



**HAL**  
open science

# A Scenario Generation Method Exploring Uncertainty and Decision Spaces for Robust Strategic Supply Chain Capacity Planning

Achille Poirier, Raphaël Oger, Cléa Martinez

► **To cite this version:**

Achille Poirier, Raphaël Oger, Cléa Martinez. A Scenario Generation Method Exploring Uncertainty and Decision Spaces for Robust Strategic Supply Chain Capacity Planning. IN4PL 2023 - 4th International Conference on Innovative Intelligent Industrial Production and Logistics, Nov 2023, Rome, Italy. pp.126-148, 10.1007/978-3-031-49339-3\_8. hal-04428323

**HAL Id: hal-04428323**

<https://imt-mines-albi.hal.science/hal-04428323v1>

Submitted on 26 Apr 2024

**HAL** is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L'archive ouverte pluridisciplinaire **HAL**, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.

# A Scenario Generation Method Exploring Uncertainty and Decision Spaces for Robust Strategic Supply Chain Capacity Planning

Achille Poirier , Raphaël Oger, and Cléa Martinez

Centre Génie Industriel, IMT Mines Albi, Université de Toulouse, Albi, France  
{achille.poirier,raphael.oger,clea.martinez}@mines-albi.fr

**Abstract.** Strategic Supply Chain Capacity Planning (SSCCP) is an essential activity for companies to prepare their future. However, since uncertainty became an essential factor to consider in this decision-making process, existing solutions to support this process do not fully satisfy their needs anymore. Especially in terms of uncertainty space coverage while exploring and assessing scenarios associated with uncertainty sources and decision options. Therefore, this paper introduces an approach to overcome the complexity of scenario exploration and improve the uncertainty space coverage, to better support SSCCP decision-making. This approach includes a bi-objective metaheuristic that first explores a probability-impact matrix to define a relevant subspace to consider in this uncertainty space, and then uses this subspace to explore the decision space and define a relevant subspace of this decision space to assess and display to decision-makers. Then, an implementation and experiment are described and discussed, and finally avenues for future research are suggested.

## 1 Introduction

For companies, delivering products to customers is conditioned by the capacity of their associated supply chain to make the product available at the right time and place. In this regard, companies must implement decision-making processes to anticipate future capacity availabilities and requirements, and adapt either one or the other. This is called supply chain capacity planning [10]. And, when adaptation decisions require a long implementation time, for example more than a year, the associated decision-making process is usually called Strategic Supply Chain Capacity Planning (SSCCP) (or long-term supply chain capacity planning) [24]. Until a few years ago the supply chain environment was relatively stable and foreseeable [17, 19] and therefore designers of SSCCP methodologies initially based their design on this assumption [4, 19]. It is the same for the designers of software aiming at supporting these decision-making processes [22, 24]. However, this assumption has been strongly challenged during the

past few years which have shown how supply chains are sensitive to uncertainties and difficult to predict [5, 13, 23]. In addition, several authors mentioned the relationship between the consideration of uncertainty in the decision-making process with the quality of the decisions made [1, 11, 20].

As a consequence, Oger et al. [16] proposed a Decision Support System (DSS) for undertaking SSCCP in uncertain environments. This SSCCP DSS contains an information system composed of 4 modules illustrated in Fig. 1. A major contribution of this SSCCP DSS is its ability to automatically deduce an assessment model, for any supply chain, that enables the evaluation of uncertainty and decision scenarios associated for supporting SSCCP decisions. This assessment model is compatible with all possible scenarios associated with provided uncertainty sources and decision options.

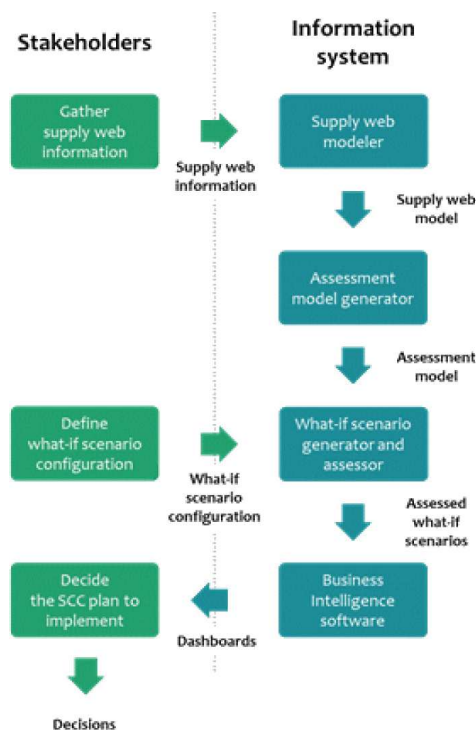


Fig. 1. SSCCP DSS proposed by Oger et al. [16].

However, a limitation of this contribution is the “what-if scenario generator and assessor” module which uses a simple cartesian product with strong filters to limit the combinatorics because the scenario space is too big to compute all scenarios for real industrial use cases. This means that the solution and uncertainty spaces are not entirely considered when defining the scenarios to assess, and that assessed scenarios are not chosen in regard to their relevance for supporting SSCCP decisions. In this paper, the “relevance for supporting SSCCP decisions” will be considered as a trade-off between the cost and the robustness of the decisions in regard to uncertainties; this will be further described in the proposition section.

Therefore, the objective of the research associated with the current paper is to propose a new approach for the “what-if scenario generator and assessor” module of the SSCCP DSS proposed by Oger et al. [16]. An approach that will enable us to better cover the solution and uncertainty spaces while assessing scenarios in regard to

their relevance for supporting SSSCP decisions. The business objective is to provide decision-makers with scenarios that will help them make robust SSSCP decisions. An important hypothesis and constraint is that the assessment model that has been deduced by the SSSCP DSS must be used as a black box that takes inputs describing a scenario configuration and returns the resulting performance indicators for the corresponding scenario.

The first section of this paper introduced the context and objective of the associated research project. The second section analyzes the existing literature related to the context and objective and formulates the research question. The third section introduces the conceptual contribution to answer the research question. The fourth section describes the technical execution of the proposition and describes the metaheuristic we use. The fifth section describes the experiment, the associated results, and the corresponding discussion. Finally, the sixth section concludes the paper and provides avenues for future research.

## **2 Literature Review**

### **2.1 Robust Supply Chain**

It has been described in the introduction that the objective of this research is to find an approach that will enable provide decision-makers scenarios that will help them make robust SSSCP decisions. Several definitions of the term robustness can be found in the literature, so it must be defined to clearly position what is meant by robustness in this paper. Therefore, a search was undertaken in the Web of Science (WoS) database to find literature reviews about robust supply chain. The used request was the following: “robust\*” (Title) and “review” OR “art” OR “advances” (Title) and “supply chain\*” (Title). It returned the following 4 papers: [8,9,12,21]

Durach et al. [8] undertook a systematic literature review on supply chain robustness, and Klibi et al. [12] and Saisridhar et al. [21] papers also contain analysis of the robustness concept. In chronological order, Klibi et al. [12] states that “a supply chain network design is robust, for the planning horizon considered, if it is capable of providing sustainable value creation under all plausible future scenarios”. Then, Durach et al. [8], who considered Klibi et al. [12] definition, defines robustness as “the ability of a supply chain to resist or avoid change”, with the resistance being the “ability of a supply chain to withstand change” and avoidance the “ability of a supply chain not to be affected by change”. Finally, Saisridhar et al. [21] defines robustness as the ability to persist without any response.

For this paper, considering previous definitions and authors’ objectives, the robustness of a strategic supply chain capacity plan should be understood as its ability to maintain performance under all plausible future scenarios without the need of making any new major adaptation decision in case uncertainties occur. It implies the choice of a performance indicator which could be different from one organization to another. Therefore, it also implies the definition of a robustness indicator to compare strategic supply chain capacity plans, that should be defined in regard to the performance indicator value for all plausible future scenarios.

## 2.2 Robust Supply Chain Planning Models

Assessing and comparing the robustness of strategic supply chain capacity plans requires appropriate tools, and more specifically appropriate assessment models. Therefore, a search was undertaken in the WoS database to find literature reviews about robust supply chain planning models. The used request was the following: “production planning” OR “supply chain planning” (Title) AND “review” OR “art” OR “advances” (Title) AND model\* (Title) AND “robust\*” OR “uncertain\*” (Title). It returned the following 4 papers: [6, 14, 15, 18]. Ciarallo et al. [6] is not a literature review but a model proposal, so not what we were looking for with this WoS request.

Among the 3 papers reviewing models for supply chain planning under uncertainty, Mula et al. [14] and Peidro et al. [18] consider all types of models while Mundi et al. [15] focuses on mathematical models. An interesting element to notice is that all 3 papers originate from the same research center. Each paper classifies planning models using a slightly different taxonomy, and we will mention Peidro et al. [18] who classify them in 4 categories:

- Analytical models (robust optimization, stochastic programming, game theory, linear programming, and parametric programming)
- Artificial intelligence-based models (reinforcement learning, genetic algorithm, evolutionary programming, fuzzy linear programming, fuzzy multi-objective programming, fuzzy goal programming, fuzzy numbers, multi-agent systems)
- Simulation models (discrete event simulation and system dynamics)
- Hybrid models

Most of the models studied by Mula et al. [14] and Peidro et al. [18] include only one or very few sources of uncertainties in the same model. While it has been highlighted that there are many sources of uncertainties regarding the demand, operations and supply aspects of supply chains. They recommend investigating new approaches to consider more uncertainties within the same model.

Some authors highlight that most of the methods for supply chain planning under uncertainty modeled the SC uncertainty by probability distributions, while uncertainty cannot always be modeled with probability distributions, especially for strategic planning, so other approaches such as fuzzy set and possibility theories should be investigated [15, 18].

Mula et al. [14] indicates that for complex supply chains the analytical approaches are often replaced by methodologies based on artificial intelligence and simulation. It is confirmed by Mundi et al. [15] who states that the most used approach to model uncertainty in the considered perimeter of their review is the scenario-based approach.

## 2.3 Research Question

Considering the context and objective described in the introduction, especially the use of an assessment model as a black box, as well as the elements found in the literature, analytical and simulation models cannot be applied to our problem. So, a solution to our problem would fall into the artificial intelligence-based models category proposed

by Peidro et al. [18]. Regarding the recommendation from Mula et al. [14] to investigate new approaches for considering more uncertainties within the same model, the assessment model proposed by Oger et al. [16] is compatible with many sources of uncertainty and decisions on operations and demand, which makes it possible to consider many scenarios, but also leads to a scenario space that cannot be entirely assessed in a reasonable amount of time.

Therefore, the business objective being to provide decision-makers with scenarios that will help them make robust SSCCP decision, the research question is the following: How to search through a scenario space including decision and uncertainty variables to recommend robust decision scenarios, while relying on a black box to assess scenarios? It also leads a sub-question: what uncertainty scenarios to consider for assessing the robustness of decision scenarios considering all uncertainty scenarios cannot be explored?

The following section will describe a conceptual proposition for answering these questions.

### **3 Conceptual Proposition**

#### **3.1 Overview**

The context of this paper is to perform an SSCCP under uncertainty by using a black box to obtain the value of a key performance indicator of the supply chain for a given scenario. The value returned by this black box is the only information we can use to help the decision-maker in his process of creating a robust supply chain. To perform an SSCCP under uncertainty, we generate scenarios to see how the supply chain reacts to the decisions and the uncertainties. The utopian way would be to generate all the possible scenarios, to then analyze them in order to determine the best decisions to take considering all the possible uncertainties. This utopic method is impossible due to the fact that doing all the combinations between the decisions and the uncertainties to generate and assess all the possible scenarios would take an unreasonable amount of time.

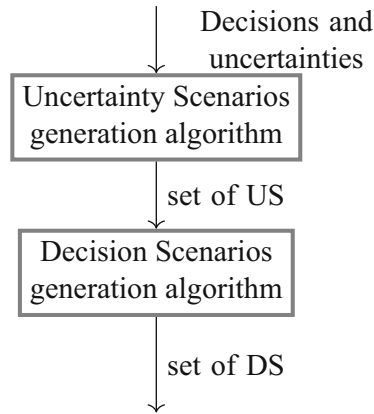
The proposed method uses multi-objective optimization to reduce the combinatorics of the problem and to give a direction to the exploration of the scenario space. It generates scenarios by maximizing or minimizing several key performance indicators of the supply chain, thus dodging scenarios that are not relevant. The relevant scenarios are the closest to the Pareto front [2]. These scenarios give us information on the impact that the uncertainties can apply to the supply chain, weighted by their likelihood, and the protection that the decisions can apply to the supply chain, weighted by the cost of setting up these decisions. We use a multi-objective metaheuristic to approximate the best set of scenarios, i.e. the Pareto front.

The scenario generation involves two steps: instead of exploring the scenario space we first explore the uncertainty space, to then explore the decision space. The exploration of the uncertainty space identifies what are the most important uncertainty combinations the supply chain should be protected from, based on a probability-impact prioritization. Then, the exploration of the decision space identifies the decision combinations having the best trade-off between cost and their robustness in regard to the previously



identified uncertainty combinations. From these two explorations we then compute scenarios to present them to the decision-maker. Figure 2 illustrates these steps.

We will first expose some definitions to clarify the elements we are working with, and the indicators that we use in this paper. Then, we will detail the exploration of both the uncertainty space and the decision space.



**Fig. 2.** Overview of our conceptual proposition.

### 3.2 Definitions

**Elements.** The SSCCP DSS creates an assessment model for the supply chain that takes a scenario configuration as input, to return the values of several key performance indicators of the supply chain for the considered scenario. The elements we work with to describe the configuration of a scenario are:

- **Decisions** that impact the supply chain, with a cost for their implementation.
- **Uncertainties** that impact the supply chain, with a given likelihood to happen.

Both the decisions and the uncertainties can impact several parameters in the supply chain: demand forecast, resources available time, bill of material, routing, supply strategies, etc. For example, a decision can be to buy equipment for a production site in order to increase the resources available time, or the launching of a marketing campaign to increase the demand forecast of a certain product. An uncertainty can be an unexpected increase in the demand forecast for a certain product, or an earthquake located near a production site leading to the decrease of its resources available time. The data behind the decisions and the uncertainties are not visible to the method. The only information we work with are whether a decision is taken or not, whether an uncertainty happens or not, and the key performance indicator of the supply chain for a given scenario. So we can generate a scenario by selecting decisions to take among the possible decisions, with happening uncertainties among the possible uncertainties. This scenario is then assessed with the assessment model generated, which returns the key performance indicators of the supply chain.

Since we work on a multi-period horizon, uncertainties can occur and decisions can be taken at different times. The proposal abstracts the temporal aspect by generating a

decision and an uncertainty per period of time. It means that an uncertainty becomes several uncertainties, representing one period of time each. The same principle applies to the decisions.

From these elements, we define the following concepts to describe scenarios:

- An **Uncertainty Scenario (US)** is a combination of uncertainties that happen, among all the possible uncertainties.
- A **Decision Scenario (DS)** is a combination of taken decisions, among all the possible decisions.
- A **scenario** is a combination of a **DS** and a **US**.

The following section defines the indicators that assess the US and the DS.

**Indicators.** In this paper, we consider the following indicators: **impact** and **probability** to assess the US, and robustness and cost to assess the DS.

The definition of robustness given in the literature review section stated that a performance indicator should be chosen to define the robustness on. This performance indicator is called **impact** in this paper. The choice of this indicator depends on the decision-maker's goal. This proposal can take any indicator for the impact and perform the exploration accordingly, as long as it is provided as an output of the assessment model used and that it makes sense to assess the US and the DS with it. The performance indicator chosen to represent the impact is the one that the decision-maker wants to minimize or maximize. This impact is then used to define the **robustness** indicator, which represents the ability of a DS to maintain the supply chain's impact indicator to a good value under all plausible future scenarios.

In this paper, these two indicators rely on the saturation of the resources that compose the supply chain. The impact of a scenario is defined as the number of resources of the supply chain that have to be used for longer than they are available. In other words, if the utilization rate of a resource (ratio of required time over available time) is greater than 100%, then it is saturated. The required time and the available time are both obtained with the SSCCP DSS. This utilization rate indicator has been chosen for defining the impact because it is a typical indicator used for capacity planning [3]. Figure 3 displays a small example of how saturation is computed. The formula of the robustness indicator is described in the decision scenario generation section.

The **cost** indicator is the sum of the costs of all taken decisions that compose the scenario, this value is obtained with the SSCCP DSS.

The **probability** indicator assesses the likelihood of a scenario. As it is easier for the decision-maker to tell if an uncertainty is very likely, likely, or unlikely to happen, instead of giving a precise probability, each uncertainty is characterized by a probability rank, expressing its likelihood. Each probability rank has a corresponding probability value in order to compute the value of the probability indicator. These values are an approximation of each probability rank. It is important to note that this indicator does not aim at giving a precise probability for each scenario, but rather an indication of the plausibility of a scenario in order to compare the them with each other. Below is an example of 5 uncertainties with their corresponding probability rank (Table 1) and the probability value for each rank (Table 2). The number of ranks defined as well



Resources	R1	R2	R3
Required time	350h	350h	200h
Available time	400h	300h	500h
Utilization rate	87.5%	116.6%	40%
Saturated	no	yes	no

**Fig. 3.** Example of the computation of resource saturation.

as the associated probability value will have an impact on the results of the proposed methodology as it impacts the probability of a scenario. The choice should depend on how discriminating we want the ranks to be. For the example given in Table 2, a scenario with 3 uncertainties of rank 1 will still be more likely than one with 1 uncertainty of rank 2.

**Table 1.** Example for the probability ranks.

uncertainty	1	2	3	4	5
probability rank	1	1	2	2	3

**Table 2.** Example of probability values for the probability ranks.

probability rank	<b>1</b> (very likely)	<b>2</b> (likely)	<b>3</b> (unlikely)
probability value	80%	50%	20%

So with this example, a scenario with uncertainties 1, 2, 4 and 5 will have a probability indicator of  $80\% * 80\% * 50\% * 20\% = 6.4\%$ .

To compare the impact of two scenarios, we compare their number of saturated resources. Since it is an integer value, a lot of different scenarios can have the same impact in terms of saturation. To break the tie of two scenarios with the same number of saturated resources, we look at the sum of utilization rates of each resource. We call this new indicator the 'relaxed saturation', it is used to tell if a scenario has a higher or a lower saturation potential than another scenario with an equal saturation. An example is portrayed in Fig. 4, where we compare the utilization rates of each resource for two different scenarios. For both scenarios, only one resource (R3) is saturated. We can compute the relaxed saturation of the scenario on the right and the scenario on the left:

$$S_{left} : 0.95 + 0.65 + 1.10 + 0.48 = 3.18$$

$$S_{right} : 0.87 + 0.65 + 1.05 + 0.55 = 3.12$$

The scenario on the right has a higher relaxed saturation, so it is considered as having a higher saturation potential.

Comparison of two scenarios utilization rates

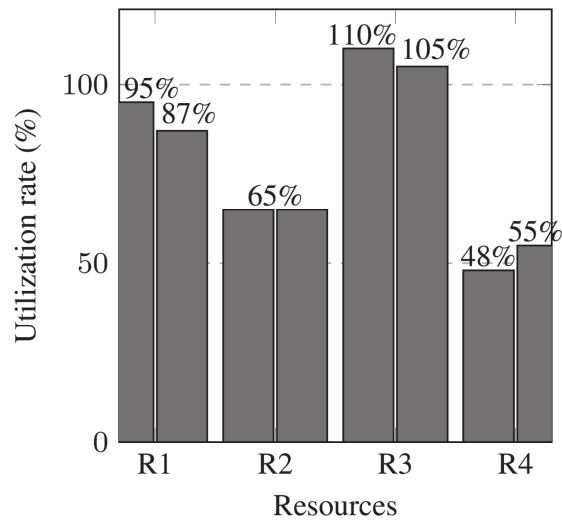


Fig. 4. Comparison of the utilization rate of each resource for two different scenarios.

### 3.3 Uncertainty Scenario Generation

The uncertainty space is composed of all of the US. This exploration is done to build a restricted set of US that gives a representative view of the risks the supply chain may be exposed to, i.e. a Pareto front of US with a good distribution (this is detailed below). To that extent, we want to find a US set that both covers probable and impactful US. Figure 5 represents the subspace we want to cover because it contains the US that will best assess the robustness of the supply chain.

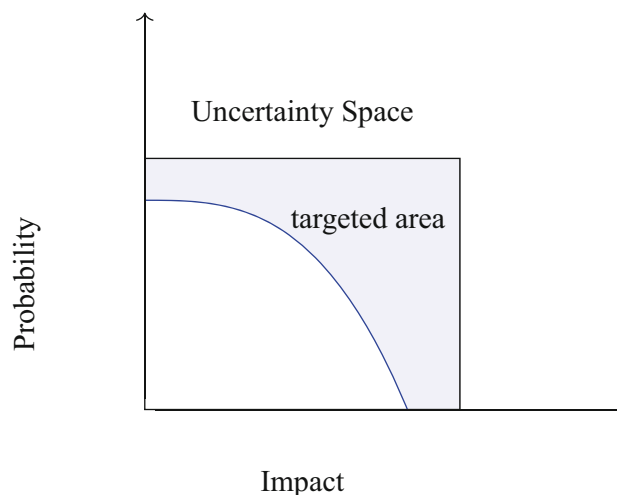


Fig. 5. Targeted area of the uncertainty space for the Uncertainty Scenarios generation.

Each US of the US set presents a trade-off between its likelihood and its impact, and to represent the uncertainty space accurately we need to have a good distribution between US that maximize the likelihood at the expense of the impact, and US that maximize the impact at the expense of the likelihood. Having too much low-impact

high-probability US means that we underestimate the risks the supply chain is exposed to, and having too much high-impact low-probability US means that we overestimated the risks to which the supply chain is exposed. This is why we aim at generating a set of US whose impact gradually increases and whose probability gradually decreases. If the density of the US found is unbalanced between several zones of the targeted uncertainty subspace (i.e. the distribution is bad), it means that the zone with the lower density of US is underrepresented, leading to a DS generation that takes less account of this zone. For example, if we find a majority of US with a high impact and a low probability, and a minority of US with low impact and a high probability, the DS set generated in regard to this US set will accord a low degree of importance to the low impact US. This would lead to the generation of a DS set that are designed mostly to prevent catastrophic scenarios, whereas we want the DS to be designed in regard to the entire targeted uncertainty subspace. Neglecting the low-impact US promotes the generation of DS with more decisions to set up in order to negate the impact of the over-represented high-impact US, thus raising the cost of the DS. This distribution criterion also allows the reduction of the size of the US set. Since the DS generation is done according to the US set, the larger the US set is, the heavier the computation of the DS generation is (more detail in Sect. 3.4). To reduce the runtime, we fix a limited size for the US set. The selection of the US in the final set is done with the goal to have the best possible distribution. (more detail in Sect. 4.1: Selection and Crowding)

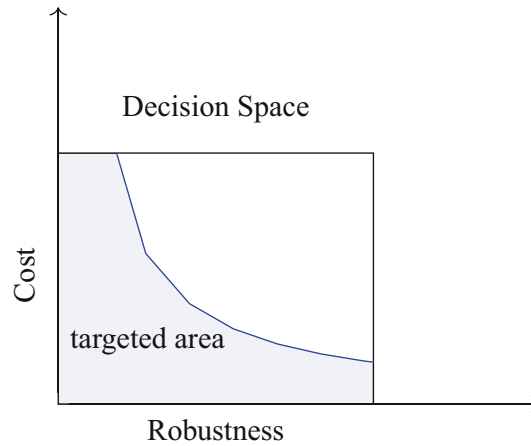
In the US generation, we want to maximize the likelihood and the impact of the US in order to find the most challenging set of US for the DS that will be generated at the next step. The US generation is constrained by an assessment model taking a scenario as input (i.e., a US combined with a DS). To address that, we first explore the scenario space to find scenarios with a maximal likelihood and/or maximal impact. From each scenario, we then extract their corresponding US and add it to the US set. In the end, the US set contains US that are very likely to happen and/or with a high impact, (i.e., a Pareto front of the US). We could explore a subspace of the scenario space by considering only the uncertainties with the nominal DS (i.e. working with scenarios without decisions), but it would mean that we build a set of US that assesses the robustness of the supply chain without any decisions taken. By allowing the decisions to be included in the search we consider all the forms that the supply chain can take with the given decisions. For example, if there is a high risk of a hurricane where a supplier is located but this supplier is not used in the nominal scenario, then considering only the nominal decision scenario would mean that the hurricane does not impact the supply chain, which is false. Indeed, if the decision to highly use this supplier is taken, the uncertainty of the hurricane becomes very impacting. Thus it must be taken into account before making the decision to rely heavily on this supplier.

### **3.4 Decision Scenario Generation**

Now that we have generated a set of US with maximal impact and/or maximal likelihood, we need to generate a set of DS to be returned to the decision-makers for recommendations. This set of DS should be a subspace of the decision space composed of all the DS.

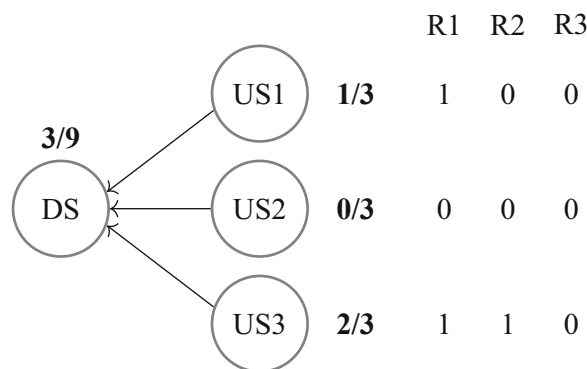
The idea is to search for DS that presents a trade-off between maximizing the robustness of the supply chain (according to the US subspace) and minimizing the costs of

the decisions. Considering that, the robustness of the DS is defined in regard to the set of US defined in the previous step. Figure 6 represents the subspace we want to cover because it contains the DS that will best protect the supply chain. It is important to note that with our robustness indicator, the lower the indicator is, the more robust the supply chain is.



**Fig. 6.** Targeted area of the decision space for the Decision Scenarios generation.

In order to evaluate a specific DS, we combine it with every US of the US set (generated in the previous step), creating a set of scenarios. Once this set of scenarios is assessed, we aggregate their results to get an indicator for the studied DS to express how much it protects the supply chain against the US set. This is portrayed in Fig. 7, where the saturation of each resource is expressed by a 1 (saturated) or a 0 (non-saturated) for each scenario that represents a US of the US set. These saturations are then summed to obtain the saturation of each scenario representing a US, then the saturation of each of these scenarios is summed to obtain the saturation of the DS. We can note that there are two levels of aggregation, one at the US level, and one at the DS level.



**Fig. 7.** Example of the aggregation of saturation for a DS, with 3 US and 3 resources.

We now have the aggregated saturation of a given DS, in relation to each US of the US set. The lower the aggregated saturation is for a DS, the better protection it offers to the supply chain from the risks approximated by the US set.

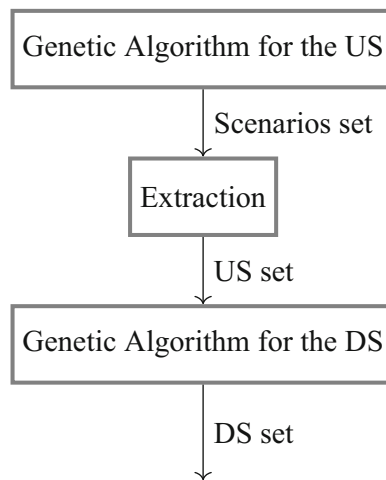
Now that we can compare two DS, we can explore the decision space to find the set of DS that will be presented to the decision-makers.

## 4 Technical Proposition

To explore the uncertainty space and the decision space, we use a bi-objective Genetic Algorithm (GA). It uses the dominance criterion to compare two solutions and find the Pareto front of a given space [2,7].

### 4.1 Exploration Method

**Overview.** Both the uncertainty space and the decision space are explored with a Genetic Algorithm (GA), differentiated by what a solution represents and how the objective values of a solution are computed. Figures 8 sum up how we explore these spaces with a genetic algorithm.



**Fig. 8.** How the genetic algorithm is used in the scenario generation.

For the uncertainty space, a solution is a scenario represented by a vector of binary variables, of which one part represents whether or not a decision is taken, and the part of which represents whether or not an uncertainty happens. The objective values are computed by assessing the solution scenario with the SSCCP DSS to obtain the saturation indicator and by a simple computation for the probability indicator. (full details in Sect. 3.2)

For the decision space, a solution is a DS, represented by a vector of binary variables that represent whether or not a decision is taken. The objective values are computed by generating and assessing the scenarios that result from the combination of the solution (i.e. a DS) with every US of the US set. These assessments are then aggregated to obtain the saturation indicator for the solution. The cost indicators are obtained with a simple computation. (full details in Sect. 3.2: Indicators)

Whether a solution represents a scenario or a DS, it will be treated as a binary vector for the rest of this section.

**Genetic Algorithm.** A GA takes as **parameters** a number of generations, a population size, and a mutation probability.

It generates an **initial population** of solutions, then improves it to create a new generation of solutions. This improvement is repeated until the number of generations specified in the parameters is reached.

To **improve** a population of solutions, an elite pool of solutions is selected to become the parent solution. This pool of parent solutions is then crossed to generate child solutions, with a chance at each crossing to have a mutation when generating a child solution.

To **select** the solutions for the next generation population, a ranking is applied to candidate solutions, composed of the parent and child solutions. It is done by computing a succession of Pareto fronts: the first Pareto front is computed to find the rank 1 solutions, these solutions are then removed from the candidates. The Pareto front is then computed for the remaining candidates to find the solutions of rank 2, and remove it from the candidates. This is repeated until all solutions are ranked.

The **next generation** population is formed of the  $i$  best rank solutions, where the  $i^{th}$  rank is the last rank of solutions we can insert into the new generation population without exceeding the population size. Solutions from the  $i + 1^{th}$  rank are selected with a crowding operator: it ranks the solution of a front by their distance from their neighbor solutions. To ensure a good distribution, the solutions with the higher distance are selected. We add crowded solutions to the new generation population until the new generation population is full.

The following subsection will go into detail about the choices made in our genetic algorithm and explain the reasons for these choices.

**Initial Population.** The initial population is generated in two steps: the generation of a mandatory solution, and the generation of random solution. The mandatory solutions are the unitary solution and the empty solution. The unitary solutions are all the possible solutions with only one variable to one and every other one to 0, and the empty solution is the solution with every variable to 0. These mandatory solutions are here to make sure that every decision and/or uncertainty has a chance to be considered. The random solutions are simply solutions with every bit set to 1 or 0 randomly, their role is to add diversity.

Because of the mandatory initial solutions, the size of the population must be higher than the number of variables. The gap between the number of mandatory initial solutions and the size of the population defines the number of random solutions.

**Elitism.** We use the ranking operator to select the elite solutions. The elite solutions are the solutions in the  $n^{th}$  Pareto front and below. The more Pareto fronts (called **ranks** from now on) are included, the more diverse the solution can be. On the other hand, the fewer ranks included, the faster it converges. We have made the choice to select the 3 best ranks to select the elite solutions.

**Crossover and Mutation.** The crossover operator generates child solutions from two parent solutions. We use a random “cut” in the parent binary vectors. A cut index is chosen at random, we first generate a child whose variables take the value of parent



1 if they are before the cut index, and take the value of parent 2 if they are after the cut index. We then generate a second child with the same principle but in reverse: its variables take the value of parent 2 if they are before the cut index, and take the value of parent 1 if they are after the cut index. Here is an illustrative example:

parent 1 [0, 0, 0, 0, 0, 0, 0, 0]  
parent 1 [1, 1, 1, 1, 1, 1, 1, 1]  
cut at index 6  
child 1 [0, 0, 0, 0, 0, 1, 1, 1]  
child 2 [1, 1, 1, 1, 1, 0, 0, 0]

Every time a child solution is generated, there is a probability that a mutation occur, this probability is given in the parameters. To mutate, a variable is chosen at random among the binary vector, then its value is inverted.

The generation of a child solution can lead to an unfeasible solution because some decisions cannot be chosen simultaneously in practice (for example two different supply strategies for the same company and product). In that case, the SSCCP DSS can't assess the scenario and does not return any value for the scenario's key performance indicators. To avoid that, we generate compatible solutions from this unfeasible solution. A compatible solution is a solution where all the variables at 1 that should not be at 1 at the same time are set to 0, with the exception of at most one of those variables. Here is an illustrative example, with incompatible variables in bold:

unfeasible solution:

[0, **1**, **1**, **1**, 0, 0, 1, 1]

compatibles solutions:

[0, **0**, **0**, **0**, 0, 0, 1, 1]

[0, **1**, **0**, **0**, 0, 0, 1, 1]

[0, **0**, **1**, **0**, 0, 0, 1, 1]

[0, **0**, **0**, **1**, 0, 0, 1, 1]

The crossover generates two child solutions unless a child solution is unfeasible, in this case, we generate as many solutions as incompatible variables (plus one for the solutions with all the incompatible variables at 0).

**Selection and Crowding.** The parent solutions and the child solutions are candidates for the new generation population. These candidate solutions are ranked, and the solutions of the  $i$  best ranks (Pareto fronts) are added, where the  $i^{th}$  rank is the last rank of solutions we can insert into the new generation population without exceeding the population size. To complete the population we select solutions from the  $i+1^{th}$  rank to add it to the new generation population. This selection is made with a crowding operator. This operator computes a distance value for each solution as described in [7], at the difference that the distance value is not computed with a cuboid, but by adding the difference of the values of the neighbor solutions (Manhattan distance). Here is an example with 3 solutions with their objectives values:

Solution 1: (23, 6)  
Solution 2: (12, 14)  
Solution 3: (5, 19)

We compute the distance value of solution 2 thanks to its neighbor solutions 1 and 3. The difference between solutions 1 and 3 on the first objective is  $|23 - 5| = 18$ , and  $|6 - 19| = 13$  on the second objective. By adding these two differences we obtain the distance value  $18 + 13 = 31$  for the solution 2.

The distance is computed for each solution, except for the two extreme solutions since they have only one neighbor solution. These extreme solutions are always kept to maintain the size of the front, so an infinite distance value is assigned to these solutions.

The candidate solutions are then added to the new generation population by ascending distance value.

This crowding is also applied to the US set for two reasons: to limit the size of the US set in order to reduce the runtime, and to ensure a good distribution between the US that maximize the likelihood at the expense of the impact, and the US that maximize the impact at the expense of the likelihood (see Sect. 3.3).

## 4.2 Dashboard

After the GA returns a set of US and a set of DS, we add some data to the result to ease the understanding, these data are described in the following subsection. After this adjustment, we display those results as clearly as possible to the decision-maker.

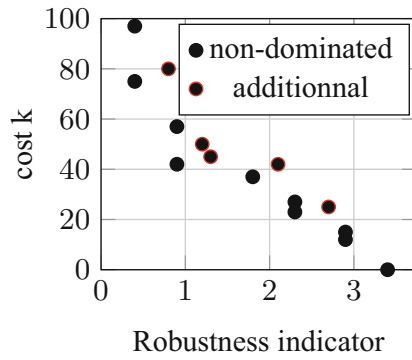
**Additional Data.** Before displaying the result, we choose to add every unitary US and the empty US to the US set in order to give visibility to the decision-maker on the impact of each uncertainty independently. We can do this without worrying about the distribution between US, because the DS are already computed with a set of US that has this property. This will only affect how they are displayed, and not how they are obtained.

To give the decision-maker room for manoeuvre, we define a visualization tool that we call a Pareto “band”: this is a Pareto front plus a few non-optimal but close solutions. An example is displayed in Fig. 9. In our proposal, we choose to take a number of best ranks within the final solutions of the genetic algorithm to compose the Pareto band.

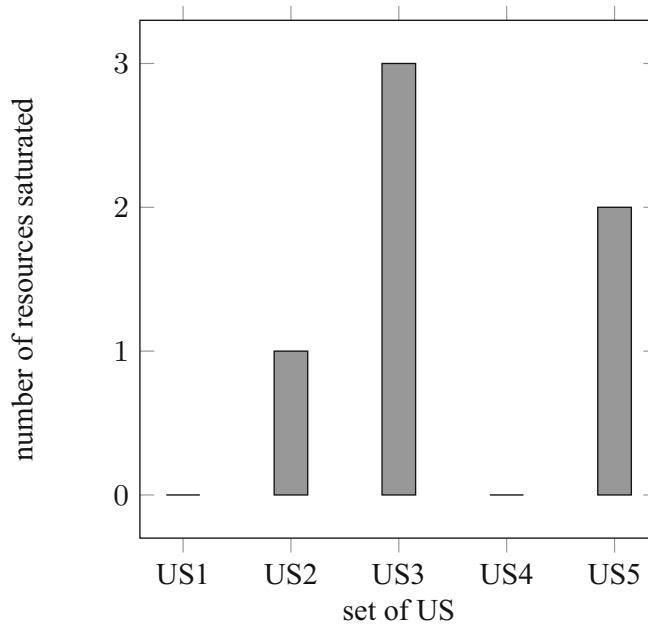
**Displaying the Results.** The resulting DS are displayed on a graph that represents their cost and their robustness indicator assessed with the US set (Fig. 9).

For each DS presented, the decision-makers can view how each US of the US set affects the supply chain with the selected DS (Fig. 10).

In addition to viewing how a DS is affected by the US of the US set, the decision-maker can view how the utilization rate of each resource of the supply for each US (Fig. 11)



**Fig. 9.** Example of a Pareto “band” of Decision Scenario (DS).



**Fig. 10.** Example of how the uncertainty scenario (US) are displayed to the decider for a given decision scenario (DS).

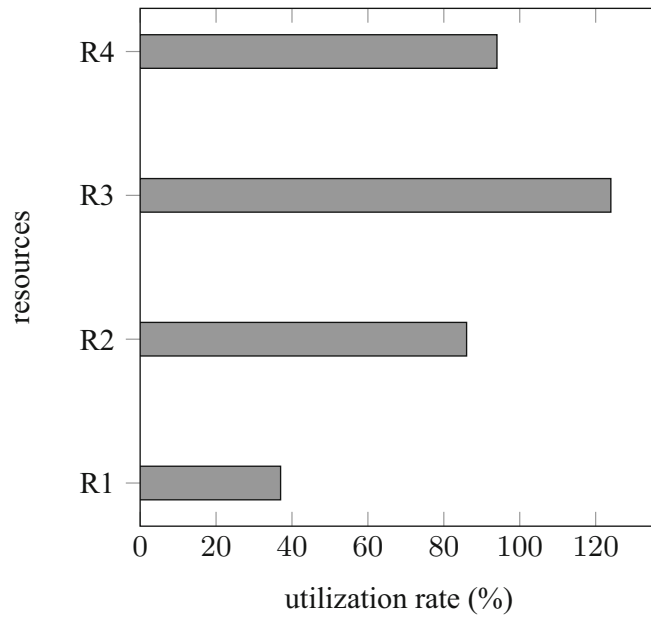
## 5 Results and Discussion

### 5.1 Use Case

The proposal is tested on an illustrative use case. The associated model, following the metamodel described in Oger et al. [16], is composed of 3 periods of time, 7 organizations, 5 product categories, 4 resource categories, 6 resources, 12 abilities, 4 demand forecast, 9 decision options (including 2 supply strategies) and 7 uncertainties. After the abstraction of the periods, we have 27 decisions and 21 uncertainties for the decision and uncertainty spaces.

The parameters for the algorithm are as follows:

- Population size: 10 random solutions + number of mandatory solutions (see 4.1: Initial Population). So  $10+27+21 = 48$  for the search of the US, and  $10+27 = 37$  for the search of the DS.



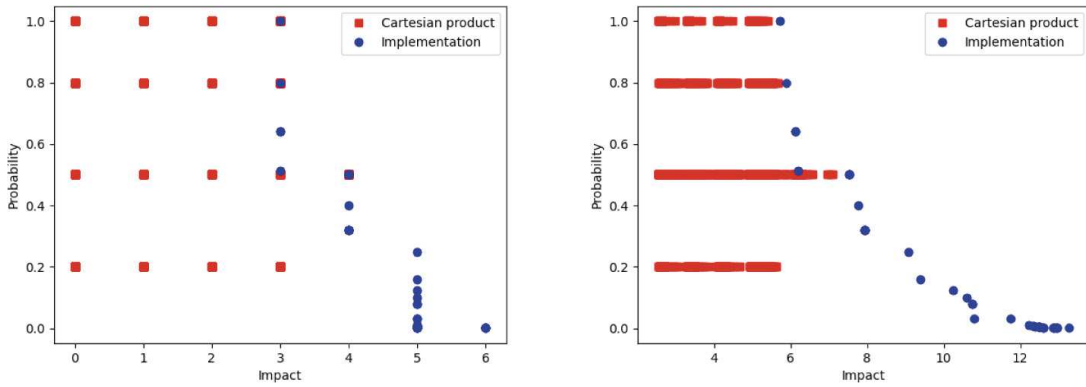
**Fig. 11.** Example of how the resources are displayed to the decider for a given scenario (a decision scenario with one of the uncertainty scenarios).

- Number of generations: 10
- Mutation probability: 5%
- Number of ranks selected for the elitism: 3
- Number of US in the US set: 10
- Number of ranks chosen for the Pareto band: 4

The values of these parameters were determined empirically for the considered use case. The addition of 10 random solutions is enough to help the search to reach better solutions without adding too much computing time. The algorithm converges within about 9 generations, so setting the number of generation to 10 ensure the algorithm converges and potentially find better solutions, the fewer generations the less computational time. A mutation probability of 5% is small enough to not perturb the convergence of the algorithm, but still allow him to escape local optima and find better solutions. The selection of 3 ranks for elitism enables a fast convergence and still provides good solutions. 10 US in the US set is enough to have a good distribution of probability within the selected US, the less US in the US set the less computational time. 4 ranks for the Pareto band seems a nice trade-off to have some non-optimal points beneath the Pareto front while remaining close to the front.

## 5.2 Uncertainty Scenarios Generation

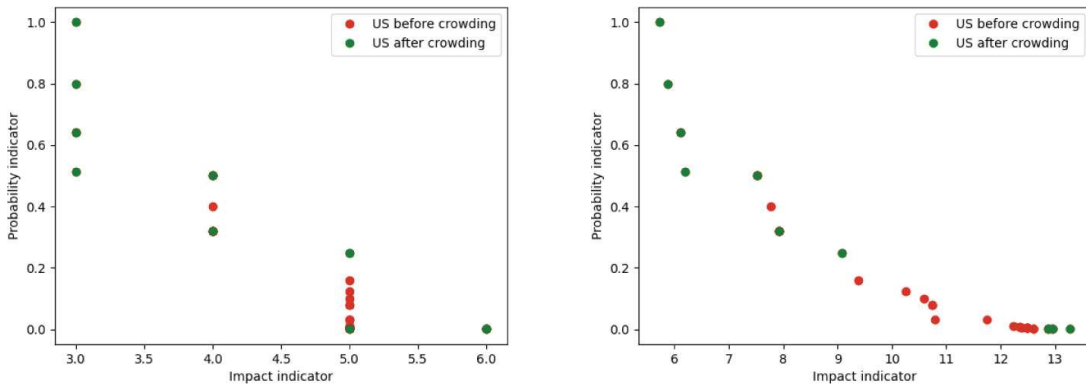
Figures 12 compare the results of the US generation between the proposal and the initial method (filtered cartesian product). Each point is a scenario, from which the US are extracted to form the US set that will assess the DS. We have a good distribution for the high probabilities, but the density increases as the probability decreases and the impact increases. With the relaxed saturation we can see that there is a concentration of high-impacting scenarios with an almost zero probability. We expect that a portion of these scenarios will be pruned with the selection of the final US set.



**Fig. 12.** Comparison of the scenarios obtained with the GA and the filtered cartesian product (left), and the same scenarios with relaxed saturation (right).

We observe that the proposed approach finds a lot more interesting scenarios than the initial method, i.e. scenarios that present a good trade-off between saturation and probability. All the scenarios from the initial method are worse or equivalent to the scenarios from the proposal. The high density of high-impacting scenarios with a low probability is due to the fact that there may be a lot of scenarios that combine several impacting uncertainties, while high-probability scenarios are obtained with the activation of a unique high probable uncertainty. A way to avoid that would be to add constraints to limit the number of activated uncertainties for the lower probability ranks.

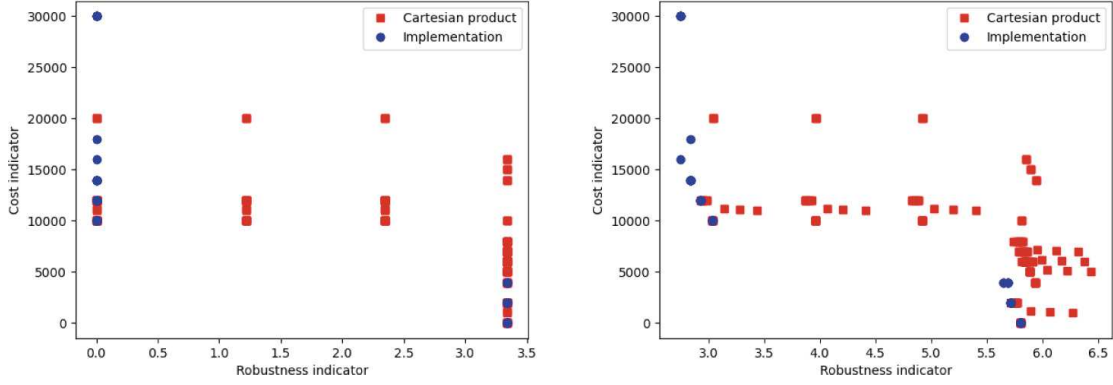
The selection of the US to form the US set is depicted in Fig. 13. We see that the selection of scenarios corresponds to our expectation since the distribution of scenarios probability is greater than before, in addition to limiting the size of the US set. We see with the relaxed saturation that the group of impacting scenarios with an almost zero probability is well pruned, as expected.



**Fig. 13.** Crowded Pareto front of scenarios for the Uncertainty Scenarios generation with relaxed saturation (left), and the same result with relaxed saturation (right).

### 5.3 Decision Scenarios Generation

Figures 14 compare the result of the DS generation between the proposal and the initial method, after the recombination. These DS are assessed with the set of US previously



**Fig. 14.** Comparison of the Decision Scenarios obtained with the proposal and the initial method (left), and the same result with relaxed saturation (right).

generated and selected. We see in Fig. 14 that the scenarios are well distributed within the cost, with only 2 values for the saturation and a clear gap between them. This is due to a critical point being reached with a given combination of decisions, allowing the saturation to be nullified. This is highlighted with the relaxed saturation that barely changes within the 2 group of DS. This is because the data set is a small test case.

We observe in Fig. 14 that the DS found with the implementation are equivalent or better than the ones found with the initial method. The initial method finds a lot of DS with the same cost and various saturation, whereas the implementation finds only DS that offer a good trade-off between cost and saturation. The implementation finds only better DS for a cost under 10 000. With the relaxed saturation we see that the DS generated by the implementation are better than the one generated with the initial method, because even the DS that have the same cost and the same saturation have a lower relaxed saturation, meaning that they are potentially less inclined to have saturation.

#### 5.4 Runtime and Number of Scenarios Generated

The Table 3 below shows the runtime of the initial method compared to the proposal. It presents the total runtime of the initial method, and the runtime of the proposal US generation, DS generation, and total time (US generation and DS generation). The computer used to perform these has 16 go of RAM, and a processor Intel Core i7-7700HQ 2.80 GHz.

**Table 3.** Runtime comparison (in sec).

Method	Runtime	Assessment model runtime
Initial method	24.65	27.17
US generation	9.73	9.6
DS generation	29.09	28.98
Total	38.82	38.58



The Table 4 below shows the number of scenarios generated and assessed by the initial method compared to the proposal. It presents the number of scenarios for the initial method, and for the proposal US generation, DS generation, and total (US generation and DS generation). We observe that the assessment model takes most of the runtime and that the runtime of the proposal and the initial method are close for this use case.

**Table 4.** Comparison of the number of scenarios generated and assessed.

Method	number of scenarios
Initial method	3388
US generation	773
DS generation	3200
Total	3973

We observe that the DS generation is the part that requires the assessment of the highest number of scenarios. The initial method and the proposal generate an equivalent amount of scenarios for this use case.

## 5.5 Discussion

The runtime of our proposal is equivalent to the runtime of the initial method. Further tests are needed to see how the runtime scales with the size of the data set. Besides, the quality of the results obtained with the proposal is better than the results from the initial method. First, the SD are better assessed because the US generated cover a wider space in the probability-impact space, with an even distribution. In addition to this, the generated DS offer a better protection against the saturation for a similar cost, plus some high-cost DS.

These results are slightly better than the initial method, but we need to keep in mind that this is only for a small data set. We make the hypothesis that it will perform better than the initial method in terms of quality, because the larger the data set gets, the more the combinatorics become mandatory to cover the probability-impact space and the saturation/cost space efficiently.

## 6 Conclusion

This paper proposes a contribution to support decision-makers in making robust Strategic Supply Chain Capacity Planning (SSCCP) decisions. The literature review section positions it in the frame of the robust supply chain planning field. It is in the continuity of the research undertaken by Oger et al. [16], which proposes a system that automatically deduces an SSCCP assessment model from supply chain data. In the initial proposal, a scenario generation and assessment module that uses a simple cartesian product with strong filters to limit the combinatorics has been implemented to illustrate the use of the assessment model. However, assessed scenarios are not chosen in regard

to their relevance for assessing the robustness of the decisions. Therefore, the objective of this paper is to automate the recommendation of robust SSCCP decisions by proposing a new scenario generation approach that takes advantage of the deduced assessment model. An objective that implies the constraint of using the assessment model as a black box, and so limits possible approaches, but has the advantage of being able to automate the recommendation of robust SSCCP decisions from business data. This led to the following research question: how to search through a scenario space including decision and uncertainty variables to recommend robust decision scenarios, while relying on a black box to assess scenarios?

A conceptual proposal to answer this question has been described, it is composed of two steps: a first step to define a set of uncertainty scenarios that will be considered as the most important to be protected from, based on a probability-impact approach. A second step to define the decision scenarios to recommend, based on their performance in regard to all the uncertainty scenarios defined in the previous step. Then, this conceptual proposal has been completed with a technical proposal that uses genetic algorithms to implement the proposed approach. Finally, it has been experimented on an illustrative use case to evaluate the ability of the approach to answer the research question. During the experiment, the results have been compared with the ones obtained from the filtered cartesian product, and it has shown that the approach finds more robust decision scenarios.

This contribution should help companies gain better visibility about the robustness of their possible decisions in regard to the multitude of uncertainties. One of the difficulties of this domain is that the scenario space cannot be entirely explored because of its size, and this contribution is a step towards the improvement of the quality of the coverage (and so the visibility) to help companies make robust decisions. An important element to note is that this proposition can take into account any key performance indicators to be robust to as long as the assessment model can provide it.

However, there are still a number of limitations to be addressed to fully conclude about the proposal. First, the experiment has been performed on a small illustrative use case, testing the proposal on several real use cases would allow us to conclude on the applicability and scalability (in terms of runtime and quality of result) of the proposal. Second, a source of approximation is the choice of the probability indicator, an alternative indicator could be proposed. Finally, in the US generation step, the approach to select the US set in the crowding part is based on the Manhattan distance, a choice that could be discussed. We could study the influence of this choice by trying another approach such as Hamming distance. Third, the impact of the parameters of the algorithm could be studied in order to find a setting that can be applied to any use case, or to find a way to automate the setting of the parameters. The impact of the value for the probability ranks could also be studied to characterize the impact of these ranks and associated values on the proposed methodology.

## References

1. Boonyathan, P., Power, D.: The impact of supply chain uncertainty on business performance and the role of supplier and customer relationships: comparison between product and service organizations. In: DSI Mini Conference on Services Management, Pittsburgh, USA, pp. 391–402 (2007)
2. Coello, C.A., Lamont, G.B., Van Veldhuizen: Pareto multi objective optimization. In: Evolutionary Algorithms for Solving Multi-Objective Problems (2002). <https://doi.org/10.1007/978-0-387-36797-2>
3. Chapman, S., Arnold, T., Gatewood, A., Clive, L.: Introduction to Materials Management, 8th edn. Pearson, Boston (2016)
4. Christopher, M., Holweg, M.: “Supply chain 2.0”: managing supply chains in the era of turbulence. *Int. J. Phys. Distrib. Logist. Manag.* **41**(1), 63–82 (2011)
5. Christopher, M., Holweg, M.: Supply chain 2.0 revisited: a framework for managing volatility-induced risk in the supply chain. *Int. J. Phys. Distrib. Logist. Manag.* **47**(1), 2–17 (2017)
6. Ciarallo, F.W., Akella, R., Morton, T.E.: A periodic review, production planning model with uncertain capacity and uncertain demand—optimality of extended myopic policies. *Manage. Sci.* **40**(3), 320–332 (1994)
7. Deb, K., Agrawal, S., Pratap, A., Meyarivan, T.: A fast elitist non-dominated sorting genetic algorithm for multi-objective optimization: NSGA-II. In: Schoenauer, M., et al. (eds.) PPSN 2000. LNCS, vol. 1917, pp. 849–858. Springer, Heidelberg (2000). [https://doi.org/10.1007/3-540-45356-3\\_83](https://doi.org/10.1007/3-540-45356-3_83)
8. Durach, C.F., Wieland, A., Machuca, J.A.: Antecedents and dimensions of supply chain robustness: a systematic literature review. *Int. J. Phys. Distrib. Logist. Manag.* **45**, 118–137 (2015)
9. Govindan, K., Cheng, T.: Advances in stochastic programming and robust optimization for supply chain planning. *Comput. Oper. Res.* **100**, 262–269 (2018)
10. Hopp, W.J., Spearman, M.L.: *Factory Physics*. Waveland Press (2011)
11. Hult, G.T.M., Craighead, C.W., Ketchen, D.J., Jr.: Risk uncertainty and supply chain decisions: a real options perspective. *Decis. Sci.* **41**(3), 435–458 (2010)
12. Klibi, W., Martel, A., Guitouni, A.: The design of robust value-creating supply chain networks: a critical review. *Eur. J. Oper. Res.* **203**(2), 283–293 (2010)
13. Kohl, M., et al.: Managing supply chains during the COVID-19 crisis: synthesis of academic and practitioner visions and recommendations for the future. *Int. J. Logist. Manag.* (ahead-of-print) (2022)
14. Mula, J., Poler, R., García-Sabater, J.P., Lario, F.C.: Models for production planning under uncertainty: a review. *Int. J. Prod. Econ.* **103**(1), 271–285 (2006)
15. Mundi, I., Alemany, M., Poler, R., Fuertes-Miquel, V.S.: Review of mathematical models for production planning under uncertainty due to lack of homogeneity: proposal of a conceptual model. *Int. J. Prod. Res.* **57**(15–16), 5239–5283 (2019)
16. Oger, R., Lauras, M., Montreuil, B., Benaben, F.: A decision support system for strategic supply chain capacity planning under uncertainty: conceptual framework and experiment. *Enterp. Inf. Syst.* **16**(5), 1793390 (2022)
17. Olhager, J.: Evolution of operations planning and control: from production to supply chains. *Int. J. Prod. Res.* **51**(23–24), 6836–6843 (2013)
18. Peidro, D., Mula, J., Poler, R., Lario, F.C.: Quantitative models for supply chain planning under uncertainty: a review. *Int. J. Adv. Manuf. Technol.* **43**, 400–420 (2009)
19. Pinon, D., Oger, R., Lauras, M.: Supply chain evolution and supply chain capability planning methodologies: a review and gap identification. In: ILS 2018–7th International Conference on Information Systems, Logistics and Supply Chain (2018)

20. Sáenz, M.J., Revilla, E.: Creating more resilient supply chains. *MIT Sloan Manag. Rev.* (2014)
21. Saisridhar, P., Thürer, M., Avittathur, B.: Assessing supply chain responsiveness, resilience and robustness (triple-R) by computer simulation: a systematic review of the literature. *Int. J. Prod. Res.* 1–31 (2023)
22. de Santa-Eulalia, L.A., D'Amours, S., Frayret, J.M., Menegusso, C.C., Azevedo, R.C.: Advanced supply chain planning systems (aps) today and tomorrow. In: *Supply Chain Management Pathways for Research and Practice*, pp. 171–200. InTech, Rijeka (2011)
23. Simangunsong, E., Hendry, L.C., Stevenson, M.: Supply-chain uncertainty: a review and theoretical foundation for future research. *Int. J. Prod. Res.* **50**(16), 4493–4523 (2012)
24. Stadtler, H., Stadtler, H., Kilger, C., Kilger, C., Meyr, H., Meyr, H.: *Supply Chain Management and Advanced Planning: Concepts, Models, Software, and Case Studies*. Springer, Heidelberg (2015). <https://doi.org/10.1007/978-3-642-55309-7>