Targeted Evidence Collection for Uncertain Supplier Selection

Document Version
Accepted author manuscript

Published in:
Expert Systems with Applications

Targeted Evidence Collection for Uncertain Supplier Selection

Edward Abel, Julio César Cortés Ríos, Norman W. Paton, John A. Keane, Alvaro A.A. Fernandes

Department of Computer Science, School of Engineering, University of Manchester, Manchester M13 9PL, United Kingdom

Abstract

The problem of selecting which suppliers, and how much of different items to order from each, involves multiple, often conflicting, criteria such as costs and delivery times. Within real world multi-criteria supplier selection problems there is inherent uncertainty involved, and consideration of its impacts and mitigation is a current and important research direction going forward within the field of supplier selection.

Uncertainty within multi-criteria supplier selection may be in relation to (i) a decision maker's ambiguous preferences, such as the importance between criteria, (ii) the suppliers' supply capacities of, and demand for, different items, and (iii) known information about suppliers with respect to the set of criteria, such as each supplier's delivery times or their average defect ratios. Whilst previous work has explored the first two of these, less work has explored uncertainty pertaining to information about suppliers in terms of the criteria and, specifically, how it could be efficiently reduced. Such uncertainty is an important problem to address, as it may have a large impact upon an order regarding its *perceived* quality compared to its *realised* quality, so reducing such uncertainty can have a significant impact.

---

*Corresponding author

Email addresses: edabelcs@gmail.com (Edward Abel),
juliocesar.cortesrios@manchester.ac.uk (Julio César Cortés Ríos),
norman.paton@manchester.ac.uk (Norman W. Paton), john.keane@manchester.ac.uk (John A. Keane), alvaro.a.fernandes@manchester.ac.uk (Alvaro A.A. Fernandes)

This paper presents a Targeted Evidence Collection (TEC) approach for efficient reduction of uncertainty, pertaining to suppliers, by looking to efficiently collect additional evidence. The approach looks to utilise and gather evidence intelligently and dynamically – by considering both the likelihood that each supplier will be part of a solution, along with a decision maker's preferences between criteria – to reduce the uncertainty and efficaciously move towards the most appropriate solution given no uncertainty. The approach is able to handle scenarios for which there are both certain and uncertain criteria present, and can take into account any number of criteria.

The TEC strategy is evaluated against alternative approaches, including an active learning based approach, for varying numbers of uncertain criteria, numbers of suppliers, and variations in a decision maker's preferences. The experimentation highlights how TEC efficiently reduces uncertainty, relating to information about suppliers with respect to the set of criteria, requiring up to three times less evidence than its competitors. In this way, TEC helps to effectively mitigate the uncertainty's adverse effects, and reduce the risks inherent within a supplier selection problem.

Keywords: supply chain management, supplier selection, uncertainty, multi-criteria decision analysis, multi-objective optimization

1. Introduction

Within Supply Chain Management (SCM), it is estimated that purchasing of goods and services constitutes up to 70% of total product costs (Ghobadian et al., 2016), rising to 80% for high-technology firms (Ghodsypour & O'Brien, 2001). Therefore, a critical decision in SCM is selecting the most appropriate suppliers to achieve the objectives of the supply chain, and ultimately help to obtain competitive advantage. Within Multi-Criteria Supplier Selection (MCSS) the problem is to determine the amounts of various items to acquire from various suppliers, such that the overall quality of the solution is most aligned with the Decision Maker (DM)'s preferences regarding multiple, often conflicting, criteria

such as price, quality and delivery times, whilst also satisfying the demand for each item-type, and additional constraints regarding budget and supply capacities.

Typical supplier selection models, such as (Ting & Cho, 2008), look to solve the MCSS problem via optimisation, under the assumption that all information is crisp and certain. For example, given a high-technology firm manufacturing motherboards, which require the purchase of varying amounts of a set of 7 major items, from a possible set of many suppliers, each supplying one of more of the items. It will be the case that each supplier will vary with regards to a set of criteria such as, costs, quality, and delivery; different suppliers will sell the same item but at different costs, some suppliers will sell the same items but with different levels of quality. In such a scenario the DM(s) can define preferences regarding the importance between the criteria, which are then used along with additional constraints regarding budget and supply capacities, such that the overall quality of the solution is most aligned to the DM's preferences.

However, within real world Supplier Selection problems there is inherent uncertainty (Sawik, 2011); consequently there is a trend in current research to cater to more practical supplier selection problems by looking to construct effective decision models that address uncertainty factors (Chai et al., 2013), instead of formulating the problem in terms of only crisp and certain conditions. Further, a significant emerging trend in supplier selection research, defined in a recent review surveying accomplishments and what lies ahead within supplier selection, is that of transiting from handling uncertainty towards risk analysis (Chai & Ngai, 2019). Here, the authors identified that although many studies seek to model uncertainty, such as modelling imprecise information via fuzzy integrated methodologies, a transition is occurring towards greater consideration of risk factors, exploring, for example, loss aversion and diminishing uncertainty's adverse effects.

Uncertainty within MCSS could be in relation to (i) a DM's subjective preferences, (ii) the demand for, and the suppliers' supply capabilities of, items, and (iii) information regarding suppliers' values (utility) in terms of the set of

criteria. For example, in our high-technology firm scenario, it may be the case that there is (i) uncertainty in relation to a DM's preference of importance between the criteria, for example, in terms of how much more important cost is compared to delivery time. Equally, there may be (ii) uncertainty in terms of the exact demand for each of the 7 item components, as well as uncertainty in terms of the quantity of different items that each supplier is able to fulfil. There may also be (iii) uncertainty in relation to information the firm has about each supplier, in relation to each criterion, such as the delivery times of the different items from a supplier or the quality of the items.

Various work has looked at uncertainty within supplier selection regarding (i) user preferences (Guneri et al., 2009, Arikan, 2013, Igoulalene et al., 2014) as well as (ii) supply and demand uncertainty (Kazaz & Webster, 2015, Zhang & Chen, 2013, Sounderpandian et al., 2008); there has been less work on (iii) uncertainty regarding suppliers' utility in terms of the criteria and specifically how such uncertainties could be efficiently reduced. However, uncertainty surrounding suppliers' utility regarding the criteria, can have a major impact upon the realised overall utility of a supplier selection solution, compared to its perceived overall utility. For example, if our high-technology firm orders items from a supplier with perceived higher item quality than in reality, this would result in an actual solution with reduced utility when realised, as some item's quality will be less than expected and desired. Equally, if our firm perceives that a supplier has higher item costs than in reality, this supplier may be overlooked within a solution, despite it representing in reality a suitable supplier to be part of a solution to best meet the DM's preferences. Identification of such uncertainty relating to suppliers, can inform a DM of the amount of risk, due to the uncertainty, associated with a given supplier selection solution. Moreover, its adverse risk effects could be diminished through actively looking to reduce this uncertainty. For example, for a supplier for which there is a lot of uncertainty in relation to its item defect ratio, we could aim to reduce the uncertainty by collecting more evidence about the supplier and its items. Such evidence could be obtained in a variety of ways such as mock orders, customer surveys and/or interview infor-

4

mation, market assessments, experts' opinions, or on-site visits to evaluate the supplier's process. Such evidence could help reduce the amount of uncertainty regarding the supplier and help mitigate any adverse effects, and consequently diminish the risk associated with a solution. Therefore, actively addressing and tackling such uncertainty is a critical issue within supplier selection and one that can have large impact and benefit to a DM.

Given that a plethora of such additional evidence could be sought, due to a large number of possible suppliers, the choice of what evidence to collect next can have a large impact upon the effectiveness of the collection, and its usefulness in impacting on reducing uncertainty. Therefore, it is crucial to consider how to collect such evidence, to be most impactful for uncertainty reduction.

This paper presents a *Targeted Evidence Collection* (TEC) strategy, for efficient reduction of uncertainty, pertaining to suppliers, within an MCSS model such as in (Ting & Cho, 2008). The TEC strategy looks to intelligently target evidence collection by dynamically considering the likelihood that each supplier will be part of a solution along with a DM's preferences regarding the set of criteria. In this way TEC looks to reduce uncertainty by swiftly gravitating towards what would be the most appropriate solution given no uncertainty. In this way, the TEC strategy embodies the ethos of an expert system in the sense that it relies on a body of knowledge - available evidence in our case - to make informed decisions about where to focus the evidence collection to reduce the uncertainty up to a point where the optimisation can provide effective outcomes.

The paper makes the following contributions:

- We introduce the TEC strategy for efficient reduction of uncertainty, within MCSS, using targeted evidence.

- The strategy is applicable to a large range of amounts of evidence collection, and we demonstrate its applicability across a varying range.

- For many real-world practical scenarios which mix uncertain and certain criteria – for example in a scenario where item costs are known with cer-

tainty but item quality is not – the strategy is able to handle both any number of criteria and the combination of uncertain and certain criteria.

- TEC is able to scale to a large number of suppliers and we demonstrate how its effectiveness only becomes more pronounced as the number of suppliers increases.

- We facilitate more informed decision making, through calculating (and visualising) the regret associated with a current solution, after any given amount of evidence collection, to help a DM ascertain the risk associated with a solution.

- The effectiveness, adaptability and scalability of TEC is evaluated through an experimental comparison against alternative approaches, including those based on active learning. The results show that TEC can achieve efficient reduction in uncertainty in the criteria, to support improved solutions using less evidence than comparable approaches.

The rest of the paper is structured as follows: related literature is presented in Section 2; Section 3 defines and describes our strategy; experimentation results and evaluation are presented in Section 4; and Section 5 concludes.

2. Literature review

MCSS has been explored by approaches that have tackled the problem within deterministic conditions and without consideration of uncertainty: Ting & Cho (2008) tackle supplier selection considering criteria of cost, quality and delivery times, assuming crisp information relating to suppliers regarding these criteria; similarly, Ayhan & Kilic (2015) tackle supplier selection, using criteria of price, quality, delivery time, and after sales performance, where past data is used to gather crisp (percentile) information about suppliers in relation to the criteria. However, within such MCSS problems there is inherent uncertainty which may be associated with (i) subjective DM preferences and opinions, (ii) supply and demand, and (iii) the values for the suppliers' utility regarding the criteria.

6

Multiple approaches have looked to address uncertainty in supplier selection in connection with subjective DM preferences and opinions. Guneri et al. (2009) use fuzzy numbers, within single item-type supplier selection, to represent uncertainty within multiple DMs' preferences between criteria, and their subjective opinions of alternative suppliers. Similarly, uncertainty relating to the aspiration levels of a DM for each criterion (as well as the demand for each item-type) has been modelled via a fuzzy mathematical model and then solved using multi-objective linear programming (Arikan, 2013). Fuzzy trapezoidal membership functions have been used in (Igoulalene et al., 2014) to model uncertainty within a group of DMs' preferences and seek a consensus on selection of a (single) supplier. The vagueness and ambiguity of experts' opinions pertaining to sustainability factors within a Sustainable Supplier Selection Problem (S-SSP) was tackled by Rabbani et al. (2019). Here, interval-valued fuzzy sets are used to represent the vague and ambiguous preferences from a group of experts. An interval-valued fuzzy decision model then calculates novel separation measures, which are used to rank the set of candidate suppliers. Group supplier selection has also been explored by Sanayei et al. (2010) for when several decision makers representing different interests are involved in defining preferences for the supplier selection problem. In this work, alternatives are evaluated by eliciting preferences from multiple DMs, via linguistic values as part of a fuzzy representation of the subjective preferences. The fuzzy preferences are then utilised within the VIKOR decision methodology to arrive at a solution. Similarly, Stević et al. (2019) use triangular fuzzy numbers, when tackling group supplier selection, to represent a group of expert's subjective opinions relating to the set of criteria. These are then aggregated into single triangle Fuzzy numbers, to be used as part of a Fuzzy Evaluation based on Distance from Average Solution (Fuzzy EDAS) to evaluate and select suppliers. An approach titled D-AHP has been proposed in (Deng et al., 2014) to tackle supplier selection using D numbers – an extension of fuzzy preference relation – to represent DMs' subjective preferences. The uncertainty involved in a DM's preferences, within Multi-criteria global supplier selection, is tackled in (Chan & Kumar,

2007), using an approach termed Fuzzy Extended Analytic Hierarchy Process (FEAHP). The approach uses triangular fuzzy numbers, to represent a DM's comparison judgements, to seek to capture the vagueness within such preferences. An approach for tackling green supplier selection problems is proposed in (Haeri & Rezaei, 2019). Here, the approach looks to consider interdependencies between criteria, such as for example, 'technology capability' having significant influence on 'pollution production', as well as take into account uncertainties associated with DMs' subjective opinions. The approach proposes a novel weight assignment model, combining a best-worst method and fuzzy grey cognitive maps to capture the interdependencies, and grey relational analysis for processing DMs' subjective opinions, to calculate a final ordinal ranking of the set of suppliers. The ramifications of a DM's subjective judgements within supplier selection are explored in (Rashidi & Cullinane, 2019), through the use of different decision methodologies. Both fuzzy DEA and fuzzy TOPSIS are used to tackle sustainable supplier selection, with criteria such as cost, quality and social responsibility. The comparison shows that the different methods resulted in differing solutions, and, although TOPSIS performed better regarding calculation complexity and sensitivity to changes, they recommend initial use of both methods to aid DM comprehension.

Multiple approaches have also looked to address uncertainty in supplier selection pertaining to supply and demand. Memon et al. (2015) look to tackle group supplier selection uncertainty in relation to demand and lead-time, as well as subjective DM preferences. Here, an approach is proposed that uses combined grey systems theory, for DM opinions regarding criteria and supplier attributes, and uncertainty theory to help model DM uncertainty and demand and lead time uncertainty, characterised as being stochastic in nature. This is used within an uncertain goal programming model to determine supplier selection and order allocation. Both supply and demand uncertainty within supplier selection is tackled in (Moghaddam, 2015). The uncertainties are handled through a fuzzy multi-objective mathematical model that looks to determine the best set of suppliers and the optimal order allocation to the selected suppliers.

In tackling the multi-criteria problem instead of eliciting subjective DM weighting of the criteria, the approach utilises monte carlo random weights over many runs to map out a set of trade-off solutions. Uncertainty pertaining to supply is tackled in (Kellner et al., 2019). Here, supply risk is defined as an objective to be minimised, within a multi-objective supplier selection problem, alongside other objectives such as cost. In considering this objective the approach looks to minimise the covariance of the logistics service between suppliers so as to mitigate, for example, suppliers in the same region all failing due to a natural disaster. The approach looks to solve the multi-objective problem by seeking to find a Pareto front of trade-off solutions between the objectives, allowing a DM to get an overall picture of the trade-offs involved and compare solutions. Kazaz & Webster (2015) analyse the differences in risk aversion, when the source of uncertainty is demand and/or supply, and how observed differences have an effect upon tractability and optimal decisions. Supplier selection and procurement with stochastic demand is tackled in (Zhang & Chen, 2013), with a nonlinear mixed integer programming model. The model looks to optimally select suppliers and allocate the ordering quantity among the selected suppliers, in order to minimise the total cost that additionally considers holding and shortage costs, along with purchase costs. Optimisation of order quantity decisions, from suppliers from developing countries, considering the loss risks due to supply uncertainty, is explored in (Sounderpandian et al., 2008). In this approach, a genetic algorithm is utilised to solve a stochastic programming model which considers the uncertainties of supply. They found that just increasing the order quantities may not be the appropriate approach when considering changes due to demand uncertainties. Uncertainty relating to demand and supply capacities of suppliers is tackled via contingency strategies in (Bilsel & Ravindran, 2011). Probability distributions are used to determine and assign backup suppliers who could be used if there is an issue, due to uncertainty, with a primary supplier selected as part of a solution. In this way, the approach looks to facilitate contingency plans to handle, for example, unexpected higher demand or supply issues from a primary supplier.

Some work has explored uncertainty of information pertaining to suppliers' utility in terms of the criteria. Uncertainty relating to criteria such as defect risk and reliability of suppliers is considered in (Sawik, 2010). Here a mixed integer model is utilised, and consideration of the risk associated with the uncertain values of such criteria – calculated as expected values using past observations – is considered via DM defined thresholds of acceptability. These are then incorporated into the model to consider purchasing costs of additional items required to compensate for defective and delayed items. The approach is looking to incorporate, and compensate for, the risk associated with the available evidence into the model, whereas in TEC we seek to actively reduce the uncertainty through efficient evidence collection.

The problem of uncertainty relating to suppliers and the item-types they supply, has been addressed via a mixed integer linear programming model (Amorim et al., 2016) focusing on food supply chain management. Here, various sources of supplier uncertainty are considered, such as raw materials availability, lead time and spot price (as well as uncertainty of demand for final products). The objective function seeks to maximise both profit and the conditional value-at-risk (Rockafellar & Uryasev, 2002). In this way, risk measures are incorporated within the supplier selection model such that worst-case scenarios are drastically reduced. In the approach a solution is sought that considers the risks of uncertainty when looking to calculate a solution, whereas in TEC our objective is to actively reduce the uncertainty.

The problem of supplier selection under conditions of total quantity discount, where suppliers offer potentially fluctuating and uncertain discounts based on the quantity of products purchased, along with upfront fixed discount activation costs, is explored in (Manerba et al., 2018). The approach tackles the scenario where uncertainty, relating to product prices is initially present, through a general two-stage Stochastic Programming formulation. The first stage looks to determine which suppliers to involve and product quantities, so as to determine in which discount interval the quantity is expected to lie. Next, given that discount intervals are then assumed locked by the first-stage decisions, the

second stage consists of recourse decisions that look to optimally adjust the quantities in each chosen supplier. Here uncertainty is only considered for the single criterion of price and is always characterised as stochastic, implying it is due to random processes. Conversely, in TEC we look to tackle uncertainty relating to multiple criteria within problems where uncertainty results from a lack of evidence (as opposed to being purely stochastic in nature). The authors also explore an alternative scenario, of the supplier selection under conditions of total quantity discount work problem, with only demand uncertainty (and without cost uncertainty). This is found to be a more complex scenario and so in subsequent work a more efficient solution to demand uncertainty is proposed (Manerba & Perboli, 2019), exploiting the structural properties of the problem with a progressive Hedging PH-based heuristic approach.

A hybrid ensemble and AHP approach is used in (Hosseini & Al Khaled, 2016) to consider uncertainty regarding suppliers to calculate a resilience value of potential suppliers, for criteria relating to their absorptive, adaptive, and restorative capacities. These are calculated using past evidence, e.g., delivery performances from suppliers, so unlike TEC there is no notion of targeting for efficient evidence collection. They conclude that robustness, reliability and rerouting are the most important enablers of supplier reliance, the ranking of which is then used, along with user preferences relating to the criteria, for selection of the best (single) supplier.

To improve quality performance of suppliers – looking to improve the utility of some aspect – under uncertainty, a Poisson-gamma model within a Bayesian framework has been utilised in (Quigley et al., 2017). The framework seeks to reduce non-conformance rates by considering the trade-off between obtaining more information verses seeking improvement based on current information. Thus, the approach helps a DM ascertain if it is more valuable to obtain additional information about a supplier (through, for example, more supplier visits and audits) before attempting improvements, or making decisions directly aimed at quality improvements (such as training and development of buyer resources into the supplier) in light of the current state of knowledge. Therefore, although

tackling a somewhat different problem – looking to actually improve the quality within suppliers, whereas in our TEC approach we focus on efficient selection (targeting) decisions for uncertainty reduction – both approaches adopt a similar philosophy of looking to decide what to target so as to be most useful.

In summary, previous work has either tackled uncertainty within MCSS with a focus on uncertainty in user preferences or supply and demand, or tackled uncertainty of supplier information without direct consideration of how to efficiently seek to reduce it. Conversely, TEC tackles uncertainty of supplier information, pertaining to the criteria, through looking to target and reduce the uncertainty in an efficient way.

3. The TEC Strategy

Our TEC approach seeks to efficiently reduce uncertainty associated with supplier information relating to the criteria within MCSS problems. Figure 1 shows the components of TEC, highlighted as those within the dotted red rectangle, and their interaction with other components. The targeted evidence collection strategy requires a set of parameters (step **1** in Figure 1) that include the candidate suppliers to consider, the criteria for the supplier selection weighted according to the user's preferences, the number of items required of each type and the allocated budget for the evidence collection. Taking into account these initial parameters, the next step is to collect an initial sample (step **2**) from the available evidence to bootstrap the targeting approach. This initial evidence is used by the TEC strategy to calculate the initial estimates for the required criteria (step **3**), which in turn are used by the optimisation process (MCSS) to obtain a preliminary selection of suppliers (step **4**). This selection is used by the TEC strategy to focus the evidence collection on those suppliers that were selected or, based on their margins of error, are candidates to be part in the selection. Having reduced the problem space to a subset of the suppliers, the TEC strategy keeps collecting additional evidence (step **5**) to refine the criteria estimates that are, again, used by the optimisation process (step **6**).
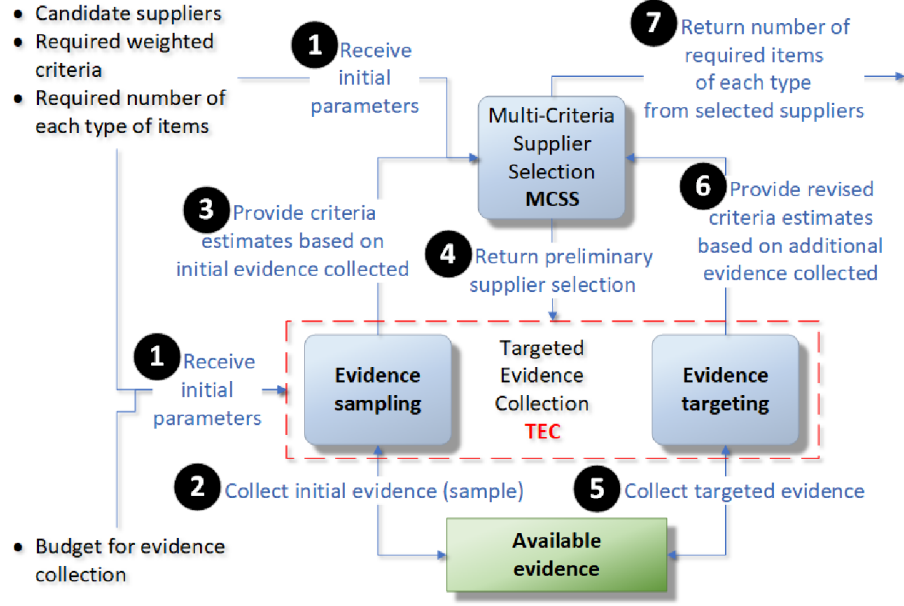
Figure 1: TEC components and operation

Steps ④ to ⑥ are repeated until certain conditions (explained in Subsection 3.3) are met and the final selection of suppliers (that includes the number of each type of items from each supplier) is generated by the optimisation process (step ⑦). TEC seeks to reduce uncertainty through iterative MCSS optimisation and evaluation, to target additional evidence to gravitate towards the most appropriate solution given no uncertainty. Next, we first briefly outline the MCSS component, then detail our evidence targeting strategy, before then discussing TEC's operation.

3.1. Multi-Criteria Supplier Selection

The base Multi-Criteria Supplier Selection (MCSS) problem, such as that defined in (Ting & Cho, 2008), which TEC utilises as shown in Figure 1, can be defined as: *Given a collection of suppliers* $\mathbf{S} = \{s_1, \ldots, s_m\}$ *from whom we wish to acquire varied types of items* $\mathbf{I} = \{i_1, \ldots, i_p\}$, *a set of weighted user criteria* $\mathbf{C} = \{c_1, \ldots, c_n\}$ *used to evaluate the candidate suppliers, and a number of items* $\mathbf{R}$ *required from each type in* $\mathbf{I}$ *, identify the number of items in* $\mathbf{I}$ *that should be*

13

purchased from each supplier in **S**, *while optimising the criteria in* **C**. For full operational details and mathematical definitions of the model's optimisation and constraints see (Ting & Cho, 2008). If multiple DMs are tasked with giving their preferences regarding the criteria then a group aggregating method such as (Abel et al., 2015) can be utilised to aggregate the group's views, with consideration of possible unequal weights of importance between the DMs, into a single set of weights for the criteria.

3.2. Evidence Targeting

The process to target evidence collection for supplier selection problems, given that there is uncertainty about the value for each supplier for each criteria of concern (i.e., real values for **C** over **S**), is to identify for which suppliers additional evidence should be collected to improve solutions. Hence, the question is: what evidence should we collect next to reduce uncertainty to enable improved solutions? The TEC approach identifies suppliers for whom additional evidence needs to be collected by analysing the overlap of the confidence intervals of each criterion for each candidate supplier. This is then used to target those suppliers that contribute to the solution or non-contributing suppliers whose confidence intervals overlap with confidence intervals of contributing suppliers. This statistical technique is utilised within TEC as it facilitates both exploration (coverage) and exploitation (focusing) of the objective space. Formulas for calculating confidence interval values for a criterion are given in Appendix A.

For a given criterion, the confidence intervals of two suppliers can overlap to an extent that may cause uncertainty about which one should contribute to the solution. This is because the real value of the criterion may lie in the overlap zone and, hence, there is no certainty about which supplier outperforms the other for that criterion. The approach, following (Knezevic, 2008), is to determine not only whether two confidence intervals overlap, but also the extent of this overlap by comparing the difference between the means of those confidence intervals. If the result indicates that the confidence intervals overlap and their means are

14

significantly different, then, for the purpose of the TEC strategy, the intervals are considered to *significantly overlap*, which is taken as strong evidence of whether suppliers are in contention to contribute to the solution. For example, assume that the estimated value of a criterion applied to supplier 1, $\hat{c}_{s_1}$, is higher than that of supplier 2, $\hat{c}_{s_2}$, then, the confidence intervals of s_1 and s_2 for that criterion *overlap* if the lower bound of the confidence interval of s_1 is less than or equal to the upper bound of the confidence interval of s_2. This can be expressed as an inequality where e is the margin of error (Equation A.2) for criterion c applied to suppliers s_1 and s_2:

$$\hat{c}_{s_1} - e_{\hat{c}_{s_1}} \leq \hat{c}_{s_2} + e_{\hat{c}_{s_2}} \text{ or } \hat{c}_{s_1} - \hat{c}_{s_2} \leq (e_{\hat{c}_{s_1}} + e_{\hat{c}_{s_2}}) \tag{1}$$

If there is overlap, then the following inequality determines if the means of two confidence intervals are not significantly different (Bulmer, 1979):

$$\hat{c}_{s_1} - \hat{c}_{s_2} \leq \sqrt{e_{\hat{c}_{s_1}}^2 + e_{\hat{c}_{s_2}}^2} \tag{2}$$

If the result of the right hand side of Equation 2 is greater or equal than the difference between the estimated values, then the means are *not significantly different*. If the means are not significantly different and there is overlap between two confidence intervals, we regard this as a *significant interval overlap* between them, and therefore that there is benefit from collecting additional evidence on both suppliers in order to determine which is the most suitable.

For example, consider four candidate suppliers $S = \{s_1, s_2, s_3, s_4\}$ that need to be evaluated on their defect ratio. Assume that some evidence has already been collected in the form of mock orders. The initial evidence is used to calculate the estimated defect ratio of these suppliers and their margins of error; the resulting confidence intervals are shown in Figure 2. In this example, assume that s_1 contributes to the solution. The overlap analysis between s_1 and the other three suppliers shows that: (i) there is no overlap between suppliers 1 and 2 as the lower bound of the confidence interval of s_1 is greater than the upper
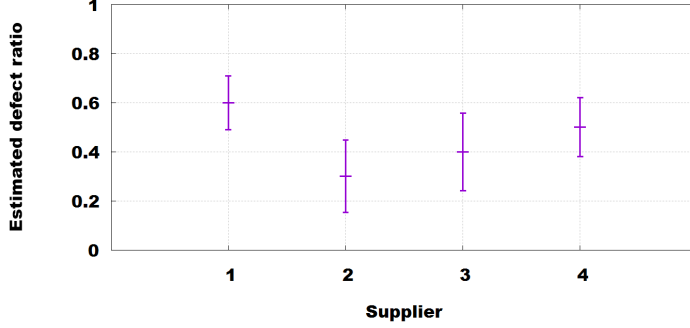
15

Figure 2: Analysis of overlapping between confidence intervals

bound of the interval of s_2, hence, this does not satisfy the inequality in Equation 1, (ii) there is overlap between suppliers 1 and 3 (satisfying Equation 1) and the difference between the means is significant (not satisfying Equation 2), and (iii) there is overlap between suppliers 1 and 4 and the difference between the means is not significant (satisfying Equation 2), thus this interval overlap can be considered as significant for TEC and, hence, further evidence collection is required to determine which supplier (s_1 or s_4) should contribute to the solution. Therefore, only s_1 and s_4 are considered for additional evidence collection; s_1 as it contributes to the solution and s_4 as its confidence interval significantly overlaps with a confidence interval of s_1.

3.3. Description of the TEC strategy

When applying the TEC strategy to the MCSS problem we have a set of suppliers $\mathbf{S} = \{s_1, \ldots, s_m\}$; a set of item types $\mathbf{I} = \{i_1, \ldots, i_p\}$; a set of criteria $\mathbf{C} = \mathbf{cC} \cup \mathbf{uC}$, where $\mathbf{cC} = \{c_1, \ldots, c_l\}$ are *crisp criteria* and $\mathbf{uC} = \{\hat{c}_1, \ldots, \hat{c}_o\}$ are *uncertain criteria*; a set of criteria weights $\mathbf{W} = \{w_1, \ldots, w_n\}$; a set of procuring budgets $\mathbf{B}_i$ per item type; a matrix of suppliers' capacities $\mathbf{Sc}_{s,i}$ with the item type maximum amounts each supplier has committed to provide; a set $\mathbf{R}_i$ with the item types required; a confidence level cL and margin of error e; and, for evidence collection, a budget b.

The TEC algorithm is presented as pseudo-code in Figure 3; we explain it using an example similar to one presented in (Ting & Cho, 2008) with seven

16

suppliers and seven item types, with an additional consideration that there is now uncertainty regarding criteria values.

We use a single cost (as a crisp criterion) that already includes the transportation, unit price and ordering costs, and two uncertain criteria, viz., the product quality in terms of the defect ratio and the delivery reliability in terms of the delayed orders ratio.

The TEC approach evaluates the overlapping between the confidence intervals surrounding the estimated values for each supplier and each criterion in contention. This approach aims to focus the evidence collection on those suppliers for which there is the highest uncertainty on whether they should contribute, and to what extent, or not to a solution to the supplier selection problem. Furthermore, the additional evidence collected reduces the uncertainty and the number of candidate suppliers that require more evidence. This process is repeated while there is still uncertainty on the number of items each supplier should contribute and while there is budget available to collect more evidence. As a consequence, the algorithm finds the number of items each supplier should contribute to satisfy the optimisation problem and includes, for informative purposes, the level of regret associated with the proposed solution.

The first step in the TEC algorithm computes the sample size for the number of pieces of evidence required to achieve a statistically representative view of candidate suppliers (Line 4). The sample size is calculated using Equation A.3 taking into account the suppliers in $\mathbf{S}'$ that are in contention for evidence collection, the budget b left for evidence collection and the given confidence level and margin of error.

Function *collectEvidence* (Line 5) obtains sS pieces of evidence, e.g., a mock order of 10 items. The evidence is collected on the criteria in C applied to the suppliers in $\mathbf{S}'$ and for the required item types $\mathbf{I}$. This function randomly selects at most sS pieces of evidence for collection. The remaining budget b is updated depending on the number sS of pieces of evidence identified (Line 6).

The values of the uncertain criteria in $\mathbf{uC}$ are estimated for each candidate supplier in $\mathbf{S}$ once the evidence is collected and assimilated. Function *estCri-*

17

Input: set $\mathbf{S}$ of suppliers

Input: set $\mathbf{I}$ of item types

Input: set $\mathbf{C} = \mathbf{cC} \cup \mathbf{uC}$ of crisp and uncertain criteria

Input: set W of weights for the criteria

Input: sets $\mathbf{B}_i$ of procuring budgets per item type

Input: matrix $\mathbf{Sc}_{s,i}$ of suppliers' capacities

Input: sets $\mathbf{R}_i$ with required item types

Input: confidence level cL

Input: margin of error e

Input: budget size b for evidence collection

Output: matrix $\mathbf{X}_{s,i}$ of items contributed per supplier

Output: vector $\mathbf{Q}_c$ of regret for each uncertain criterion

1:   $\mathbf{X} \leftarrow \{\}\{\}$

2:   $\mathbf{S}' \leftarrow \mathbf{S},\, sS \leftarrow 1$

3:   **while** $b > 0$ and $sS > 0$ and $\mathbf{S}' <> \{\}$ **do**

4:      $sS \leftarrow \mathrm{computeSampleSize}(\mathbf{S}', b, cL, e)$

5:      $\mathbf{E} \leftarrow \mathrm{collectEvidence}(\mathbf{S}', \mathbf{I}, sS)$

6:      $b \leftarrow b - sS$

7:      $\mathbf{u\hat{C}} \leftarrow \mathrm{estCriteriaValues}(\mathbf{S}, \mathbf{E}, \mathbf{I}, \mathbf{uC}, cL, e)$

8:      $\mathbf{X} \leftarrow \mathrm{solveMCSS}(\mathbf{S}, \mathbf{u\hat{C}}, \mathbf{W}, \mathbf{B}_i, \mathbf{Sc}_{s,i}, \mathbf{R}_i)$

9:      $\mathbf{S}' \leftarrow \{\}$

10:      **for** $s \in \mathbf{S}$ **do**

11:        **for** $i \in \mathbf{I}$ **do**

12:          $oL \leftarrow false$

13:          **for** $\hat{c} \in \mathbf{uC}$ **do**

14:            $oL \leftarrow \mathrm{overlapAnalysis}(\hat{c}_{s,i}, \mathbf{X}, \mathbf{W})$ or oL

15:          **end for**

16:          **if** $\mathbf{X}_{s,i} > 0$ or oL **then**

17:            $\mathbf{S}' \leftarrow \mathbf{S}' \cup \mathbf{s}$

18:          **end if**

19:        **end for**

20:      **end for**

21:   **end while**

22:   $\mathbf{Q} \leftarrow \mathrm{computeCriteriaRegret}(\mathbf{S}, \mathbf{uC}, \mathbf{X})$

23:   **return** $\mathbf{X}, \mathbf{Q}$

Figure 3: The TEC algorithm

18

teriaValues (Line 7) uses the candidate suppliers in **S**, the set E of collected pieces of evidence, the required item types in **I**, the collection of uncertain criteria **uC**, and a given confidence level cL and margin of error e, to obtain the set **uĈ** of estimated criteria values for the suppliers in **S**. These estimates rely on the evidence collected and are computed using Equation A.5. The margin of error around each criterion estimate (and for each supplier and item type) is computed by Equation A.2 and used to build the confidence intervals surrounding these estimates by solving Equation A.4. The confidence intervals are then analysed for overlap by checking one criterion at a time (as explained in Subsection 3.2). We use this approach to handle multiple criteria without considering them simultaneously for the statistical computations. An example of how the confidence intervals may look at this early stage of the process is shown in Figure 4 (a), where s_2, s_4 and s_6 are contributing to the solution, s_1 and s_3 have significant overlap with s_2 and s_4, respectively, hence, these five suppliers are considered for further evidence collection to refine their estimates. In contrast, s_5 and s_7 are not targeted for additional evidence collection as there is no statistical possibility of them contributing to the solution.

The next step is to solve the MCSS problem by applying an optimisation model such as Min-max, as in (Ting & Cho, 2008), to obtain a solution. This step is expressed as function *solveMCSS* (Line 8) that requires a collection S of candidate suppliers, the set **uĈ** of estimated uncertain criteria, the set **W** of
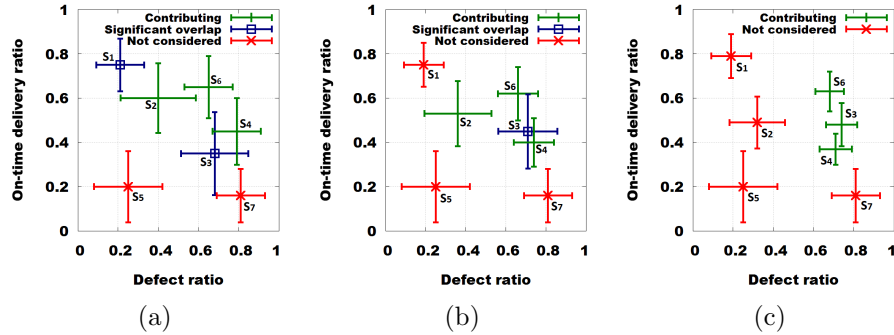

Figure 4: Confidence intervals for different amounts of evidence collected through TEC.

weights representing the user preferences, the set $\mathbf{B}_i$ of procuring budgets for each item type i, the matrix $\mathbf{Sc}_{s,i}$ of suppliers capacities for each supplier and item type, and the total number of user requested items of each type $\mathbf{R}_i$. The result of the optimisation is a matrix $\mathbf{X}_{s,i}$ with the number of items of each type contributed by each supplier. The set $\mathbf{S}'$, of suppliers for whom evidence needs to be collected is initialised before processing the overlap analysis.

Having the confidence intervals for each criterion and supplier, and the number of items of each type that each supplier contributes to a preliminary solution $\mathbf{X}$, the overlap between these intervals can be analysed. Function *overlapAnalysis* represents this step (Line 14), by solving Equations 1 and 2 (defined in Subsection 3.2). The overlap of confidence intervals is evaluated at supplier level (not criterion or item levels), hence, if at least one criterion for any item type is evaluated to have significant overlap, then the supplier is considered for evidence collection – condition *or oL* in Line 14.

In the next step, suppliers contributing to the solution along with non-contributing suppliers that have significant overlap with contributing suppliers (Line 16), are added to the set $\mathbf{S}'$, which holds the suppliers for whom evidence needs to be collected (Line 17). $\mathbf{S}'$ is used in the next cycle to compute a new sample size sS. After a few rounds of evidence collection, the scenario can be visualised as in Figure 4 (b), where there remains overlap between the confidence intervals of s_3 and s_4, but the suppliers contributing to the solution are now mostly identified. Suppliers 1, 5 and 7 are no longer considered for evidence collection.

The iteration continues whilst any of the following conditions hold (Line 3): (i) there is overlap between confidence intervals of suppliers contributing to the solution and suppliers that do not contribute; (ii) the sample size sS is greater than zero, i.e., there is still some evidence that can be collected; (iii) a budget b remains available for evidence collection.

When the algorithm finishes, the regret $\mathbf{Q}$ (Savage, 1951, Inuiguchi & Sakawa, 1995) is computed for the last obtained solution for each criteria under consideration. The regret of solutions can be compared by calculating the difference

20

between the last estimated value for each criterion $\hat{c}$ in the current solution $\mathbf{X}$ and that of the best computed solution using single-criteria optimisation (positive ideal) $\mathbf{X}^+$. This difference is then divided by the range of possible solutions to obtain the regret of criterion c in solution $\mathbf{X}$ as follows (Hsu et al., 2014):

$$\mathbf{Q}_{c_Z} = \frac{c_{Z^+} - \hat{c}_Z}{c_{Z^+} - c_{Z^-}} \qquad (3)$$

This regret provides a measure of the opportunity loss due to the remaining uncertainty (Line 22), and finally, the set of regrets Q and the solution matrix $\mathbf{X}$ are returned (Line 23). The solution $\mathbf{X}$ contains the number of items of each type that each supplier contributes. Figure 4 (c) presents a potential scenario at this stage, where no overlap exists between confidence intervals of those suppliers contributing (viz., s_3, s_4 and s_6) and those not contributing to the solution (viz., s_1, s_2, s_5 and s_7). This lack of overlap indicates that the contributing suppliers should not change if additional evidence is collected, providing a well-founded approach to deciding when additional evidence is unlikely to be fruitful.

4. Experimentation

We now present an experimental evaluation of the TEC strategy investigating (i) its properties and effectiveness for varying amounts of evidence, (ii) its effectiveness with different numbers and ratios of certain and uncertain criteria, (iii) varying numbers of suppliers, and (iv) varying criteria weights.

4.1. Comparison approaches

To the best of our knowledge, there is no directly comparable approach that tackles the problem of efficient reduction of uncertainty, pertaining to suppliers, within supplier selection; hence, we evaluate between TEC and more generic baselines defined below.

4.1.1. Random

Random sampling is used as a base case where evidence is collected randomly. This approach has the advantage that evidence collection is distributed evenly among candidate suppliers; however, it ignores specific user requirements and so may collect evidence on suppliers clearly excluded from the solution.

4.1.2. Uncertainty Sampling

Active learning techniques (Settles, 2012) are alternatives for evidence collection. Such techniques learn from collected evidence to determine which additional piece of information would most reduce associated uncertainty. In this way, the largest reductions in uncertainty within the problem can be sought for the smallest amount of evidence collection (Lewis & Gale, 1994). Among such techniques, *Uncertainty Sampling* was chosen based on its applicability to supplier selection, as it only requires a measure of uncertainty (the margin of error in our case) to decide which pieces of information to target. Uncertainty Sampling is expressed using a heuristic that takes into account user preferences (in the form of weights) and the margins of error, as a measure of uncertainty over the estimated criteria for a supplier. Evidence is collected first on suppliers who, for the overall weighted criteria, have the largest margin of weighted error, taking into account the importance of each criterion to the user. The uncertainty value u of a piece of evidence t belonging to a supplier s is calculated as:

$$u_{t,s} = \max(w(\hat{c}_{k_s}) \cdot e(s_t, \hat{c}_{k_s})); k = 1, 2, \ldots, n \qquad (4)$$

where w is the weight associated with each criterion; e is the margin of error; $\hat{c}_{k_s}$ is the estimated value of criterion k for supplier s; and n is the number of criteria for the scenario being considered.

Uncertainty Sampling is effective for maximising overall uncertainty reduction, through targeting pieces of evidence belonging to suppliers with the highest uncertainty, without consideration of the likelihood that they will contribute to

a solution.

4.2. Evaluation metrics

For comparison we use metrics to measure a supplier selection solution's quality in relation to current evidence and the associated regret:

Maximum deviation (MD): A measure of the quality of a supplier selection solution (in relation to current evidence) considering both overall utility and the faithfulness of a solution with regards to the DM's preferences, see (Ting & Cho, 2008)

Regret (Q): To assess a potential solution in the scenario where there is uncertainty the regret criterion (Inuiguchi & Sakawa, 1995, Savage, 1951) can be used to evaluate solutions in terms of risk or opportunity loss that the DM may have if he/she decides to choose a particular solution as presented in Equation 3.

4.3. Experimental setup

For the following experiments, a data set with synthetic pieces of evidence for up to 7 criteria on up to 117 suppliers is gradually sampled to simulate the evidence collection mechanism. The first step for each experiment is to calculate a representative sample (See Appendix A.3) of pieces of evidence to obtain for each criteria and each candidate supplier, in order to compute the initial estimated criteria values (See Appendix A.1) required by the TEC algorithm.

The experiments evaluate the performance of the TEC approach against the Uncertainty Sampling and Random competitors, and evaluate varied dimensions of the problem scenario, such as, the amount of evidence collected (Experiment 1), the number and nature of criteria (Experiment 2), the number of suppliers (Experiment 3) and the criteria weights (Experiment 4). The results for the three approaches are compared in terms of which approach provides a solution closer to the reference solution with less evidence collected, the final actual solution achieved MD, the regret Q (opportunity loss) associated with that

solution, and the number of pieces of evidence needed to reach the reference solution.

The default parameters used, except when a specific parameter is varied during an experiment, are: 117 suppliers with varied criteria values, 7 types of items produced by those suppliers, 3 criteria where one of them is certain (total purchasing cost - C1) and the rest uncertain (number of defects - C2, and delivery delays - C3), criteria weights for 3 criteria (0.22, 0.44, 0.34), and up to 12,000 pieces of evidence produced. In all the experiments, results have been averaged over 25 runs and the confidence level and margin of error were 95% and 0.05 respectively.

4.4. Experiment 1 - Varying the amounts of evidence collection

We firstly explore effects on the solution of varying amounts of evidence. For Experiment 1, the number of pieces of evidence collected was varied and the three approaches compared (viz., TEC, Random, Uncertainty Sampling). Evidence was collected up to the number of pieces required for each approach to find the reference solution, i.e, the solution computed when there is no uncertainty in the criteria values. For each approach the MD of each solution, for both actual and estimated criteria values, was calculated, as was Q per criterion for each solution.

The MD values for the three approaches for varying amounts of evidence, up to 12,000 pieces, and the reference solution line, are shown in Figure 5. The reference solution represents the MD that would be found with no uncertainty present. From this we observe that for each approach, as more evidence is collected the actual solutions (solid lines) converge towards the reference solution, as would be expected. Similarly, the estimated solutions (dotted lines) converge towards the reference solution; note they are closer to it, as the algorithm is using estimates to obtain solutions which, given a partial knowledge based on the evidence collected, seem better than the reality. In addition, this plot highlights that, after some initial evidence collection calibration, TEC improves more swiftly than the other approaches; this suggests that TEC is able to reduce
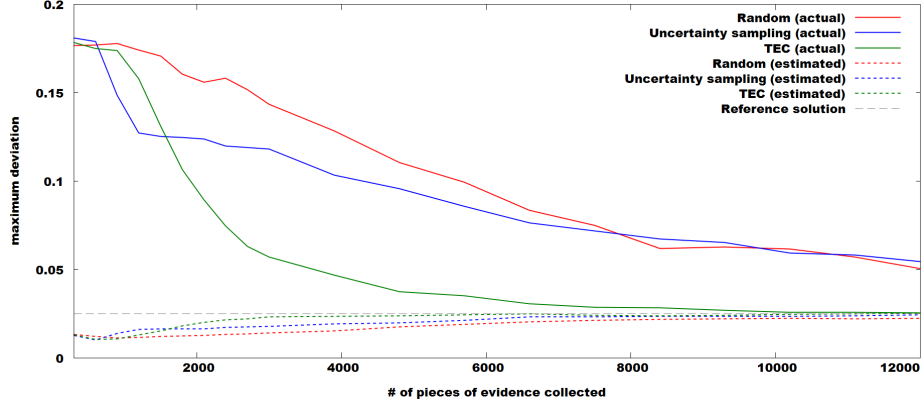
24

Figure 5: MD over varying amounts of evidence

uncertainty in ways that are more helpful to the optimisation. Initially, Uncertainty Sampling with its focus on the largest reduction in overall uncertainty, without much knowledge of which suppliers may be part of good solutions, is able to gain the most improvement; however, as more evidence is collected this focus hinders the efficiency of subsequent collections.

We further compare the approaches by analysing their solutions broken down by criteria. The values of the 3 criteria for solutions found by the three approaches after collection of 12,000 pieces of evidence, along with the reference solution, are shown in Figure 6:Left. From this we observe that TEC has found a solution which, for the uncertain criteria, is close to that of the reference solution. The Q for each criterion can also be calculated, as shown in Figure 6:Right. In such a plot, the bulls-eye represents the reference solution (a solution with 0 regret for any of the criteria) and the amount of regret associated with each criterion compared to this can be examined by a DM. From this, we observe that by using TEC the solution found has, for each criterion, significantly lower regret than that obtained by the other approaches. Such plots of a solution assist a DM to make a more informed decision in understanding, with respect to each criterion, the amount of risk attached to the solution.

Finally, comparison can be made between the solution found through TEC and that found before any evidence collection, in terms of the regret values of
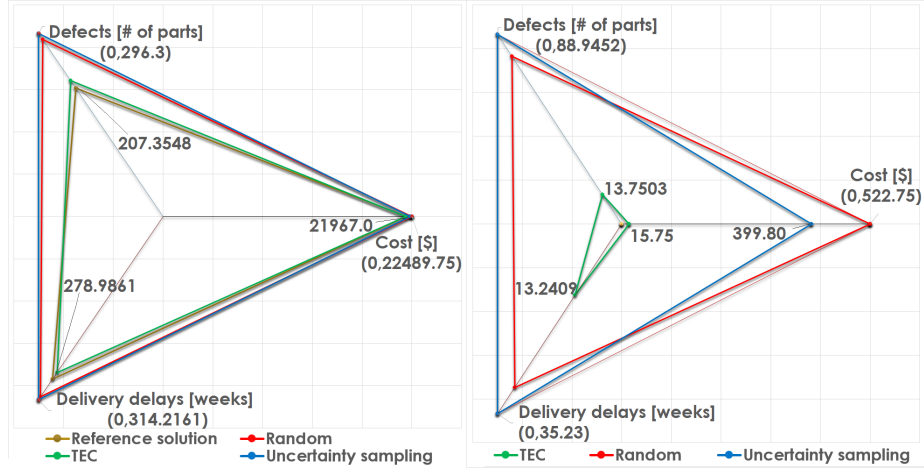
Figure 6: After 12000 pieces of evidence collected. Left: Breakdown of solution per criterion, Right: Regret per criterion

each criterion, as shown in Figure 7. The left of this plot shows the impact that different quantities chosen in the solutions have in terms of the overall values for each criterion. The regret shown on the right of this plot highlights the large difference between the amount of regret associated with a solution found without any evidence collection and one found with evidence collection.

This experiment has shown that, over varying amounts of evidence collection, TEC is able to efficiently reduce uncertainty, and find improved solutions utilising less evidence than both Uncertainty Sampling and Random approaches.

4.5. Experiment 2 - Varying the number and nature of criteria

Next we explore the impact of increasing the amount of uncertainty, by varying the number of criteria that are uncertain. For Experiment 2, the number of certain criteria was set to one, and the number of uncertain criteria was varied from two to six. Therefore, five variants with sets of criteria from three to seven were performed. For each variant the three approaches (viz., TEC, Random and Uncertainty Sampling) were compared and evidence collected until each approach reached the reference solution. In each case, the set of criteria weights were set as equal (for the set of three criteria each weight was 1/3, for set of four criteria each weight was 1/4 and so forth).
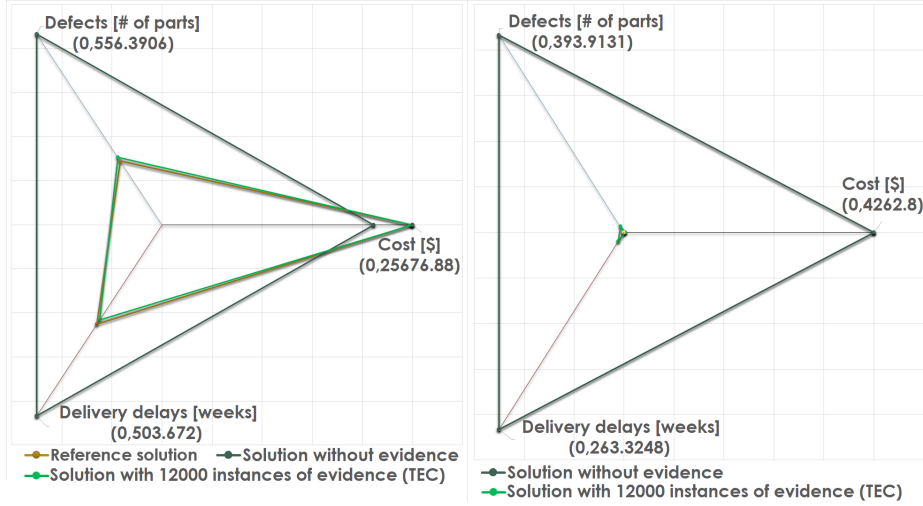
26

Figure 7: Solution (left) and regret (right) comparison with and without evidence collected

For each variant set of criteria, the amount of evidence collection required to reach the reference solution for each approach is shown in Figure 8. Here we observe that, for each set of criteria, TEC finds the reference solution using less evidence than Uncertainty Sampling and Random. We further observe that, as the number of uncertain criteria increases, our approach shows little variation in the amount of evidence required to reach the reference solution. Uncertainty Sampling requires similar amounts of evidence for each step as it does not directly consider the optimisation problem to target the most uncertain suppliers' criteria. Conversely, Random requires less evidence, as the number of uncertain criteria goes up because, as the number of constraints increases, the number of potential solutions that can satisfy all constraints is reduced; therefore a solution to the optimisation problem is found earlier in the evidence collection process. This behaviour can be further appreciated in Figure 9, where the MD of the solutions, found by the three approaches with 12,000 pieces of evidence, for the different sizes of criteria sets is shown. This plot highlights that with an increasing number of criteria, the MD increases because the amount of compromise required increases; however, the TEC approach, for each set of criteria, finds a solution with lower MD than the other approaches.
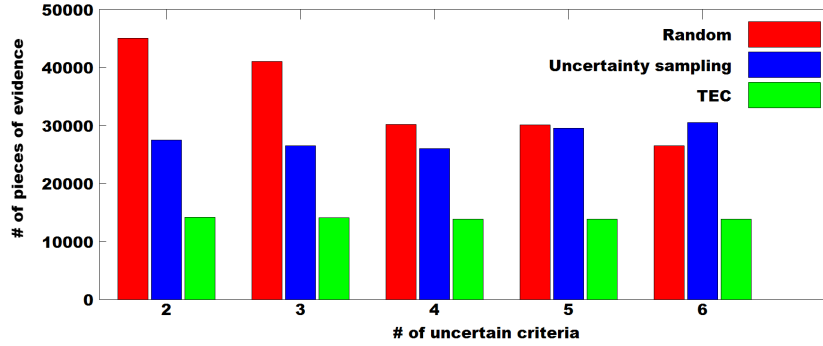
Figure 8: Evidence needed to reach reference solution per number of uncertain criteria
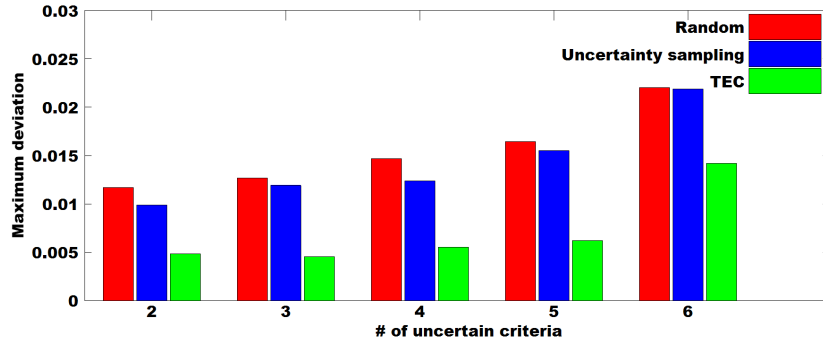

Figure 9: MD from solution per number of uncertain criteria

We also analyse solutions broken down with respect to each criterion. Figure 10:Left shows the solutions found by the three approaches (and the reference solution) for the set of 7 criteria (1 certain and 6 uncertain), with 12,000 pieces of evidence. Here we observe how, through TEC, the values obtained for each criterion are close to that of the reference solution, whereas Uncertainty Sampling and Random have obtained less satisfactory values for the uncertain criteria. In this plot, we further observe that for the crisp criterion, values have been found that are better than the reference solution. This is because the reference solution is an optimal trade-off solution with respect to MD within the multi-dimensional space. Therefore, a solution may have better values than the reference solution for some criteria whilst having worse values for other criteria, such that overall it represents a less optimal trade-off solution than the reference
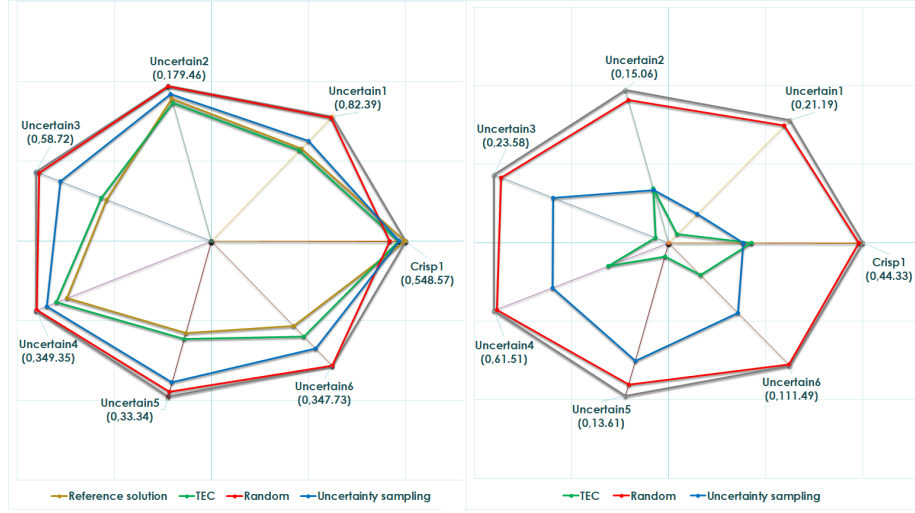
Figure 10: One Certain, six Uncertain Criteria after 12000 evidence Left: Breakdown of solution per criterion, Right: Regret per criterion

solution.

The breakdown of regret for each criterion for each solution is shown in Figure 10:Right, with the reference solution represented as the bulls-eye. In this plot, we observe that TEC finds a solution with less regret for each criterion than Random, and a solution with less regret for most criteria compared to Uncertainty Sampling (such that the overall regret from our approach is much less than that from Uncertainty Sampling). From such a plot, a DM is able to obtain a clear breakdown of the regret and associated risk attached to each criterion within a solution, leading again to better informed decisions.

This experiment has shown that TEC is able to perform efficient uncertainty reduction as the amount of uncertainty within the problem increases, with regards to the number of uncertain criteria.

4.6. Experiment 3 - Varying the number of suppliers

Here we explore the effect of varying the number of suppliers. The number of suppliers was varied in steps of 20, from 20 to 100.

The results of MD for the three approaches, over a range of evidence collected, for the steps of 40 suppliers to 100 are shown in Figure 11. From these
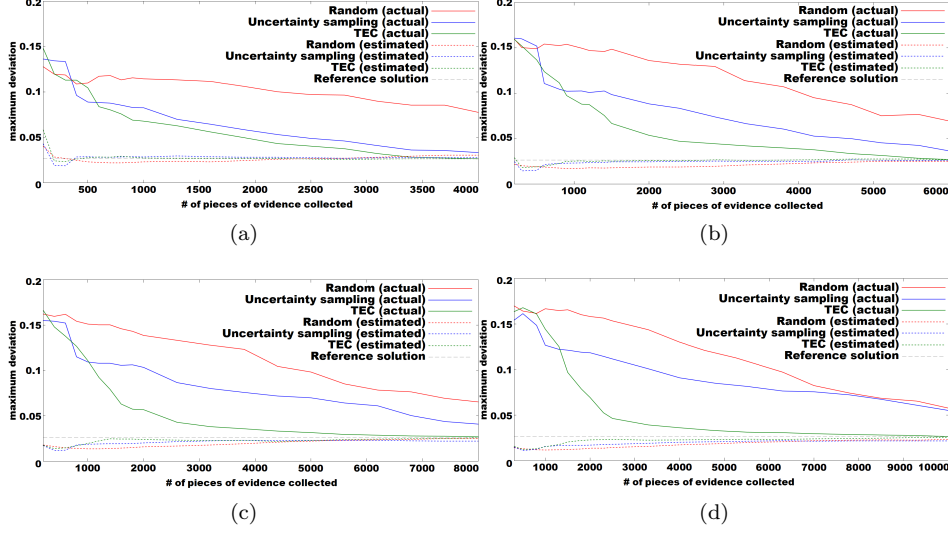
Figure 11: MD from solution. Number of suppliers (a) 40 (b) 60 (c) 80 (d) 100

plots, we observe that, as the number of suppliers increases, TEC's superior performance becomes more prominent - this highlights that as the number of suppliers increases, and thus the size of the space of possible choices of evidence to collect increases, the advantage of more targeted evidence collection increases. The amount of evidence collection required by each approach to reach the reference solution in each step is shown in Figure 12. From this, we see that as the number of suppliers about whom evidence can be collected increases, the approaches take longer to reach the reference solution. However, at each step TEC is able to reach the reference solution before Random and Uncertainty Sampling, and the gap increases as the number of suppliers increases.

4.7. Experiment 4 - Varying criteria weights

Finally, we explore the effects of varying DM preference weights. For Experiment 4, four runs were performed, each increasing the weight of the third (an uncertain) criterion. Weights for the four runs were: (1/3,1/3,1/3), (0.3,0.3,0.4), (0.2,0.2,0.6), (0.1,0.1,0.8).

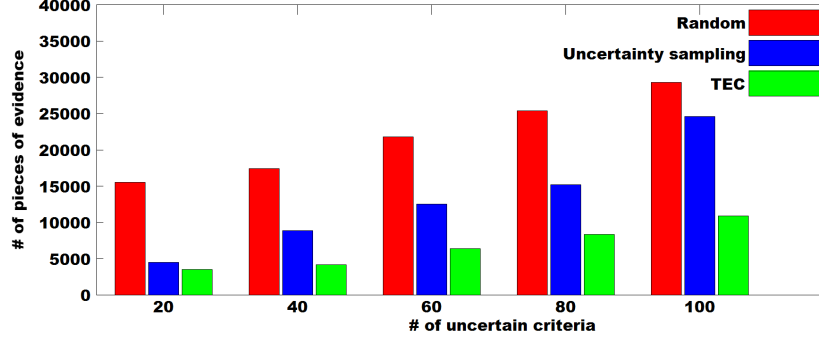The MD values for the three approaches over the range of evidence collec-

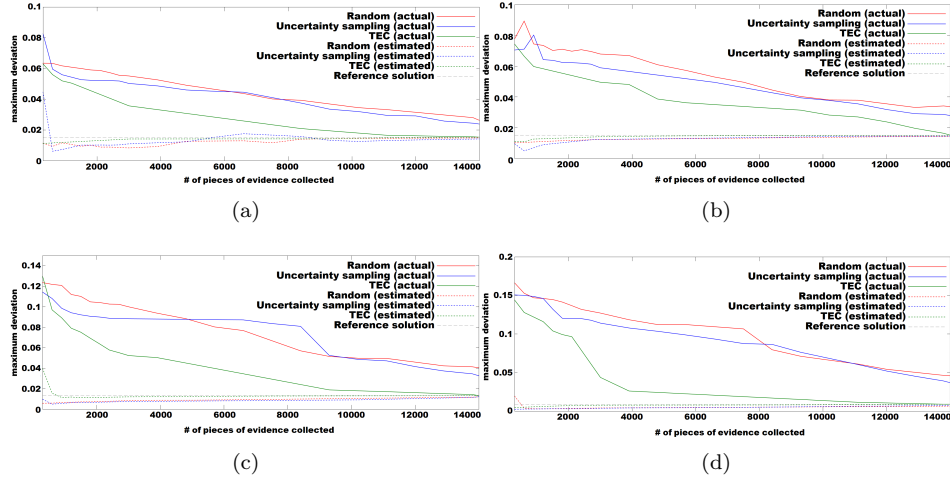Figure 12: Evidence required to reach reference solution per number of suppliers

Figure 13: MD with $\mathbf{W}$: (a) 1/3,1/3,1/3 (b) 0.3,0.3,0.4 (c) 0.2,0.2,0.6 (d) 0.1,0.1,0.8

tion for each run are shown in Figure 13. From this set of plots the impact of making one criterion more important is highlighted; we see that TEC more efficiently utilises a DM's preference within evidence collection than Random and Uncertainty Sampling. TEC more efficiently takes into account the user preferences to determine which suppliers should be considered for evidence collection. Therefore, as the solution favours suppliers that have higher estimated values for highly-weighted criteria, TEC will consistently target those suppliers to more efficiently collect evidence aligned with the user's preferences.

We can see this further through analysis of the breakdown of the solutions
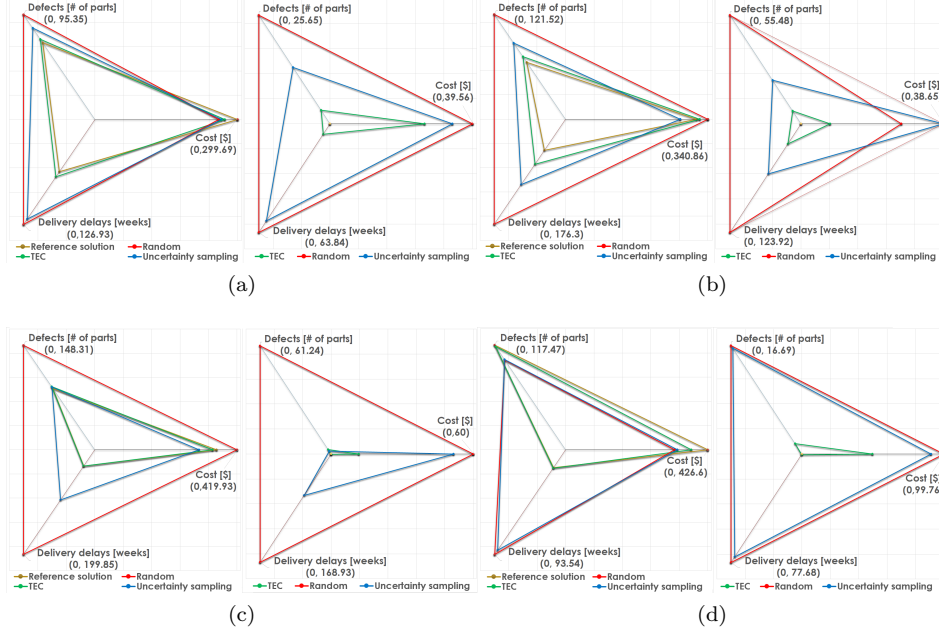
31

Figure 14: Solution values (left) and regret (right) after 12000 instances of evidence collected, **W**: (a) 1/3,1/3,1/3 (b) 0.3,0.3,0.4 (c) 0.2,0.2,0.6 (d) 0.1,0.1,0.8

found after 12,000 pieces of evidence were collected for each run, see Figure 14. This set of plots highlights that, as a single criterion is given more and more importance by the DM, TEC consistently targets those suppliers that contribute to the solution (lines 13 and 14 in Figure 3), such that the uncertainty in the solution value, for the most important criterion, is reduced the most. Moreover, the regret of the more important criterion is also most reduced by TEC, thus decreasing the amount of associated risk within a solution, in line with what is most important to the DM.

This experiment has shown how TEC is able to efficiently reduce uncertainty under diverse user preference weights.

5. Conclusions

Within MCSS problems there is inherent uncertainty that could relate to a DM's subjective preferences between criteria, supply and demand, and/or infor-

mation regarding suppliers' utility values in terms of the set of criteria. Whereas previous work has explored the first two of these, there has been less work on uncertainty regarding suppliers' utility in terms of the criteria, and specifically how such uncertainties could be efficiently reduced to look to mitigate their adverse effects. It is important to tackle such uncertainty as it can have a major impact upon the *realised* utility of a supplier selection solution, compared to its *perceived* utility.

This paper proposes the TEC strategy for efficient uncertainty reduction relating to supplier information, for criteria for which uncertainty is inherent, by looking to collect additional evidence in an efficient manner. By dynamically considering the likelihood that each supplier will be part of a solution, along with consideration of the DM's preferences between criteria, the strategy reduces uncertainty, enabling swift gravitation towards the solution that would be produced in the absence of uncertainty. TEC is able to scale to a large number of suppliers, and our experimentation highlights how its effectiveness becomes more pronounced as the number of suppliers (and the amount of overall uncertainty) increases. Moreover, the strategy can handle scenarios where there is a mix of both uncertain and certain criteria, and it is applicable and effective for a large range of amounts of evidence collection.

After a given amount of evidence collection, the level of regret associated with the current solution is calculated. This allows a DM to ascertain the risk associated with a solution, due to remaining uncertainty, in terms of each criterion, and so better inform his/her decision making. This can be calculated each time additional evidence has been collected, informing a DM of the new level of regret achieved after the additional evidence, helping them triangulate how much evidence they wish to collect, informed by its ongoing productiveness.

The effectiveness, adaptability and scalability of TEC has been compared through experimentation against alternative approaches, including an approach based on active learning, and we have shown how TEC can achieve efficient reduction in uncertainty, requiring up to three times less evidence than comparable approaches.

Regarding the practicalities of applying the TEC strategy, one implication is in terms of its requirement for a sufficient amount of evidence to be effective. Although, as we showed within the experimentation, the TEC strategy can be effective over a range of evidence amounts, for scenarios in which limited evidence can be collected TEC's capacity and aptitude is limited. For example, for a new firm that makes infrequent but expensive orders, of few items, from suppliers that are not forthcoming with detailed performance history, it could be difficult to collect sufficient evidence for the TEC strategy to be effective. As considered in the discussion on Figure 5, it is the case that initially, while the amounts of evidence are small, TEC can be outperformed.

The more uncertainty that is present, the more effective is the TEC strategy; with its effectiveness diminishing as the amount of uncertainty falls. We see this through the gradient of the TEC line in Figure 5; where initially, with more uncertainty present, the more dramatically TEC is able to tackle and reduce it; this tails off as the amount of remaining uncertainty becomes smaller, and tends towards zero. Therefore, in scenarios for which there is very little initial uncertainty, TEC would be less effective than in scenarios for which there is a lot. Conversely, it is likely that such scenarios, with modest uncertainty, may not require a targeted approach, that focuses on uncertainty reduction, such as TEC.

Future work will investigate how to expand the applicability of TEC to a wider set of scenarios, as well as to explore some of the scenarios discussed in the practicalities above. For example, a variation of the TEC strategy will be explored for use when it is know *a priori* that only a small amount of evidence is or can be collected. When it is known that a large enough amount of evidence is to be collected for TEC to suitably warm up, we can be assured, as can be seen from inspecting a plot such as Figure 5, that TEC will be effective. In such cases, it would be irrelevant that TEC takes time to get up to speed; however, if it were known that constraints dictated that only a small amount of evidence is to be collected, a variation of the strategy could explore how to modify its operation so as to be more effective in such scenarios.

We will also consider how TEC could be extended to scenarios for which evidence collection happens over a much longer time-frame than we currently envisage. Potentially, evidence that is collected in the distant past may lose some of its currency. For example, evidence collected from a supplier a number of years ago but who has spend the intervening time since its collection looking to actively improve its quality will no longer be as useful to us as more recent evidence about the supplier. Therefore, further work will explore how TEC could be extended to best handle such a scenario. Considering longer term scenarios could also allow exploration of interesting conditions, such as how to best handle the amalgamation of evidence collected over a long time for separate suppliers who have subsequently merged and become one entity.

References

Abel, E., Mikhailov, L., & Keane, J. (2015). Group aggregation of pairwise comparisons using multi-objective optimization. *Information Sciences*, *322*, 257–275.

Amorim, P., Curcio, E., Almada-Lobo, B., Barbosa-Póvoa, A. P., & Grossmann, I. E. (2016). Supplier selection in the processed food industry under uncertainty. *European Journal of Operational Research*, *252*, 801–814. doi:10.1016/j.ejor.2016.02.005.

Arikan, F. (2013). A fuzzy solution approach for multi objective supplier selection. *Expert Systems with Applications*, *40*, 947–952. doi:10.1016/j.eswa.2012.05.051.

Ayhan, M. B., & Kilic, H. S. (2015). A two stage approach for supplier selection

problem in multi-item/multi-supplier environment with quantity discounts. *Computers & Industrial Engineering*, *85*, 1–12.

Bilsel, R. U., & Ravindran, A. (2011). A multiobjective chance constrained programming model for supplier selection under uncertainty. *Transportation Research Part B: Methodological*, *45*, 1284–1300.

Bulmer, M. G. (1979). *Principles of Statistics*. Dover Publications.

Chai, J., Liu, J. N., & Ngai, E. W. (2013). Application of decision-making techniques in supplier selection: A systematic review of literature. *Expert systems with applications*, *40*, 3872–3885.

Chai, J., & Ngai, E. W. (2019). Decision-making techniques in supplier selection: Recent accomplishments and what lies ahead. *Expert Systems with Applications*, (p. 112903).

Chan, F. T., & Kumar, N. (2007). Global supplier development considering risk factors using fuzzy extended ahp-based approach. *Omega*, *35*, 417–431.

Corder, G. W., & Foreman, D. I. (2009). *Nonparametric Statistics for Non-Statisticians: A Step-by-Step Approach*. New Jersey: Wiley.

Deng, X., Hu, Y., Deng, Y., & Mahadevan, S. (2014). Supplier selection using ahp methodology extended by d numbers. *Expert Systems with Applications*, *41*, 156–167.

Ghobadian, A., Stainer, A., Liu, J., & Kiss, T. (2016). A Computerised Vendor Rating System. In *Developments in Logistics and Supply Chain Management* (pp. 103–112). London: Palgrave Macmillan UK. doi:`10.1057/9781137541253_10`.

Ghodsypour, S., & O'Brien, C. (2001). The total cost of logistics in supplier selection, under conditions of multiple sourcing, multiple criteria and capacity constraint. *International Journal of Production Economics*, *73*, 15–27. doi:`10.1016/S0925-5273(01)00093-7`.

Guneri, A. F., Yucel, A., & Ayyildiz, G. (2009). An integrated fuzzy-lp approach for a supplier selection problem in supply chain management. *Expert Systems with Applications*, *36*, 9223–9228. doi:`10.1016/j.eswa.2008.12.021`.

Haeri, S. A. S., & Rezaei, J. (2019). A grey-based green supplier selection model for uncertain environments. *Journal of cleaner production*, *221*, 768–784.

Hosseini, S., & Al Khaled, A. (2016). A hybrid ensemble and ahp approach for resilient supplier selection. *Journal of Intelligent Manufacturing*, (pp. 1–22).

Hsu, C. W., Kuo, R. J., & Chiou, C. Y. (2014). A multi-criteria decision-making approach for evaluating carbon performance of suppliers in the electronics industry. *International Journal of Environmental Science and Technology*, *11*, 775–784.

Igoulalene, I., Benyoucef, L., & Tiwari, M. K. (2014). Novel fuzzy hybrid multi-criteria group decision making approaches for the strategic supplier selection problem. *Expert Systems with Applications*, *42*, 3342–3356. doi:`10.1016/j.eswa.2014.12.014`.

Inuiguchi, M., & Sakawa, M. (1995). Minimax regret solution to linear programming problems with an interval objective function. *European Journal of Operational Research*, *86*, 526–536. doi:`10.1016/0377-2217(94)00092-Q`.

Kazaz, B., & Webster, S. (2015). Price-setting newsvendor problems with uncertain supply and risk aversion. *Operations Research*, *63*, 807–811.

Kellner, F., Lienland, B., & Utz, S. (2019). An a posteriori decision support methodology for solving the multi-criteria supplier selection problem. *European Journal of Operational Research*, *272*, 505–522.

Knezevic, A. (2008). Overlapping confidence intervals and statistical significance. *StatNews, Cornell University, Cornell Statistical Consulting Unit*, *73*.

Lewis, D. D., & Gale, W. A. (1994). A sequential algorithm for training text classifiers. In *Proceedings of the 17th annual international ACM SIGIR con-*

ference on Research and development in information retrieval (pp. 3–12). Springer-Verlag.

Manerba, D., Mansini, R., & Perboli, G. (2018). The capacitated supplier selection problem with total quantity discount policy and activation costs under uncertainty. *International Journal of Production Economics*, *198*, 119–132.

Manerba, D., & Perboli, G. (2019). New solution approaches for the capacitated supplier selection problem with total quantity discount and activation costs under demand uncertainty. *Computers & Operations Research*, *101*, 29–42.

Memon, M. S., Lee, Y. H., & Mari, S. I. (2015). Group multi-criteria supplier selection using combined grey systems theory and uncertainty theory. *Expert Systems with Applications*, *42*, 7951–7959.

Moghaddam, K. S. (2015). Fuzzy multi-objective model for supplier selection and order allocation in reverse logistics systems under supply and demand uncertainty. *Expert Systems with Applications*, *42*, 6237–6254.

Quigley, J., Walls, L., Demirel, G., MacCarthy, B. L., & Parsa, M. (2017). Supplier quality improvement: The value of information under uncertainty. *European Journal of Operational Research*, *264*, 932–947. doi:`10.1016/j.ejor.2017.05.044`.

Rabbani, M., Foroozesh, N., Mousavi, S. M., & Farrokhi-Asl, H. (2019). Sustainable supplier selection by a new decision model based on interval-valued fuzzy sets and possibilistic statistical reference point systems under uncertainty. *International Journal of Systems Science: Operations & Logistics*, *6*, 162–178.

Rashidi, K., & Cullinane, K. (2019). A comparison of fuzzy dea and fuzzy topsis in sustainable supplier selection: Implications for sourcing strategy. *Expert Systems with Applications*, *121*, 266–281.

Rockafellar, R. T., & Uryasev, S. (2002). Conditional value-at-risk for general loss distributions. *Journal of Banking & Finance*, *26*, 1443–1471. doi:`10.1016/S0378-4266(02)00271-6`.

Sanayei, A., Mousavi, S. F., & Yazdankhah, A. (2010). Group decision making process for supplier selection with vikor under fuzzy environment. *Expert Systems with Applications*, *37*, 24–30.

Savage, L. J. (1951). The theory of statistical decision. *Journal of the American Statistical Association*, *46*, 55–67.

Sawik, T. (2010). Single vs. multiple objective supplier selection in a make to order environment. *Omega*, *38*, 203–212.

Sawik, T. (2011). Selection of supply portfolio under disruption risks. *Omega*, *39*, 194–208.

Settles, B. (2012). Active Learning. *Synthesis Lectures on Artificial Intelligence and Machine Learning*, *6*, 1–114. doi:`10.2200/S00429ED1V01Y201207AIM018`.

Sounderpandian, J., Prasad, S., & Madan, M. (2008). Supplies from developing countries: optimal order quantities under loss risks. *Omega*, *36*, 122–130.

Stevens, S. S. (1946). On the Theory of Scales of Measurement. *Science (New York, N.Y.)*, *103*, 677–80. doi:`10.1126/science.103.2684.677`.

Stević, Ž., Vasiljević, M., Puška, A., Tanackov, I., Junevičius, R., & Vesković, S. (2019). Evaluation of suppliers under uncertainty: a multiphase approach based on fuzzy ahp and fuzzy edas. *Transport*, *34*, 52–66.

Ting, S.-C., & Cho, D. I. (2008). An integrated approach for supplier selection and purchasing decisions. *Supply Chain Management: An International Journal*, *13*, 116–127. doi:`10.1108/13598540810860958`.

Zhang, J.-l., & Chen, J. (2013). Supplier selection and procurement decisions with uncertain demand, fixed selection costs and quantity discounts. *Computers & Operations Research*, *40*, 2703–2710.

Appendix A: Calculating criteria confidence intervals

To evaluate performance of suppliers *ratio measures* (Stevens, 1946) can be utilised to indicate the proportion of entities belonging to categories or groups. Criteria such as on-time delivery and product/service quality, which rank highly in importance according to "The Supplier Performance Measurement Benchmark report"[1], are supportable within TEC through the form that the evidence to be collected can take (e.g., the proportion of items delivered on time can be updated by using the results of more orders). Specifically, the evidence needs to be classified according to user expectations, i.e., on time or not on time.

To improve estimates on the suppliers' criteria additional evidence can be collected from different sources depending on the criterion being considered. Examples of evidence sources that may help reduce uncertainty in TEC are: (i) Customer interviews: collect information by face-to-face or remote interactions about a particular criterion, (ii) Mock orders: submit a relatively small but representative request to the supplier, (iii) Market assessment (reviews): perform a survey or search among diverse supplier's reviews over a particular item to compare criterion values, and (iv) On-site visits: assess a particular process, policy or task to be measured where the products or services are generated.

Uncertainty associated with criteria values used to evaluate the performance of candidate suppliers can be modelled through statistical concepts such as sample size, sampling errors and confidence intervals (Bulmer, 1979). Sampling error can be reduced by collecting additional evidence. TEC uses criterion errors and their confidence intervals *surrounding* the estimated value to determine which suppliers need more evidence to reduce uncertainty and improve solutions.

---

[1]`http://www.lyonsinfo.com/_resources/aberdeen_spms_report.pdf`

The formula to compute the estimated value for a criterion c (ratio) is:

$$\hat{c} = \frac{tp_c}{tp_c + fp_c} \tag{A.1}$$

where tp is the amount of evidence expected for criterion c (e.g., on-time deliveries) or true positives, and fp is the amount of evidence unexpected for criterion c (e.g., delayed deliveries) or false positives.

The sampling error e is computed using the following equation:

$$e_{\hat{c}} = z_{cL} \cdot \sqrt{\frac{\hat{c} \cdot (1 - \hat{c})}{n}} \tag{A.2}$$

The sample size ss, required to apply the statistical analysis over a sample providing a significant representation of the total population, is computed as:

$$ss_{\hat{c}} = (z_{cL})^2 \cdot \frac{\hat{c} \cdot (1 - \hat{c})}{(e_{\hat{c}})^2} \tag{A.3}$$

The confidence interval CI around the estimated criterion value $\hat{c}$ is:

$$CI_{\hat{c}} = \hat{c} \pm e_{\hat{c}} \tag{A.4}$$

where $\hat{c}$ is the estimated value of c, n is the amount of evidence collected for c, tp_c is an expected item for criterion c, fp_c is an unexpected item for criterion c, z is the z–score for a given confidence level cL (assuming a normal distribution), and $e_{\hat{c}}$ is the sampling error for criterion c.

To ensure that there is a minimum population size to apply the statistical formulae and to distinguish the type of data distribution (Corder & Foreman, 2009), additional constraints of $(n \cdot \hat{c}) \geq 15$, and $(n \cdot (1 - \hat{c})) \geq 15$ must be fulfilled. The notation $\hat{c}_{h_{j,k}}$ is used to represent the estimated value of criterion h evaluated for supplier j and item-type k.

From a set of estimated item-type values of a criterion for a supplier, an estimated value of the criterion for the supplier can be computed. The estimated

average value of criterion h for supplier j can be calculated via:

$$\hat{c}_{h_j} = \frac{1}{l} \cdot \left(\sum_{k=1}^{l} i_k \right) \tag{A.5}$$

where l is the number of item-types.