Luján-Mora, Alejandro Maté, and Juan Trujillo. Current state of linked data in digital libraries. *Journal of Information Science*, 42(2):117–127, 2016.

[16] Nikolaos Papakostas, James O'Connor, and Gerald Byrne. Internet of things technologies in manufacturing: Application areas, challenges and outlook. In *Information Society (i-Society), 2016 International Conference on.*, pages 126–131. IEEE, 2016.

[17] Michael Rüßmann, Markus Lorenz, Philipp Gerbert, Manuela Waldner, Jan Justus, Pascal Engel, and Michael Harnisch. Industry 4.0: The future of productivity and growth in manufacturing industries. *Boston Consulting Group*, 9, 2015.

[18] Shiyong Wang, Jiafu Wan, Di Li, and Chunhua Zhang. Implementing smart factory of industrie 4.0: an outlook. *International Journal of Distributed Sensor Networks*, 12(1):3159805, 2016.

[19] Reiner Anderl. Industrie 4.0-advanced engineering of smart products and smart production. In *19th International seminar on high technology, technological innovations in the product development, Piracicaba, Brazil*, pages 1–14, 2014.

[20] Zahir Irani, Amir M Sharif, and Peter ED Love. Transforming failure into success through organisational learning: an analysis of a manufacturing information system. *European Journal of Information Systems*, 10(1):55–66, 2001.

[21] Douglas M Lambert and Martha C Cooper. Issues in supply chain management. *Industrial Marketing Management*, 29(1):65–83, 2000.

[22] Vijay Gurbaxani and Seungjin Whang. The impact of information systems on organizations and markets. *Communications of the ACM*, 34(1):59–73, 1991.

[23] Jay Lee, Edzel Lapira, Behrad Bagheri, and Hung-An Kao. Recent advances and trends in predictive manufacturing systems in big data environment. *Manufacturing Letters*, 1(1):38–41, 2013.

[24] Fei Tao, Jiangfeng Cheng, Qinglin Qi, Meng Zhang, He Zhang, and Fangyuan Sui. Digital twin-driven product design, manufacturing and service with big data. *The International Journal of Advanced Manufacturing Technology*, pages 1–14, 2017.

[25] Hyun-Jun Shin, Kyoung-Woo Cho, and Chang-Heon Oh. Svm-based dynamic reconfiguration cps for manufacturing system in industry 4.0. *Wireless Communications and Mobile Computing*, 2018, 2018.

[26] Mehmet Karaköse and Hasan Yetiş. A cyberphysical system based mass-customization approach with integration of industry 4.0 and smart city. *Wireless Communications and Mobile Computing*, 2017, 2017.

[27] Ari-Pekka Hameri. Production flow analysis—cases from manufacturing and service industry. *International Journal of Production Economics*, 129(2):233–241, 2011.

[28] Dimitris Mourtzis and Michalis Doukas. Decentralized manufacturing systems review: challenges and outlook. *Logistics Research*, 5(3-4):113–121, 2012.

[29] Jeffrey Prince and JM Kay. Combining lean and agile characteristics: creation of virtual groups by enhanced production flow analysis. *International Journal of Production Economics*, 85(3):305–318, 2003.

[30] George Chryssolouris, Konstantinos Efthymiou, Nikolaos Papakostas, Dimitris Mourtzis, and Aris Pagoropoulos. Flexibility and complexity: is it a trade-off? *International Journal of Production Research*, 51(23-24):6788–6802, 2013.

[31] Konstantinos Efthymiou, Dimitris Mourtzis, A Pagoropoulos, Nikolaos Papakostas, and George Chryssolouris. Manufacturing systems complexity analysis methods review. *International Journal of Computer Integrated Manufacturing*, 29(9):1025–1044, 2016.

[32] John L Burbidge. Production flow analysis for planning group technology. *Journal of Operations Management*, 10(1):5–27, 1991.

[33] John L Burbidge. Production flow analysis. *Production Engineer*, 50(4.5):139–152, 1971.

[34] JL Burbidge. A manual method of production flow analysis. *Production Engineer*, 56(10):34–38, 1977.

[35] AR Rahani and Muhammad Al-Ashraf. Production flow analysis through value stream mapping: a lean manufacturing process case study. *Procedia Engineering*, 41:1727–1734, 2012.

[36] Anne-Françoise Cutting-Decelle, Line Pouchard, BP Das, RI Young, and JJ Michel. Utilising standards based approaches to information sharing and interoperability in manufacturing decision support. In *international conference FAIM04, Toronto, Canada*, pages 1–12, 2004.

[37] HK Lin and Jenny A Harding. A manufacturing system engineering ontology model on the semantic web for inter-enterprise collaboration. *Computers in Industry*, 58(5):428–437, 2007.

[38] Jennifer Frangos Barabási Albert-László. *Linked: The new science of networks.* AAPT, 2003.

[39] Stefano Boccaletti, Ginestra Bianconi, Regino Criado, Charo I Del Genio, Jesús Gómez-Gardenes, Miguel Romance, Irene Sendina-Nadal, Zhen Wang, and Massimiliano Zanin. The structure and dynamics of multilayer networks. *Physics Reports*, 544(1):1–122, 2014.

[40] Manojit Chattopadhyay, Pranab K Dan, and Sitanath Mazumdar. Application of visual clustering properties of self organizing map in machine–part cell formation. *Applied Soft Computing*, 12(2):600–610, 2012.

[41] Orhan Torkul, Ismail Hakki Cedimoglu, and AK Geyik. An application of fuzzy clustering to manufacturing cell design. *Journal of Intelligent & Fuzzy Systems*, 17(2):173–181, 2006.

[42] Simon Li and Houman Mehrabadi. Generation of block diagonal forms using hierarchical clustering for cell formation problems. *Procedia CIRP*, 17:44–49, 2014.

[43] Edurne Aguirre and Benôıt Raucent. Performances of wire harness assembly systems. In *Industrial Electronics, 1994. Symposium Proceedings, ISIE'94, 1994 IEEE International Symposium on.*, pages 292–297. IEEE, 1994.

[44] George Michalos, Sotiris Makris, Loukas Rentzos, and George Chryssolouris. Dynamic job rotation for workload balancing in human based assembly systems. *CIRP Journal of Manufacturing Science and Technology*, 2(3):153–160, 2010.

[45] Hisup Park, Soo Hong Lee, and Mark R Cutkosky. Computational support for concurrent engineering of cable harnesses. *Computers in Engineering*, 1, 1992.

[46] Andrew B Conru. A genetic approach to the cable harness routing problem. In *Evolutionary Computation, 1994. IEEE World Congress on Computational Intelligence, Proceedings of the First IEEE Conference on.*, pages 200–205, 1994.

[47] NS Ong. Activity-based cost tables to support wire harness design. *International Journal of Production Economics*, 29(3):271–289, 1993.

[48] B Agard and M Tollenaere. Design of wire harnesses for mass customization. *Recent Advances in Integrated Design and Manufacturing in Mechanical Engineering*, page 53, 2003.

[49] IGK Ei-Essawy and J Torrance. Component flow analysis-an effective approach to production systems' design. *Production Engineer*, 51(5):165–170, 1972.

[50] Tamás Csiszér. *Use of network approach in the improvement of operation efficiency on the field of quality development in Hungarian Ph.D. thesis.* Magyarországi Egyetem Cziráki József Doktori Iskola, 2014.

[51] Ian Robinson, Jim Webber, and Emil Eifrem. *Graph databases.* " O'Reilly Media, Inc.", 2013.

[52] Ian Robinson, Jim Webber, and Emil Eifrem. *Graph databases: new opportunities for connected data*. O'Reilly Media, Inc., 2015.

[53] Sean Bechhofer. Owl: Web ontology language. In *Encyclopedia of database systems*, pages 2008–2009. Springer, 2009.

[54] Amit Aherwar. An implementation of group technology using 'c' on a problem of machine sequencing. *Recent Trends in Information Technology (RTIT)*, 2010.

[55] Yongli Li, Peng Luo, and Chong Wu. A new network node similarity measure method and its applications. *arXiv preprint arXiv:1403.4303*, 2014.

[56] Mark EJ Newman. Modularity and community structure in networks. *Proceedings of the National Academy of Sciences*, 103(23):8577–8582, 2006.

[57] Michael J Barber. Modularity and community detection in bipartite networks. *Physical Review E*, 76(6):066102, 2007.

[58] Usha Nandini Raghavan, Réka Albert, and Soundar Kumara. Near linear time algorithm to detect community structures in large-scale networks. *Physical Review E*, 76(3):036106, 2007.

[59] Cesar O Flores, Timothée Poisot, Sergi Valverde, and Joshua S Weitz. *BiMat: Start Guide*. Wiley Online Library, 2015.

[60] Tai-Hsi Wu, Chin-Chih Chang, and Jinn-Yi Yeh. A hybrid heuristic algorithm adopting both boltzmann function and mutation operator for manufacturing cell formation problems. *International Journal of Production Economics*, 120(2):669–688, 2009.

[61] Xin Liu and Tsuyoshi Murata. Community detection in large-scale bipartite networks. *Information and Media Technologies*, 5(1):184–192, 2010.

[62] M Rosvall and CT Bergstrom. Maps of information flow reveal community structure in complex networks. *arXiv preprint physics.soc-ph/0707.0609*, 2007.

[63] Martin Rosvall, Daniel Axelsson, and Carl T Bergstrom. The map equation. *The European Physical Journal Special Topics*, 178(1):13–23, 2009.

[64] Manlio De Domenico, Mason A Porter, and Alex Arenas. Muxviz: a tool for multilayer analysis and visualization of networks. *Journal of Complex Networks*, 3:159–176, 2015.

[65] Manlio De Domenico, Vincenzo Nicosia, Alex Arenas, and Vito Latora. Layer aggregation and reducibility of multilayer interconnected networks. *arXiv preprint arXiv:1405.0425*, 2014.

[66] Santo Fortunato. Community detection in graphs. *Physics reports*, 486(3-5):75–174, 2010.

[67] Frits K. Pil and Matthias Holweg. Linking Product Variety to Order-Fulfillment Strategies, 2004.

[68] Noel M. Tichy, Michael L. Tushman, and Charles Fombrun. Social network analysis for organizations. *The Academy of Management Review*, 4(4):507, Oct 1979.

[69] David Krackhardt and Daniel J Brass. *Intraorganizational networks: The micro side.* Sage Publications, Inc, 1994.

[70] M. Magnani and L. Rossi. The ml-model for multi-layer social networks. In *2011 International Conference on Advances in Social Networks Analysis and Mining*, page 5–12. IEEE, Jul 2011.

[71] Annalisa Socievole, Floriano De Rango, and Antonio Caputo. *Wireless contacts, Facebook friendships and interests: Analysis of a multi-layer social network in an academic environment*, page 1–7. IEEE, 2014.

[72] Liu Weiyi, Chen Lingli, and Hu Guangmin. Mining essential relationships under multiplex networks. *arXiv preprint arXiv:1511.09134*, 2015.

[73] Mark E Dickison, Matteo Magnani, and Luca Rossi. *Multilayer social networks.* Cambridge University Press, 2016.

[74] Yohsuke Murase, János Török, Hang-Hyun Jo, Kimmo Kaski, and János Kertész. Multilayer weighted social network model. *Physical Review E*, 90(5), Nov 2014.

[75] Gabriel Lodewijks. Two decades dynamics of belt conveyor systems. *Bulk Solids Handling*, 22(2):124–132, 2002.

[76] David W Cash, William C Clark, Frank Alcock, Nancy M Dickson, Noelle Eckley, David H Guston, Jill Jäger, and Ronald B Mitchell. Knowledge systems for sustainable development. *Proceedings of the national academy of sciences*, 100(14):8086–8091, 2003.

[77] Abdoulmohammad Gholamzadeh Chofreh, Feybi Ariani Goni, and Jiří Jaromír Klemeš. Development of a framework for the implementation of sustainable enterprise resource planning. *Chemical Engineering Transactions*, 61:1543–1548, 2017.

[78] Sarah Cornell, Frans Berkhout, Willemijn Tuinstra, J. David Tàbara, Jill Jäger, Ilan Chabay, Bert de Wit, Richard Langlais, David Mills, Peter Moll, Ilona M. Otto, Arthur Petersen, Christian Pohl, and Lorrae van Kerkhoff. Opening up knowledge systems for better responses to global environmental change. *Environmental Science & Policy*, 28:60–70, 2013.

[79] Qinghua Zhu, Ying Qu, Yong Geng, and Tsuyoshi Fujita. A comparison of regulatory awareness and green supply chain management practices among chinese and japanese manufacturers. *Business Strategy and the Environment*, 26(1):18–30, 2017.

[80] Albert-László Barabási and Márton Pósfai. *Network science*. Cambridge university press, 2016.

[81] Jing Du. The “weight” of models and complexity. *Complexity*, 21(3):21–35, 2016.

[82] Jay Wright Forrester. *World dynamics*, volume 59. Wright-Allen Press Cambridge, MA, 1971.

[83] Donella H. Meadows, Dennis L. Meadows, Jørgen Randers, and William W. Behrens III. *The Limits to Growth: A Report of the Club of Rome’s Project on the Predicament of Mankind*. Universe Books, 1972.

[84] Slobodan P Simonovic. World water dynamics: global modeling of water resources. *Journal of Environmental Management*, 66(3):249–267, 2002.

[85] Roberto Pasqualino, Aled W Jones, Irene Monasterolo, and Alexander Phillips. Understanding global systems today—a calibration of the world3-03 model between 1995 and 2012. *Sustainability*, 7(8):9864–9889, 2015.

[86] Donella H Meadows, Dennis L Meadows, and Jorgen Randers. *Beyond the Limits: Confronting Global Collapse: Envisioning a Sustainable Future.* Chelsea Green Publishing, London,UK, 1992.

[87] Dennis Meadows and Jorgan Randers. *The limits to growth: the 30-year update.* Chelsea Green Publishing, London, UK, 2012.

[88] Steve Bankes, Warren E Walker, and Jan H Kwakkel. *Exploratory modeling and analysis*, pages 532–537. Springer, 2013.

[89] Stephanie Albin, Jay Wright Forrester, and Lucia Breierova. *Building a system dynamics model: Part 1: conceptualization.* MIT, 2001.

[90] James K Doyle and David N Ford. Mental models concepts for system dynamics research. *System dynamics review: the journal of the System Dynamics Society*, 14(1):3–29, 1998.

[91] Ehsan Nabavi, Katherine A Daniell, and Husain Najafi. Boundary matters: the potential of system dynamics to support sustainability? *Journal of cleaner production*, 140:312–323, 2017.

[92] Hazhir Rahmandad and Nelson Repenning. Capability erosion dynamics. *Strategic Management Journal*, 37(4):649–672, 2016.

[93] John D Sterman, Rebecca Henderson, Eric D Beinhocker, and Lee I Newman. Getting big too fast: Strategic dynamics with increasing returns and bounded rationality. *Management Science*, 53(4):683–696, 2007.

[94] J David Tàbara and Ilan Chabay. Coupling human information and knowledge systems with social–ecological systems change: Reframing research, education, and policy for sustainability. *Environmental Science & Policy*, 28:71–81, 2013.

[95] John D. Sterman. All models are wrong: reflections on becoming a systems scientist. *System Dynamics Review*, 18(4):501–531, 2002.

[96] Eleonora Bottani, Maria Carmen Gentilotti, and Marta Rinaldi. A fuzzy logic-based tool for the assessment of corporate sustainability: A case study in the food machinery industry. *Sustainability*, 9(4), 2017.

[97] Ji Chen, Jinsheng Wang, Tomas Baležentis, Fausta Zagurskaitė, Dalia Streimikiene, and Daiva Makutėnienė. Multicriteria approach towards the sustainable selection of a teahouse location with sensitivity analysis. *Sustainability*, 10(8), 2018.

[98] Yi Lu and Bin Chen. Urban ecological footprint prediction based on the markov chain. *Journal of Cleaner Production*, 163:146 – 153, 2017. Achieving Low/no Fossil-carbon Economies based upon the Essential Transformations to Support them.

[99] George A Miller. The magical number seven, plus or minus two: Some limits on our capacity for processing information. *Psychological review*, 63(2):81, 1956.

[100] Ehsan Nabavi, Katherine A. Daniell, and Husain Najafi. Boundary matters: the potential of system dynamics to support sustainability? *Journal of Cleaner Production*, 140:312–323, 2017.

[101] G. Dörgő, G. Honti, and J. Abonyi. Automated analysis of the interactions between sustainable development goals extracted from models and texts of sustainability science. *Chemical Engineering Transactions*, 70:781–786, 2018.

[102] Michelle Girvan and Mark EJ Newman. Community structure in social and biological networks. *Proceedings of the national academy of sciences*, 99(12):7821–7826, 2002.

[103] M. E. J. Newman. Modularity and community structure in networks. *Proceedings of the National Academy of Sciences*, 103(23):8577–8582, 2006.

[104] Matthias Dehmer, Veronika Kraus, Frank Emmert-Streib, and Stefan Pickl. *Quantitative graph theory.* CRC press, 2014.

[105] Albert-László Barabási and Márton Pósfai. *Network science.* Cambridge University Press, Cambridge, 2016.

[106] Linton C Freeman. Centrality in social networks conceptual clarification. *Social networks*, 1(3):215–239, 1978.

[107] Lawrence Page, Sergey Brin, Rajeev Motwani, and Terry Winograd. The pagerank citation ranking: Bringing order to the web. Report, Stanford InfoLab, 1999.

[108] Hartmut Bossel. *System Zoo 3 Simulation Models: Economy, Society, Development*, volume 3. BoD, 2007.

[109] Alexandra Milik, Alexia Prskawetz, Gustav Feichtinger, and Warren C Sanderson. Slow-fast dynamics in wonderland. *Environmental Modeling & Assessment*, 1(1-2):3–17, 1996.

[110] Peder Hjorth and Ali Bagheri. Navigating towards sustainable development: A system dynamics approach. *Futures*, 38(1):74–92, 2006.

[111] Irene Pluchinotta, Alessandro Pagano, Raffaele Giordano, and Alexis Tsoukiàs. A system dynamics model for supporting decision-makers in irrigation water management. *Journal of Environmental Management*, 223:815–824, 2018.

[112] Krystyna A. Stave. A system dynamics model to facilitate public understanding of water management options in las vegas, nevada. *Journal of Environmental Management*, 67(4):303–313, 2003.

[113] Yongtao Tan, Liudan Jiao, Chenyang Shuai, and Liyin Shen. A system dynamics model for simulating urban sustainability performance: A china case study. *Journal of Cleaner Production*, 199:1107–1115, 2018.

[114] Jay W Forrester. Urban dynamics. *IMR; Industrial Management Review (pre-1986)*, 11(3):67, 1970.

[115] Chihhao Fan, Shu-Kai S. Fan, Chen-Shu Wang, and Wen-Pin Tsai. Modeling computer recycling in taiwan using system dynamics. *Resources, Conservation and Recycling*, 128:167–175, 2018.

[116] Josef Maroušek. Study on agriculture decision-makers behavior on sustainable energy utilization. *Journal of Agricultural and Environmental Ethics*, 26(3):679–689, Jun 2013.

[117] Cataldo Musto, Fedelucio Narducci, Pasquale Lops, Marco de Gemmis, and Giovanni Semeraro. Linked open data-based explanations for transparent recommender systems. *International Journal of Human-Computer Studies*, 121:93–107, 2019.

[118] Jose Emilio Labra Gayo, Johan Jeuring, and Jose María Álvarez Rodríguez. Inductive representations of RDF graphs. *Science of Computer Programming*, 95:135–146, 2014.

[119] Nahla Mohammed Elzein, Mazlina Abdul Majid, Ibrahim Abaker Targio Hashem, Ibrar Yaqoob, Fadele Ayotunde Alaba, and Muhammad Imran. Managing big RDF data in clouds: Challenges, opportunities, and solutions. *Sustainable Cities and Society*, 39:375–386, 2018.

[120] Klyne Graham, Carroll Jeremy J, and Brian McBride. RDF 1.1 Concepts and Abstract Syntax, 2014.

[121] Patrick J. Hayes and Peter F. Patel-Schneider. RDF 1.1 Semantics.

[122] Maria-Evangelia Papadaki, Nicolas Spyratos, and Yannis Tzitzikas. Towards Interactive Analytics over RDF Graphs. *Algorithms*, 14(2):34, February 2021.

[123] Jonathan Hayes and Claudio Gutierrez. Bipartite graphs as intermediate model for RDF. In *International Semantic Web Conference*, pages 47–61. Springer, 2004.

[124] Graham Klyne and Jeremy J Carroll. Resource description framework (RDF): Concepts and abstract syntax. *W3C Recommendation*, 2005.

[125] Nigel Shadbolt, Tim Berners Lee, and Wendy Hall. The semantic web revisited. *IEEE Intelligent Systems*, 21(3):96–101, 2006.

[126] Stefan Decker, Sergey Melnik, Frank Van Harmelen, Dieter Fensel, Michel Klein, Jeen Broekstra, Michael Erdmann, and Ian Horrocks. The semantic web: The roles of xml and rdf. *IEEE Internet Computing*, 4(5):63–73, 2000.

[127] E. Kalampokis, D. Zeginis, and K. Tarabanis. On modeling linked open statistical data. *Journal of Web Semantics*, 55:56–68, 2019. cited By 6.

[128] Nigel Shadbolt and Kieron O'Hara. Linked data in government. *IEEE Internet Computing*, 17(4):72–77, 2013.

[129] Alison Callahan, José Cruz-Toledo, Peter Ansell, and Michel Dumontier. Bio2rdf release 2: improved coverage, interoperability and provenance of life science linked data. In *Extended Semantic Web Conference*, pages 200–212. Springer, 2013.

[130] Anja Jentzsch, Jun Zhao, Oktie Hassanzadeh, Kei-Hoi Cheung, Matthias Samwald, and Bosse Andersson. Linking open drug data. In *I-SEMANTICS*, pages 1–4, 2009.

[131] Philipp Cimiano, Christian Chiarcos, John P McCrae, and Jorge Gracia. Representing annotated texts as rdf. In *Linguistic Linked Data*, pages 61–87. Springer, 2020.

[132] Ivan Ermilov, Michael Martin, Jens Lehmann, and Soren Auer. Linked open data statistics: Collection and exploitation. In *International Conference on Knowledge Engineering and the Semantic Web*, pages 242–249. Springer, 2013.

[133] Nicolas Marie and Fabien Gandon. Survey of linked data based exploration systems. In *IESD 2014-Intelligent Exploitation of Semantic Data*, pages 1–13, 2014.

[134] Dieter Fensel, Frank Van Harmelen, Ian Horrocks, Deborah L McGuinness, and Peter F Patel-Schneider. Oil: An ontology infrastructure for the semantic web. *IEEE Intelligent Systems*, 16(2):38–45, 2001.

[135] Albert-laszlo Barabasi and Jennifer Frangos. *Linked: The New Science of Networks.* American Association of Physics Teachers, 2002.

[136] Lei Zou and M Tamer Özsu. Graph-based rdf data management. *Data Science and Engineering*, 2(1):56–70, 2017.

[137] Rosa Gil, Roberto García, and Jaime Delgado. Measuring the semantic web. *AIS SIGSEMIS Bulletin*, 1(2):69–72, 2004.

[138] Francesco Bellomi and Roberto Bonato. Network analysis for wikipedia. In *Proceedings of Wikimania*, 2005.

[139] Peter Mika. Flink: Semantic web technology for the extraction and analysis of social networks. *Web Semantics: Science, Services and Agents on the World Wide Web*, 3(2-3):211–223, 2005.

[140] Edmundo-Pavel Soriano-Morales, Julien Ah-Pine, and Sabine Loudcher. Hypergraph modelization of a syntactically annotated english wikipedia dump. In *International Conference on Language Resources and Evaluation (LREC 2016)*, 2016.

[141] Gergely Palla, Illés J Farkas, Péter Pollner, Imre Derényi, and Tamás Vicsek. Fundamental statistical features and self-similar properties of tagged networks. *New Journal of Physics*, 10(12):123026, 2008.

[142] Péter Pollner, Gergely Palla, and Tamás Vicsek. Clustering of tag-induced subgraphs in complex networks. *Physica A: Statistical Mechanics and its Applications*, 389(24):5887–5894, 2010.

[143] Gergely Palla, Gergely Tibély, Enys Mones, Péter Pollner, and Tamás Vicsek. Hierarchical networks of scientific journals. *Palgrave Communications*, 1:15016, 2015.

[144] Alexandre Passant. Measuring semantic distance on linking data and using it for resources recommendations. In *AAAI spring symposium: Linked Data Meets Artificial Intelligence*, volume 77, page 123, 2010.

[145] G Sudha Sadasivam, KG Saranya, and KG Karrthik. Hypergraph-based wikipedia search with semantics. *International Journal of Web Science*, 2(1-2):66–79, 2013.

[146] Roberto Mirizzi, Azzurra Ragone, Tommaso Di Noia, and Eugenio Di Sciascio. Ranking the linked data: The case of dbpedia. In *International Conference on Web Engineering*, pages 337–354. Springer, 2010.

[147] Vincenzo Nicosia and Vito Latora. Measuring and modeling correlations in multiplex networks. *Physical Review E*, 92(3):032805, 2015.

[148] H. Zhaohui, C. Huajun, Y. Tong, S. Hao, Z. Luo, and M. Yuxin. Semantic text mining with linked data. pages 338–343, 2009.

[149] Qaiser Mehmood, Muhammad Saleem, Ratnesh Sahay, Axel-Cyrille Ngonga Ngomo, and Mathieu D'Aquin. Qppds: Querying property paths over distributed rdf datasets. *IEEE Access*, 7:101031–101045, 2019.

[150] Alexandru Iosup, Tim Hegeman, Wing Lung Ngai, Stijn Heldens, Arnau Prat-Pérez, Thomas Manhardto, Hassan Chafio, Mihai Capotă, Narayanan Sundaram, Michael Anderson, et al. Ldbc graphalytics: A benchmark for large-scale graph analysis on parallel and distributed platforms. *Proceedings of the VLDB Endowment*, 9(13):1317–1328, 2016.

[151] M.-E. Papadaki, Y. Tzitzikas, and N. Spyratos. Analytics over rdf graphs. *Communications in Computer and Information Science*, 1197 CCIS:37–52, 2020.

[152] Zhi-yun Zheng, Chen-yu Wang, Yang Ding, Lun Li, and Dun Li. Research on partitioning algorithm based on rdf graph. *Concurrency and Computation: Practice and Experience*, pages 5600–5612, 2019.

[153] Theofilos Mailis, Yannis Kotidis, Vaggelis Nikolopoulos, Evgeny Kharlamov, Ian Horrocks, and Yannis Ioannidis. An efficient index for rdf query containment. In *Proceedings of the 2019 International Conference on Management of Data*, pages 1499–1516, 2019.

[154] M. Morzy, A. Ławrynowicz, and M. Zozuliński. Using substitutive itemset mining framework for finding synonymous properties in linked data. *Lecture Notes in Computer Science (including subseries Lecture Notes in Artificial Intelligence and Lecture Notes in Bioinformatics)*, 9202:422–430, 2015.

[155] S. Boytcheva, G. Angelova, Z. Angelov, D. Tcharaktchiev, and V. Vodenicharov. Enrichment of ehr with linked open data for risk factors identification. pages 84–90, 2019.

[156] Andreas Harth, Katja Hose, Marcel Karnstedt, Axel Polleres, Kai-Uwe Sattler, and Jürgen Umbrich. Data summaries for on-demand queries over linked data. In *Proceedings of the 19th international conference on World wide web*, pages 411–420, 2010.

[157] W3C SPARQL Working Group. SPARQL 1.1 Overview, 2013.

[158] Hans Peter Hertig. *Universities, Rankings and the Dynamics of Global Higher Education. Perspectives from Asia, Europe and North America.* Springer, 2016.

[159] Tero Erkkila and Ossi Piironen. *Rankings and Global Knowledge Governance: Higher Education, Innovation and Competitiveness.* Springer, 2018.

[160] Boanerges Aleman-Meza, C Halaschek-Weiner, I Budak Arpinar, Cartic Ramakrishnan, and Amit P Sheth. Ranking complex relationships on the semantic web. *IEEE Internet computing*, 9(3):37–44, 2005.

[161] Juyong Park and Albert-László Barabási. Distribution of node characteristics in complex networks. *Proceedings of the National Academy of Sciences*, 104(46):17916–17920, 2007.

[162] Michael Färber. The microsoft academic knowledge graph: A linked data source with 8 billion triples of scholarly data. In *International Semantic Web Conference*, pages 113–129. Springer, 2019.

[163] Alfio Ferrara, Lorenzo Genta, Stefano Montanelli, and Silvana Castano. Dimensional clustering of linked data: techniques and applications. In *Transactions on Large-Scale Data-and Knowledge-Centered Systems XIX*, pages 55–86. Springer, 2015.

[164] Rakesh Agrawal and Ramakrishnan Srikant. Fast algorithms for mining association rules. In *Proc. 20th int. conf. very large data bases, VLDB*, volume 1215, pages 487–499, 1994.

[165] Mohammed J Zaki and Ching-Jui Hsiao. CHARM: An efficient algorithm for closed itemset mining. In *Proceedings of the 2002 SIAM international conference on data mining*, pages 457–473. SIAM, 2002.

[166] Gösta Grahne and Jianfei Zhu. Fast algorithms for frequent itemset mining using fp-trees. *IEEE transactions on knowledge and data engineering*, 17(10):1347–1362, 2005.

[167] Jiawei Han, Jian Pei, Yiwen Yin, and Runying Mao. Mining frequent patterns without candidate generation: A frequent-pattern tree approach. *Data mining and knowledge discovery*, 8(1):53–87, 2004.

[168] Chin-Hoong Chee, Jafreezal Jaafar, Izzatdin Abdul Aziz, Mohd Hilmi Hasan, and William Yeoh. Algorithms for frequent itemset mining: a literature review. *Artificial Intelligence Review*, 52(4):2603–2621, 2019.

[169] Giulia Menichetti, Daniel Remondini, Pietro Panzarasa, Raúl J Mondragón, and Ginestra Bianconi. Weighted multiplex networks. *PloS one*, 9(6), 2014.

[170] Fan Zhang, Yan Zhang, and Jason Bakos. Gpapriori: Gpu-accelerated frequent itemset mining. In *2011 IEEE International Conference on Cluster Computing*, pages 590–594. IEEE, 2011.

[171] Y. Xun, J. Zhang, X. Qin, and X. Zhao. Fidoop-dp: Data partitioning in frequent itemset mining on hadoop clusters. *IEEE Transactions on Parallel and Distributed Systems*, 28(1):101–114, 2017.

[172] R. Joy and K. K. Sherly. Parallel frequent itemset mining with spark rdd framework for disease prediction. In *2016 International Conference on Circuit, Power and Computing Technologies (ICCPCT)*, pages 1–5, 2016.

[173] Tim Stock and Günther Seliger. Opportunities of sustainable manufacturing in industry 4.0. *Procedia Cirp*, 40:536–541, 2016. ISBN: 2212-8271 Publisher: Elsevier.

[174] Thomas Bauernhansl, B Diegner, Johannes Diemer, M Dümmler, C Eckert, W Herfs, H Heyn, C Hilger, M Ten Hompel, J Kalhoff, et al. Industrie 4.0-whitepaper fuethemen. *Berlin: Bundesministerium für Wirtschaft und Energie-Plattform Industrie*, 4, 2014.

[175] Gergely Marcell Honti and Janos Abonyi. A review of semantic sensor technologies in internet of things architectures. *Complexity*, 2019, 2019.

[176] S. Bautista, A. Espinoza, P. Narvaez, M. Camargo, and L. Morel. A system dynamics approach for sustainability assessment of biodiesel production in colombia. baseline simulation. *Journal of Cleaner Production*, 213:1–20, 2019.

[177] S.A. Narayana, R.K. Pati, and S.S. Padhi. Market dynamics and reverse logistics for sustainability in the indian pharmaceuticals industry. *Journal of Cleaner Production*, 208:968–987, 2019.

[178] M.H. Ahmadi and M. Zarghami. Should water supply for megacities depend on outside resources? a monte-carlo system dynamics simulation for shiraz, iran. *Sustainable Cities and Society*, 44:163–170, 2019.

[179] N. Chalise, P. Kumar, P. Priyadarshini, and G.N. Yadama. Dynamics of sustained use and abandonment of clean cooking systems: Lessons from rural india. *Environmental Research Letters*, 13(3), 2018.

[180] Andra Blumberga, Gatis Bazbauers, Pål I Davidsen, Dagnija Blumberga, Armands Gravelsins, and Toms Prodanuks. System dynamics model of a biotechonomy. *Journal of Cleaner Production*, 172:4018–4032, 2018.

[181] K. Chaudhary and P. Vrat. Circular economy model of gold recovery from cell phones using system dynamics approach: a case study of india. *Environment, Development and Sustainability*, pages 1–28, 2018.

[182] F. Cosenz and G. Noto. A dynamic business modelling approach to design and experiment new business venture strategies. *Long Range Planning*, 51(1):127–140, 2018.

[183] A. Moldavska and T. Welo. Testing and verification of a new corporate sustainability assessment method for manufacturing: A multiple case research study. *Sustainability (Switzerland)*, 10(11), 2018.

[184] I.S. Piri, O. Das, M.S. Hedenqvist, T. Väisänen, S. Ikram, and D. Bhattacharyya. Imparting resiliency in biocomposite production systems: A system dynamics approach. *Journal of Cleaner Production*, 179:450–459, 2018.

[185] C. Prouty, S. Mohebbi, and Q. Zhang. Socio-technical strategies and behavior change to increase the adoption and sustainability of wastewater resource recovery systems. *Water Research*, 137:107–119, 2018.

[186] D. Sandor, S. Fulton, J. Engel-Cox, C. Peck, and S. Peterson. System dynamics of polysilicon for solar photovoltaics: A framework for investigating the energy security of renewable energy supply chains. *Sustainability (Switzerland)*, 10(1), 2018.

[187] P. Setyohadi, K.B. Artana, D. Manfaat, and R.O.S. Gurning. Dynamic response of risk management model to mitigate impact of maritime regulatory changes: Oil tanker owners perspective. *Operations and Supply Chain Management*, 11(3):118–127, 2018.

[188] L. Shen, L. Du, X. Yang, X. Du, J. Wang, and J. Hao. Sustainable strategies for transportation development in emerging cities in china: A simulation approach. *Sustainability (Switzerland)*, 10(3), 2018.

[189] K. Täuscher and N. Abdelkafi. Scalability and robustness of business models for sustainability: A simulation experiment. *Journal of Cleaner Production*, 170:654–664, 2018.

[190] R. Raven and B. Walrave. Overcoming transformational failures through policy mixes in the dynamics of technological innovation systems. *Technological Forecasting and Social Change*, 2018.

[191] J. Wang, W. Li, and N. Mishima. A study of formal post-products collected to reuse business structures on resource sustainability through system dynamics. volume 69, pages 833–837, 2018.

[192] T. Wang, Q. He, Y. Lu, and D. Yang. How does organizational citizenship behavior (ocb) affect the performance of megaprojects? insights from a system dynamic simulation. *Sustainability (Switzerland)*, 10(6), 2018.

[193] X. Zhang, L. Zhou, and H.A. Antwi. The impact of china's latest population policy changes on maternity insurance—a case study in jiangsu province. *International Journal of Health Planning and Management*, 2018.

[194] B. Cavicchi. The burden of sustainability: Limits to sustainable bioenergy development in norway. *Energy Policy*, 119:585–599, 2018.

[195] E. Jin and J.W. Sutherland. An integrated sustainability model for a bioenergy system: Forest residues for electricity generation. *Biomass and Bioenergy*, 119:10–21, 2018.

[196] Z. Jokar and A. Mokhtar. Policy making in the cement industry for co2 mitigation on the pathway of sustainable development- a system dynamics approach. *Journal of Cleaner Production*, 201:142–155, 2018.

[197] Guilherme Luis Luís Roehe Vaccaro, Alexsander Longhi, Marcelo Moutinho, Annibal Scavarda, Catia M Lopes, Adriana Neves dos Reis, Fabiano Nunes, and Debora Azevedo. Interrelationship among actors in ethanol production chain as a competitive and sustainable factor: the case of associative production and family-farming in southern brazil. *Journal of Cleaner Production*, pages 1239–1255, 2018.

[198] J. Venkateswaran, C.S. Solanki, K. Werner, and G.N. Yadama. Addressing energy poverty in india: A systems perspective on the role of localization, affordability, and saturation in implementing solar technologies. *Energy Research and Social Science*, 40:205–210, 2018.

[199] J.E. Givens, J. Padowski, C.D. Guzman, K. Malek, R. Witinok-Huber, B. Cosens, M. Briscoe, J. Boll, and J. Adam. Incorporating social system dynamics in the columbia river basin: Food-energy-water resilience and sustainability modeling in the yakima river basin. *Frontiers in Environmental Science*, 6(SEP), 2018.

[200] A. Ibragimov, S.F. Sidique, and Y.S. Tey. Productivity for sustainable growth in malaysian oil palm production: A system dynamics modeling approach. *Journal of Cleaner Production*, 213:1051–1062, 2019.

[201] J.D.O. Musse, A.S. Homrich, R. de Mello, and M.M. Carvalho. Applying backcasting and system dynamics towards sustainable development: The housing planning case for low-income citizens in brazil. *Journal of Cleaner Production*, 193:97–114, 2018.

[202] M.C.P. Sing, P.E.D. Love, and H.J. Liu. Rehabilitation of existing building stock: A system dynamics model to support policy development. *Cities*, 2018.

[203] A. Tenza, J. Martínez-Fernández, I. Pérez-Ibarra, and A. Giménez. Sustainability of small-scale social-ecological systems in arid environments: trade-off and synergies of global and regional changes. *Sustainability Science*, 2018.

[204] F. Anastasiadis, N. Tsolakis, and J.S. Srai. Digital technologies towards resource efficiency in the agrifood sector: Key challenges in developing countries. *Sustainability (Switzerland)*, 10(12), 2018.

[205] E. Aivazidou, N. Tsolakis, D. Vlachos, and E. Iakovou. A water footprint management framework for supply chains under green market behaviour. *Journal of Cleaner Production*, 197:592–606, 2018.

[206] O. Sahin, E. Bertone, C. Beal, and R.A. Stewart. Evaluating a novel tiered scarcity adjusted water budget and pricing structure using a holistic systems modelling approach. *Journal of Environmental Management*, 215:79–90, 2018.

[207] J. Song, B. Tang, J. Zhang, X. Dou, Q. Liu, and W. Shen. System dynamics simulation for optimal stream flow regulations under consideration of coordinated development of ecology and socio-economy in the weihe river basin, china. *Ecological Engineering*, 124:51–68, 2018.

[208] N. Tsolakis, J.S. Srai, and E. Aivazidou. Blue water footprint management in a uk poultry supply chain under environmental regulatory constraints. *Sustainability (Switzerland)*, 10(3), 2018.

[209] Yejing Zhou and Jingxuan Zhou. Urban atmospheric environmental capacity and atmospheric environmental carrying capacity constrained by gdp–pm 2.5. *Ecological indicators*, 73:637–652, 2017.

[210] Reinis Azis, Andra Blumberga, and Gatis Bazbauers. The role of forest biotechonomy industry in the macroeconomic development model of the national economy of latvia: a system dynamics approach. *Energy Procedia*, 128:32–37, 2017.

[211] Shahin Dabirian, Mostafa Khanzadi, and Reza Taheriattar. Qualitative modeling of sustainability performance in construction projects considering productivity approach. *International Journal of Civil Engineering*, 15(8):1143–1158, 2017.

[212] Emad A Edaibat, Jason Dever, and Steven MF Stuban. System dynamics simulation modeling of health information exchange (hie) adoption and policy intervention: A case study in the state of maryland. *Operations Research for Health Care*, 12:60–70, 2017.

[213] Elham Faham, Ahmad Rezvanfar, Seyed Hamid Movahed Mohammadi, and Meisam Rajabi Nohooji. Using system dynamics to develop education for sustainable development in higher education with the emphasis on the sustainability competencies of students. *Technological Forecasting and Social Change*, 123:307–326, 2017.

[214] Wei Fang, Haizhong An, Huajiao Li, Xiangyun Gao, Xiaoqi Sun, and Weiqiong Zhong. Accessing on the sustainability of urban ecological-economic systems by means of a coupled emergy and system dynamics model: A case study of beijing. *Energy Policy*, 100:326–337, 2017.

[215] Verônica Ghisolfi, Gisele de Lorena Diniz Chaves, Renato Ribeiro Siman, and Lucia Helena Xavier. System dynamics applied to closed loop supply chains of desktops and laptops in brazil: A perspective for social inclusion of waste pickers. *Waste Management*, 60:14–31, 2017.

[216] Erika K Palmer. Structural disadvantage: Evidence of gender disparities in the norwegian pension system. *Social Sciences*, 6(1):22, 2017.

[217] Vinícius P Rodrigues, Daniela CA Pigosso, and Tim C McAloone. Simulation-based business case for pss: a system dynamics framework. *Procedia CIRP*, 64:283–288, 2017.

[218] Luhe Wan, Yuwei Zhang, Shaoqun Qi, Haiyan Li, Xiaohong Chen, and Shuying Zang. A study of regional sustainable development based on gis/rs and sd model—case of hadaqi industrial corridor. *Journal of cleaner production*, 142:654–662, 2017.

[219] Widhianthini Widhianthini. A dynamic model for sustainable tourism village planning based on local institutions. *Journal of Regional and City Planning*, 28(1):1–15, 2017.

[220] A Espinoza, S Bautista, PC Narváez, M Alfaro, and M Camargo. Sustainability assessment to support governmental biodiesel policy in colombia: A system dynamics model. *Journal of cleaner production*, 141:1145–1163, 2017.

[221] David T Tan, Yi Gong, and José Gabriel Siri. The impact of subsidies on the prevalence of climate-sensitive residential buildings in malaysia. *Sustainability*, 9(12):2300, 2017.

[222] Yuzhuo Zhang, Xingang Zhao, Yi Zuo, Lingzhi Ren, and Ling Wang. The development of the renewable energy power industry under feed-in tariff and renewable portfolio standard: a case study of china's photovoltaic power industry. *Sustainability*, 9(4):532, 2017.

[223] Jelena Ziemele, Armands Gravelsins, Andra Blumberga, and Dagnija Blumberga. Combining energy efficiency at source and at consumer to reach 4th generation district heating: Economic and system dynamics analysis. *Energy*, 137:595–606, 2017.

[224] Tayyab Ahmad, Ajibade Aibinu, and Muhammad Jamaluddin Thaheem. The effects of high-rise residential construction on sustainability of housing systems. *Procedia engineering*, 180:1695–1704, 2017.

[225] L Cox, A Bassi, J Kolling, A Procter, N Flanders, N Tanners, and R Araujo. Exploring synergies between transit investment and dense redevelopment: A scenario analysis in a rapidly urbanizing landscape. *Landscape and Urban Planning*, 167:429–440, 2017.

[226] Aline R Fernandes and Carlos Arthur B da Silva. Ten years later: A comparison of results of simulation scenarios under a systems dynamic approach and the actual economic performance of small-scale agro-industries supported by brazilian agro-industrial development programs. *International Journal on Food System Dynamics*, 8(2):106–129, 2017.

[227] Apostolos Giannis, Miaoju Chen, Ke Yin, Huanhuan Tong, and Andrei Veksha. Application of system dynamics modeling for evaluation of different recycling scenarios in singapore. *Journal of Material Cycles and Waste Management*, 19(3):1177–1185, 2017.

[228] Andrew Procter, Andrea Bassi, Jenna Kolling, Llael Cox, Nicholas Flanders, Nadav Tanners, and Rochelle Araujo. The effectiveness of light rail transit in achieving regional co 2 emissions targets is linked to building energy use: insights from system dynamics modeling. *Clean Technologies and Environmental Policy*, 19(5):1459–1474, 2017.

[229] David Rees, Janet Stephenson, Debbie Hopkins, and Adam Doering. Exploring stability and change in transport systems: combining delphi and system dynamics approaches. *Transportation*, 44(4):789–805, 2017.

[230] Naoum Tsolakis and Jagjit Singh Srai. A system dynamics approach to food security through smallholder farming in the uk. *Chemical Engineering*, pages 2023–2028, 2017.

[231] Francesca Varia, Giovanni Dara Guccione, Dario Macaluso, and Danilo Marandola. System dynamics model to design effective policy strategies aiming at fostering the adoption of conservation agriculture practices in sicily. *Chemical Engineering*, 58:763–768, 2017.

[232] Jorge A Duran-Encalada, Alberto Paucar-Caceres, ER Bandala, and GH Wright. The impact of global climate change on water quantity and quality: A system dynamics approach to the us–mexican transborder region. *European Journal of Operational Research*, 256(2):567–581, 2017.

[233] Alireza Ghasemi, Bahram Saghafian, and Saeed Golian. System dynamics approach for simulating water resources of an urban water system with emphasis on sustainability of groundwater. *Environmental Earth Sciences*, 76(18):637, 2017.

[234] Nuri Onat, Murat Kucukvar, Anthony Halog, and Scott Cloutier. Systems thinking for life cycle sustainability assessment: A review of recent developments, applications, and future perspectives. *Sustainability*, 9(5):706, 2017.

[235] Yong Zhang, Tingsheng Zhao, Zhengzhu Zhang, Jun Wan, Xiaonan Feng, Xiangmin Liang, and Aijiao Zhou. Modeling and dynamic assessment on sustainable development of drainage enterprise: Application of a coupled system dynamics-comprehensive assessment model. *Journal of cleaner production*, 141:157–167, 2017.

[236] A. Oumer, S.M. Atnaw, J.K. Cheng, and L. Singh. Modeling energy efficiency as a green logistics component in vehicle assembly line. volume 160, 2016.

[237] I. Abdulla, F. Mohamed Arshad, B.K. Bala, N.L. Bach, and S. Mohammadi. Management of beef cattle production in malaysia: A step forward to sustainability. *American Journal of Applied Sciences*, 13(9):976–983, 2016.

[238] S.-H. Chen. The influencing factors of enterprise sustainable innovation: An empirical study. *Sustainability (Switzerland)*, 8(5), 2016.

[239] V. de Gooyert, E. Rouwette, H. van Kranenburg, E. Freeman, and H. van Breen. Sustainability transition dynamics: Towards overcoming policy resistance. *Technological Forecasting and Social Change*, 111:135–145, 2016.

[240] G. Ozcan-Deniz and Y. Zhu. A system dynamics model for construction method selection with sustainability considerations. *Journal of Cleaner Production*, 121:33–44, 2016.

[241] George Papachristos and Emmanuel Adamides. A retroductive systems-based methodology for socio-technical transitions research. *Technological Forecasting and Social Change*, 108:1–14, 2016.

[242] L.M. Safiih, M.N. Afiq, A.R. Naeim, A.M. Ikhwanuddin, H. Madzli, Z. Syerrina, and I. Marzuki. A system dynamics model for analyzing the eco-aquaculture system with policy recommendations: Case study on integrated aquaculture park (i-sharp), setiu terengganu, malaysia. volume 1750, 2016.

[243] R. Sinha, R. Laurenti, J. Singh, M.E. Malmström, and B. Frostell. Identifying ways of closing the metal flow loop in the global mobile phone product system: A system

dynamics modeling approach. *Resources, Conservation and Recycling*, 113:65–76, 2016.

[244] J.P. Walters, D.W. Archer, G.F. Sassenrath, J.R. Hendrickson, J.D. Hanson, J.M. Halloran, P. Vadas, and V.J. Alarcon. Exploring agricultural production systems and their fundamental components with system dynamics modelling. *Ecological Modelling*, 333:51–65, 2016.

[245] S.W. Yoon and S.J. Jeong. Implementing coordinative contracts between manufacturer and retailer in a reverse supply chain. *Sustainability (Switzerland)*, 8(9), 2016.

[246] N.C. Onat, M. Kucukvar, O. Tatari, and G. Egilmez. Integration of system dynamics approach toward deepening and broadening the life cycle sustainability assessment framework: a case for electric vehicles. *International Journal of Life Cycle Assessment*, 21(7):1009–1034, 2016.

[247] B. Cavicchi. Sustainability that backfires: the case of biogas in emilia romagna. *Environmental Innovation and Societal Transitions*, 21:13–27, 2016.

[248] F. Nasiri, F. Mafakheri, D. Adebanjo, and F. Haghighat. Modeling and analysis of renewable heat integration into non-domestic buildings - the case of biomass boilers: A whole life asset-supply chain management approach. *Biomass and Bioenergy*, 95:244–256, 2016.

[249] L. Shipeng and D. Castro-Lacouture. Holistic modeling of microalgae for powering residential communities. volume 88, pages 788–793, 2016.

[250] T. Ahmad, M.J. Thaheem, A. Anwar, and Z.U. Din. Buildings and their integration in communities: Case study of a parking plaza. volume 145, pages 88–95, 2016.

[251] Vanessa Armendáriz, Stefano Armenia, and Alberto Stanislao Atzori. Systemic analysis of food supply and distribution systems in city-region systems—an examination of fao's policy guidelines towards sustainable agri-food systems. *Agriculture*, 6(4):65, 2016.

[252] B.L. Turner, H.M. Menendez, R. Gates, L.O. Tedeschi, and A.S. Atzori. System dynamics modeling for agricultural and natural resource management issues: Review of some past cases and forecasting future roles. *Resources*, 5(4), 2016.

[253] W. Gao, B. Hong, D.P. Swaney, R.W. Howarth, and H. Guo. A system dynamics model for managing regional n inputs from human activities. *Ecological Modelling*, 322:82–91, 2016.

[254] K. Neely and J.P. Walters. Using causal loop diagramming to explore the drivers of the sustained functionality of rural water services in timor-leste. *Sustainability (Switzerland)*, 8(1):1–18, 2016.

[255] I. Nikolaou, K. Evangelinos, and W. Leal Filho. A system dynamic approach for exploring the effects of climate change risks on firms' economic performance. *Journal of Cleaner Production*, 103:499–506, 2015.

[256] S. Bhattacharjee and J. Cruz. Economic sustainability of closed loop supply chains: A holistic model for decision and policy analysis. *Decision Support Systems*, 77:67–86, 2015.

[257] I. Baur and C.R. Binder. Modeling and assessing scenarios of common property pastures management in switzerland. *Ecological Economics*, 119:292–305, 2015.

[258] A. Blumberga, L. Timma, F. Romagnoli, and D. Blumberga. Dynamic modelling of a collection scheme of waste portable batteries for ecological and economic sustainability. *Journal of Cleaner Production*, 88:224–233, 2015.

[259] V. Faghihi, A.R. Hessami, and D.N. Ford. Sustainable campus improvement program design using energy efficiency and conservation. *Journal of Cleaner Production*, 107:400–409, 2015.

[260] E. Fatma. Development of sustainable tuna processing industry using system dynamics simulation. *Procedia Manufacturing*, 4:107–114, 2015.

[261] S.R. Golroudbary and S.M. Zahraee. System dynamics model for optimizing the recycling and collection of waste material in a closed-loop supply chain. *Simulation Modelling Practice and Theory*, 53:88–102, 2015.

[262] I.J. Orji and S. Wei. An innovative integration of fuzzy-logic and systems dynamics in sustainable supplier selection: A case on manufacturing industry. *Computers and Industrial Engineering*, 88:1–12, 2015.

[263] S. Sudarto, K. Takahashi, and K. Morikawa. Efficient flexible long-term capacity planning for optimal sustainable dimensions performance of reverse logistics social responsibility: A system dynamics approach. 2015.

[264] P. Vugteveen, E. Rouwette, H. Stouten, M.M. van Katwijk, and L. Hanssen. Developing social-ecological system indicators using group model building. *Ocean and Coastal Management*, 109:29–39, 2015.

[265] J. Zhang, M. Ji, and Y. Zhang. Tourism sustainability in tibet - forward planning using a systems approach. *Ecological Indicators*, 56:218–228, 2015.

[266] J. Zhou and Y.J. Liu. The method and index of sustainability assessment of infrastructure projects based on system dynamics in china. *Journal of Industrial Engineering and Management*, 8(3):1002–1019, 2015.

[267] M.-R. Yan. Project-based market competition and policy implications for sustainable developments in building and construction sectors. *Sustainability (Switzerland)*, 7(11):15423–15448, 2015.

[268] A. Blumberga, L. Timma, and D. Blumberga. System dynamic model for the accumulation of renewable electricity using power-to-gas and power-to-liquid concepts. *Environmental and Climate Technologies*, 16(1):54–68, 2015.

[269] S. Eker and E. van Daalen. A model-based analysis of biomethane production in the netherlands and the effectiveness of the subsidization policy under uncertainty. *Energy Policy*, 82(1):178–196, 2015.

[270] C.J. Franco, M. Castaneda, and I. Dyner. Simulating the new british electricity-market reform. *European Journal of Operational Research*, 245(1):273–285, 2015.

[271] F.M. Arshad, B.K. Bala, E.F. Alias, and I. Abdulla. Modelling boom and bust of cocoa production systems in malaysia. *Ecological Modelling*, 309-310:22–32, 2015.

[272] Y.-H. Cheng, Y.-H. Chang, and I.J. Lu. Urban transportation energy and carbon dioxide emission reduction strategies. *Applied Energy*, 157:953–973, 2015.

[273] L.S.K. Abadi, A. Shamsai, and H. Goharnejad. An analysis of the sustainability of basin water resources using vensim model. *KSCE Journal of Civil Engineering*, 19(6):1941–1949, 2015.

[274] E. Aivazidou, N. Tsolakis, D. Vlachos, and E. Iakovou. Water footprint management policies for agrifood supply chains: A critical taxonomy and a system dynamics modelling approach. *Chemical Engineering Transactions*, 43:115–120, 2015.

[275] F.J. Alcalá, J. Martínez-Valderrama, P. Robles-Marín, F. Guerrera, M. Martín-Martín, G. Raffaelli, J.T. de León, and L. Asebriy. A hydrological-economic model for sustainable groundwater use in sparse-data drylands: Application to the amtoudi oasis in southern morocco, northern sahara. *Science of the Total Environment*, 537:309–322, 2015.

[276] W. Xiao-jun, Z. Jian-yun, S. Shamsuddin, H. Rui-min, X. Xing-hui, and M. Xin-li. Potential impact of climate change on future water demand in yulin city, northwest china. *Mitigation and Adaptation Strategies for Global Change*, 20(1):1–19, 2015.

[277] B.K. Bala, F.M. Arshad, E.F. Alias, S.F. Sidique, K.M. Noh, M.K. Rowshon, Q.M.M. Islam, and M.M. Islam. Sustainable exploitation of hilsa fish (tenualosa ilisha) population in bangladesh: Modeling and policy implications. *Ecological Modelling*, 283:19–30, 2014.

[278] H. Chen, Y.-C. Chang, and K.-C. Chen. Integrated wetland management: An analysis with group model building based on system dynamics model. *Journal of Environmental Management*, 146:309–319, 2014.

[279] C.G. Choong and A. McKay. Sustainability in the malaysian palm oil industry. *Journal of Cleaner Production*, 85:258–264, 2014.

[280] Wissam El Hachem, Ramy Harik, and Joseph Khoury. Systemic assessment framework of a learning organization's competitive positioning. *Journal of Industrial Engineering and Management (JIEM)*, 7(4):891–918, 2014.

[281] G.F. Mubiena and A. Ma'Ruf. Development of an assessment model for sustainable supply chain management in batik industry. volume 319, 2018.

[282] H.U. Sverdrup, K.V. Ragnarsdottir, and D. Koca. On modelling the global copper mining rates, market supply, copper price and the end of copper reserves. *Resources, Conservation and Recycling*, 87:158–174, 2014.

[283] H. Sverdrup, D. Koca, and K.V. Ragnarsdottir. Investigating the sustainability of the global silver supply, reserves, stocks in society and market price using different approaches. *Resources, Conservation and Recycling*, 83:121–140, 2014.

[284] D.J. Van Dyk and L. Pretorius. A systems thinking approach to the sustainability of quality improvement programmes. *South African Journal of Industrial Engineering*, 25(1):71–84, 2014.

[285] Y. Wang, H. Yi, and M. Fang. Developing a sustainability performance assessment tool for public funded projects according to policies and stakeholders' perceptions. *Open Construction and Building Technology Journal*, 8(1):52–62, 2014.

[286] J. Xu, Y. Deng, and L. Yao. Sustainable development-oriented industrial restructuring modeling and analysis: A case study in leshan. *Clean Technologies and Environmental Policy*, 16(2):267–279, 2014.

[287] X. Zhang, Y. Wu, L. Shen, and M. Skitmore. A prototype system dynamic model for assessing the sustainability of construction projects. *International Journal of Project Management*, 32(1):66–76, 2014.

[288] G. Papachristos. Transition inertia due to competition in supply chains with remanufacturing and recycling: A systems dynamics model. *Environmental Innovation and Societal Transitions*, 12:47–65, 2014.

[289] M.-C. Chen and K. Chang. Reasoning the causality of city sprawl, traffic congestion, and green land disappearance in taiwan using the cld model. *International journal of environmental research and public health*, 11(11):11464–11480, 2014.

[290] N.P. Bueno. Misperceptions of feedbacks and the resilience of common-pool resource systems: A discussion for irrigation systems based on loop dominance analysis. *International Journal of the Commons*, 8(1):79–106, 2014.

[291] S. Dawadi and S. Ahmad. Evaluating the impact of demand-side management on water resources under changing climatic conditions and increasing population. *Journal of Environmental Management*, 114:261–275, 2013.

[292] M. Faezipour and S. Ferreira. Assessing water sustainability related to hospitals using system dynamics modeling. volume 36, pages 27–32, 2014.

[293] R. Rehan, M.A. Knight, A.J.A. Unger, and C.T. Haas. Financially sustainable management strategies for urban wastewater collection infrastructure - development of a system dynamics model. *Tunnelling and Underground Space Technology*, 39:116–129, 2014.

[294] Marshall Duer-Balkind, Kasey R Jacobs, Burak Güneralp, and Xavier Basurto. Resilience, social-ecological rules, and environmental variability in a two-species artisanal fishery. *Ecology and Society*, 18(4):50, 2013.

[295] P. Georgiadis. An integrated system dynamics model for strategic capacity planning in closed-loop recycling networks: A dynamic analysis for the paper industry. *Simulation Modelling Practice and Theory*, 32:116–137, 2013.

[296] S.H. Moosavirad, S. Kara, and M.Z. Hauschild. Long term impacts of international outsourcing of manufacturing on sustainability. *CIRP Annals - Manufacturing Technology*, 63(1):41–44, 2014.

[297] Y.Y. Feng, S.Q. Chen, and L.X. Zhang. System dynamics modeling for urban energy consumption and co2 emissions: A case study of beijing, china. *Ecological Modelling*, 252(1):44–52, 2013.

[298] H. Qudrat-Ullah. Understanding the dynamics of electricity generation capacity incanada: A system dynamics approach. *Energy*, 59:285–294, 2013.

[299] O. Sahin and S. Mohamed. A spatial temporal decision framework for adaptation to sea level rise. *Environmental Modelling and Software*, 46:129–141, 2013.

[300] K. Qaiser, S. Ahmad, W. Johnson, and J.R. Batista. Evaluating water conservation and reuse policies using a dynamic water balance model. *Environmental Management*, 51(2):449–458, 2013.

[301] R Rehan, MA Knight, AJA Unger, and C Tl Haas. Development of a system dynamics model for financially sustainable management of municipal watermain networks. *Water research*, 47(20):7184–7205, 2013.

[302] X. Xi and K.L. Poh. Using system dynamics for sustainable water resources management in singapore. volume 16, pages 157–166, 2013.