PhD Dissertation

ANALYSIS AND SYNTHESIS OF PROCESSES WITH UNCERTAIN PARAMETERS

DOI:10.18136/PE.2020.751

Éva König

Supervisor:
Dr. Botond Bertók

University of Pannonia
Faculty of Information Technology
Doctoral School Information Science

2020

ANALYSIS AND SYNTHESIS OF PROCESSES WITH UNCERTAIN PARAMETERS

Thesis for obtaining a PhD degree in the Doctoral School of Information Science of the University of Pannonia

in the branch of Information Sciences

Written by Éva König

Supervisor: Dr. Botond Bertók

propose acceptance (yes / no) ……………………….

(supervisor/s)

The PhD-candidate has achieved .... % in the comprehensive exam,
Veszprém, ……………………….

(Chairman of the Examination Committee)

As reviewer, I propose acceptance of the thesis:

Name of Reviewer: …...................... …................. yes / no

……………………….

(reviewer)

Name of Reviewer: …...................... …................. yes / no

……………………….

(reviewer)

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

Veszprém, ……………………….

(Chairman of the Committee)

The grade of the PhD Diploma …..................................... (…….. %)
Veszprém,

……………………….

(Chairman of UDHC)

Contents

Acknowledgement

First of all, I would like to thank my supervisor Dr. Botond Bertók for leading my research all along in the past 8 years by his knowledge and ideas and by introducing me to highly qualified scientists. I am also grateful to Professor Heriberto Cabezas for introducing me to the field of Information Theory and for leading our brand new research regarding systems' resilience. Without their knowledge, ideas and guidance I would have never been capable of accomplishing this work.

I thank Dr. Csaba Fábián for giving me strong base knowledge in stochastic programming and stochastic optimization.

I highly appreciate the support of Audrey L. Mayer in the systems' resilience research, she has provided crucial information regarding the Isle Royale National Park's ecosystem and regarding ecosystems in general.

I would like to thank Dr Zoltán Süle for his great help in the P-graph related approaches.

I am grateful to my former and present colleagues who are always there for helping me.

Finally and most importantly, I would like to thank my family and friends for continuously believing in me, celebrating smaller milestones with me and helping me getting up after every defeat and getting through hopeless situations.

Abstract

One of the key components in industry and market supply is the freight transportation, besides the expanding market relies on the efficiency of shipments even more. In transportation, uncertainty cannot be neglected, e.g., traffic or navigability of rivers, seas, or weather conditions at airports and their stochastic nature can affect contracts between a firm and a transport company. Analysis of contracts' conditions is essential, whether they have short or medium term validity, e.g., the first costs less while the latter can offer more flexibility. Medium term scenario analysis can be formulated as a two-stage decision problem, which are typically managed by methods applying decision trees. Since realistic problems are complex enough to result in a decision tree with enormous size, application of these kind of methods is limited in practice.

In this thesis there is a computer aided algorithmic method presented based on P-graph framework, which is able to implicitly involve and enumerate all feasible scenarios instead of explicitly enumerate the possibilities, and the problem formulation is still kept compact and transparent. For supporting long term decisions regarding complex systems, the direct calculation of the resilience of a system's current regime has the highest importance. Concerning ecosystems, resilience (the system's resistance to disturbances) is a key concern for managing human impacts on them and managing their risk of collapse. Approaches applying statistics or information theory have confirmed utility in identifying regime boundaries. In this thesis, Fisher information is used to establish the limits of the resilience of a dynamic regime of a predator-prey system. The importance of this technique lays in the focus of the approach. Previous studies using Fisher information focused on detecting whether a regime change has occurred, whereas here the attention is on determining how much an ecological system can vary its properties without a regime change occurring. The theory and the method are illustrated with simple two species systems; first it is applied to a predatory-prey model system and then to a 60-year wolf-moose population dataset from Isle Royale National Park in Michigan, USA. The resilience boundaries and the operating range of a system's parameters are assessed without a regime change from entirely new criteria for Fisher information, oriented towards regime stability. The approach provides possibility to use system measurements to determine the shape and depth of the stability "cup" as defined by the broader resilience concept.

Kivonat

Az ipar és a piaci kínálat egyik kulcseleme a teherfuvarozás, emellett a bővülő piac még inkább támaszkodik a szállítmányozás hatékonyságára. A szállítással kapcsolatban nem lehet figyelmen kívül hagyni a bizonytalanságot, például a forgalmat, a folyók, tengerek vagy a repülőterek időjárási körülményeit vagy a hajózhatóságot; ezek sztochasztikus jellege befolyásolhatja a vállalkozás és a szállítmányozó cég közötti szerződéseket. A szerződések feltételeinek elemzése elengedhetetlen, függetlenül attól, hogy azok rövid vagy középtávon érvényesek, például az előbbi költsége alacsonyabb, míg az utóbbi nagyobb rugalmasságot nyújthat. A középtávú forgatókönyv elemzás megfogalmazható kétlépcsős döntési problémaként, amelyet általában döntési fákat alkalmazó módszerekkel közelítenek meg. Mivel a valós problémák elég bonyolultak ahhoz, hogy hatalmas méretű döntési fát eredményezzenek, az ilyen módszerek alkalmazhatósága a gyakorlatban korlátozott. Ebben a dolgozatban számítógépes algoritmikus módszer kerül bemutatásra, amely a P-gráf keretrendszeren alapul, és amely képes implicit módon bevonni és leszámlálni az összes lehetséges forgatókönyvet, ahelyett, hogy explicit módon sorolná fel az összes lehetőséget, amellett, hogy a probléma megfogalmazása továbbra is kompakt és átlátható marad. A komplex rendszerekkel kapcsolatos hosszú távú döntések támogatása szempontjából a legfontosabb a rendszer aktuális állapotának ellenálló képessége és annak közvetlen kiszámítása. Az ökoszisztémák vonatkozásában az ellenálló képesség (a rendszer zavarokkal szembeni ellenálló képessége) kulcsfontosságú az őket érintő emberi hatások és a rendszerösszeomlás kockázatának kezelése szempontjából. A statisztikát vagy az információelméletet alkalmazó megközelítések igazolták a rezsimhatárok azonosításának fontosságát. Ebben a dolgozatban a Fisher információ kerül alkalmazásra a zsákmány-ragadozó dinamikus rendszer ellenálló képesség korlátainak meghatározására. Ennek a technikának jelentősége a megközelítésben rejlik. A Fisher információt alkalmazó korábbi tanulmányok arra irányultak, hogy észleljék a rendszerek közötti váltásokat, míg itt a figyelem annak meghatározására irányul, hogy egy ökológiai rendszer tulajdonságai, paraméterei mennyiben változhatnak anélkül, hogy rendszerváltozás következne be. Az elméletet és a módszert egyszerű, két fajból álló rendszer szemlélteti; először egy ragadozó-zsákmány modell rendszer, majd egy 60 éves farkas-jávorszarvas populáció-adatállomány, az Egyesült Államok Michigan állambeli Isle Royale Nemzeti Parkból. Az ellenálló képesség határait és a rendszer paramétereinek működési tartományát a módszer úgy képes megbecsülni a Fisher-információ teljesen új, adatokra vonatkozó kritériumai alapján, hogy a rendszer közben nem változik meg, azaz a rendszer stabilitására koncentrál. Ezzel a megközelítéssel lehetővé válik az ellenállóképesség tágabb megfogalmazása szerinti "csésze" mélységének és szélességének meghatározása a rendszer mérőszámainak felhasználásával.

Abstracto

Uno de los componentes clave en la industria y el suministro del mercado es el transporte de carga, además el mercado en expansión depende aún más de la eficiencia de los envíos. En el transporte, no se puede descuidar la incertidumbre. Por ejemplo, el tráfico o la navegabilidad de ríos, mares o las condiciones climáticas en los aeropuertos y su naturaleza estocástica puede afectar los contratos entre una empresa y una empresa de transporte. El análisis de las condiciones de los contratos es esencial, ya sea que tengan validez a corto o mediano plazo. Por ejemplo, el primero cuesta menos mientras que el segundo puede ofrecer más flexibilidad. El análisis de escenarios a mediano plazo se puede formular como un problema de decisión en dos etapas, que generalmente se manejan mediante métodos que aplican árboles de decisiónes. Dado que los problemas realistas son lo suficientemente complejos como para generar un árbol de decisiónes con un tamaño enorme, la aplicación de este tipo de métodos es limitada en la práctica. En esta tesis sepresenta un método algorítmico asistido por computadora basado en el marco de P-graph. Este es capaz de involucrar y enumerar implícitamente todos los escenarios factibles en lugar de enumerar explícitamente las posibilidades, y la formulación del problema aún se mantiene compacta y transparente. Para soportar decisiones a largo plazo con respecto a sistemas complejos, el cálculo directo de la resistencia del régimen actual de un sistema tiene la mayor importancia. Con respecto a los ecosistemas, la resiliencia (la resistencia del sistema a las perturbaciones) es una interés clave para manejar los impactos humanos sobre ellos y manejar su riesgo de colapso. Los enfoques que aplican estadísticas o teoría de la información han confirmado su utilidad para identificar los límites del régimen. En esta tesis, la información de Fisher se utiliza para establecer los límites de la resistencia de un régimen dinámico de un ecosistema predador-presa. La importancia de esta técnica radica en el enfoque del método. Estudios previos que utilizaron información de Fisher se centraron en detectar si se produjo un cambio de régimen, mientras que aquí la atención está en determinar cuánto puede variar un sistema ecológico sus propiedades sin que ocurra un cambio de régimen. La teoría y el método se ilustran con sistemas simples de dos especies; primero se aplica a un sistema modelo de presas-predadoras y luego a un conjunto de datos de población de lobos y alces de 60 años del Parque Nacional Isle Royale en Michigan, EE.UU. Los límites de resiliencia y el rango operativo de los parámetros de un sistema se evalúan sin un cambio de régimen de criterios completamente nuevos para la información de Fisher, orientados hacia la estabilidad del régimen. El método brinda la posibilidad de utilizar mediciones del sistema para determinar la forma y la profundidad de la "copa" de estabilidad tal como se define en el concepto de resiliencia más amplio.

Chapter 1

Introduction

There are several ways to support decision making depending on the decision maker, the type of decision, the system itself that is affected by the certain decision, or the time frame that specifies the decision. In this thesis, the focus is on two different methods that are able to support decisions practically with strong mathematical bases. The first presented method is a multi-stage optimization approach considering uncertainty assisting short- and midterm decisions for companies. The second technique presented in this work is much more relevant for long-term maintenance of huge and complex systems by using the Fisher information and providing the possibility to the direct calculation of the system's resilience, the ability of the system to persist within a certain regime in the presence of disturbances. In other words, the first method is for optimal design of the structure of complex systems, while the other approach is able to define the dynamic systems' limits and boundaries within the system can remain in a certain regime.

The thesis is written in passive voice but the new results are highlighted as I have actively worked on them all along my PhD studies.

In the following sections, the aim of this thesis is determined, then comes the brief overview of the background and relevant literature for both above introduced methods for decision support. In the succeeding chapters, the new techniques are described in detail, first the multi-stage stochastic optimization method and then the approach for calculating systems' resilience using Fisher information. Each method is illustrated by a representative example. After summarizing the presented work, the new scientific results and the related publications are highlighted in separate chapters.

1.1 Aim

Optimal processes are essential in all sectors of business and industry so that a company can stay competitive and efficient in the market. However, uncertainty cannot be neglected when speaking of making optimal decisions. Several robust and reliable process network optimization algorithms and software have been developed and implemented on the basis of the P-graph framework in the last three decades, e.g., Algorithm SSG [33], Algorithm ABB [34], Software PNS-Studio [9]. The approach based on the P-graph framework is capable of generating mathematical model automatically, and providing the n-best networks for process synthesis. All the steps involved are mathematically proven, such as comprehensive superstructure generation, mathematical model construction, optimization and the solution interpretation. An optimization problem with uncertain parameters can be solved by the P-graph framework in several ways depending on which parameters are uncertain or how the unpredictability is taken into consideration.

If the uncertainty is considered only in the available flow of resources, there is a P-graph based technique that is able to provide the optimal structure with minimal cost and expected reliability. For this method, the original algorithms have been extended to consider the reliability of the raw materials' availability and to guarantee the predefined level of reliability for the overall process design. Another P-graph based approach is capable of identifying the least sensitive among all the feasible solution structures, if the cost or the available flow of the resources, the activity cost, and the required flow of the product can be stochastic. This methodology determines the optimal structure with the initial parameter set then recalculates the best solution for a large set of possible parameters with uniform distribution and proposes the structure most often identified as optimal. The most complex P-graph based technique for managing uncertain parameters where the structure is separated into two stages. Decisions regarding the investments (fix costs) are represented in the first stage and decisions about the operations (proportional costs) in the second stage, where different scenarios can be considered [58].

In the following chapter of this thesis, the focus is on the third method of the above mentioned ones, where the aim is to find the structure with the most promising expected behavior. In process network synthesis (PNS), there are two major classes of decisions, one is about investments and another one about the operation. Various modes of operating units for complex structures can be investigated. If there is a failure of some operating units in the structure, the optimization remains possible. For the calculation of the expected behavior, each potential scenario has to be considered in order to evaluate possible investments. All the cost parameters of the operating units are sorted out from the basic (i.e., single stage) structure in order to get solutions for different scenarios. In the first stage, all the major decisions are made, e.g., investments. In the second stage volumes of the activities are determined according to the actual situations, i.e., scenarios. Consequently, the first stage has effect on investment costs while the second stage on the operational costs. The scenarios are weighted according to the probabilities of their occurrence. An example without details only as a general impression can be seen in Figure 1.1, the detailed description of this method is presented in the following chapter via a real-life transportation problem of bio-fuels.

The resilience of the systems is a fundamental concept for large and complex systems where long-continued reliable operation has the highest importance. Modeling and assessing the resilience of systems, which is in nature complex and large-scale, has raised remarkable interest among both practitioners and researchers in the past decade. Due to this recent popularity of the topic, several definitions and numerous approaches appeared regarding the concept of resilience and the measurement of it. In this work, the resilience is defined as the system's ability to

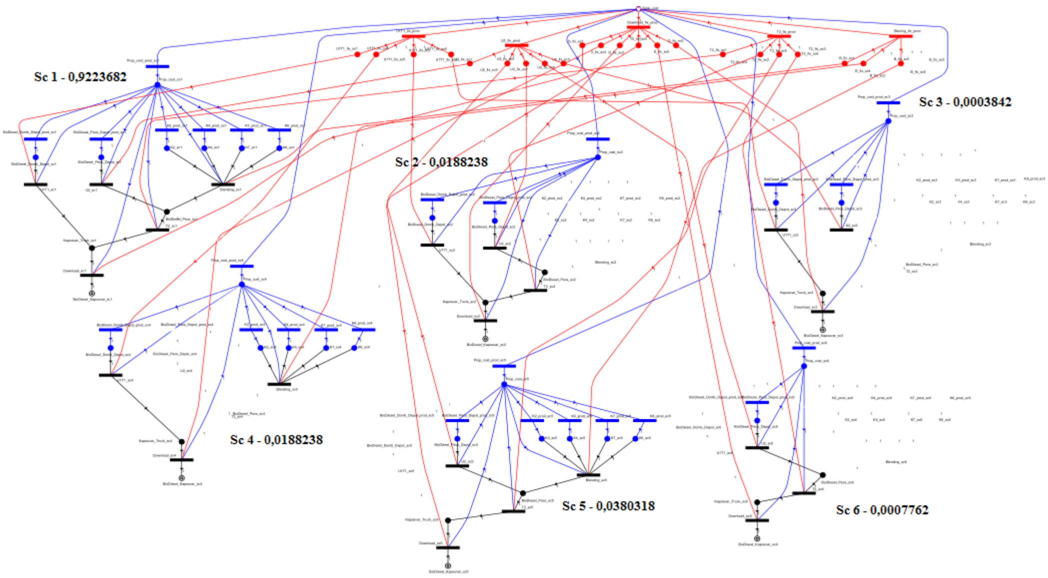

Figure 1.1: Maximal structure of two-stage model with 3 unreliable operating units and 6 possible scenarios

remain within a certain regime in the presence of disturbances. It determines how and to what magnitude systems will change in response to these perturbations ([49], [39], [16], [22], [35]).

The human-nature relationship gets probably the greatest attention from natural scientists these days, therefore ecosystems are of high priority among large-scale and complex systems ([95] [94]). The direct measurement of the resilience of an ecosystem and identification of its thresholds remains a key concern for managing human impacts on these ecosystems and the risk of their brake down. There are numerous approaches utilizing statistics or information theory that demonstrate some utility to identifying these thresholds or transition zones between one dynamic regime and another. In this thesis, Fisher information is used to measure the size of the dynamic regime existing between thresholds of different regimes. This approach has been first developed on a simplistic predatory-prey model, and then applied to the 60-year wolf-moose population dataset from Isle Royale National Park in Michigan, USA. The developed method makes it possible to calculate where a stable system has its bounds, and what the ranges of the interacting parameters are where the system keeps its stable regime independently of the perturbations. This last point has high importance since perturbations are difficult to foresee. This approach can be applied in its present form to larger, more complicated systems as well. Hence, Fisher information demonstrates an early promise to directly measure the resilience of a dynamic regime.

The aim of the third chapter of this thesis is to demonstrate the above mentioned two methodologies in details by two illustrative examples, which are complex enough to highlight the advantages and main features of the methods but simple enough to make it easy to understand these two techniques.

1.2 General introduction to the supply chain optimization methods under uncertainties

There are numerous examples where optimization methodologies and decomposition techniques were applied together. An accelerated Benders' decomposition with a sampling strategy is presented in Santoso et al. (2005) to design supply chain networks with uncertain parameters [87]. Bidhandi and Yusuff (2011) present an approach, where again an accelerated Benders' decomposition method is involved, it is integrated into a mixed-integer linear programming (MILP) solution phase to solve a two-stage stochastic supply chain network design model where the two stages correspond to the strategic and the tactical decisions [11]. A stochastic two-stage Branch and Fix Coordination algorithmic approach has been developed to manage supply chains by determining the production topology, plant sizing, product selection, product allocation among plants and vendor selection for raw materials [2]. The goal here was the expected profit maximization over time subtracting the investment depreciation and the operational costs. Uncertainty appears in numerous properties, the net price and demand of the product, the raw material supply cost and the production cost.

It is common to define more than one objectives when optimizing supply chains. In the work of Sabri and Beamon (2000), the optimization objectives include cost, customer service levels, and flexibility. This supply chain model is for simultaneous strategic and operational planning, where the uncertainty is in the demand [86]. Three complex objectives are defined in Azaron et al. (2008), i.e., minimization of the sum of current investment and the expected costs of processing, transportation, shortage and capacity expansion; minimization of the variaty of the total cost; and minimization of the financial risk in other words the probability of not meeting a certain budget [6]. It is a stochastic model where the uncertainty appears in the demands, supplies, processing, transportation, shortage and capacity expansion costs. A novel method has been presented in Goh et al. (2007) applying the Moreau-Yosida regularization considering two objectives, maximal profit with minimal risk. The approach is applied to a multi-stage global supply chain network problem [38]. A little bit different, an integrated model has been developed in order to optimize logistics and production costs associated with the supply chain members. The demand is uncertain and the manufacturing setting is flexible. Binary decision variables select companies to form the supply chain and continuous decision variables determine volumes of the production flows. It is a robust optimization model with three objectives, minimal expected total cost, minimal cost variability due to demand uncertainty and minimal expected penalty for demand unmet at the end of the planning horizon [81]. Marufuzzaman et al. (2014) developed a two-stage stochastic programming model for designing and managing biodiesel supply chains. The model has two objectives minimizing the cost together with the emission of the supply chain. The proposed technique is an extension of a MILP and the classical two-stage stochastic location-transportation model [69].

There are so many different techniques that are able to consider uncertain parameters. The most likely source of uncertainty is the stochastic demand. A two-stage, stochastic programming approach for planning multisite midterm supply chains under demand uncertainty is presented in the works of Gupta and Maranas (2000 and 2003). Decisions about the production are made "here-and-now" prior to the appearance of the uncertainty; and the supply-chain decisions are in a "wait-and-see" mode [41] and [42]. There is another extended stochastic LP model to take demand uncertainty and cash flow into consideration for medium term [96]; and a MILP model that integrates financial consideration with supply chain design decisions by uncertain demands [66].

The source of uncertainty can be altered for example as it is in Chen and Lee (2004) where the sales prices are uncertain [18]. It is a multi-product, multi-stage and multi-period production and distribution model to reach the maximal total profit of the whole network. The environment can be stochastic as well like in the work of Leung et al. (2006), which presents a stochastic programming approach to optimize medium-term production loading plans [64].

The sources of uncertainty also can be multiple, there are several examples in the literature. A two-stage stochastic model has been built up to analyze the strategic planning of an oil supply chain [15]. It is a scenario-based approach with three sources of uncertainty namely, oil supply, demand of the final product and the prices of the oil and the product. The goal here is to maximize the expected net present value. Significant differences appeared in the results, which demonstrates that considering uncertainties is a fundamental step in decision-making processes. Another two-stage mixed integer stochastic approach is presented in Kim et al. (2011) where the objective is to maximize the expected profit of a biofuel supply chain by several sources of uncertainty. The first stage decisions are about the capital investments including the size and location of the processing plants, while the flows of the biomass and product in each scenario are decided in the second stage. The model is formulated and implemented in GAMS [55]. A hybrid robust-stochastic approach is introduced in [54], where the focus is on profit maximization for closed-loop supply chain networks considering uncertainty in the transport, in the demands and returns. The solution method is based on a stochastic accelerated Benders' decomposition.

Related to supply chain networks, there are tremendous other aspects and approaches developed. For example, a MILP optimization problem has been built up to design multiproduct and multi-echelon supply chain network where the network consists of a number of manufacturing sites and a number of costumer zones at fixed locations and a number of warehouses and distribution centers of unknown locations (selected from a potential location set). The objective is the minimization of the total annual cost of the network and decisions are made to determine the number, location and capacity of warehouses and distribution centers, the transportation links, as well as the flows and production rates of materials [106]. A multi-criteria genetic algorithm has been applied to a distribution problem among a number of sources and a number of destinations. The method combines analytic hierarchy processes with genetic algorithms and there is the possibility to give weights for criteria using pairwise comparison approach [17]. Another warehouse location problem has been solved considering the variability of the demand is the only uncertain parameter [1]. A robust network design model has been developed to optimize location-allocation problem by the minimal overall cost [52]. In the work of Bertsimas and Youssef (2019), a novel robust optimization approach is detailed that is to analyze and optimize the expected performance of supply chain networks considering uncertainty in the demand ([10]).

Parallelly with the development of new technologies and methods, numerous simulation tools evolved in order to analyze supply chains. Petrovic (2001) details simulation tool SCSIM applicable to study supply chain behavior and performance if the costumer demand, external supply of raw materials and lead times to the facilities are uncertain [82]. An iterative hybrid analytic and simulation model has been developed in order to solve the integrated production-distribution problem in supply chain management, where operation time is considered as a dynamic factor in the work of Lee and Kim (2002) [63].

Understanding the contractual forms and their economic implications is a crucial part of supply chain performance evaluation. These contracts define the independent parties coordinating the whole supply chain and answers the questions who controls what decisions and how parties would be compensated [62].

More then 85% of the world energy needs is covered by fossil-based fuels that are finite and unsustainable [12]. The importance of environmental protection and the tightness of its regularization is increasing therefore in the latest decades new alternative sources for fossil-based fuel have been widely studied [59]. One of the alternative to petroleum-based diesel fuel is the biodiesel that is considered as a renewable and natural energy source since it is made of vegetable oils and animal fats. It is a cleaner-burning diesel replacement fuel that operates in compression-ignition engines or Diesel engines and has very similar physical properties to conventional diesel fuel [24]. Besides, in the beginning of the 21st century, the traditional supply chain network of procurement, production, distribution and sales was extended to the whole lifecycle of the product by the business processes [27]. Currently, logistics and supply chain management are regarded as critical business concerns and if they are optimal, they can provide huge advantage in the competition among businesses [20]. Numerous methods and techniques in the latest decades have been developed to tackle the problem designing supply chain networks or identifying and handling the uncertainties of such systems. Researchers have viewed this issue from several aspects and have restricted the field to many specific applications and case studies.

An accelerated stochastic Benders' decomposition technique has been developed for planning the investments of petroleum products supply chain represented by a stochastic two-stage model [79]. Another stochastic planning model for a biofuel supply chain under demand and price uncertainties is presented in Awudu and Zhang (2013) [5]. It is a stochastic LP model for maximizing the expected profit where the products' demands are uncertain but with known distribution. The applied technique comprises Benders' decomposition and Monte Carlo simulation. For strategic planning of bioenergy supply chain systems and optimal feedstock resource allocation under supply and demand uncertainties stochastic MILP models have been applied e.g., a two-stage model developed with a Lagrange relaxation based decomposition algorithm [19]. Awudu and Zhang (2012) presented the general structure of the biofuel supply chain with three type of decisions – strategic, tactical and operational [4]. The supposed sources of uncertainty are the biomass supply, transportation, production and operation, demand and prices. They studied different modelling techniques, like analytical and simulation methods with respect to sustainability considering environmental, economic and social aspects. Another related research is presented in Gebreslassie et al. (2012) where a multiperiod, bicreterion stochastic MILP model has been developed to design optimal hydrocarbon biorefinery supply chains where the demand and supply are uncertain [36]. A two-stage stochastic model has been built up to achieve maximal expected profit in a bioethanol supply chain under jointly appearing uncertainties, such as switchgrass yield, crop residue purchase price, bioethanol demand and sales price [80]. Shabani and Sowlati (2016) introduced a hybrid multi-stage stochastic programming robust optimization model to simultaneously include uncertainty in biomass quality and biomass availability [93].

Since the beginning of the 21st century, the importance of thinking "green" and therefore the significance of green supply chains has been increasing. Mirzapour Al-e-hashem et al. (2013) have developed a stochastic programming approach for a multi-period multi-product multi-site aggregate production problem in a green supply chain where uncertainty appears in the demand. Their model is a MILP converted into an LP by applying some theoretical and numerical techniques [75]. Another two-stage stochastic approach has been built up in order to design green supply chains considering carbon trading environment. The uncertainty lays in the product demand and the carbon price [84].

A dynamic, spatially explicit and multi-echelon MILP modelling framework is detailed in the work of Dal-Mas et al. (2011) to help assessing economic performances risk on investment of the entire biomass-based ethanol supply chain [23]. A multi-period and multi-echelon MILP model has been developed to design and plan bioethanol upstream supply chain considering that the market is uncertain. The approach has an economic value to the overall GHG emission implemented through an emissions allowances trading scheme [37]. A slightly different approach has been built up to define the set of all Pareto-optimal configurations of the supply chain simultaneously taking into consideration the efficiency and the risk. The latter is measured by the standard deviation of the efficiency. The approach is an extended branch-and-reduce algorithm that applies optimality cuts and upper bounds to eliminate parts of the infeasible region and the non-Pareto-optimal region [51]. A similar approach is introduced in Bernstein and Federgruen (2005) where a two-echelon supply chain model is presented with a single supplier servicing a network of retailers [7]. Retailers face uncertain (random) demands and the distribution may depend only on each the retailer's own price (noncompeting) or on its own price as well as those of the other retailers (competing).

An example is presented in Tan & Aviso (2016) that is closely related to method to be presented herein [101]. It proposes an extension and generalization of the multi-period P-graph framework [48]. It suggests that the multi-period approach may be applied to robust network synthesis involving multiple scenarios instead of time periods.

1.3 General introduction to resilience in ecosystems

Internal and external drivers, like climate change, human activity, species extinction and several other causes constantly interact with dynamic ecosystems.([107], [98], [92]). The resilience of an ecosystem, as defined by the system's ability to remain within a particular regime in the presence of disturbances, determines how and to what magnitude ecosystems will change in response to these drivers ([49], [39], [16], [22], [35]). It is essential to understand the mechanisms of ecological resilience to natural and anthropogenic disturbances if the vulnerability of systems to regime-changing disturbances is to be measured ([109], [67], [99], [98], [65]) and managed.

The movement of a system from one regime (or alternative stable state) to another is called regime change, and can be triggered by either exogenous disturbances (such as fire or the introduction of disease), or internal causes (e.g., loss of species, increased mortality, etc.; [97]). The system's resilience to that certain disturbance determines the likelihood of regime change; in other words, its ability to maintain itself in that regime through internal feedbacks and interactions ([89], [29]). Note that in this work, the focus is on one regime as the measure of resilience, and not multiple regimes or the recovery of a system to a previous regime after disturbance (where recovery time is an alternative measure of resilience; see Grimm and Wissel (1997). The identification of the location of regime boundaries, also known as thresholds or tipping points, is of critical importance as early warning systems for the management and sustainability of coupled human-environment systems ([43], [90], [88], [50], [97], [99]).

Holling (1973) adopted a quantitative view of the behavior of ecological systems. Since then perspectives on ecosystem resilience have been expanded and refined to explicitly consider non-linear dynamics, boundaries, uncertainty and unpredictability, and how such dynamics interact across different time and spatial scales ([16], [28], [13], [88], [109], [91]). Generally, resilience may be estimated by computing the eigenvalues of the system at its equilibrium ([60]), but this approach does not provide any information about the behavior of a system close to its limits, right before the patterns decay.

Neubert and Caswell (1997) investigated several measures of a transient response, such as the biggest proportional deviation that can be generated by any perturbations, the maximal possible growth rate that directly follows the perturbation, and the time at which the amplification occurs. Scheffer et al. (2015) presented methods based on the critical slowing down phenomena, which implies that recovery upon small perturbations becomes slower as a system approaches a regime threshold. In their research they also characterized the resilience of alternative regimes in probabilistic terms, measuring critical slowing down by using generic indicators related to the fundamental properties of a dynamic system ([91]). Levine et al. (2016) studied Amazon forests and reported contradictory predictions in the sensitivity and ecological resilience of them to changes in climate, sometimes resulting in biomass stability, other times in catastrophic biomass loss; transitions between regimes was continuous (no thresholds observed). Other drivers are also able to amplify climate change-driven transitions between forests and savanna globally, e.g. fire disturbances, grazing, logging or other anthropogenic activities ([70]). The key to the identification of these ecosystem transitions is the availability of long-term data, which is expensive and resource-intensive.

Information Theory has been applied to assess the sustainability of dynamic systems ([85], [25]), mainly to detect transitions from one dynamic regime to another ([71]; [53], [97], [26], [100], [108]). The "ball and cup" mental model has been central to this work ([40]). As common analogy for dynamic regimes, the ball, representing a system that moves within a cup, representing a specific regime. The ability of the ball to remain in that same cup (or basin of attraction) means the resilience of the system ([39]). To functionally relate resilience to regimes and regime change, two things must be determined 1) how large the cup is (regime resilience), and 2) whether the system is in the cup or outside of it (regime shift). In this work, Fisher information is applied to identify the boundaries of the regime (the size and depth of the cup) relative to the position of the ecological system (the ball) from actual values of system variables. It moves the state of the science beyond discussing symbolic cups meant to represent basins of attraction to working with the actual basin of attraction for the system, which is primary importance of this work. Unlike in prior studies (e.g., [100]), where boundaries were identified post-regime shift, it is possible to identify regime boundaries before the system has a regime change as it is demonstrated in this work. This is important because knowing the size and shape of the basin of attraction provides the opportunity to take remedial action to keep the system away from the regime boundaries before a shift has occurred. (Or, conversely in a restoration attempt, how far a system will need to be pushed in order to flip it into a more desirable regime.) The concept is illustrated with a simple modeled system and with a two-species predator-prey system (the wolves and moose population of Isle Royale National Park, Michigan USA). It is further shown that Fisher information can determine the range of predator-prey abundance over which the ecosystem remains in one regime, and hence exhibits resilience.

1.4 Introduction to Fisher Information Theory

The concept now known as Fisher information was first introduced by the statistician Ronald Fisher (1922) regarding fitting a parameter to data. Starting from the seminal work of Fisher, an expression for computing the Fisher information ([72]) from time series has been developed with the form of,

$$I = \int \frac{1}{p(s)} \left[\frac{dp(s)}{ds} \right]^2 \tag{1.1}$$

where $p(s)$ is now the simple probability density for observing particular values of s, and $dp(s)/ds$ is the slope of $p(s)$.

Fisher information is also closely related to the orderliness of dynamic systems. A very ordered dynamic system is where repetitive observations of the system provide about the same result. When the system has one observable variable s, this means that measuring s repeatedly gives about the same value. In that case, $p(s)$ is very narrow and sharp around the mean of s, and the slope $dp(s)/ds$ is a high value. The Fisher information is proportional to $[dp(s)/ds]^2$, therefore the Fisher information has also a correspondingly high value. In the extreme case of a system where the measurable variables are constant, the system is said to be perfectly orderly, $dp(s)/ds \to +\infty$, and the Fisher information is positive infinity. For a very disorderly dynamic system with again one observable variable s, each measurement of s provides a more or less different value. Therefore, $p(s)$ is broad and relatively flat, and the slope $dp(s)/ds$ of $p(s)$ is close to zero. Correspondingly, the Fisher information for a very disorderly dynamic system is near zero. In the extreme situation where a system completely lacking order, each measurement of s yields a different value. Then, $p(s)$ is flat, $dp(s)/ds$ is zero, and the Fisher information for this completely disorderly system is exactly zero. In summary, the Fisher information of an ordered system is high and that of a disordered system is low. One should also note the work of Al-Saffar and Kim (2017) which explored the mathematical behavior of Fisher information under different perturbations and oscillatory regimes with possible implications for small populations of one species.

The aforementioned arguments apply also for systems that have more than one observable variable, but in such case s represents an n-dimensional state of the system which depends on all of the observable variables of the system. Hence, a state of the system s for a dynamic system with n measurable variables $x_1, x_2, \ldots x_n$ is defined by a certain value of each of the n variables. Even two states that differ by the value of only one variable mean different states of the system. Note that this can lead to a very large number of states of the system, each one being unique.

In order to develop a practical and calculable expression for Fisher information, consider that for a sequence of observations of s that have been taken over a time period, there is a one to one correlation between frequency of observations and the time over which they were taken. Hence, $p(s)ds = p(t)dt$ where t is time, and $p(t)$ is the probability density for sampling at a certain time. Now, $T = \int dt$ is the total time over which the observations were made. For a cyclic system, T should generally be at least equal to one cycle, if the aim is to capture changes in system behavior. Sampling at any time point is equally probable, therefore $p(t) = 1/T$. Then, $p(s) = (1/T)/(ds/dt)$ where now ds/dt is the transit speed of the system in s space. Inserting these results into Equation (1.1) the following expression results for Fisher information after some manipulations,

$$I = \frac{1}{T} \int_{t}^{(t+T)} \frac{[R'']^2}{[R']^4} dt' \tag{1.2}$$

where $R' \equiv ds/dt$ is the speed and $R'' \equiv \frac{d^2 s}{dt^2}$ is the acceleration. For the case where $s(x_1, x_2, \ldots x_n)$ depends on n measurable variables, R' and R'' can be calculated from the Euclidean metric in a linear space where the coordinates are again time and the measurable variables $x_1, x_2, \ldots x_n$. This linear space is called the system phase space. Then R' can be calculated from,

$$R' \equiv \frac{ds}{dt} = \sqrt{\sum_{i=1}^{n} \left[\frac{dx_i}{dt}\right]^2} \tag{1.3}$$

and R'' can be calculated from,

$$R'' \equiv \frac{d}{dt}\left[\frac{ds}{dt}\right] = \frac{1}{R'} \sum_{i=1}^{n} \frac{dx_i}{dt} \frac{d^2 x_i}{dt^2} \tag{1.4}$$

where R' and R'' are the speed and acceleration tangential to the path of the system in its phase space.

In Equations (1.2), (1.3), and (1.4) are the practical expressions that can be used to calculate Fisher information. If a differential equation model is available as in the case of the prey-predator system used in this work, the derivatives dx_i/dt and d^2x_i/dt^2 can be computed directly from the model equations. In cases where a differential equation model is not available, the derivatives can be approximated with finite difference methods (see [47]). There are also many cases including this study where computing the integral in Equation (1.2) is not possible analytically, and a numerical approximation is required. For such cases the Fisher information can be approximated from,

$$I = \frac{1}{T} \sum_{t}^{(t+T)} \frac{[R'']^2}{[R']^4} \Delta t \tag{1.5}$$

Chapter 2

Two-stage Optimization by P-graph

In this chapter, a multi-stage optimization technique based on the P-graph Framework is described in detail, and applied to a real-life biodiesel transportation problem. The goal is to find that structure among all the feasible ones that has the most promising expected behavior. The two major classes of decisions are about investments and about the operation. By this approach, various modes of operating units for complex structures can be investigated. The optimization procedure is still possible, even if there is a failure of some operating units in the structure. Every potential scenario has to be regarded for the calculation of the expected behavior, so that the evaluation of possible investments can be done. All the cost parameters of the operating units are sorted out from the single stage structure so the solutions for different scenarios can be achieved. In the first stage, all the major decisions are made, and then, in the second stage, volumes of the activities are determined according to the scenarios. Consequently, the first stage has effect on investment costs while the second stage on the operational costs. The scenarios are weighted according to the probabilities of their occurrence.

The novelty of the approach is described in subsection 2.3.4 regarding the parametric cost modeling and in section 2.4 where the extended model is detailed.

2.1 Illustrative example

The problem, that is illustrating the approach presented hereinafter is demonstrated in this section. The task is to transport biodiesel from two locations — Szazhalombatta, Hungary and Bratislava, Slovakia — to a single destination — Korneuburg, Austria — by two different means of transport —- barge or cargo train [Figure 2.1].

The main difference between the two types of cargo is in their price. Barge transport is cheaper than the rail cargo, on the other hand the uncertainty is much higher in the navigability of the Danube and the availability of the docks than in the accessibility of the rail cargo. There are four possible scenarios considered regarding the above mentioned uncertainties:

- Scenario 1: All the cargo options are available

Figure 2.1: Locations of three cities in the example and the routes of the two different means of transport: the path of barges (blue line) and the track of rail cargo (red line) [https://www.google.hu/maps/]

Table 2.1: The probability of the occurrence of each scenario based on the determining factors and their probabilities

Upper reach	**Lower reach**	**Dock in Bratislava**	**Dock in SzBatta**	**Scenario** $\mathcal{T}_1$	**Scenario** $\mathcal{T}_2$	**Scenario** $\mathcal{T}_3$	**Scenario** $\mathcal{T}_4$
90 %	*90 %*	*80 %*	*85 %*	*Everything is available*	*Barge is available only from SzB*	*Barge is available only from Br*	*Barge is not available*
Y	Y	Y	Y	0.5508			
N	Y	Y	Y				0.0612
Y	N	Y	Y			0.0612	
Y	Y	N	Y		0.1377		
Y	Y	Y	N			0.0972	
N	Y	N	Y				0.0153
N	Y	Y	N				0.0108
Y	N	N	Y				0.0153
Y	N	Y	N			0.0108	
Y	Y	N	N				0.0243
N	N	N	Y				0.0017
N	N	Y	N				0.0012
N	Y	N	N				0.0027
Y	N	N	N				0.0027
N	N	N	N				0.0003
				0.5508	**0.1377**	**0.1692**	**0.1423**

- Scenario 2: Barge is available only from Szazhalombatta
- Scenario 3: Barge is available only from Bratislava
- Scenario 4: Barge is not available at all

Each scenario has an estimated probability where the assessment method is considering four related factors, namely the navigability of the upper reach of the river, the navigability of the lower reach of the river and the availability of the dock in Bratislava and the dock in Szazhalombatta. In the present instant, upper reach corresponds to the river section from Bratislava to Korneuburg and the lower reach corresponds to the river section from Szazhalombatta to Bratislava. It is also important to note, that it is assumed, if the upper reach is unnavigable, then the lower reach is unnavigable as well [Table 2.1]. The overall probability of each scenario is based on the fact that the above described four factors are independent events therefor each situation — each line of the table — has a certain probability value that is the product of the probability of the factors. The last line of the table presents the overall probability of each scenario, which is simply the sum of the corresponding probability values of the certain scenario.

The fix and proportional costs are derived from the distance between the certain repository and the destination [Table 2.2].

Table 2.2: The fix and proportional parts of the transportation costs as a function of distance

City	Distance [km]	Barge fix	Barge prop. [EUR/t]	Cargo fix	Cargo prop. [EUR/t]
SzBatta – Korneuburg	300	900	9	1200	12
Bratislava – Korneuburg	100	300	3	400	4

Table 2.3: The maximal available amount of biodiesel at the source locations

Source location	Maximum available flow [t/yr]
Biodiesel in Szazhalombatta	1800
Biodiesel in Bratislava	1100

There are limits for the maximal available amount of biodiesel at the source locations but each of them individually has a higher limit as the required flow at the destination, in the city of Korneuburg [Table 2.3 and 2.4]. There are also limits on the transportation capacity for all means of transportation but these upper bounds technically do not limit the transportation, since all the maximal capacities are higher than the required flow [Table 2.5].

The goal is to minimize the overall cost of this transportation problem considering the prices, limits, requirements and uncertainties as well.

In the upcoming subsection, the mathematical formulation of this illustrative example is presented; first, the single stage problem then the extended stochastic two-stage problem.

2.1.1 Mathematical formulation of the illustrative example

In this section the mathematical formulation of the previously described illustrative example is introduced. First, the formulation of the single stage problem is presented that is the equivalent of the P-graph representation of the problem depicted in Figure 2.3. Afterwards, the mathematical formulation of the extended problem is presented that is an extension of the single stage problem into a two-stage stochastic programming problem, which is equivalent to the P-graph representation of the extended problem shown in Figure 2.4 a).

It is important to note, that all the bellow applied costs, prices and limits are the same as it is detailed in the previous section.

Table 2.4: The required amount of biodiesel at the destination

Destination	Required flow [t/yr]
Biodiesel in Korneuburg	1000

Table 2.5: The maximal capacity of the transportation means

Means of transport	**Maximum capacity [t]**
Barge from Bratislava	1400
Cargo from Bratislava	2000
Barge from Szazhalombatta	1600
Cargo from Szazhalombatta	2000

Decision variables of the single stage problem

b_{trBr} — the binary value of usage of train from Bratislava
b_{baBr} — the binary value of usage of barge from Bratislava
b_{trSzb} — the binary value of usage of train from Szazhalombatta
b_{baSzb} — the binary value of usage of barge from Szazhalombatta
b_p — the binary value of paying the penalty
x_{trBr} — the amount of biodiesel transported by train from Bratislava
x_{baBr} — the amount of biodiesel transported by barge from Bratislava
x_{trSzb} — the amount of biodiesel transported by train from Szazhalombatta
x_{baSzb} — the amount of biodiesel transported by barge from Szazhalombatta

Objective function of the single stage problem

$$\begin{aligned}&minimize\\&\quad 400b_{trBr} + 300b_{baBr} + 1200b_{trSzb} + 900b_{baSzb} + 6666b_p +\\&\quad 4x_{trBr} + 3x_{baBr} + 12x_{trSzb} + 9x_{baSzb}\end{aligned} \tag{2.1}$$

subject to

$$x_{trBr} + x_{baBr} + x_{trSzb} + x_{baSzb} = 1000 \tag{2.2a}$$
$$x_{trBr} + x_{baBr} \leq 1100 \tag{2.2b}$$
$$x_{trSzb} + x_{baSzb} \leq 1800 \tag{2.2c}$$
$$x_{trBr} \leq 2000 \tag{2.2d}$$
$$x_{baBr} \leq 1400 \tag{2.2e}$$
$$x_{trSzb} \leq 2000 \tag{2.2f}$$
$$x_{baSzb} \leq 1600 \tag{2.2g}$$
$$x_{trBr}, x_{baBr}, x_{trSzb}, x_{baSzb} \geq 0 \tag{2.2h}$$
$$b_{trBr}, b_{baBr}, b_{trSzb}, b_{baSzb}, b_p \in \{0, 1\} \tag{2.2i}$$

This model is equivalent to the one described in subsection 2.3.3, where the P-graph representation is detailed and also illustrated by this example [Figure 2.3]. When uncertainty is taken into consideration the previously defined scenarios and their probability of occurrence [Table 2.1] appear in the mathematical formulation as well. The decision variables are the same but regarding the transported amount of biodiesel there are four of each labeled with 1,2,3,4 linking them to the first, second, third or fourth scenario respectively.

Decision variables of the extended two-stage problem

b_{trBr} — the binary value of usage of train from Bratislava
b_{baBr} — the binary value of usage of barge from Bratislava
b_{trSzb} — the binary value of usage of train from Szazhalombatta
b_{baSzb} — the binary value of usage of barge from Szazhalombatta
b_p — the binary value of paying the penalty
$x_{trBr_{1,2,3,4}}$ — the amount of biodiesel transported by train from Bratislava in Scenario 1, 2, 3 and 4
$x_{baBr_{1,2,3,4}}$ — the amount of biodiesel transported by barge from Bratislava in Scenario 1, 2, 3 and 4
$x_{trSzb_{1,2,3,4}}$ — the amount of biodiesel transported by train from Szazhalombatta in Scenario 1, 2, 3 and 4
$x_{baSzb_{1,2,3,4}}$ — the amount of biodiesel transported by barge from Szazhalombatta in Scenario 1, 2, 3 and 4

Objective function of the extended two-stage problem

$$\begin{aligned}
&minimize \\
&\quad 400b_{trBr} + 300b_{baBr} + 1200b_{trSzb} + 900b_{baSzb} + 6666b_p + \\
&\quad + 0.5508(4x_{trBr_1} + 3x_{baBr_1} + 12x_{trSzb_1} + 9x_{baSzb_1}) + \\
&\quad + 0.1377(4x_{trBr_2} + 3x_{baBr_2} + 12x_{trSzb_2} + 9x_{baSzb_2}) + \\
&\quad + 0.1692(4x_{trBr_3} + 3x_{baBr_3} + 12x_{trSzb_3} + 9x_{baSzb_3}) + \\
&\quad + 0.1423(4x_{trBr_4} + 3x_{baBr_4} + 12x_{trSzb_4} + 9x_{baSzb_4})
\end{aligned} \tag{2.3}$$

subject to

$$x_{trBr_i} + x_{baBr_i} + x_{trSzb_i} + x_{baSzb_i} = 1000 \tag{2.4a}$$
$$x_{trBr_i} + x_{baBr_i} \leq 1100 \tag{2.4b}$$
$$x_{trSzb_i} + x_{baSzb_i} \leq 1800 \tag{2.4c}$$
$$x_{trBr_i} \leq 2000 \tag{2.4d}$$
$$x_{baBr_i} \leq 1400 \tag{2.4e}$$
$$x_{trSzb_i} \leq 2000 \tag{2.4f}$$
$$x_{baSzb_i} \leq 1600 \tag{2.4g}$$
$$x_{trBr_i}, x_{baBr_i}, x_{trSzb_i}, x_{baSzb_i} \geq 0 \tag{2.4h}$$
$$b_{trBr_i}, b_{baBr_i}, b_{trSzb_i}, b_{baSzb_i}, b_{p_i} \in \{0, 1\} \tag{2.4i}$$
$$i \in \{1, 2, 3, 4\} \tag{2.4j}$$

2.2 Decision tree

Decisions are made in different levels in those processes where uncertain parameters and factors play important role. The antecedent decisions restrict the alternatives for the later situations. The complexity of the problem is highly affected by the number of decision levels and by the number of possible scenarios on each level [14]. There are two levels of decisions in this illustrative example, on first level the reservation has to be defined and on second level the decision is about the utilization of the available means of transport. In Figure 2.2, the possible combinations of the first stage decision variables are depicted with grey background (excluding penalty in order to reduce the size of the figure), while the leaves with white background represent the possibilities of second stage decision (only a some of them in order to reduce the size of the figure). Note that, penalty has to be payed if non of the transportation options satisfies the required demand of biodiesel at the end. In the mathematical formulation of the extended two-stage problem, the first stage decisions are equivalent to the binary variables of each transport option and the penalty ($b_{trBr}, b_{baBr}, b_{trSzb}, b_{baSzb}, b_p$), while the second stage decisions are the transported amount of biodiesel by different means of transport in each scenario ($x_{trBr_i}, x_{baBr_i}, x_{trSzb_i}, x_{baSzb_i}$ and $i = \{1, 2, 3, 4\}$).

Decision trees [68] are often applied as the representation of the alternative opportunities in such problems, since these structures are comprehensible. On the other hand, decision trees easily can grow to enormous and impenetrable sizes. The simplified decision tree for the above introduced example is represented in Figure 2.2 and the regarding notations are presented in Table 2.6.

The stochastic problem has 48 825 possible alternatives, in other words there are 48 825

Table 2.6: Notation regarding the simplified decision tree

Notation	Description
Tr-SzB	Rail cargo transport from Szazhalombatta
Ba-SzB	Barge transport from Szazhalombatta
Tr-Br	Rail cargo transport from Bratislava
Ba-Br	Barge transport from Bratislava
P	Penalty

combinatorially different ways to transport biodiesel to Korneuburg applying two different means of transport and having two repositories and considering the 4 different scenarios. It also means that the decision tree has 48 825 leaves on the last stage. Certainly, by applying additional conditions the number of leaves can be reduced. By excluding the redundant transportation routes within each scenario, there are only 625 leaves remain. Even more leaves can be eliminated by excluding the unfeasible solutions. The motivational problem with the previously described parameter settings has only 240 feasible alternative solutions. Although, 240 is still too much to easily evaluate each of them.

In the following sections, a graph theoretic approach by utilizing the process graph or P-graph framework is presented that is capable of automatically generating a transparent and easy-to-use two-stage model with respect to their probability of occurrence and availability of operations in multiple scenarios.

2.3 The P-graph Framework

The P-graph methodology rooted in graph theory has been developed by Friedler, Fan and their coauthors, and initially applied for solving process-network synthesis (PNS) problems in the field of chemical engineering process design, whose complexity is characterized by its combinatorial nature [31].

Unlike input-output models in engineering process design where operating units are represented by nodes and connected to each other through arcs, in P-graph outputs from an operating unit are not directly connected to an input to another operating unit, but instead to an another type of nodes, which are assigned to potential qualities of material streams. Arcs are leading from raw materials or from the nodes denoting qualities of input materials to the nodes representing operating units where they can be utilized and from the nodes of operating units to the nodes depicting their potential qualities of output materials or to products. P-graph unequivocally defines structural alternatives as material-type nodes with multiple incoming and outgoing arcs, i.e., uniquely denoting a material quality which can alternatively produced or consumed by more than one operating unit. Note that, a P-graph, where a material type node has multiple incoming or outgoing arcs, may lead to several input-output models, where outputs from and inputs to operating units that are able to produce or consume materials of the same quality are paired differently, or mixers and splitter are incorporated in the network of operations to collect

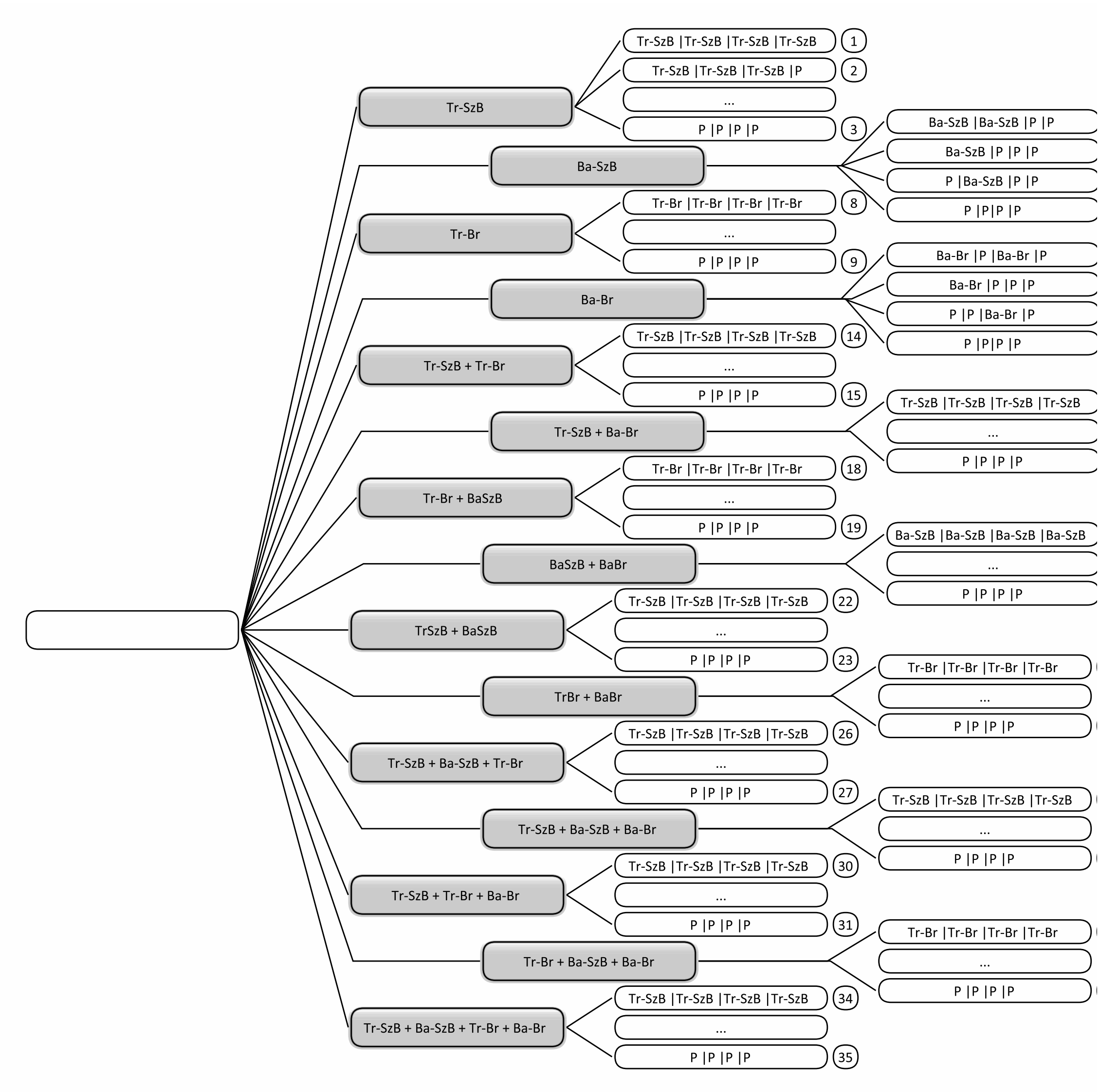

Figure 2.2: The simplified decision tree representation of the motivational example

or share the flows of materials of the same quality; see e.g. [74].

There are three cornerstones constituted by the P-graph framework, namely the graph representation of process networks, the five axioms stating the underlying properties of the combinatorially feasible solutions, and the effective algorithms that are derived from the first two cornerstones. The applied algorithms are for generating the maximal-structure (MSG) [32], the solution-structure generation (SSG) [31] and [30], and finally, the algorithm for determining the optimal structure (ABB) that is based on an accelerated branch-and-bound technique [33].

The P-graph framework has been applied on several fields of process synthesis since the beginning of the 1990's. These fields are optimization, and multiobjective evaluation closely related to problem introduced herein. It was implemented in the field of business process modeling ([21], [104], [46], [105], and [102]), as well as, supply chain modeling ([56] and [61]). The P-graph framework was applied in order to solve crisis management problems ([3] and [103]), energy supply problems [110] and to minimize waste ([44] and [45] and [57]).

In the following subsections the outset, the problem definition, basic notations, and concepts are given formally as a summary of the original papers written by Friedler, Fan, and coauthors.

2.3.1 Problem Definition

A process synthesis or a process-network problem is defined by the available raw materials, potential operating units and desired products. Various properties of the operating units and materials are also given in the problem definition. These properties include the coefficients for the functions expressing the costs of operating units depending on their load, and upper bounds on their respective capacities. It is often practical to specify prices and upper bounds on the availability of raw materials and similarly, lower bounds can be assessed on the desired products, which specifies the minimum quantity to be manufactured from the certain product by the process. In the problem specification, the relations between the materials and operating units are also included, i.e., the consumption rates of input and production rates of output materials by the operating units. The goal is to determine the optimal network where the objective can be either cost minimization or profit maximization [8].

2.3.2 Combinatorial foundation of process synthesis

There are two major steps of the mathematical programming approach to process synthesis, mathematical model generation and then solving the generated model. Both of these steps have combinatorial aspects. The first step should express the existence or absence of links among the candidate operating units; consequently, the generated mathematical model to be solved in the second step contains integer variables. Note that the value of the objective function is often affected more drastically by the integer variables than the continuous variables in the model. Moreover, the number of integer variables, i.e., the combinatorial part of the problem affects the most the computational time. Furthermore, in practice, a process synthesis problem cannot be separated into combinatorial and continuous parts: both should be taken into account simultaneously during the solution process.

Let $\mathcal{M}$ be a given finite nonempty set of all materials that are taken into consideration in the process synthesis. Clearly, the set of required products $\mathcal{P}$ and the set of available raw materials $\mathcal{R}$ are the subsets of $\mathcal{M}$. Operating units are the functional units in a process network performing various operations. Denote the set of operating units that are taken into account by $\mathcal{O}$.

Throughout this chapter of the thesis, the materials are indexed by j, and the operating units, by i. Furthermore, the number of materials, i.e., the cardinality of set $\mathcal{M}$, is denoted by k, while the number of operating units, i.e., the cardinality of set $\mathcal{O}$, is denoted by n.

2.3.3 P-graph representation

The complexity of a process synthesis problem has an exponential relation to the number of candidate operating units, n, due to the $(2^n - 1)$ possible alternative networks among which the optimal network is to be identified. Additional insights are required to eliminate redundant networks.

Each feasible process structure must conform to certain combinatorial properties. The introduction of a unique class of graphs provides the possibility to represent the structures of the process networks unambiguously and to extract these universal combinatorial properties that are inherent in all feasible processes.

Let m and o be two finite sets with

$$\mathsf{o} \subseteq \wp(\mathsf{m}) \times \wp(\mathsf{m}). \tag{2.5}$$

Then, a P-graph (Process graph) is defined to be an (m,o) pair with vertex set $\mathsf{V} = \mathsf{m} \cup \mathsf{o}$ and arc set $\mathsf{A} = \mathsf{A}_1 \cup \mathsf{A}_2$ where

$$\mathsf{A}_1 = \{(x, y) : y = (\alpha, \beta) \in o \text{ and } x \in \alpha\} \tag{2.6}$$

and

$$\mathsf{A}_2 = \{(y, x) : y = (\alpha, \beta) \in o \text{ and } x \in \beta\}. \tag{2.7}$$

P-graphs (Process graphs) are capable of representing process structures. For a $(\mathcal{P}, \mathcal{R}, \mathcal{O})$ synthesis problem, let m be a subset of $\mathcal{M}$, and o be a subset of $\mathcal{O}$. Furthermore, it is assumed that the sets m and o satisfy Eq.((2.5)). Therefore, the structure of the system with set m of materials and set o of operating units is formally defined as P-graph (m,o).

$$o_i = (\alpha_i, \beta_i) : \alpha_i, \beta_i \subseteq \mathcal{M} \tag{2.8}$$

It important to note that P-graphs are directed bipartite graphs. The sets of materials and operating units are independent by definition, i.e., there are no arcs between the same vertex types. There are two disjoint sets of arcs where the elements of set A_1 are the arcs from materials to operating units and the elements of set A_2 are the arcs from operating units to materials.

The simple P-graph structure of the illustrative example is shown in [Figure 2.3]. The two

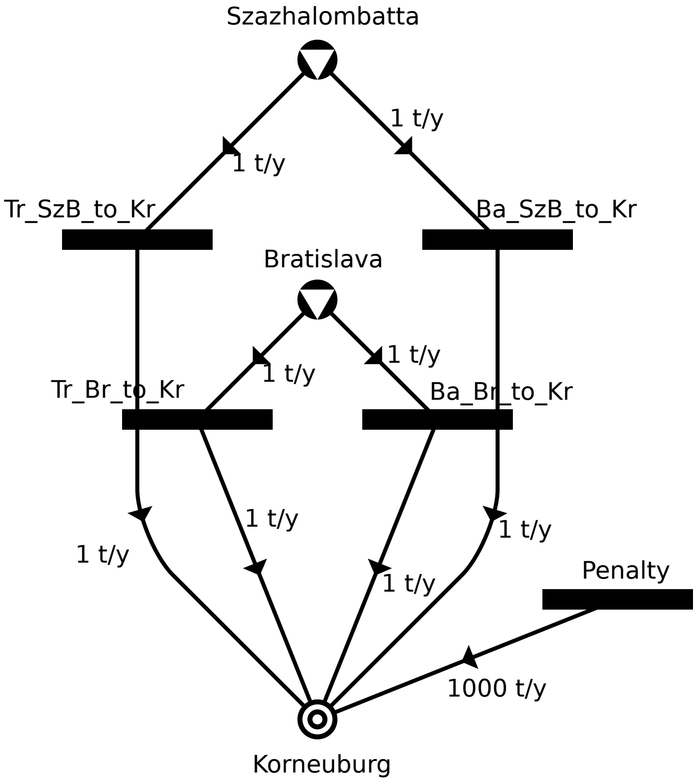

Figure 2.3: The P-graph representation of the motivational example

raw materials represent the source cities of biodiesel, the product stands for the biodiesel at the destination, and the operating units mean the two different cargo types from each source city to the destination and there is an additional operating unit to represent the penalty, if the fulfillment does not occur. The notation used in the basic graph is the same as it was in the decision tree [see Table 2.6]. The P-graph representation of the illustrative example is equivalent to the single stage problem that is defined by mathematical formulation in the first part of subsection 2.1.1.

2.3.4 Parametric cost modeling

The objective of the model is to minimize the overall cost of the network, which is equal to the total sum of the costs of the operating units and the prices of the raw materials.

The annual cost of an operating unit is considered as the sum of its yearly operating cost and its annualized investment cost:

$$\text{annual cost} = \text{operating cost} + \frac{\text{investment cost}}{\text{payout period}} \tag{2.9}$$

The optimization model is expected to provide the optimal loads of operating units beside the optimal process structure, therefore the cost is given as function of the mass load, e.g., by a linear function with a fixed charge

$$cf(o_i) + cp(o_i)x_i \tag{2.10}$$

where $cf(o_i)$ is the fixed charge, $cp(o_i)$ is the proportionality constant, and x_i is the load of the operating unit, which typically varies between 0 and 1, i.e., 0-100% of the operating unit's capacity. If both the investment and operating costs are given as functions, then the cost function is a combination of them. The parameters of the linear cost function with fixed charge are the sums of the parameters $cf^{op}(o_i)$ and $cp^{op}(o_i)$ of the function of the operating cost as well as the parameters $\frac{cf^{inv}(o_i)}{\text{payout period}}$ and $\frac{cp^{inv}(o_i)}{\text{payout period}}$ of the annualized investment cost

$$cf(o_i) = \frac{cf^{inv}(o_i)}{\text{payout period}} + cf^{op}(o_i) \tag{2.11}$$

$$cp(o_i) = \frac{cp^{inv}(o_i)}{\text{payout period}} + cp^{op}(o_i) \tag{2.12}$$

The problem is defined as $(\mathcal{P}, \mathcal{R}, \mathcal{O}, cf^{op}, cp^{op}, cf^{inv}, cp^{inv})$.
Furthermore, denote the vector of the operating units' optimal loads for the problem by $\mathbf{x}^* = [x_1, x_2, ..., x_n]^T$ and the objective value by

$$z^* = \sum_{(\alpha_i,\beta_i)=o_i \in \mathbf{o}^*} (cf(o_i) + x_i^* cp(o_i)) \longrightarrow min \tag{2.13}$$

The objective function defined by Equation (2.13) is perfectly the same as the objective function defined by Equation (2.3) in the second part of subsection 2.1.1 illustrative example. The fixed charge values $cf(o_i)$ are the coefficients of the binary decision variables, while the proportionality constants $cp(o_i)$ are the coefficients of the second stage decision variables in each scenario. In subsection 2.1.1, the objective function is already applied to the illustrative example, the biodiesel transportation problem.

In the example of this chapter, the overall cost comes from only the fix and proportional parts of the transportation cost and the penalty itself, if the fulfillment does not occur. By this manner, the deposit for reserving the barge or the railway in advance represents the "investments", while the cost of the transportation itself indicates the operating costs. The prime cost is neglected in this specific case. The transportation cost is based on the distances between the cities, and it is considered that transporting by barge is cheaper than by rail cargo [Table 2.2]. It is also considered that the fix cost is the 10 percent of the proportional cost per 1000 tons of biodiesel [Equation (2.14)].

$$Fix\ cost \quad = \quad Proportional\ cost \quad * \quad 10\% \quad * \quad 1000t \quad [\text{EUR}] \tag{2.14}$$

The P-graph of the illustrative example has been built and optimization problem has been solved by the software P-graph Studio [p-graph.org] that is represented in Figure 2.8. The capacity upper bound is set to 2000 for the operating units representing transportation, which is sufficient to satisfy the demands in this example; the fix part of the investment costs is set to 1200 EUR and the proportional part of the operating cost is 12 EUR/t, i.e., 12'000 EUR/1'000t.

The required amount of biodiesel in Korneuburg is 1000 t/year in every scenario and the available inventory in each depot — Szazhalombatta and Bratislava — is considered to be large enough to fulfill that requirement on its own [Table 2.3 and 2.4].

It is important to note that both transportation types are able to fulfill the requirement on their own.

For determining the penalty, it is necessary to ascertain the value of the product first. It is derived from the total required amount, the density of biodiesel and the price of the biodiesel [Equations (2.15) and (2.16)].
The penalty is determined as 0.3% of the product value multiplied by the time needed to accomplish the requirements [Equations (2.17) and (2.18)].

$$\textit{Product value} = \textit{Total mass of the required flow} * \textit{Density of biodiesel} * \textit{Price} \tag{2.15}$$

$$555555.56 = 1000 * \frac{1000}{0.9} * 0.5 \quad [\text{EUR}] \tag{2.16}$$

$$\textit{Penalty} = \textit{Product value} * 0.3\ \% * 4\ \textit{days} \tag{2.17}$$

$$6666.67 = 555555.56 * 0.3\% * 4 \quad [\text{EUR}] \tag{2.18}$$

2.4 Model Extension - Two-stage P-graphs

The parametric PNS model introduced in the previous subsection acts as the input for the two-stage model. It has been extended by the scenarios, $\mathcal{T}$; the probability of the scenario's incidence, practically the weight of the scenario, w; and the availability of the operation, $\mathcal{X}$, which is a binary parameter that says whether an operation is available in a specific scenario or it is not.

$$(\mathcal{P}, \mathcal{R}, \mathcal{O}, cf^{op}, cp^{op}, cf^{inv}, cp^{inv}) \quad + \quad (\mathcal{T}, w, \mathcal{X}) \tag{2.19}$$

where $\mathcal{T}$ is the set of scenarios,

$$0 \leq w_k \leq 1; \quad \sum_k w_k = 1 \tag{2.20}$$

w_k is the incidence probability of the k^{th} scenario, $\mathcal{T}_k$ and

$$\mathcal{X}_{i,k} = \begin{cases} true & \text{if} \quad o_i \in \mathcal{O} \text{ available in scenario } \mathcal{T}_k \\ false & \text{otherwise} \end{cases} \tag{2.21}$$

where $\mathcal{X}_{i,k}$ is the availability of the i^{th} operating unit, o_i presenting in the k^{th} scenario, $\mathcal{T}_k$.

The extended process-network synthesis model involving multiple scenarios is built by using the same components such as products, resources, operating units and cost parameters; however, most of them are related to alternative scenarios.

$$(\mathcal{P}', \mathcal{R}', \mathcal{O}', cf^{op'}, cp^{op'}, cf^{inv'}, cp^{inv'}) \tag{2.22}$$

$\mathcal{P}'$ represents the set of products of the two-stage model, which means the set of products contains the products of each scenario.

$$\mathcal{P}' = \{p_{j,k} : \forall p_j \in \mathcal{P}, \tau_k \in \mathcal{T}\} \tag{2.23}$$

$\mathcal{R}'$ is the set of resources in the two-stage model, which are the resources of the basic structure in each scenario.

$$\mathcal{R}' = \{r_{j,k} : \forall r_j \in \mathcal{R}, \tau_k \in \mathcal{T}\} \tag{2.24}$$

$\mathcal{M}'$ defines the set of materials in the two-stage model. It also contains all the materials in the basic model and it is extended by the materials represented in each scenario.

$$\mathcal{M}' = \{m_{j,k} : \forall m_j \in \mathcal{M}, \tau_k \in \mathcal{T}\} \tag{2.25}$$

In addition to the above redefined set, artificial materials are introduced to provide links between the investment and the utilization of an operating unit.
Let

$$\mathcal{M}' = \mathcal{M}' \bigcup \{m_{i,k}^{inv} : (cf_i^{inv} + cp_i^{inv}) > 0\} \tag{2.26}$$

The set of operating units also has to be expanded with all the operating units that are present in each scenario.

$$\mathcal{O}' = \{o_{i,k} : \; \tau_k \in \mathcal{T}, \; o_i \in \mathcal{O}, \; \mathcal{X}_{i,k} = true\} \tag{2.27}$$

Artificial operating units represents the investments in the related operating units, similarly to materials, e.g., payment of the reservation for a specific means of transportation.
Let

$$\mathcal{O}' = \mathcal{O}' \bigcup \{o_i^{inv} : \; \forall o_i = (\alpha_i, \beta_i) \in \mathcal{O}, (cp^{inv}(o_i) + cf^{inv}(o_i)) > 0\} \tag{2.28}$$

The operational cost of these artificial operating units are set to zero.

$$cp^{op'}(o_i) = 0, \;\; cf^{op'}(o_i) \tag{2.29}$$

If a certain operating unit has any kind of investment cost, the inlet streams of that certain operating unit in the second stage are expanded with the investments.

Table 2.7: Definition of scenarios

$\mathcal{T}$	w	$\mathcal{X}$	
		Ba-SzB	*Ba-Br*
Scenario $\mathcal{T}_1$	0.5508	true	true
Scenario $\mathcal{T}_2$	0.1377	true	false
Scenario $\mathcal{T}_3$	0.1692	false	true
Scenario $\mathcal{T}_4$	0.1423	false	false

Let

$$o_{i,k} = (\alpha_{i,k}, \beta_{i,k}) : \tag{2.30}$$

$$\alpha_{i,k} = \begin{cases} \{m_{j,k} : m_j \in \alpha_j; \tau_k \in \mathcal{T}\} \bigcup, \{m_{i,k}^{inv}\} & \text{if} \quad (cf^{inv}(o_i) + cp^{inv}(o_i)) > 0 \\ \{m_{j,k} : m_j \in \alpha_j; \tau_k \in \mathcal{T}\} & \text{otherwise} \end{cases} \tag{2.31}$$

$$\beta_{i,k} = \{m_{j,k} : m_j \in \beta_j; \tau_k \in \mathcal{T}\} \tag{2.32}$$

The operational cost of a certain operating unit in one of the scenarios is weighted by the incidence probability of that scenario.

$$cp^{op'}(o_{i,k}) = cp^{op}(o_i) \cdot w_k \tag{2.33}$$

And finally, the investment costs of the operating units in each scenario are set to zero since those have been regarded in the first stage.

$$cp^{inv'}(o_{i,k}) = 0 \tag{2.34}$$

The extended P-graph of the illustrative example was generated by the above detailed method. It can be seen in Figure 2.4 a). The model, provided in the second part of subsection 2.1.1 can be matched to this P-graph structure. The top part of the graph — colored with red — represents the first stage decisions, more precisely the red operating units (horizontal bars) are equivalent to the binary decision variables and the bottom part of the graph — colored with black — represents the four scenarios and in each scenario the operating units are equivalent to the second stage decision variables. There are missing operating units in the second, third and fourth scenarios compared to the original single structure because those operating units represent those means of transportation that are not available in the certain scenario.

It is important to note, that the penalty is present in the lower section of the graph, but in this specific case it does not cause any mistake, since the cost of the penalty is either 6666 euro or zero, therefore there is no necessity to introduce an extra block in the top red section in the graph for the penalty.

The optimal process structure for the illustrative example is presented in Figure 2.4 b).

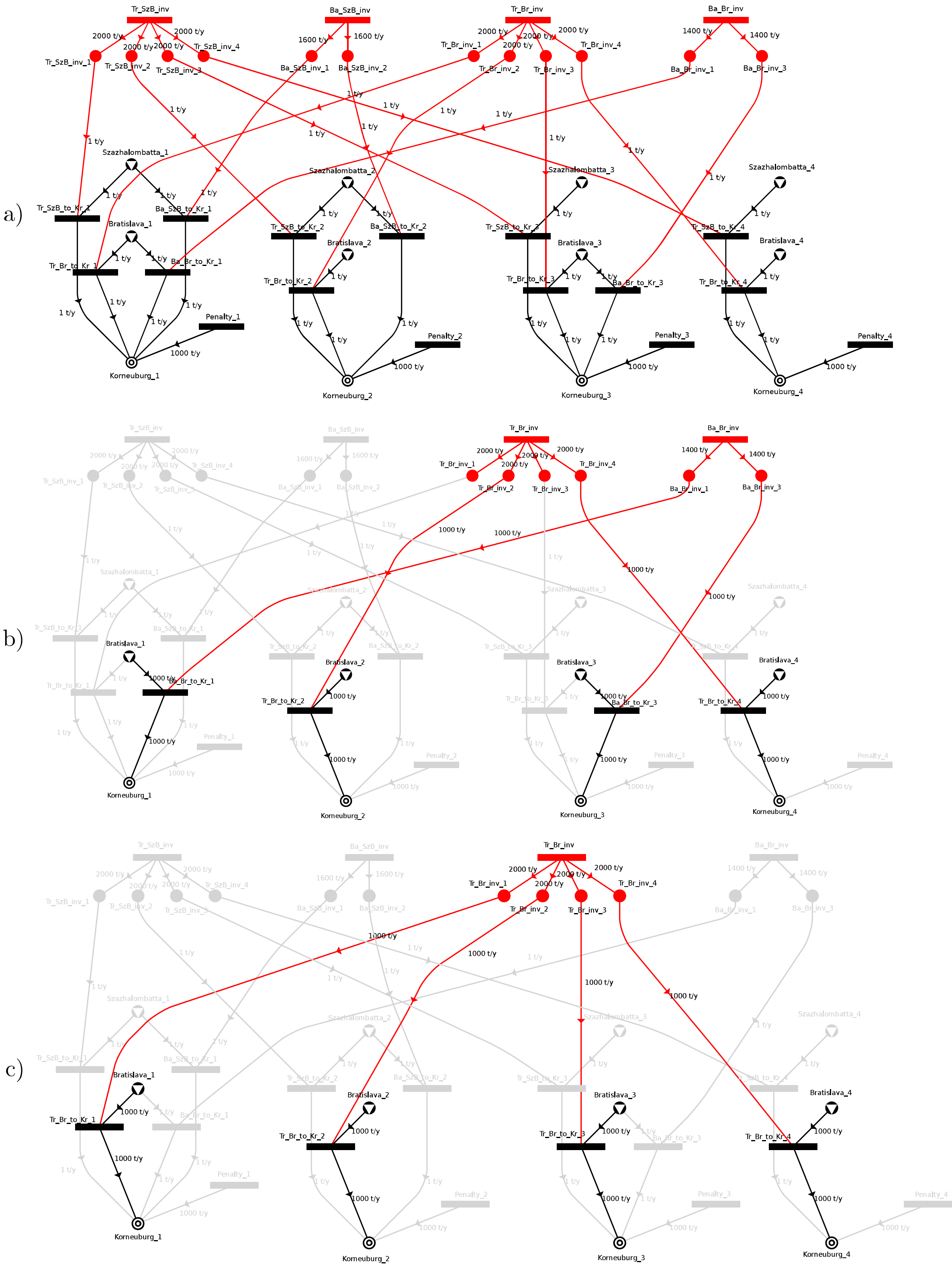

Figure 2.4: a) The extended P-graph representation of the motivational example b) The optimal solution of the motivational example c) The optimal solution of the motivational example with modified scenario probabilities

Table 2.8: Summary of the alternative solutions of the motivational example

Solution	Investment	Operation	Total cost [EUR]
1	Ba-Br, Tr-Br	Ba-Br($\mathcal{T}_1, \mathcal{T}_3$), Tr-Br($\mathcal{T}_2, \mathcal{T}_4$)	3980.00
2	Ba-Br, Tr-Br	Ba-Br($\mathcal{T}_1$), Tr-Br($\mathcal{T}_2, \mathcal{T}_3, \mathcal{T}_4$)	4149.20
3	Ba-Br	Ba-Br($\mathcal{T}_1, \mathcal{T}_3$), Penalty($\mathcal{T}_2, \mathcal{T}_4$)	4326.48
...	...	...	...
6	Tr-Br	Tr-Br($\mathcal{T}_1, \mathcal{T}_2, \mathcal{T}_3, \mathcal{T}_4$)	4400.00
...	...	...	...
24	Ba-Br, Ba-SzB	Ba-Br($\mathcal{T}_1, \mathcal{T}_3$), Ba-SzB($\mathcal{T}_2$), Penalty($\mathcal{T}_4$)	5547.87
25	Ba-SzB, Ba-Br, Tr-Br	Ba-Br($\mathcal{T}_1, \mathcal{T}_3$), Ba-SzB($\mathcal{T}_2$), Tr-Br($\mathcal{T}_4$)	5568.50
...	...	...	...
37	Ba-Br, Tr-SzB	Ba-Br($\mathcal{T}_1, \mathcal{T}_3$), Tr-SzB($\mathcal{T}_2$), Penalty($\mathcal{T}_4$)	6260.97
...	...	...	...
51	nothing	Penalty($\mathcal{T}_1, \mathcal{T}_2, \mathcal{T}_3, \mathcal{T}_4$)	6666.00
...	...	...	...
240	Tr-SzB, Ba-SzB	Tr-SzB($\mathcal{T}_1, \mathcal{T}_3, \mathcal{T}_4$), Ba-SzB($\mathcal{T}_2$)	13686.90

It can be seen accordingly that the best decision is to invest into both barge and cargo from Bratislava and transport the required amount of biodiesel to Korneuburg by barge if it is possible depending on the navigability or transport by rail cargo if the barge is not available. In other words, comparing this solution to the mathematical formulation presented in the second part of subsection 2.1.1, from the binary (first stage) decision variables, only the barge and the cargo from Bratislava is one, while all the others are zero; and in scenario 1 and 3, the required amount of biodiesel is transported by barge from Bratislava and in scenario 2 and 4, the required amount of biodiesel is transported by rail cargo from Bratislava.

The summary of all the alternative feasible solutions of the extended biodiesel transportation problem is presented in Table 2.8. As it has already been mentioned, the optimal solution is to invest both into the barge and the rail cargo from Bratislava and then transport the biodiesel on the river either scenario $\mathcal{T}_1$ or scenario $\mathcal{T}_3$ occur, and transport by rail cargo otherwise. It is important to note, that the optimal and the second best solution are strategically the same. In both cases, the barge and the cargo are invested in the first stage, only the utilization of the transportation types is different. Namely, the transportation is served by rail cargo also in scenario $\mathcal{T}_3$ even though the barge would be available. On the other hand the third best solution is a strategically different structure, since only the barge from Bratislava is invested and utilized if it is possible; if the barge is unavailable (scenarios $\mathcal{T}_2$ and $\mathcal{T}_4$), penalty has to be paid.

In Table 2.8, all the strategically different alternative structures are listed. Solution structure #51, where nothing is invested, but the penalty is paid in each scenario. There are 240 feasible solution structures and the last solution has 3.5 times higher total cost than the optimal structure.

The value of stochastic solution (VSS) is frequently calculated for stochastic programming problems in order to get deeper understanding of the relationship between the deterministic solution and its stochastic counterpart. For the illustrative example presented in this work, the deterministic solution can be calculated by modifying the single stage structure of the problem [Figure 2.3]. There should be upper limitations introduced to the transportation capacity of the barge both from Bratislava and from Szazhalombatta and this limitation must be proportional to the probability of their availability. For example, if the barge from Bratislava is available with 72%, it should be considered that maximum transportation capacity of it is the 72% of the required amount, so 720 [t]; and similarly for the other means of transport. Then the modified single stage problem can be solved.

For this illustrative example, the solution provided by the above detailed method is exactly the same as the optimal solution derived from the two-stage stochastic model, so the value of the stochastic solution is zero. However, there are numerous examples in the literature where the VSS indicates that the stochastic solution describes more the reality than the deterministic one ([83], [78]). These confirm the benefits of stochastic programming for that the extended P-graph methodology presented in this work is a combinatorially accelerated approach.

It is important to note that, if the weight of the scenarios is modified, the optimal solution structure and the order of the feasible solutions structures may change; however having the P-graph model in hand, the parameters can be altered, optimal and alternative best solutions can be calculated again by the P-graph Studio software within seconds. For example, the probability of scenarios $\mathcal{T}_1$ and $\mathcal{T}_3$ can drop to 10% due to some trouble in the docks of Bratislava and the probability of the other two scenarios can become 40%, the optimal medium term decision is to invest only into the rail cargo from Bratislava and transport the biodiesel by cargo in all scenarios, which was only the 6^{th} best solution with the previous settings [see Figure 2.4 c)].

2.5 Introduction to the P-graph Studio

This section is to briefly introduce the software, the P-graph Studio that has been used to design the graph structures and models for the previously described illustrative example. The optimization, the determination of the number combinatorially possible alternatives and the calculation of the value of stochastic solution was done by this software and it all provided basis for the evaluation of the new method, the two-stage extension of the P-graph framework. The software was specially developed in order to support the P-graph framework.

The design of the graph structure is pretty easy, all the vertices — either materials or operating units — can be placed by drag and drop method, and the links among them can be simply defined by simply drawing a line from a vertex to another. This also defines the direction of the arrow between the two vertices, besides, the software does not allow forbidden links, e.g. a link between two material type of vertices.

By highlighting one of the graph object of the model of interest, all the corresponding properties can be set, name, costs, limits or even the style and the layout of the certain object [Figure 2.5].

When the structure is ready, the solution algorithm can be chosen and the number of solutions

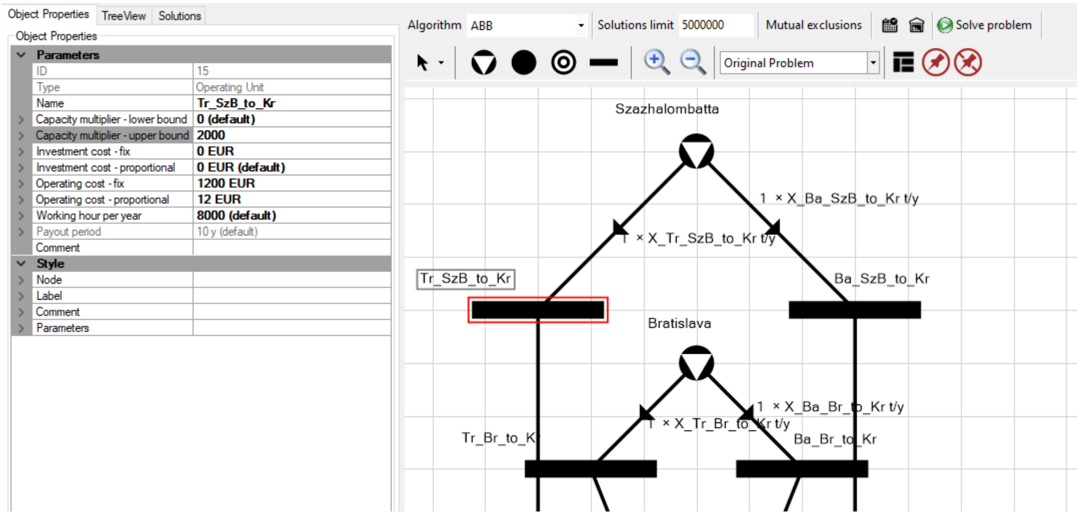

Figure 2.5: General interface for designing a model in the P-graph Studio /p-graph.org/

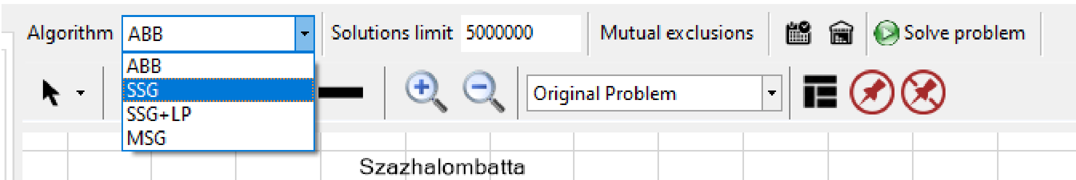

Figure 2.6: The set of algorithms to be chosen in the P-graph Studio /p-graph.org/

can be limited [Figure 2.6] and there is a possibility to define mutual exclusions if necessary for the problem [Figure 2.7].
It is important to note, that the number of all the combinatorially possible alternatives can be determined by the SSG algorithm, which can be reduced by defining mutual exclusions that mean further restrictions regarding the combinatorially possible alternatives. The ABB and SSG+LP algorithms are responsible for generating the optimal and the n-best solutions of the problem.

When the optimization algorithm is executed, the optimal and all the other feasible solutions can be analyzed individually [Figure 2.8] and there are several export options for further usage or further analysis of the problem and its feasible solutions [Figure 2.9].

2.6 Summary

Herein has been presented a superstructure approach for multistage stochastic optimization. An initial structure is built graphically in a way where each scenario is achievable through a series of decisions from any stage, besides each of them is part of the so called superstructure. First, the potential activities of stages are defined formally, then the overall model is generated algorithmically, and the resultant model is analyzed by P-graph algorithms originally conceived for process synthesis. Algorithms MSG is verifying the integrity of the model, algorithm SSG is

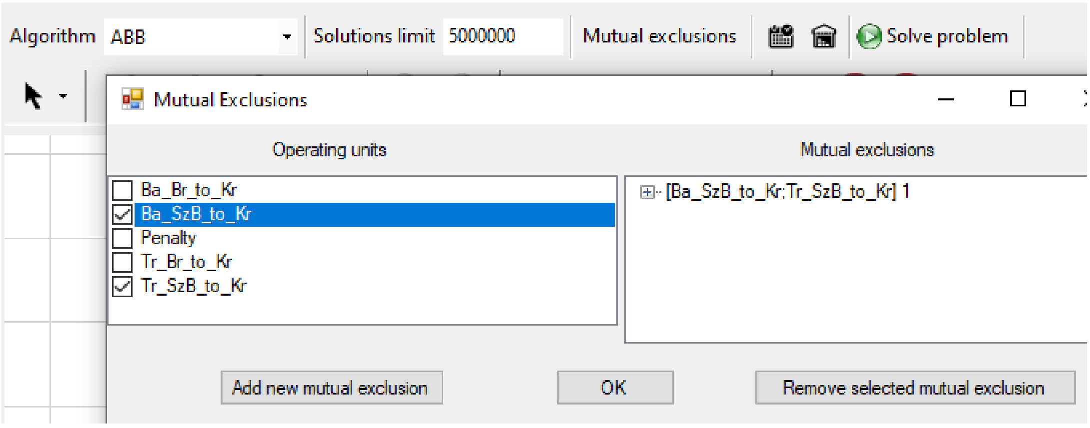

Figure 2.7: Defining mutual exclusions for the model if necessary in the P-graph Studio /p-graph.org/

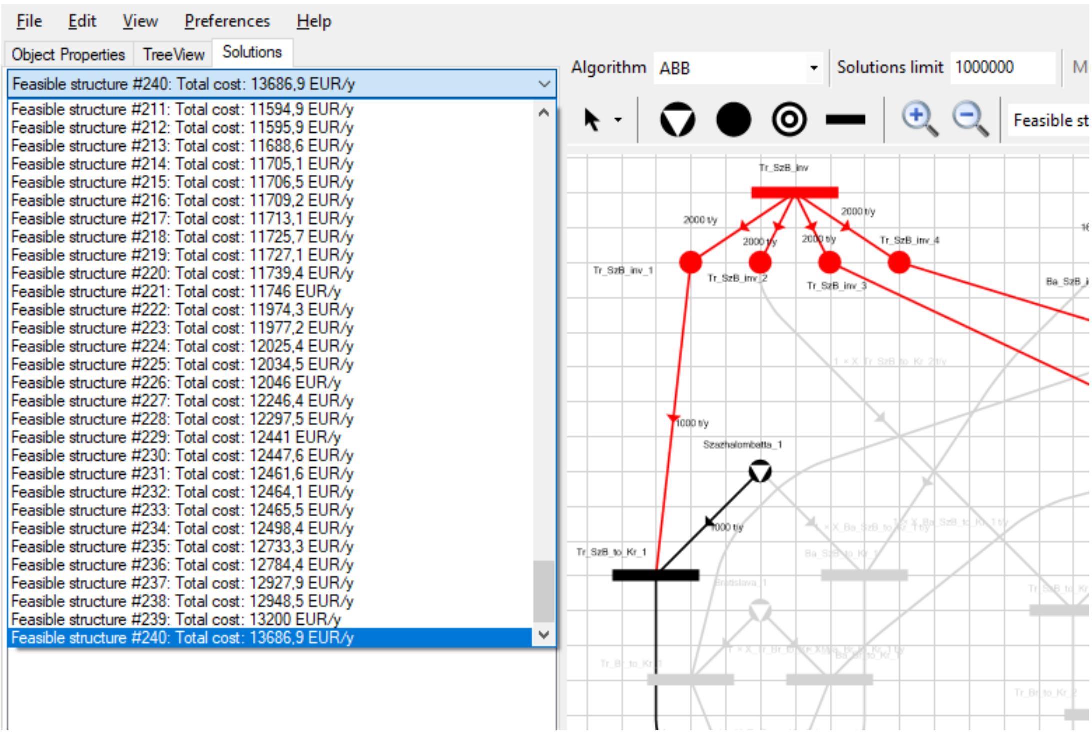

Figure 2.8: Software P-Graph Studio lists all the alternative feasible solutions structures /p-graph.org/

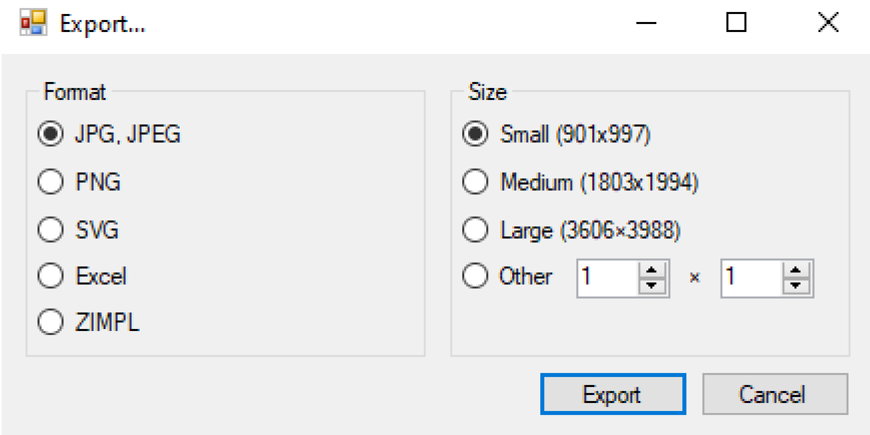

Figure 2.9: Export options in the P-Graph Studio /p-graph.org/

enumerating all the potential scenarios algorithmically, i.e., encounter the leaves of the decision tree, which are required to be constructed manually in other approaches. Finally, algorithm ABB lists the N-best combinations of decisions aiming to minimize the expected costs and maximize the expected profit.

The method presented in this chapter was illustrated by automatically generating a two-stage decision problem from a single stage process model. The first stage decisions are made on the investments, and the operation is determined in the second stage according to the scenario that takes place. It is rewarding to construct a model applying the above detailed method, because the proposed software implementation is capable of calculating and visualizing the optimal and alternative decisions; moreover, the results of any change in the parameters can be seen immediately. Therefore, sensitivity analysis of alternative decision strategies is fast and simple.

Further research and development is to be performed to support both the multistage model generation and sensitivity analysis of the calculated best process structures by computer aid. These features are to be included in the later versions of software P-graph Studio.

2.7 Related publication

Refereed Journal Paper

- E. Konig, B. Bertok. Process graph approach for two-stage decision making: Transportation contracts. *Computers & Chemical Engineering*, 121:1-11, 2019.
  IF: 3.334
  Published: 2 February 2019

Conference Papers and Presentations

- E. Konig, K. Kalauz, B. Bertok, Synthesizing Flexible Process Networks by Two stage P-graphs, presented at the ESCAPE-24, Budapest, June 15-18, 2014.
- E. Konig, B. Bertok, Cs. Fabian, Scaling power generation and storage capacities by P-graphs, presented at ECSP 2014 (European Conference on Stochastic Programming and Energy Applications), Paris, France, September 24-26, 2014.
- E. Konig, Z. Sule, B. Bertok, Comparison of optimization techniques in the P-graph framework for the design of supply chains under uncertainties, presented at the VOCAL 2014 (ASCONIKK - Annual Scientific Conference of NIKK), Veszprem December 14-17, 2014.
- E. Konig, Z. Sule, B. Bertok, Design of transportation networks under uncertainties by the P-graph framework, presented at the P-graph Conference, Balatonfured January 22-25, 2015.
- E. Konig, Z. Sule, B. Bertok, Design of Supply Chains under uncertainties by the Two Stage Model of the P-Graph Framework, presented at PRES'15 International Conference, Kuching, Malaysia August 22-25, 2015.
- E. Konig, B. Bertok, Z. Sule, Planning Optimal River Transport of Petrochemicals Concerning Uncertainties of Water Levels By Two Stage P-Graph, presented at AIChE Spring Meeting and 12th Global Congress on Process Safety, Houston, TX, USA, April 10-12, 2016.
- E. Konig, A. Bartos, B. Bertok, Free Software for the Education of Supply Chain Optimization, presented at VOCAL 2016 (VOCAL Optimization Conference: Advanced Algorithms), Esztergom December 12-15, 2016.
- E. Konig, J. Baumgartner, Z. Sule, Optimizing examination appointments focusing on oncology protocol, presented at the 8th Annual Conference of the European Decision Science Institute (EDSI 2017): Information and Operational Decision Sciences, Granada, Spain May 29 - June 1, 2017.
- E. Konig, B. Bertok, Automated Scenario Generation by P-Graph, presented at SEEP 2017 (10th International Conference on Sustainable Energy and Environmental Protection: Mechanical Engineering), Bled, Slovenia, June, 27-30, 2017.

Chapter 3

Fisher information for resilience of ecosystems

In this chapter it is explained how the Fisher information is used to establish the limits of the resilience of the dynamic regime of a predator-prey system. Previous studies using Fisher information focused on detecting whether a regime change has occurred, whereas by the method presented in this work the goal is to determine how much an ecological system can vary its properties without a regime change occurring. The theory is illustrated with a simple two species systems. It is first applied to a predatory-prey model and then to a 60-year wolf-moose population dataset from Isle Royale National Park in Michigan, USA. The resilience boundaries and the operating range of a system's parameters without a regime change are assessed from entirely new criteria for Fisher information, oriented towards regime stability. The approach provides the possibility to use system measurements to determine the shape and depth of the "cup" of stability as defined by the broader resilience concept.

3.1 Calculating resilience from Fisher information

Holling (1973) has defined the resilience of an ecological system as the ability of the system to continue functioning within the same dynamic regime despite externally inflicted perturbations. Within the same regime, the system can be very resilient to some kinds of disturbances over a long period of time, and not resilient at all to others. The resilience of an ecological system in a regime can vary over time, such as with the loss of species or gradually changing external conditions, at the same time that stability can appear constant (the system does not change regimes). Regime shift occurs when one or more borders have been reached (e.g., the loss of too many species, or a catastrophic disturbance). In previous research, Fisher information has been used retroactively, to identify regime thresholds after regime shifts have occurred ([71], [100], [108]). Identifying the thresholds of a regime without first observing a regime shift is a different problem.

Consider that computing Fisher information for an ecosystem is possible as a function of any of its characteristic parameters (species mortality, growth rate, etc.). Perturbations can be represented as changes in the characteristic parameters - note that the characteristic parameters of an ecosystem can change for other reasons as well. However, within the range of parameter values consistent with the existence of a functioning ecosystem the Fisher information would be relatively low since the system is dynamic, and the Fisher information would have a relatively high value for the range of parameter values leading to a non-functional or static and dead system. A Fisher information calculation, however, is an observational process. It provides information about the system dynamic regimes and the changes in those regimes. It can provide hints at what changes in the system parameters may be driving the changes, but determining cause and effect is not its primary purpose. That requires either an explicit mathematical model of the system such as the prey-predator model, or an implicit model such as the observations for the moose-wolf population data for Isle Royale, both of which are discussed later.

If only one system parameter is being perturbed, a two-dimensional plot of Fisher information versus the parameter values would appear as a ”cup” with steep walls [Figure 3.1]. The systems with parameter values at the bottom of the ”cup” are dynamic and functioning, and the ones on the steep wall have very low resilience as they can ”flip” into a different regime. If two parameters are being simultaneously perturbed, a three-dimensional plot of Fisher information versus the two parameters would generally appear as a ”canyon” with steep walls, and again the systems with parameter values at the bottom of the ”canyon” are dynamic and functioning systems and the ones near the steep walls have low resilience [Figure 3.2]. In the transition phase where the system has lost resilience and therefore it is not functioning well, the observable variables of the system would fluctuate beyond the values normally seen in a healthy functioning ecosystem. This means that the measurable values of the system variables would fluctuate more widely around their mean leading to a broadening and flattening of $p(t)$, and a Fisher information lower than that of a resilient and orderly system and much lower than that of a system with very low resilience. Hence, if the Fisher information is computed continuously as a system transitions from resilient to less so, the Fisher information of the resilient system would have a non-zero value, a much lower value for the system in transition, and a high value after the ecosystem has shifted out of the regime and into a new one. This is important, because it can be ”seen” how the system is moving towards a new regime before it has done so. Such a detailed calculation requires either a model capable of representing the transition or finely grained data capturing the transition. However, consistent with the *Sustainable Regimes Hypothesis* of Fath et al. (2003), the following criteria can be stated:

$$\langle I \rangle|_h > 0 \qquad \text{and} \qquad \left.\frac{d\langle I \rangle}{dt}\right|_h \cong 0 \tag{3.1}$$

$$\langle I \rangle|_f \ll \langle I \rangle|_h \qquad \text{and} \qquad \left.\frac{d\langle I \rangle}{dt}\right|_f \neq 0 \tag{3.2}$$

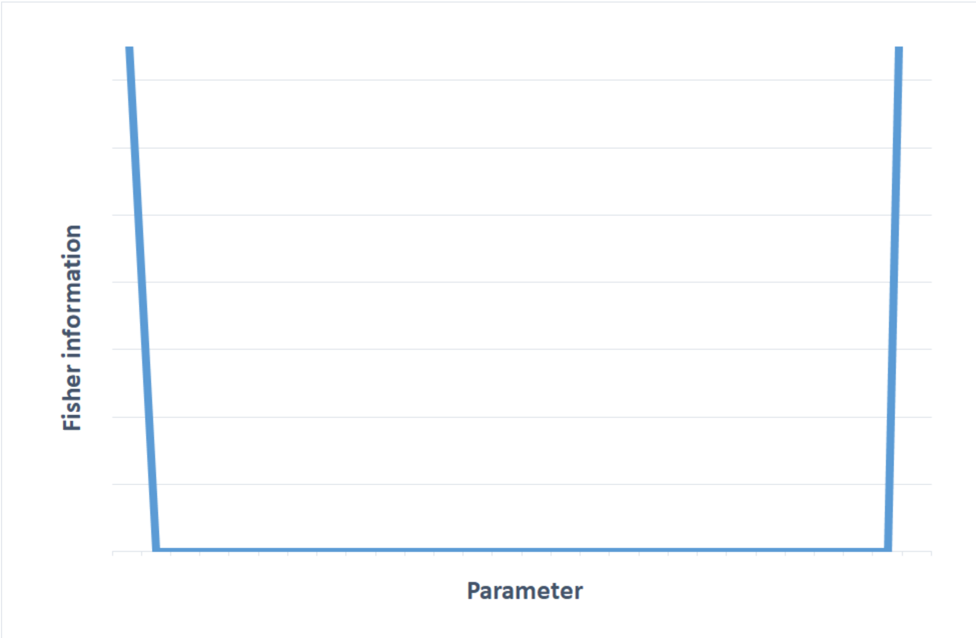

Figure 3.1: Illustrating the "cup" with steep walls. The Fisher information as the function of one perturbed parameter (the transition phase is not depicted)

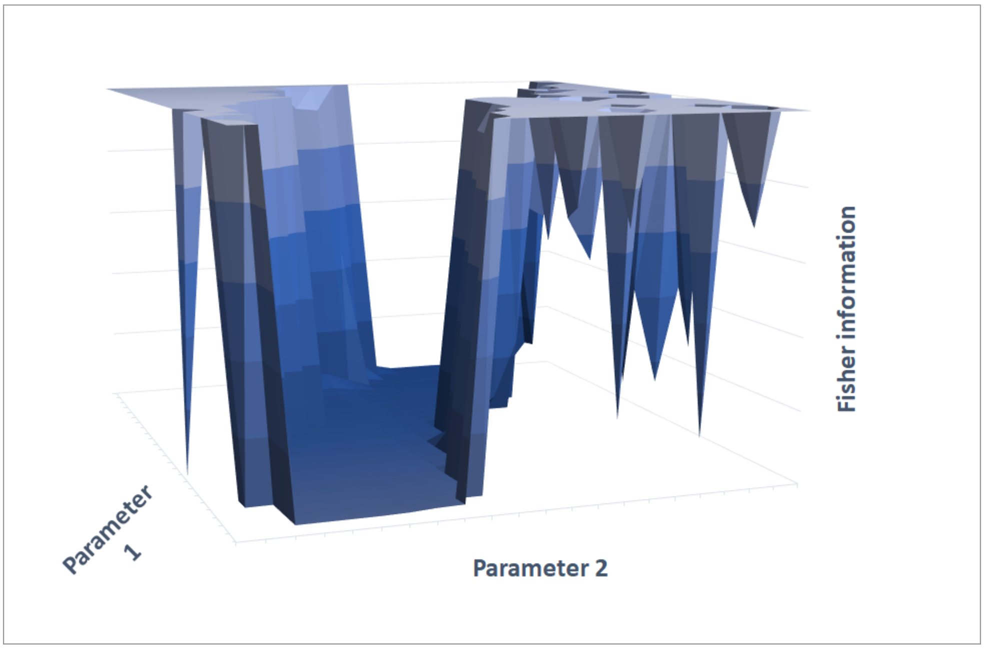

Figure 3.2: Illustrating the "canyon" with steep walls. The Fisher information as the function of two simultaniously perturbed parameters (the transition phase is not depicted)

$$\langle I \rangle|_d \gg \langle I \rangle|_h \qquad \text{and} \qquad \left.\frac{d\langle I \rangle}{dt}\right|_d = 0 \tag{3.3}$$

where $\langle I \rangle$ is the average Fisher information over some time interval t defined by $\langle I \rangle \equiv 1/T \int_0^T I(t)dt$ and the subscripts h, f, and d refer to ecosystems that are healthy, in flux or transition, and totally dysfunctional, respectively. It is important to note that the prey-predator model system that is described later is unable to represent the transition since it is too simple of a model, and because the Heaviside step function is applied to the model system in order to eliminate the part where the system recovers from biologically unsustainably low population.

The above mentioned suppositions can be summarized mathematically by proposing the hypothesis that the averaged Fisher information of a stable system does not significantly change with changes in the value of the system parameters under a perturbation. Considering the case of one system parameter (α), for example the mortality rate of a species, being perturbed this can be expressed by:

$$\frac{d\langle I \rangle}{d\alpha} \cong 0 \tag{3.4}$$

For the case when two system parameters (α and β) are under perturbation, for example the mortality rate and the growth rate, the general expression would be:

$$\left.\frac{d\langle I \rangle}{d\alpha}\right|_\beta \cong 0 \tag{3.5}$$

$$\left.\frac{d\langle I \rangle}{d\beta}\right|_\alpha \cong 0 \tag{3.6}$$

Finally, for the general case when an arbitrary number (n) of ecosystem parameters (α_i) are being perturbed, the corresponding expression is:

$$\left.\frac{d\langle I \rangle}{d\alpha_i}\right|_{\alpha_{j\neq i}} \cong 0 \qquad\qquad i = 1, 2, \ldots n \tag{3.7}$$

where $\langle I \rangle$ is now the average Fisher information defined for the one perturbed parameter case of Equation (3.4) by $\langle I \rangle \equiv \int [I(\alpha)\, d\alpha] \,/\int d\alpha$, for the two parameter case of Equations (3.5) and (3.6) by $\langle I \rangle \equiv \iint [I(\alpha, \beta)\, d\alpha d\beta] \,/\iint d\alpha d\beta$, and for the general case of Equation (3.7) by $\langle I \rangle \equiv \iint\cdots\int [I(\alpha_1\alpha_2\ldots\alpha_n)\, d\alpha_1 d\alpha_2 \ldots d\alpha_n] \,/\iint\cdots\int d\alpha_1 d\alpha_2 \ldots d\alpha_n$.

It is very difficult to visualize the Fisher information as a function of three or more model parameters since it would lie in a four or higher dimensional space, which is unfortunately outside the range of human perception. But the mathematical approach is still valid. The algorithm that would be carried out for the investigation of such a system would be similar to the one presented here for one and two parameter systems. Hence, in this work firstly parameter α_1 is varying over the range of interest while all parameters $\alpha_{i\neq 1}$ are kept constant at some predetermined value. For the following step of the method parameter α_2 is varying while all parameters $\alpha_{i\neq 2}$ are constant. At the end there will be set of Fisher information values that depend on the aforementioned n parameters, i.e. $I(\alpha_1\alpha_2\ldots\alpha_n)$. In order to identify the parameter range over which the system is resilient, it is required to search for regions where the Fisher information is flat in this n parameter space. These are ranges of parameter values where the Fisher information does not significantly vary as given by Equations (3.4),(3.5),(3.6) and (3.7).

The result of these conjectures derived from Fisher information considerations is that of providing the mathematical machinery that is required to estimate how much the system parameters can vary without generating a change in the dynamic regime of the system. Then it can be argued that the wider the range of parameter variation that can be tolerated without a regime change, the more resilient the system.

3.2 Prey-Predator model system

The simple ecological system model is a predator-prey model adopted from the work of Fath et al. (2003). The initial values of the relevant parameters are also those used by Fath et al. (2003). The population fluctuation in time is natural behavior, and the populations are depending on one another as well as on other parameters, like the growth or density rate of the prey and the mortality rate or reproduction rate of the predator. The system is defined by a Lotka-Volterra type mathematical model. The model variables and parameters are as follows:

y_1	population mass of the prey [mass]
y_2	population mass of the predator [mass]
g_1	growth rate of prey [1/time]
l_{12}	loss rate to prey due to predator feeding [1/time]
g_{21}	feeding rate of predator [1/time]
m_2	mortality rate of predator [1/time]
k	density dependence of prey [mass]
β	reproduction rate of predator [mass/mass]

Definition of the population fluctuation:

$$\frac{dy_1}{dt} = g_1\left(1-\frac{y_1}{k}\right)y_1 - \frac{l_{12}y_1y_2}{1+\beta y_1} \tag{3.8}$$

$$\frac{dy_2}{dt} = \frac{g_{21}y_1y_2}{1+\beta y_1} - m_2y_2 \tag{3.9}$$

Pure mathematics has no biological restrictions for reproduction of species therefore this pure mathematical model is able to increase the population even from an infinitesimally small population number, which is biologically impossible. It is required to force the model to set the population exactly to zero when it reaches a lower limit where the system is biologically not sustainable any more. Hence, the values of y_1 and y_2 is set to zero when they became 1 or less. This is reflected in Equations (3.10) and (3.11) where a Heaviside step function is applied to both y_1 and y_2.

$$y_1 = \begin{cases} 0 & \text{if } (y_1 - 1) < 0 \\ y_1 & \text{if } (y_1 - 1) \geq 0 \end{cases} \tag{3.10}$$

$$y_2 = \begin{cases} 0 & \text{if } (y_2 - 1) < 0 \\ y_1 & \text{if } (y_2 - 1) \geq 0 \end{cases} \tag{3.11}$$

Solving Equations (3.8) and (3.9) applying the logic statements from Equations (3.10) and (3.11) yields to values for y_1 and y_2 in each time step. Replacing the values of the population into the Equations (3.8) and (3.9), the values for $(dy_1)/dt$ and $(dy_2)/dt$ can be calculated also in each time step. In order to be able to calculate the Fisher information, the values of $(d^2y_1)/dt^2$ and $(d^2y_2)/dt^2$ are also required. Therefore, we need to express the second time derivative of equations (3.8) and (3.9).

$$\frac{d^2y_1}{dt^2} = g_1\frac{dy_1}{dt} - 2\frac{g_1}{k}y_1\frac{dy_1}{dt} - \left(\frac{1}{1+\beta H(y_1-1)y_1}\right)\left[l_{12}y_2\frac{dy_1}{dt} - l_{12}y_1\frac{dy_2}{dt}\right] + l_{12}y_1y_2\left(\frac{1}{1+\beta y_1}\right)^2\beta\frac{dy_1}{dt} \tag{3.12}$$

$$\frac{d^2y_2}{dt^2} = g_{21}y_2\frac{dy_1}{dt}\left(\frac{1}{1+\beta y_1}\right) + g_{21}y_1\frac{dy_2}{dt}\left(\frac{1}{1+\beta y_1}\right) - g_{21}y_1y_2\left(\frac{1}{1+\beta y_1}\right)^2\beta\frac{dy_1}{dt} - m_2\frac{dy_2}{dt} \tag{3.13}$$

In summary, for purposes of the study of a model prey-predator system presented in this work, the Fisher information is calculated from Equation (1.5) setting $\Delta t = 1$ and using y_1 and y_2 computed as a function of time from Equations (3.8),(3.9) and Equations (3.12),(3.13).

3.2.1 Results of the predator-prey model system

The default position of the calculation presented in this work was a parameter set where the system has stable limit cycle behavior for the populations of the prey and the predator species. The initial values were set based on the research of Fath et al. (2003). See Figure 3.3, where the initial values are as follows:

g_1 = 1 [1/time]
l_{12} = 0.01 [1/time]
g_{21} = 0.01 [1/time]
m_2 = 1 [1/time]
k = 625 [mass]
β = 0.005 [mass/mass]

Equations (3.10) and (3.11) were implemented in MATLAB version 2016b and solved by the ODE15 solver for 300 time steps, with initial values 5 and 15 for y_1 and y_2 respectively. Note that the system has the populations independently from the initial values as it migrates to its

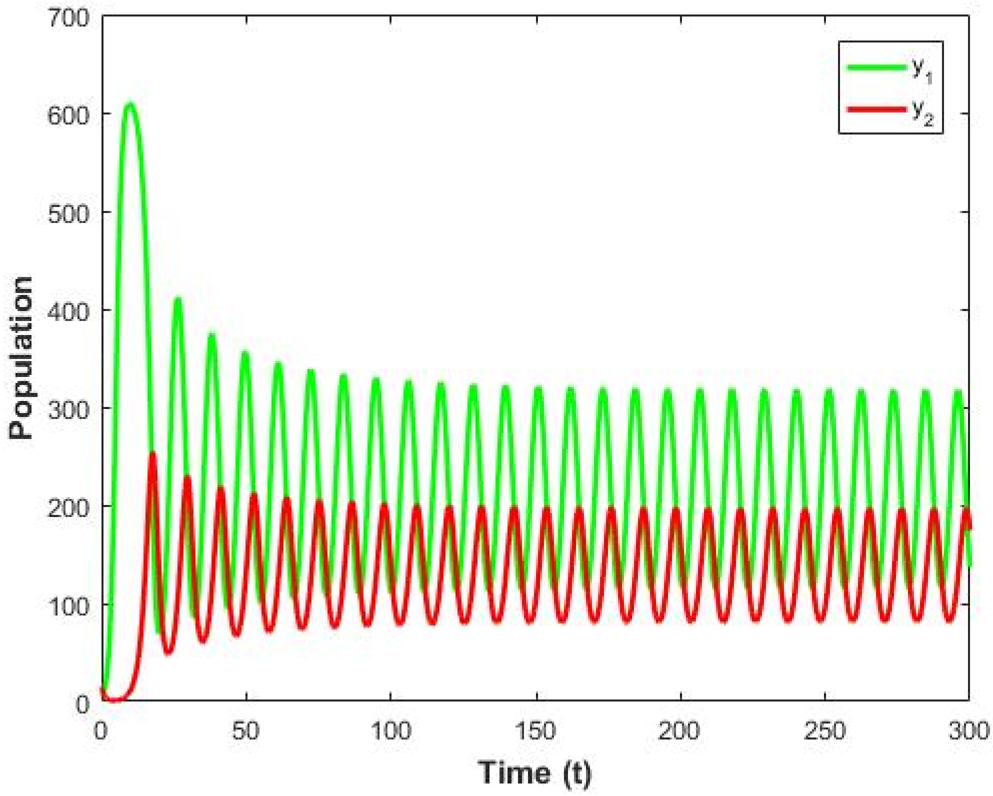

Figure 3.3: The fluctuation of the model populations of prey y_1 and predator y_2 in time with default parameter values of $k = 625$, $m_2 = 1$.

steady-state regime therefore the initial population values can be arbitrary. Then the resulting values of y_1 and y_2 were imported into the Excel spreadsheet software, and all further calculations, namely the values of the first and second derivatives as well as the numerical assumption of the Fisher information, were executed in Excel. The value of the Fisher information for this specific parameter set and model system in its steady-state regime is around 0.00015. However, it is the relative values of Fisher information and the relative changes in Fisher information values that are critical here, not the value itself.

It is a typical living and functioning system in ecology that is depicted in Figure 3.3; both species are present and the value of its Fisher information is finite and steady. If the value of the parameter k were changed enough – increased and decreased – the fluctuation of the populations eventually ceases because one of the species became extinct. The lower limit of k is around 395 and there the predator y_2 immediately dies out and the prey population grows to its upper limit [Figure 3.4]. If the other extreme case when parameter k is increased until it reaches its upper border ($k \approx 1325$), the same phenomena is perceptible but it is delayed; that is, after one period of fluctuation, the predator dies out and the prey population grows to its upper limit [Figure

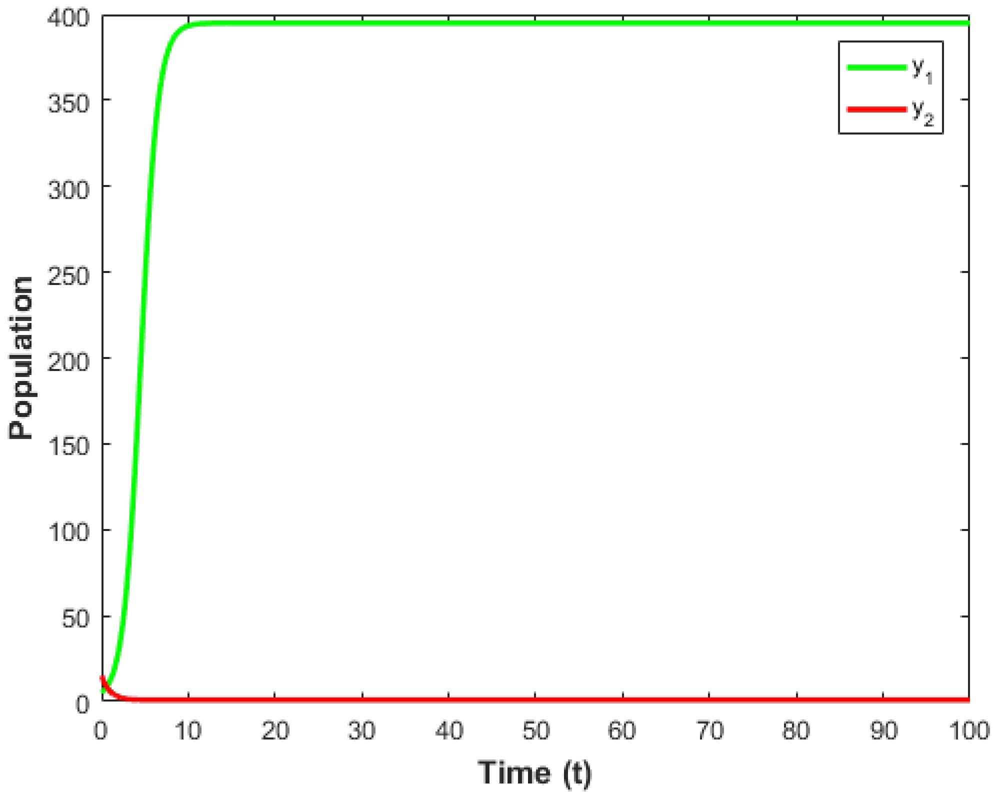

Figure 3.4: The fluctuation of the populations ceases, the predator population y_2 immediately dies out and the prey population y_1 stops growing when k reaches its lower limit ($k = 395$, $m_2 = 1$).

3.5]. The value of the Fisher information grows to a relatively high value in both edges [Figure 3.6]. It is because with $m_2 = 1$, the system becomes a static ecosystem [Figure 3.4] when the value of k is above the upper limit ($k > 1325$) or below the lower limit ($k < 395$). As it was explained before, as Fisher information is a measure of order, a static system has very high order and high Fisher information.

Now the stable range of the parameter k had been defined and the next step is to vary the m_2 parameter (the mortality of the predator) in the middle of the stable k range when $k = 860$. It was found that the system is much more sensitive to variation in mortality; it has a much narrower stability range. The parameter m_2 can be varied between 0.38 and 1.045 without getting a static, dead system state. If m_2 reaches its lower end and k is in its middle, the prey dies out earlier, therefore the predator also dies out soon afterwards [Figure 3.7]. These kinds of collapses occur where one of the species dies out on the edges, immediately or after one or two periods. This study showed that if the system has stable dynamics, the order of the Fisher information is around 10-3, and it grows suddenly when the system collapses as species populations start

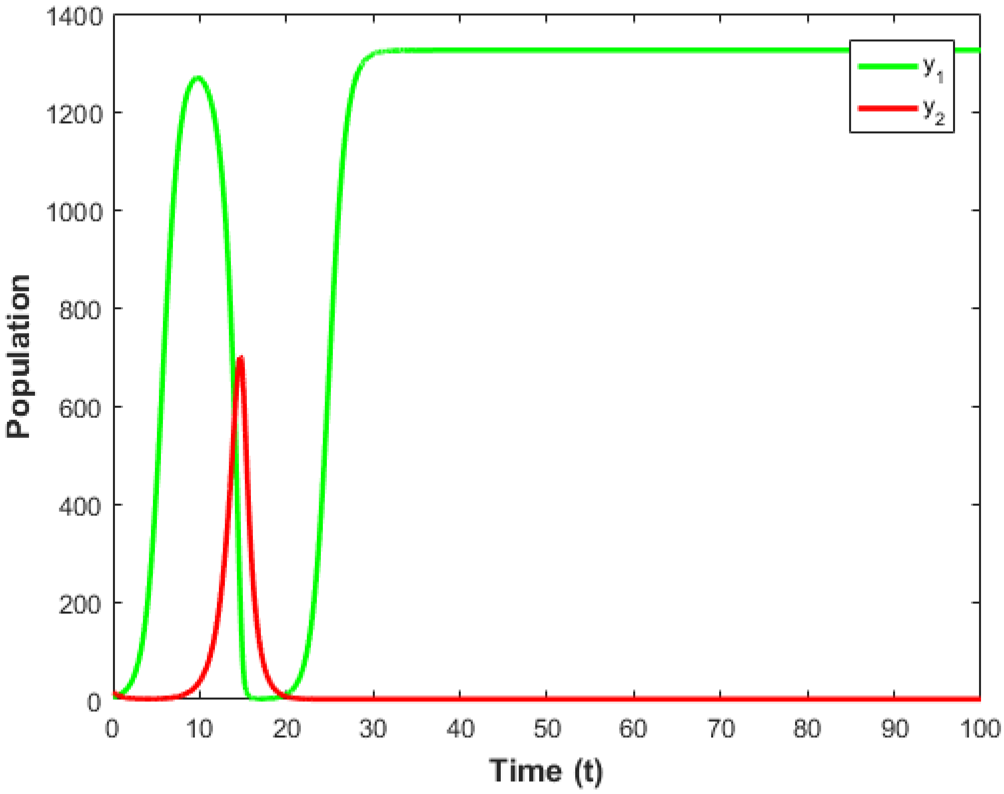

Figure 3.5: The fluctuation of the model populations of prey y_1 and predator y_2 in time with default parameter values at its upper limit of $k = 1325$, $m_2 = 1$.

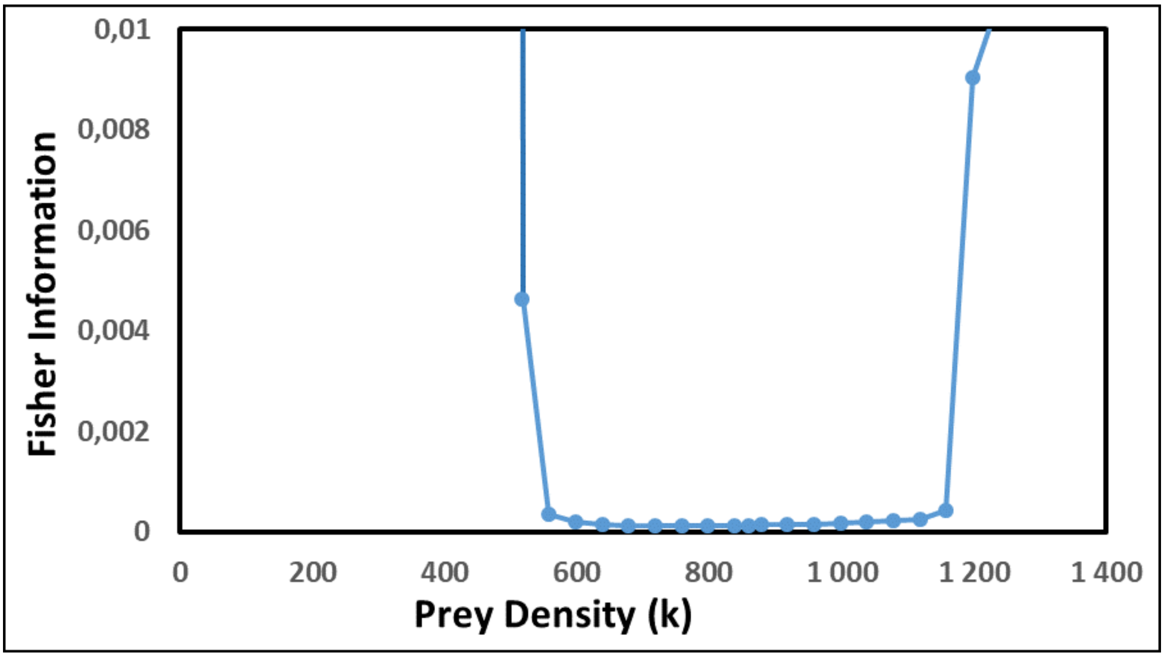

Figure 3.6: Fisher information for a prey-predator model system where the prey density parameter k is varied while the predator mortality m is held constant at $m = 0.9996$. Note that the systems where $518 \leq k \geq 1158$ are functioning systems with two species, and systems where $k < 518$ and $k > 1158$ are dysfunctional systems where at least one species has gone extinct. Note that the vertical scale has been truncated so that the more important details around $500 \leq k \geq 1300$ become easier to visualize.

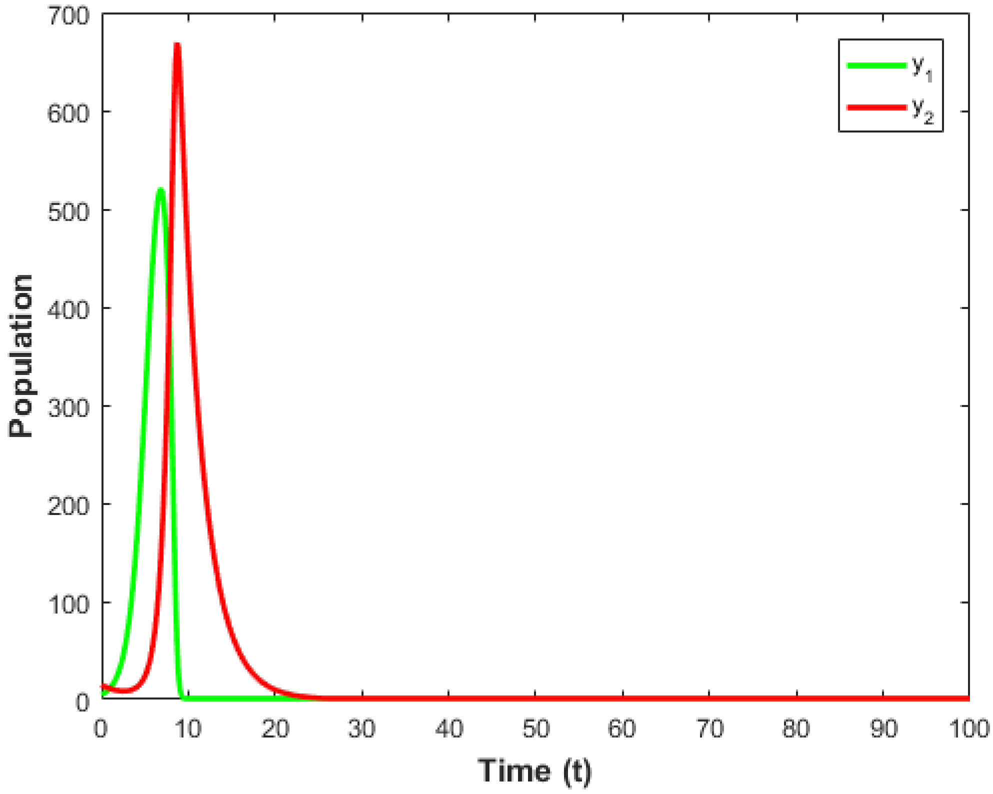

Figure 3.7: The prey dies out after a half of a period, therefore the predator also dies out afterwards ($k = 860$, $m_2 = 0.38$).

going to zero. Out of the stability range the value of the Fisher information is over the order of 1015 [Figure 3.8] and [Figure 3.9]. It is important to note that these system collapses define a different system ([71]), one lacking at least one of the two species.

On the right side of Figure 3.8 or on the left side of Figure 3.9, strange inverse peaks appear outside the stability range. (And another one appears on the other side of the canyon.) These peaks are due to numerical problems with the calculation method. Since the time steps are discrete as well as the values of y_1 and y_2 in each time step, technically Equation (1.3) becomes a sum instead of an integral. This can be seen in Equation (3.1),(3.2) and (3.3); note that the time step is defined as $\Delta t = 1$. These peaks appear in a state where the system is dysfunctional, namely the prey population dies out after one period. Therefore the predator population dies out as well after this first period [see Figure 3.10]. Practically, in these cases R' becomes exactly zero, but in mathematics dividing by zero has no meaning. Therefore, while calculating the Fisher information, only those time steps can be considered where the division is valid, i.e. while the value of R' is over zero. In the specific case shown in the Figure 3.10, the division is valid until $t \approx 26$, and the system is functional between $t = 0$ and $t = 26$.

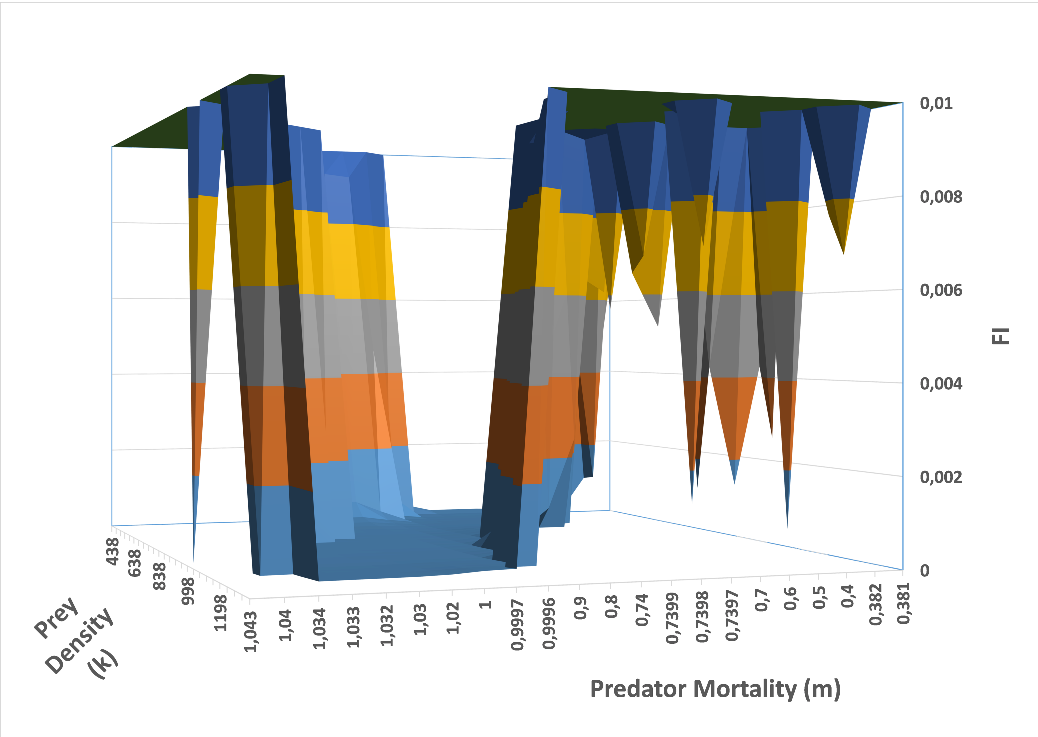

Figure 3.8: The value of the Fisher information as a function of prey density k and predator mortality rate $m2$, from a side view. Note that a functioning ecosystem with two species present exists only for combined values of k and $m2$ within the confines of the bottom of the "canyon".

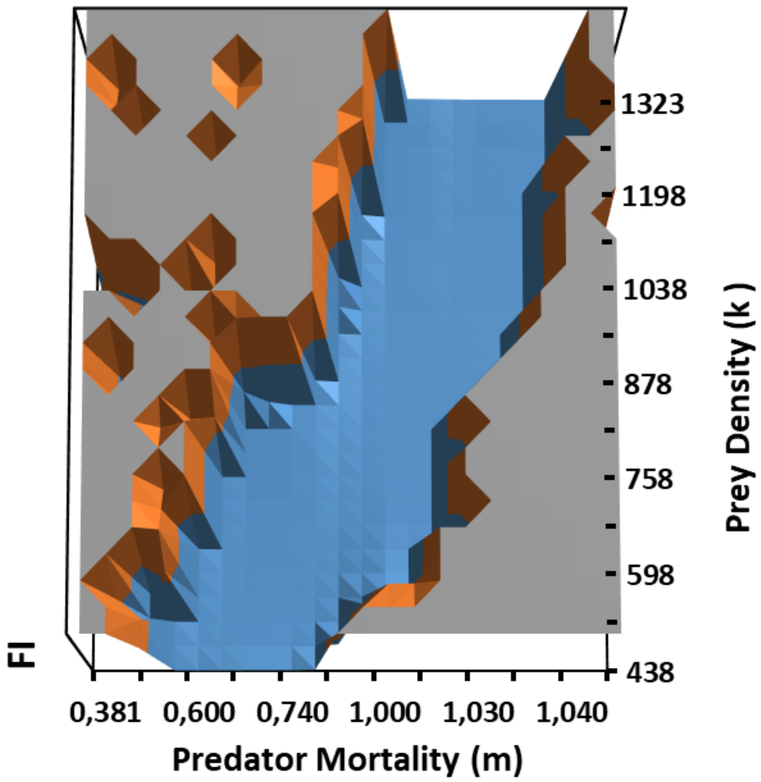

Figure 3.9: The value of the Fisher information as a function of prey density k and predator mortality rate $m2$, from a side view. Note that a functioning ecosystem with two species present exists only for combined values of k and $m2$ within the confines of the bottom of the "canyon".

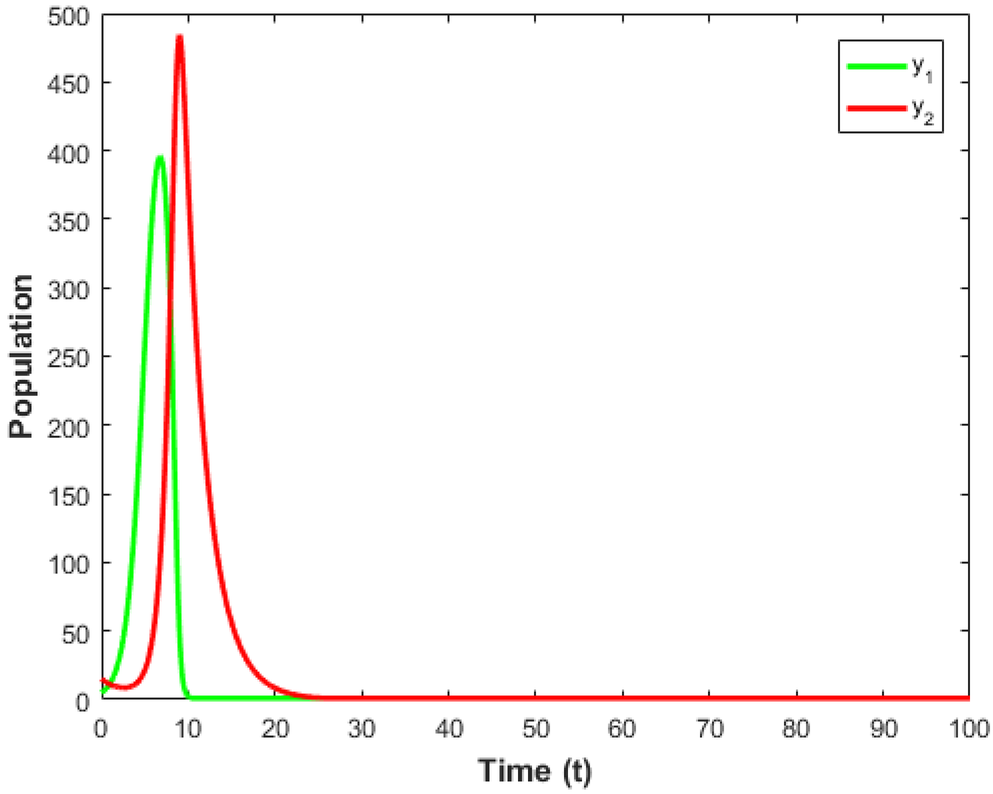

Figure 3.10: The prey dies out after a half of a period, therefore the predator also dies out afterwards ($k = 598$, $m_2 = 0.381$)

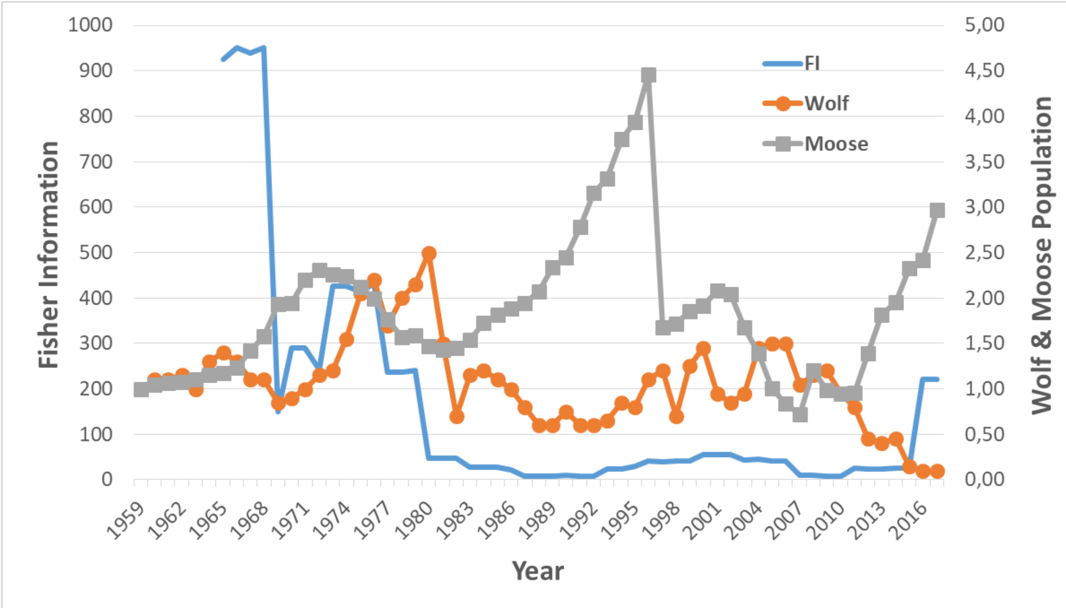

Figure 3.11: The trend of Fisher information in a 6-year-long moving time window together with the normalized population of wolf and moose in Isle Royale National Park.

3.3 Results of a real predator-prey ecosystem

To study the application of the methodology, a real predator-prey ecosystem has been involved into the analysis. This real ecosystem is represented by the wolf-moose (Canis lupis, Alces alces) system from Isle Royale National Park in the United States. There is a 540 km^2 remote island, Isle Royale in the Lake Superior where the wolf and moose population (and their impacts on the vegetation) has been monitored and the research project have provided a 60-year long (1957-2017) data ([111]; [76]). The population sizes of wolves and moose are surveyed each winter; the dataset contains the precise number of wolves and estimated number of moose. The system has been in the news in the past several years after the wolf population began an unsustainable decline in abundance; as of 2017, only one inbred pair of wolves lived on the island, and the moose population was increasing rapidly in the absence of sufficient predation ([76]).

For the Fisher information calculation, the 1/10th of the length of the 60-year long data, i.e. a 6-year-long moving time window has been applied, which is plotted in Figure 3.11. The wolf and moose population values (normalized so that both fit on the second y-axis) are also plotted in Figure 3.11. All population values are dimensionless in Figure 3.11; values are divided by the first value (in 1959) for each species.

Comparing the Fisher information trend to the population trends, a brief delay is perceptible, but as expected, Fisher information has high values when population fluctuations are low and drops when the fluctuations intensify. The Fisher information calculated here indicates that there is, perhaps, a functional state with relatively high dynamic order that persisted in the

1970s, where wolf populations were around 40 individuals and moose around 1000. However, this region may not be entirely resilient, as since that time this system has spent the bulk of its time in a low Fisher information region of less than 20 wolves and well over 1000 moose. The sharp decline of wolf population in 1981 (echoed in a decline in Fisher information) was due to the accidental introduction of canine parvovirus to the island ([112]). It is notable that the wolf an moose populations show similar dynamic changes to the healthy behavior of the model system in the 1970s and between 2000 and 2007, however the Fisher information is different in these periods. That can occur for several reasons, since the dataset only contains the population of the wolf and the moose, there is explicit information regarding the parameters that are described for the model system, besides the noise is eliminated in the model system due to its purity.

However, this resilience degraded as the wolf population entered a sharp decline after 2009. The behavior of Fisher information for this real-life system is consistent with the behavior observed for the model system, although the impact of the noise in the real system can be easily recognized on the clarity of Fisher information behavior. However, broadly speaking, Fisher information indicates that some event (internal or external) occurred in the early 1980s, despite the appearance of some stability in population numbers in the early 2000s, which set this system on a less-resilient pathway from which it has not yet recovered.

3.4 Summary

The previous works related to Fisher information and system regimes are focusing mostly on regime changes when a system shifts from one regime into another. The goal of this research was to develop a method to calculate where a resilient system has its borders, and to identify those ranges of the interacting parameters where the system is capable of persisting in one regime independently of the perturbations. By the criterion formulated as Equations (3.4), it is possible to decide if a dynamic system is in a healthy, dynamically changing state, in a dysfunctional and therefore static state, or in transition from a healthy state into a dysfunctional one. The criterion, defined by Equations (3.5),(3.6),(3.7), (3.8) and (3.9) tells where a system is resilient when there is only one, two or more varying system parameters respectively.

The theory of Fisher information is well known and frequently applied in several scientific fields, but it has not been utilized for measuring system resilience directly. The method described in this thesis provides a technique to measure the resilience of a dynamic system by checking the criteria defined by Equations (3.4),(3.5),(3.6),(3.7),(3.8) and (3.9). The Fisher information remains highly sensitive to the quality of the data as it was seen in previous iterations ([71]); accordingly, the selected variables must be relevant to characterizing changes in the condition of the system, otherwise the Fisher information results are meaningless. However, Fisher information may provide valuable information to the management of the resilience of the wolf-moose system on Isle Royale National Park. For example, in 2016-2017 the National Park Service debated about several management options in order to stabilize the wolf and moose populations. One option was to doing nothing and waiting to see if wolves return via an ice bridge over Lake Superior, or reintroducing several wolf packs from Canada over a period of 3 years (81 Federal

Register 91192 2016; [73]). In 2018, the National Park Service has come to a consensus and decided to slowly introduce very small numbers of wolves each year, releasing the first four in October 2018 ([77]). With better refinement, Fisher information could help park managers and wildlife biologists in determining whether this management option is having the desired effect (increasing the resilience of the wolf and moose populations). For example, Fisher information suggests that the island system with parvovirus present is not likely to allow for a resilient wolf-moose regime, and a policy prescription of vaccinations against parvo for all wolves may be warranted.

The theory has been illustrated via the predator-prey model system and the wolf-moose population data, but it can be applied in its present form to larger, more complicated systems as well. It should also be noted that the theory in its present form is applicable to any dynamic system if the model differential equations or time series data are available for the system variables. The system can be biological, social, economic, or technological. This means that it is possible to generally assess the resilience of a system by assessing the impact of changes in system parameters on the value of Fisher information. It is easy to represent the Fisher information as a function of two varying parameters since a line or a surface is easy to visualize (as it can be seen in Figure 3.6 or Figure 3.8 and 3.9). But the plot becomes four or higher dimensional if there are three or more varying variable, which is outside the range of human perception but the method is still valid. Further work will need to develop methods to interpret Fisher information accurately in these higher dimensions.

The Fisher information of any system is a fundamental and calculable property that is a measure of order. When applied to ecological systems, it was found and presented in this thesis that living functioning systems have a relatively low but steady Fisher information, while dysfunctional ecosystems can have either very high or very low Fisher information, depending on the variability in the system parameters. Fisher information is very sensitive to the dynamic behavior of complex systems which makes it a good indicator of regime changes. Here, it was used to measure the range of system parameter values over which a system remains within the same regime; larger range indicate higher resilience. Resilience defined and measured in this manner can be accomplished irrespective of the specific perturbation affecting the ecosystem; the change was measured without having information on the perturbation causing it. It would be ideal to know which disturbance caused the observed resilience loss, but this information is not always available. This form of resilience is, therefore, a measure of robustness or ruggedness in the face of often unpredictable perturbations. While much work remains to understand its strengths and limitations, the index shows promise as a way to characterize an important aspect of resilience in ecological systems and other dynamic systems generally.

3.5 Related publication

Refereed Journal Paper

- E. Konig, H. Cabezas, and A. L. Mayer. Detecting dyanimc system regime boundaries with fisher information: the case of ecosystems. *Clean Technologies and Environmental Policy*, pages 1-13, 2019.
  IF: 2.277
  Published: 13 June 2019

Chapter 4

Summary

In this thesis, there were two methods presented, which are capable of supporting decision makers regarding managing complex systems. The first technique presented herein is an optimization approach based on the P-graph framework for short- and medium term decisions, while the other method is for long term system management support by directly calculating the resilience of the system from its varying parameters using Fisher information. By applying the first method, the optimal design of a complex system's structure is possible and the other approach is capable of providing the limits and boundaries where the system remains in its certain regime.

For the P-graph based superstructure approach for two-stage stochastic optimization, an initial structure is built graphically so that each scenario is achievable through a series of decisions from any stage. Each scenario is a part of the superstructure. All the potential activities are formally defined first, then the complete model is algorithmically generated, and the resultant model is analyzed by the algorithms of P-graphs that were originally developed for process synthesis. This method was illustrated by a transportation problem with two source locations, one destination and two means of transport. The two-stage decision problem was generated from a single stage process model. The first stage decisions are made on the investments, and the operation is determined in the second stage according to the scenario that takes place. The proposed software implementation is capable of calculating and visualizing the optimal and alternative decisions; moreover, the results of any change in the parameters can be visualized immediately. Therefore, sensitivity analysis of alternative decision strategies is fast and simple.

The resilience calculation using Fisher information was illustrated via a predator-prey model system and the wolf-moose population data from Isle Royale National Park, Michigan, USA. The method was developed to calculate where, a resilient system has its limits, and to identify the ranges of the interacting parameters where, independently of the perturbations, the system is capable of remaining in one regime. By the criterion formulated in this thesis, it can be decided if a dynamic system is in a healthy, dynamically changing state, in a dysfunctional and therefore static state, or in transition from a healthy state into another, most likely into a dysfunctional one. This theory in its present form can be applied to any biological, social, economic, or technological dynamic system if the model differential equations or time series data are available for the system variables. By virtue of this, the general estimation of a system's resilience is possible by assessing the impact of changes in system parameters on the value of Fisher information. It is easy to represent the Fisher information as a function of two varying parameters since a line or a surface is easy to visualize but the plot becomes four or higher dimensional if there are three or more varying variable, which is outside the range of human perception but the method is still valid.

Chapter 5

New Scientific Results

1. I have proposed a technique based on the P-graph framework for multistage decision models where the number of scenarios in lower stages can be reduced only to those that can at all result in feasible solutions due to the axioms and combinatorial methods. Besides in upper stages, there is no need to enumerate all the possible feasible solution structures, it is enough if the algorithmically built superstructure implicitly includes them.

    - I have extended the process network synthesis model for the P-graph framework by the the scenarios and the probability of their occurrence.
    - I have introduced a new process for modeling cost parameters for the multistage decision problems.

2. I have proposed a method based on Fisher Information Theory that can be applied to calculate system's resilience directly. The approach provides the possibility to determine the borders within the system can vary its properties without a regime change occurring.

    - I have formulated criteria for the value of Fisher information by that the state of a dynamic regime (i.e., healthy, dynamically changing state; disfunctional, static state; or in transition from a healthy state into a dysfunctional one) can be determined (Equations (3.1) - (3.7)s).

Chapter 6

Publications

Refereed Journal Papers

1. E. Konig, B. Bertok. Process graph approach for two-stage decision making: Transportation contracts. *Computers & Chemical Engineering*, 121:1-11, 2019. (IF: 3.334)

2. E. Konig, H. Cabezas, and A. L. Mayer. Detecting dyanimc system regime boundaries with fisher information: the case of ecosystems. *Clean Technologies and Environmental Policy*, pages 1-13, 2019. (IF: 2.277)

International Conference Papers and Presentations

1. E. Konig, K. Kalauz, B. Bertok, Synthesizing Flexible Process Networks by Two stage P-graphs, presented at the ESCAPE-24, Budapest, June 15-18, 2014.

2. E. Konig, B. Bertok, Cs. Fabian, Scaling power generation and storage capacities by P-graphs, presented at ECSP 2014 (European Conference on Stochastic Programming and Energy Applications), Paris, France, September 24-26, 2014.

3. E. Konig, Z. Sule, B. Bertok, Comparison of optimization techniques in the P-graph framework for the design of supply chains under uncertainties, presented at the VOCAL 2014 (ASCONIKK - Annual Scientific Conference of NIKK), Veszprem December 14-17, 2014.

4. E. Konig, Z. Sule, B. Bertok, Design of transportation networks under uncertainties by the P-graph framework, presented at the P-graph Conference, Balatonfured January 22-25, 2015.

5. E. Konig, Z. Sule, B. Bertok, Design of Supply Chains under uncertainties by the Two Stage Model of the P-Graph Framework, presented at PRES'15 International Conference, Kuching, Malaysia August 22-25, 2015.

6. E. Konig, B. Bertok, Z. Sule, Planning Optimal River Transport of Petrochemicals Concerning Uncertainties of Water Levels By Two Stage P-Graph, presented at AIChE Spring Meeting and 12th Global Congress on Process Safety, Houston, TX, USA, April 10-12, 2016.

7. E. Konig, A. Bartos, B. Bertok, Free Software for the Education of Supply Chain Optimization, presented at VOCAL 2016 (VOCAL Optimization Conference: Advanced Algorithms), Esztergom December 12-15, 2016.

8. E. Konig, J. Baumgartner, Z. Sule, Optimizing examination appointments focusing on oncology protocol, presented at the 8th Annual Conference of the European Decision Science Institute (EDSI 2017): Information and Operational Decision Sciences, Granada, Spain May 29 - June 1, 2017.

9. E. Konig, B. Bertok, Automated Scenario Generation by P-Graph, presented at SEEP 2017 (10th International Conference on Sustainable Energy and Environmental Protection: Mechanical Engineering), Bled, Slovenia, June, 27-30, 2017.

Chapter 7

Appendix

The following models were evaluated for the enhanced approach of the P-graph framework. All these models are compatible with the P-graph Studio software that is available on the p-graph.org.

The models for the biodiesel transportation problem are available on the CD attachment of this thesis or can be downloaded via the following link from the p-graph.org:
P-graph models

- The single stage problem
BDTransportSingleStage.pgsx
- The two-stage problem for enumerating all the combinatorially possible alternatives
BDTransport2StageForSSG.pgsx
- The two-stage problem for optimization
BDTransport2StageForABB.pgsx
- The modified deterministic model for calculating the value of stochastic solution
BDTransportDeterministicForVSS.pgsx

Bibliography

[1] E. Aghezzaf. Capacity planning and warehouse location in supply chains with uncertain demands. *Journal of the Operational Research Society*, 56(4):453–462, 2005.

[2] A. Alonso-Ayuso, L. F. Escudero, A. Garìn, M. T. Ortuño, and G. Pérez. An approach for strategic supply chain planning under uncertainty based on stochastic 0-1 programming. *Journal of Global Optimization*, 26(1):97–124, 2003.

[3] K. Aviso, C. Cayamanda, F. Solis, A. Danga, M. Promentilla, K. Yu, J. Santos, and R. Tan. P-graph approach for gdp-optimal allocation of resources, commodities and capital in economic systems under climate change-induced crisis conditions. *Journal of Cleaner Production*, 92:308–317, 2015.

[4] I. Awudu and J. Zhang. Uncertainties and sustainability concepts in biofuel supply chain management: A review. *Renewable and Sustainable Energy Reviews*, 16(2):1359–1368, 2012.

[5] I. Awudu and J. Zhang. Stochastic production planning for a biofuel supply chain under demand and price uncertainties. *Applied Energy*, 103:189–196, 2013.

[6] A. Azaron, K. Brown, S. Tarim, and M. Modarres. A multi-objective stochastic programming approach for supply chain design considering risk. *International Journal of Production Economics*, 116(1):129–138, 2008.

[7] F. Bernstein and A. Federgruen. Decentralized supply chains with competing retailers under demand uncertainty. *Management Science*, 51(1):18–29, 2005.

[8] B. Bertok and M. Barany. Analyzing mathematical programming models of concep- tual process design by p-graph software. *Industrial & Engineering Chemistry Research*, 52: 166–171, 2013.

[9] B. Bertok, M. Barany, and F. Friedler. Generating and analyzing mathematical programming models of conceptual process design by p-graph software. *Industrial & Engineering Chemistry Research*, 52(1):166–171, 2012.

[10] D. Bertsimas and N. Youssef. Stochastic optimization in supply chain networks: averaging robust solutions. *Optimization Letters*, pages 1–17, 2019.

[11] H. M. Bidhandi and R. M. Yusuff. Integrated supply chain planning under uncertainty using an improved stochastic approach. *Applied Mathematical Modelling*, 35(6):2618–2630, 2011.

[12] BP. Bp statistical review of world energy 2019, 2019. URL `https://www.bp.com/content/dam/bp/business-sites/en/global/corporate/pdfs/energy-economics/statistical-review/bp-stats-review-2019-full-report.pdf`.

[13] F. Brand and K. Jax. Focusing the meaning (s) of resilience: resilience as a descriptive concept and a boundary object. *Ecology and society*, 12(1), 2007.

[14] B. Brehmer. Dynamic decision making: Human control of complex systems. *Acta psychologica*, 81(3):211–241, 1992.

[15] M. C. Carneiro, G. P. Ribas, and S. Hamacher. Risk management in the oil supply chain: a cvar approach. *Industrial & Engineering Chemistry Research*, 49(7):3286–3294, 2010.

[16] S. Carpenter, B. Walker, J. M. Anderies, and N. Abel. From metaphor to measurement: resilience of what to what? *Ecosystems*, 4(8):765–781, 2001.

[17] F. Chan and S. Chung. Multi-criteria genetic optimization for distribution network problems. *The International Journal of Advanced Manufacturing Technology*, 24(7):517–532, 2004.

[18] C.-L. Chen and W.-C. Lee. Optimization of multi-echelon supply chain networks with uncertain sales prices. *Journal of chemical engineering of Japan*, 37(7):822–834, 2004.

[19] C.-W. Chen and Y. Fan. Bioethanol supply chain system planning under supply and demand uncertainties. *Transportation Research Part E: Logistics and Transportation Review*, 48(1):150–164, 2012.

[20] M. Christopher. *Logistics & supply chain management*. Pearson UK, 2016.

[21] G. Csertán, A. Pataricza, P. Harang, O. Dobán, G. Biros, A. Dancsecz, and F. Friedler. Bpm based robust e-business application development. *Dependable Computing EDCC-4*, pages 659–663, 2002.

[22] L. Dai, K. S. Korolev, and J. Gore. Relation between stability and resilience determines the performance of early warning signals under different environmental drivers. *Proceedings of the National Academy of Sciences*, 112(32):10056–10061, 2015.

[23] M. Dal-Mas, S. Giarola, A. Zamboni, and F. Bezzo. Strategic design and investment capacity planning of the ethanol supply chain under price uncertainty. *Biomass and bioenergy*, 35(5):2059–2071, 2011.

[24] A. Demirbaş. Biodiesel from vegetable oils via transesterification in supercritical methanol. *Energy conversion and management*, 43(17):2349–2356, 2002.

[25] T. Eason and H. Cabezas. Evaluating the sustainability of a regional system using fisher information in the san luis basin, colorado. *Journal of environmental management*, 94(1): 41–49, 2012.

[26] T. Eason, A. S. Garmestani, C. A. Stow, C. Rojo, M. Alvarez-Cobelas, and H. Cabezas. Managing for resilience: an information theory-based approach to assessing ecosystems. 2016.

[27] G. Fandel and M. Stammen. A general model for extended strategic supply chain management with emphasis on product life cycles including development and recycling. *International journal of production economics*, 89(3):293–308, 2004.

[28] C. Folke. Resilience: The emergence of a perspective for social–ecological systems analyses. *Global environmental change*, 16(3):253–267, 2006.

[29] C. Folke, S. Carpenter, B. Walker, M. Scheffer, T. Elmqvist, L. Gunderson, and C. S. Holling. Regime shifts, resilience, and biodiversity in ecosystem management. *Annu. Rev. Ecol. Evol. Syst.*, 35:557–581, 2004.

[30] F. Friedler, K. Tarjan, Y. Huang, and L. Fan. Combinatorial algorithms for process synthesis. *Computers & chemical engineering*, 16:S313–S320, 1992.

[31] F. Friedler, K. Tarjan, Y. Huang, and L. Fan. Graph-theoretic approach to process synthesis: axioms and theorems. *Chemical Engineering Science*, 47(8):1973–1988, 1992.

[32] F. Friedler, K. Tarjan, Y. Huang, and L. Fan. Graph-theoretic approach to process synthesis: polynomial algorithm for maximal structure generation. *Computers & Chemical Engineering*, 17(9):929–942, 1993.

[33] F. Friedler, J. Varga, and L. Fan. Decision-mapping: a tool for consistent and complete decisions in process synthesis. *Chemical engineering science*, 50(11):1755–1768, 1995.

[34] F. Friedler, J. Varga, E. Feher, and L. Fan. Combinatorially accelerated branch-and-bound method for solving the mip model of process network synthesis. In *State of the art in global optimization*, pages 609–626. Springer, 1996.

[35] J. Gao, B. Barzel, and A.-L. Barabási. Universal resilience patterns in complex networks. *Nature*, 530(7590):307, 2016.

[36] B. H. Gebreslassie, Y. Yao, and F. You. Design under uncertainty of hydrocarbon biorefinery supply chains: multiobjective stochastic programming models, decomposition algorithm, and a comparison between cvar and downside risk. *AIChE Journal*, 58(7):2155–2179, 2012.

[37] S. Giarola, N. Shah, and F. Bezzo. A comprehensive approach to the design of ethanol supply chains including carbon trading effects. *Bioresource technology*, 107:175–185, 2012.

[38] M. Goh, J. Y. Lim, and F. Meng. A stochastic model for risk management in global supply chain networks. *European Journal of Operational Research*, 182(1):164–173, 2007.

[39] V. Grimm and C. Wissel. Babel, or the ecological stability discussions: an inventory and analysis of terminology and a guide for avoiding confusion. *Oecologia*, 109(3):323–334, 1997.

[40] L. H. Gunderson. Ecological resilience—in theory and application. *Annual review of ecology and systematics*, 31(1):425–439, 2000.

[41] A. Gupta and C. D. Maranas. A two-stage modeling and solution framework for multisite midterm planning under demand uncertainty. *Industrial & Engineering Chemistry Research*, 39(10):3799–3813, 2000.

[42] A. Gupta and C. D. Maranas. Managing demand uncertainty in supply chain planning. *Computers & Chemical Engineering*, 27(8):1219–1227, 2003.

[43] V. Guttal and C. Jayaprakash. Spatial variance and spatial skewness: leading indicators of regime shifts in spatial ecological systems. *Theoretical Ecology*, 2(1):3–12, 2009.

[44] I. Halim and R. Srinivasan. Systematic waste minimization in chemical processes. 1. methodology. *Industrial & engineering chemistry Research*, 41(2):196–207, 2002.

[45] I. Halim and R. Srinivasan. Design synthesis for simultaneous waste source reduction and recycling analysis in batch processes. *Computer Aided Chemical Engineering*, 20:1513–1518, 2005.

[46] I. Halim and R. Srinivasan. An intelligent system for green process design. *International Journal of Environment and Sustainable Development*, 8(1):1–9, 2009.

[47] R. Hamming. *Numerical methods for scientists and engineers*. Courier Corporation, 2012.

[48] I. Heckl, L. Halász, A. Szlama, H. Cabezas, and F. Friedler. Process synthesis involving multi-period operations by the p-graph framework. *Computers & Chemical Engineering*, 83:157–164, 2015.

[49] C. S. Holling. Resilience and stability of ecological systems. *Annual review of ecology and systematics*, 4(1):1–23, 1973.

[50] R. D. Horan, E. P. Fenichel, K. L. Drury, and D. M. Lodge. Managing ecological thresholds in coupled environmental–human systems. *Proceedings of the National Academy of Sciences*, 108(18):7333–7338, 2011.

[51] E. Huang and M. Goetschalckx. Strategic robust supply chain design based on the pareto-optimal tradeoff between efficiency and risk. *European Journal of Operational Research*, 237(2):508–518, 2014.

[52] A. Jabbarzadeh, B. Fahimnia, and S. Seuring. Dynamic supply chain network design for the supply of blood in disasters: a robust model with real world application. *Transportation Research Part E: Logistics and Transportation Review*, 70:225–244, 2014.

[53] A. T. Karunanithi, H. Cabezas, B. R. Frieden, and C. W. Pawlowski. Detection and assessment of ecosystem regime shifts from fisher information. *Ecology and society*, 13(1), 2008.

[54] E. Keyvanshokooh, S. M. Ryan, and E. Kabir. Hybrid robust and stochastic optimization for closed-loop supply chain network design using accelerated benders decomposition. *European Journal of Operational Research*, 249(1):76–92, 2016.

[55] J. Kim, M. J. Realff, J. H. Lee, C. Whittaker, and L. Furtner. Design of biomass processing network for biofuel production using an milp model. *Biomass and bioenergy*, 35(2):853–871, 2011.

[56] Y. Kim and S. Park. Supply network modeling using process graph theory: a framework for analysis. In *SICE-ICASE, 2006. International Joint Conference*, pages 1726–1729. IEEE, 2006.

[57] R. Kollmann, S. Maier, K. Shahzad, F. Kretschmer, G. Neugebauer, G. Stoeglehner, T. Ertl, and M. Narodoslawsky. Waste water treatment plants as regional energy cells–evaluation of economic and ecologic potentials in austria. *Chemical Engineering*, 39, 2014.

[58] E. Konig, Z. Sule, and B. Bertók. Comparison of optimization techniques in the p-graph framework for the design of supply chains under uncertainties. *INFORMATION SECURITY*, page 35.

[59] M. Kouzu, T. Kasuno, M. Tajika, S. Yamanaka, and J. Hidaka. Active phase of calcium oxide used as solid base catalyst for transesterification of soybean oil with refluxing methanol. *Applied Catalysis A: General*, 334(1):357–365, 2008.

[60] S. J. Lade and S. Niiranen. Generalized modeling of empirical social-ecological systems. *Natural Resource Modeling*, 30(3):e12129, 2017.

[61] H. L. Lam, P. S. Varbanov, and J. J. Klemeš. Optimisation of regional energy supply chains utilising renewables: P-graph approach. *Computers & Chemical Engineering*, 34(5): 782–792, 2010.

[62] M. A. Lariviere. Supply chain contracting and coordination with stochastic demand. *Quantitative models for supply chain management*, 17(2):235–268, 1999.

[63] Y. H. Lee and S. H. Kim. Production–distribution planning in supply chain considering capacity constraints. *Computers & industrial engineering*, 43(1):169–190, 2002.

[64] S. Leung, Y. Wu, and K. Lai. A stochastic programming approach for multi-site aggregate production planning. *Journal of the Operational Research Society*, 57(2):123–132, 2006.

[65] N. M. Levine, K. Zhang, M. Longo, A. Baccini, O. L. Phillips, S. L. Lewis, E. Alvarez-Dávila, A. C. S. de Andrade, R. J. Brienen, T. L. Erwin, et al. Ecosystem heterogeneity determines the ecological resilience of the amazon to climate change. *Proceedings of the National Academy of Sciences*, 113(3):793–797, 2016.

[66] P. Longinidis and M. C. Georgiadis. Integration of financial statement analysis in the optimal design of supply chain networks under demand uncertainty. *International journal of production economics*, 129(2):262–276, 2011.

[67] A. S. MacDougall, K. S. McCann, G. Gellner, and R. Turkington. Diversity loss with persistent human disturbance increases vulnerability to ecosystem collapse. *Nature*, 494 (7435):86, 2013.

[68] J. F. Magee. *Decision trees for decision making*. Harvard Business Review, 1964.

[69] M. Marufuzzaman, S. D. Eksioglu, and Y. E. Huang. Two-stage stochastic programming supply chain model for biodiesel production via wastewater treatment. *Computers & Operations Research*, 49:1–17, 2014.

[70] A. L. Mayer and A. H. Khalyani. Grass trumps trees with fire. *Science*, 334(6053):188–189, 2011.

[71] A. L. Mayer, C. W. Pawlowski, and H. Cabezas. Fisher information and dynamic regime changes in ecological systems. *Ecological modelling*, 195(1-2):72–82, 2006.

[72] A. L. Mayer, C. Pawlowski, B. D. Fath, and H. Cabezas. Applications of fisher information to the management of sustainable environmental systems. In *Exploratory data analysis using Fisher information*, pages 217–244. Springer, 2007.

[73] L. D. Mech, S. Barber-Meyer, J. C. Blanco, L. Boitani, L. Carbyn, G. Delgiudice, S. H. Fritts, D. Huber, O. Liberg, B. Patterson, et al. An unparalleled opportunity for an important ecological study. *Bioscience*, 67(10):875–876, 2017.

[74] S. Mirzapour Al-e Hashem, A. Baboli, and Z. Sazvar. Integrated synthesis of process and heat exchanger networks: algorithmic approach. *Applied Thermal Engineering*, 21(13–14): 1407–1427, 2001.

[75] S. Mirzapour Al-e Hashem, A. Baboli, and Z. Sazvar. A stochastic aggregate production planning model in a green supply chain: Considering flexible lead times, nonlinear purchase and shortage cost functions. *European Journal of Operational Research*, 230(1):26–41, 2013.

[76] C. Mlot. Two wolves survive in world's longest running predator-prey study, April 2017.

[77] C. Mlot. Classic wolf-moose study to be restarted on isle royale, 2018.

[78] I. B. Mohamed, W. Klibi, and F. Vanderbeck. Designing a two-echelon distribution network under demand uncertainty. *European Journal of Operational Research*, 280(1):102–123, 2020.

[79] F. Oliveira, I. E. Grossmann, and S. Hamacher. Accelerating benders stochastic decomposition for the optimization under uncertainty of the petroleum product supply chain. *Computers & Operations Research*, 49:47–58, 2014.

[80] A. Osmani and J. Zhang. Stochastic optimization of a multi-feedstock lignocellulosic-based bioethanol supply chain under multiple uncertainties. *Energy*, 59:157–172, 2013.

[81] F. Pan and R. Nagi. Robust supply chain design under uncertain demand in agile manufacturing. *Computers & operations research*, 37(4):668–683, 2010.

[82] D. Petrovic. Simulation of supply chain behaviour and performance in an uncertain environment. *International Journal of Production Economics*, 71(1):429–438, 2001.

[83] J. Ranisau, E. Ogbe, A. Trainor, M. Barbouti, M. Elsholkami, A. Elkamel, M. Fowler, et al. Optimization of biofuel production from corn stover under supply uncertainty in ontario. *Biofuel Research Journal*, 4(4):721–729, 2017.

[84] A. Rezaee, F. Dehghanian, B. Fahimnia, and B. Beamon. Green supply chain network design with stochastic demand and carbon price. *Annals of Operations Research*, 250(2): 463–485, 2017.

[85] V. Rico-Ramirez, M. A. Reyes-Mendoza, P. A. Quintana-Hernandez, J. A. Ortiz-Cruz, S. Hernandez-Castro, and U. M. Diwekar. Fisher information on the performance of dynamic systems. *Industrial & engineering chemistry research*, 49(4):1812–1821, 2010.

[86] E. H. Sabri and B. M. Beamon. A multi-objective approach to simultaneous strategic and operational planning in supply chain design. *Omega*, 28(5):581–598, 2000.

[87] T. Santoso, S. Ahmed, M. Goetschalckx, and A. Shapiro. A stochastic programming approach for supply chain network design under uncertainty. *European Journal of Operational Research*, 167(1):96–115, 2005.

[88] M. Scheffer. Complex systems: foreseeing tipping points. *Nature*, 467(7314):411, 2010.

[89] M. Scheffer and S. R. Carpenter. Catastrophic regime shifts in ecosystems: linking theory to observation. *Trends in ecology & evolution*, 18(12):648–656, 2003.

[90] M. Scheffer, J. Bascompte, W. A. Brock, V. Brovkin, S. R. Carpenter, V. Dakos, H. Held, E. H. Van Nes, M. Rietkerk, and G. Sugihara. Early-warning signals for critical transitions. *Nature*, 461(7260):53, 2009.

[91] M. Scheffer, S. R. Carpenter, V. Dakos, and E. H. van Nes. Generic indicators of ecological resilience: inferring the chance of a critical transition. *Annual Review of Ecology, Evolution, and Systematics*, 46:145–167, 2015.

[92] A. W. Seddon, M. Macias-Fauria, P. R. Long, D. Benz, and K. J. Willis. Sensitivity of global terrestrial ecosystems to climate variability. *Nature*, 531(7593):229, 2016.

[93] N. Shabani and T. Sowlati. A hybrid multi-stage stochastic programming-robust optimization model for maximizing the supply chain of a forest-based biomass power plant considering uncertainties. *Journal of Cleaner Production*, 112:3285–3293, 2016.

[94] Y. Shastri, U. Diwekar, and H. Cabezas. Optimal control theory for sustainable environmental management. *Environmental science & technology*, 42(14):5322–5328, 2008.

[95] Y. Shastri, U. Diwekar, H. Cabezas, and J. Williamson. Is sustainability achievable? exploring the limits of sustainability with model systems. *Environmental science & technology*, 42(17):6710–6716, 2008.

[96] M. S. Sodhi and C. S. Tang. Modeling supply-chain planning under demand uncertainty using stochastic programming: A survey motivated by asset–liability management. *International Journal of Production Economics*, 121(2):728–738, 2009.

[97] T. L. Spanbauer, C. R. Allen, D. G. Angeler, T. Eason, S. C. Fritz, A. S. Garmestani, K. L. Nash, and J. R. Stone. Prolonged instability prior to a regime shift. *PLoS One*, 9 (10):e108936, 2014.

[98] K. Suding, E. Higgs, M. Palmer, J. B. Callicott, C. B. Anderson, M. Baker, J. J. Gutrich, K. L. Hondula, M. C. LaFevor, B. M. Larson, et al. Committing to ecological restoration. *Science*, 348(6235):638–640, 2015.

[99] K. N. Suding and R. J. Hobbs. Threshold models in restoration and conservation: a developing framework. *Trends in ecology & evolution*, 24(5):271–279, 2009.

[100] S. M. Sundstrom, T. Eason, R. J. Nelson, D. G. Angeler, C. Barichievy, A. S. Garmestani, N. A. Graham, D. Granholm, L. Gunderson, M. Knutson, et al. Detecting spatial regimes in ecosystems. *Ecology letters*, 20(1):19–32, 2017.

[101] R. R. Tan and K. B. Aviso. An extended p-graph approach to process network synthesis for multi-period operations. *Computers & Chemical Engineering*, 85:40–42, 2016.

[102] R. R. Tan, C. D. Cayamanda, and K. B. Aviso. P-graph approach to optimal operational adjustment in polygeneration plants under conditions of process inoperability. *Applied Energy*, 135:402–406, 2014.

[103] R. R. Tan, M. F. D. Benjamin, C. D. Cayamanda, K. B. Aviso, and L. F. Razon. P-graph approach to optimizing crisis operations in an industrial complex. *Industrial & Engineering Chemistry Research*, 55(12):3467–3477, 2015.

[104] J. Tick, Z. Kovacs, and F. Friedler. Synthesis of optimal workflow structure. *J. UCS*, 12 (9):1385–1392, 2006.

[105] J. Tick, C. Imreh, and Z. Kovács. Business process modeling and the robust pns problem. *Acta Polytechnica Hungarica*, 10(6):193–204, 2013.

[106] P. Tsiakis, N. Shah, and C. C. Pantelides. Design of multi-echelon supply chain networks under demand uncertainty. *Industrial & Engineering Chemistry Research*, 40(16):3585–3604, 2001.

[107] A. Valiente-Banuet and M. Verdú. Human impacts on multiple ecological networks act synergistically to drive ecosystem collapse. *Frontiers in Ecology and the Environment*, 11 (8):408–413, 2013.

[108] L. Vance, T. Eason, H. Cabezas, and M. E. Gorman. Toward a leading indicator of catastrophic shifts in complex systems: Assessing changing conditions in nation states. *Heliyon*, 3(12):e00465, 2017.

[109] A. J. Veraart, E. J. Faassen, V. Dakos, E. H. van Nes, M. Lürling, and M. Scheffer. Recovery rates reflect distance to a tipping point in a living system. *Nature*, 481(7381):357, 2012.

[110] P. Voll, C. Klaffke, M. Hennen, and A. Bardow. Automated superstructure generation and optimization of distributed energy supply systems. *ECOS*, 2012:1–13, 2012.

[111] J. A. Vucetich and R. O. Peterson. The population biology of isle royale wolves and moose: an overview, 2012. URL `https://isleroyalewolf.org/data/data/home.html`.

[112] C. C. Wilmers, E. Post, R. O. Peterson, and J. A. Vucetich. Predator disease out-break modulates top-down, bottom-up and climatic effects on herbivore population dynamics. *Ecology letters*, 9(4):383–389, 2006.