



Designing a robust index for WWTP energy efficiency: The ENERWATER water treatment energy index

Miguel Mauricio-Iglesias, Stefano Longo, Almudena Hospido

Accepted Manuscript

How to cite:

Science of The Total Environment, Volume 713, 15 April 2020, 136642. Doi:
[10.1016/j.scitotenv.2020.136642](https://doi.org/10.1016/j.scitotenv.2020.136642)

Copyright information:

© 2020 Elsevier B.V. This manuscript version is made available under the CC-BY-NC-ND 4.0 license (<http://creativecommons.org/licenses/by-nc-nd/4.0/>)

Designing a robust index for WWTP energy efficiency: the ENERWATER Water Treatment Energy Index

Miguel Mauricio-Iglesias^{*}, Stefano Longo[‡], Almudena Hospido

CRETUS Institute, Department of Chemical Engineering, Universidade de Santiago de Compostela, Spain.

[‡] Currently at HERA, Central Innovation Department, via Cesare Diana 40, 44121 Ferrara, Italy

* Corresponding author

Abstract

The development and use of composite indexes has exploded in the last 15 years as a tool to summarise the large amount of information available nowadays. To ensure that composite indicators reflect faithfully the purpose of evaluation and are widely accepted and used, the users must understand the relationship between individual sub-indicators and the result (transparency) and the evaluation should not depend on weights, lack of information, etc. (robustness). It is proposed here for the first time a composite index to measure energy efficiency in a wastewater treatment plant, from the definition of the individual sub-indicators to the assignation of an energy label, easy to communicate to a broad public. Using as a demonstration real data from 44 wastewater treatment plants, the index robustness is tested and improved by uncertainty and sensitivity analysis results, finally achieving a robust algorithm which can be used by the large majority of wastewater treatment plants.

Keywords: Composite Indicators; Energy Efficiency; Robustness; Wastewater Treatment; Data Analysis; Uncertainty and Sensitivity Analysis

Nomenclature

Symbol	Meaning
BAP	Budget allocation process
BOD	Benefit of doubt
COD	Chemical oxygen demand
d	Ranking distance
E	Expected value
KPI	Key performance indicator
m	Number of indicators
max	Maximum value of a set
min	Minimum value of a set
N_{plants}	Number of units (WWTPs)
N	Number of iterations in uncertainty analysis
P	Percentile
s	Sample standard deviation
S	Sensitivity
TN	Total Nitrogen
TP	Total Phosphorous
V	Variance
WTEI	Water Treatment Energy Index
w	Weight
\bar{x}	Sample mean
α	Parameter in Gumbel's distribution
β	Parameter in Gumbel's distribution

μ	Distribution mean
σ	Standard deviation
ρ	Spearman's rank order correlation coefficient
Super- & subscripts	
norm	Normalized
agg	Aggregate
i	Index for unit (WWTP)
j	Index for individual KPI/ index for composite index construction step
k	Index for iteration in uncertainty analysis

1. Introduction

The spread of sensors and monitoring techniques in industry has multiplied the amount of data available for operating and controlling a process. Actually, the current bottleneck to take advantage of the available data is their proper interpretation and transformation into information, and, hence, support decision making. Among the different manners to translate data into meaningful information, composite indexes or indicators have become extremely popular in the last twenty years (Greco et al., 2018). These indexes are characterised as “[...] *based on sub-indicators that have no common meaningful unit of measurement and there is no obvious way of weighting*”. They are widely used to summarise, rank and/or benchmark the performance in several dimensions of systems of different degrees of complexity: from the energy efficiency of common appliances (European Commission, 2010) to the performance of national football teams (Fédération Internationale de Football Association, 2019), universities (ARWU, 2019) and even countries welfare (United Nations Development Programme, 2016). Bandura (2008) identified 178 official composite indexes used to rank countries performance in politics, welfare or environmental aspects and many more are being proposed continuously. However, not all indexes become widely used. For an index to be successful, it must be robust, transparently built, and stakeholders and users must understand how it properly captures the performance objective (Greco et al., 2018). Otherwise, perception of an arbitrary framework and/or big influence of the choices taken by index proposers will likely jeopardise its dissemination (Liu et al., 2019).

As increasing energy efficiency is considered as one of the main actions against global warming (European Commission, 2012), composite indexes are becoming increasingly popular to evaluate energy efficiency in different fields and help decision making. These indexes are ultimately based on the definition of energy efficiency (Directive 2012/27/EU) as “the ratio of output of

performance, service, goods or energy, to input of energy”. But while applying the previous definition is more or less clear in production technologies, where the outputs are clearly defined, in environmental technologies such as wastewater treatment, there is no widely accepted definition of energy efficiency (Longo et al., 2016).

Wastewater treatment is, in particular, the largest public consumer of electricity accounting for approximately 1% of the electricity consumption in industrialised countries (Ganora et al., 2019), a steady growing trend as the treatment requirements become increasingly stricter. Reducing the energy consumption by increasing the energy efficiency of wastewater treatment plants (WWTPs) is an obviously desired goal by plant managers and operators, consumers, public water boards and agencies, given that energy accounts for the 10-40 % of the operating costs, depending on the technology used (Haslinger et al., 2016; Panepinto et al., 2016; Venkatesh and Brattebø, 2011). Furthermore, from a purely environmental perspective it is essential to prevent the risk of “impact transfer”, i.e. transferring the eutrophication risk to global warming risk by using large amounts of energy to treat the wastewaters (Guinee, 2002). From a Life Cycle Thinking perspective, any intervention entails a risk of problem shifting: the solution to a particular problem may create another one elsewhere, hence the importance of increasing the energy efficiency (as a ratio between the WWTP objective and the energy demand) instead of just reducing the energy consumption (Corominas et al., 2013). Despite the many motivations for increasing efficiency, the lack of a definition and measurement of energy efficiency in WWTPs has prevented the deployment of wide and systematic actions to tackle the challenge of reducing the energy consumption in wastewater treatment. Most of the actions have been carried out on a single plant or group of plants without a common framework available (Borzooei et al., 2019) . Therefore, this article tackles two important knowledge gaps: the proposal of a common definition of energy efficiency for a WWTP and how

this energy efficiency can be measured and communicated by a single composite index. There are many diverse stakeholders involved or interested in WWTP energy efficiency (Luo et al., 2019), including plant operators, consumers, public administrators and utility managers. Composite indexes, thanks to their simplicity, have been widely used for communication to both specialized and wide public, also in the water sector (Molinos-Senante et al., 2014). As this is the first index for WWTP energy efficiency, particular attention is paid to transparency and robustness, so that the index can be applied trustfully by all stakeholders. Other relevant methods for energy efficiency benchmarking include data envelopment analysis (Lorenzo-Toja et al., 2015) and clustering techniques (Li et al., 2018) but are hardly used for wide benchmarking and their weights change from plant to plant, which limits the perceived transparency by the users. With respect to another relevant benchmarking method, the German standard DWA-A 216 E (2015), the method proposed here relates to the actual WWTP process performance rather than to the individual equipment, allowing a more general assessment of the process energy efficiency.

The proposal of the energy efficiency index was carried out in recently finished H2020 project ENERWATER (www.enerwater.eu), which focused on developing a standard methodology to define, evaluate and diagnose energy efficiency in WWTPs. The methodology covers all the steps from energy measurement and estimation, definition of indicators, determination of the water treatment energy index (WTEI), assignment of energy label and diagnostics of inefficient processes. As part of the ENERWATER methodology (Longo et al., 2019), this article covers in detail the design of a composite index (the WTEI) to benchmark and rank the energy efficiency of wastewater treatment plants from individual indicators. Hence, the manuscript is organised as follows: first the standard steps (OECD, 2008) in the construction of a composite index are described; then the attributes and the evaluation method for the index construction are defined and, through a sensitivity and uncertainty analysis, the construction method is evaluated. The design process is

iterated based on the sensitivity analysis results and the steps responsible for most of the index uncertainty are corrected; and finally, an algorithm for calculating the composite index is proposed.

2. Description of the ENERWATER methodology

ENERWATER methodology is described in detail elsewhere (Longo et al. 2019) and briefly summarised here for completeness. There are two versions of the methodology that differ mainly on the level of detail of the information required. The Decision Support methodology provides not only the WTEI but also a diagnosis of the inefficient spots in a plant, requiring for this endeavour an intense sampling campaign and analyses of influent and effluent of different plant sections. The Rapid Audit methodology only provides the WTEI for benchmarking using as inputs routine sampling and energy measurements. Both approaches basically consist of 1) measuring energy use and WWTP performance in terms of flowrate treated, COD and nutrient removal, sludge disposal and pathogen load reduction; 2) determining or estimating key performance indicators (KPI) that represent the efficiency of the different WWTP functions; 3) expressing the KPIs in a single index (WTEI) that can be used for benchmarking energy efficiency of different WWTPs and 4) assigning a label (A, B, C...) for ease of communication to a broad public. This work will use the Rapid Audit methodology as a demonstration but the results are extendable to the Decision Support methodology.

The KPIs (table 1) have been defined in order to represent the plant energetic performance in different dimensions while being flexible to adapt to different WWTP layouts.

Table 1. Definition of KPIs used in ENERWATER Rapid Audit methodology

KPI #	KPI units	Meaning/function
KPI₁	kWh/m ³	Energy for pumping wastewater through the WWTP
KPI₂	kWh/kg TPE _{rem} *	Energy for removal of COD, nitrogen and phosphorus
KPI₃	kWh/(logReduction·m ³)	Energy for removal of pathogens
KPI₄	kWh/kg TS _{proc}	Energy for treatment of sludge and removal of solids

*Where the total pollution equivalent TPE = COD (kgCOD) + 20 TN (kgTN) + 100 TP (kgTP) as defined by (Benedetti et al., 2008)

2.2 Description of case study

The set of data used in this study is part of the ENERWATER database which includes energy consumption and removal data, with different levels of detail, of 399 WWTPs. The dataset was gathered i) from the water utilities participating in ENERWATER in Spain, Italy and Germany; ii) by web-search engines; and iii) by collecting energy data from regional water agencies (in particular from Germany, Spain and Switzerland). Only those plants for which the four KPIs defined in table 1 are available were included. The resulting dataset is thus composed by 13 demo plants studied in detail within the framework of the ENERWATER project between 2015 and 2018 (4 from Italy and 9 from Spain) and 32 plants whose data have been gathered from the literature (9 from USA in 1973 and 22 from Canada in 1994). Descriptive statistics for all variables used in the analysis are given in Table 2. Both the database and the computer code used in this case study are available upon request from the authors.

Table 2. Descriptive statistics for the dataset used

Variable	Units	Obs.	Mean	SD	Min	Max
KPI ₁	kWh/m ³	44	0.482	0.346	0.118	2.34
KPI ₂	kWh/kg TPE _{rem} *	44	0.463	0.244	0.165	1.74
KPI ₃	kWh/(logReduction·m ³)	44	$9.00 \cdot 10^{-2}$	$7.82 \cdot 10^{-2}$	$1.97 \cdot 10^{-2}$	$48.9 \cdot 10^{-2}$
KPI ₄	kWh/kg TS _{proc}	44	2.77	1.22	0.702	6.01

3. Framework to build a composite indicator

The KPIs, as formulated in table 1, can be combined following different steps, each with several possible methods, leading to a large combinatorial problem. One of the major criticisms faced by composite index relates precisely with the selection of steps, methods for each step and, ultimately, the influence of these choices on the composite index value. In this context, the OECD has issued a Handbook on the Construction of Composite Indexes (OECD, 2008) which is partially followed here in the construction of the composite index algorithm, adapting it to the specificity of wastewater treatment. Once the individual indicators have already been defined, the handbook proposes the following steps for composite index construction: normalisation, weighting and aggregation. In this work, we have added a final step where the index is assigned to a performance label such as those used in energy efficiency labelling (A, B, C ...). The assignation between the index and the label requires a distribution function given that labelling in energy efficiency is commonly based on the performance percentile of a composite index (European Commission, 2010). All the steps are mathematically defined here and, if required, their adaptation to wastewater treatment is clarified.

3.1 Normalisation

A normalisation procedure is necessary to ensure that the individual indicators (here the KPI) are comparable before aggregation. Sound normalisation methods must be robust to outliers and can serve to transform nonlinear indicators into linear ones. Commonly used normalisation methods for continuous variables are standardisation (eq. 1), min-max normalisation (eq. 2) and categorical scales (eq. 3):

$$\text{Standardisation} \quad KPI_{i,j}^{norm} = \frac{KPI_{i,j} - \overline{KPI}_j}{\sigma(KPI_j)} \quad (1)$$

$$\text{Min-max} \quad KPI_{i,j}^{norm} = \frac{KPI_{i,j} - \min(KPI_j)}{\max(KPI_j) - \min(KPI_j)} \quad (2)$$

$$\text{Categorical scale (e.g.)} \quad KPI_{i,j}^{norm} = \begin{cases} 0 & \text{if } KPI_{i,j} < P_j^{15} \\ 25 & \text{if } P_j^{15} \leq KPI_{i,j} < P_j^{40} \\ 50 & \text{if } P_j^{40} \leq KPI_{i,j} < P_j^{65} \\ 75 & \text{if } P_j^{65} \leq KPI_{i,j} < P_j^{90} \\ 100 & \text{if } P_j^{90} \leq KPI_{i,j} \end{cases} \quad (3)$$

Where index i represents a given WWTP, \overline{KPI} and $\sigma(KPI_j)$ are the mean and standard deviation value for a given KPI, and P stands for the percentile indicated as superscript.

The obvious advantage of the standardisation approach (eq.1) is that it is zero-centred in the mean value, i.e. positive normalised KPI are larger than the average and negative are lower than the average giving a straightforward interpretation. Furthermore, according to Chebyshev inequality, 89% of the data would fall between $\pm 3\sigma(KPI_j)$ which translates to a normalised value of ± 3 (and up to 99.7% fall between ± 3 if the data are close to a normal distribution). These rules of thumb give a good intuition of how a standardised KPI is close or far from the average.

The min-max approach (eq. 2) leads to values comprehended between 0 and 1. Hence it gives the user an idea of the range of the set of data; furthermore for nonskewed data the average will be close to 0.5 but, in contrast, the average can be even close to 0 or 1 for heavily tailed datasets.

A categorical scale, such as the one defined in eq. (3) provides similar information as both of the previous approaches, i.e. an intuitive interpretation of the range of the data and where the median is located. A non-negligible shortcoming is that using percentile intervals to normalise the data, there are edge-effects whereby data uncertainty can have a large impact on the normalised value.

3.2 Weights selection

Weights in composite index construction are meant to attribute a certain importance to each KPI as they are combined. As such, weights have a major effect on the composite index, prioritizing different WWTP functions or individual indicators. Equal weights (EW) implies that all KPI are given equal importance; it is the most common method thanks to its simplicity, perceived objectivity or the fact that it is used “by default”, unless theoretical and/or empirical knowledge justify a different weighting scheme (Greco et al., 2018) .

Equal weights
$$w_j = w \quad i = 1, \dots, n \quad (4)$$

Another relevant method is the *budget allocation process (BAP)* which relies on using knowledge from a pool of experts to assign different weights, reflecting the relevance of each indicator. It depends heavily on the legitimacy of the expert opinion to arrive/justify the selection of weights. A major shortcoming is that expert allocated weights can reflect specific concerns in a region which are not common elsewhere, restricting the applicability of the weights.

Overcoming the lack of objectivity of the BAP method, a number of developers have conceived methods to obtain weights from data, KPIs, etc. The contribution of the KPIs to an overall performance is a way to obtain weights. A different data-driven method to obtain weights is the *benefit of doubt* (BOD) approach, which assigns weights to each unit (WWTP) so that the overall index is maximized (being very closely related to data envelopment analysis). The advantage of such a method is that it rewards the units for their best performance. Alternatively, it leads to conformism as there is little incentive to improve in the worst performance area of each unit. Furthermore, the benefit-of-doubt weights are not equal for all the units, which may not be perceived as transparent. The weights are determined by solving the following linear program (5):

$$\begin{aligned}
 & \text{Benefit of doubt} && w_{i,j} = \arg \max \sum_{j=1}^m w_{i,j} KPI_{i,j} \quad i = 1, \dots, n \\
 & && \text{subject to} \\
 & && \sum_{j=1}^m w_{i,j} KPI_{i,j} \leq 1 \quad i = 1, \dots, n; \quad j = 1, \dots, m \\
 & && w_{i,j} \geq 0 \quad i = 1, \dots, n; \quad j = 1, \dots, m
 \end{aligned} \tag{5}$$

3.3 Aggregation methods

Aggregation refers to the combination of the information from the individual indicators into a single overall index. The two main aggregation methods are the linear additive aggregation and geometric aggregation. Linear additive aggregation, defined in eq. (6), leads to compensability between

different indicators, i.e. poor performance in a given indicator can be compensated by performing well enough in other tasks.

Linear additive aggregation

$$KPI_{agg,i} = \sum_{j=1}^m w_{i,j} KPI_{i,j} \quad (6)$$

with $0 \leq w_{i,j} \leq 1$ and $\sum_{j=1}^m w_{i,j} = 1$

In contrast, geometric aggregation (eq. 7) entails severe non-compensability and requires good performance in all the indicators (Cai et al., 2019). As a consequence, even if the weights are equal, each unit is interested in improving the worst indicator as it has the largest impact on the overall indicator.

Geometric aggregation

$$KPI_{agg,i} = \prod_{j=1}^m KPI_{i,j}^{w_{i,j}} \quad i = 1, \dots, n \quad (7)$$

3.4 Distribution functions for percentile assignment

Distribution functions are commonly used in labelling and composite indexes to compare an item with a pool of peers. After an internal discussion by the members of the ENERWATER consortium, three distributions are proposed here with different characteristics to describe a population: normal, Gumbel's and t-Student distribution. This selection tried to cover different scenarios of the distribution of WWTP performance. Therefore, Gumbel's distribution was chosen as appropriate to describe one-side heavy-tail distributions and has been used in extreme-value modelling (Beirlant et

al., 2012) such as rare climate events: it would be the case if few WWTPs had an extremely poor performance. In contrast, non-standardized t-Student can accommodate to a certain extent higher two-side heavy-tails than normal distribution (Lange et al., 1989), hence being more appropriate than normal if WWTPs performance is distributed in both good and bad performance, away from the average. The cumulative distribution functions are widely available and, for completeness, are defined in Supp. Material.

4. Proposal and evaluation of the composite index algorithm

Following the guidelines given by the Handbook on Constructing Composite Indicators (OECD, 2008) and internal discussion in the ENERWATER consortium, a flowchart for determining the composite indicator was agreed and is represented in Figure 1.

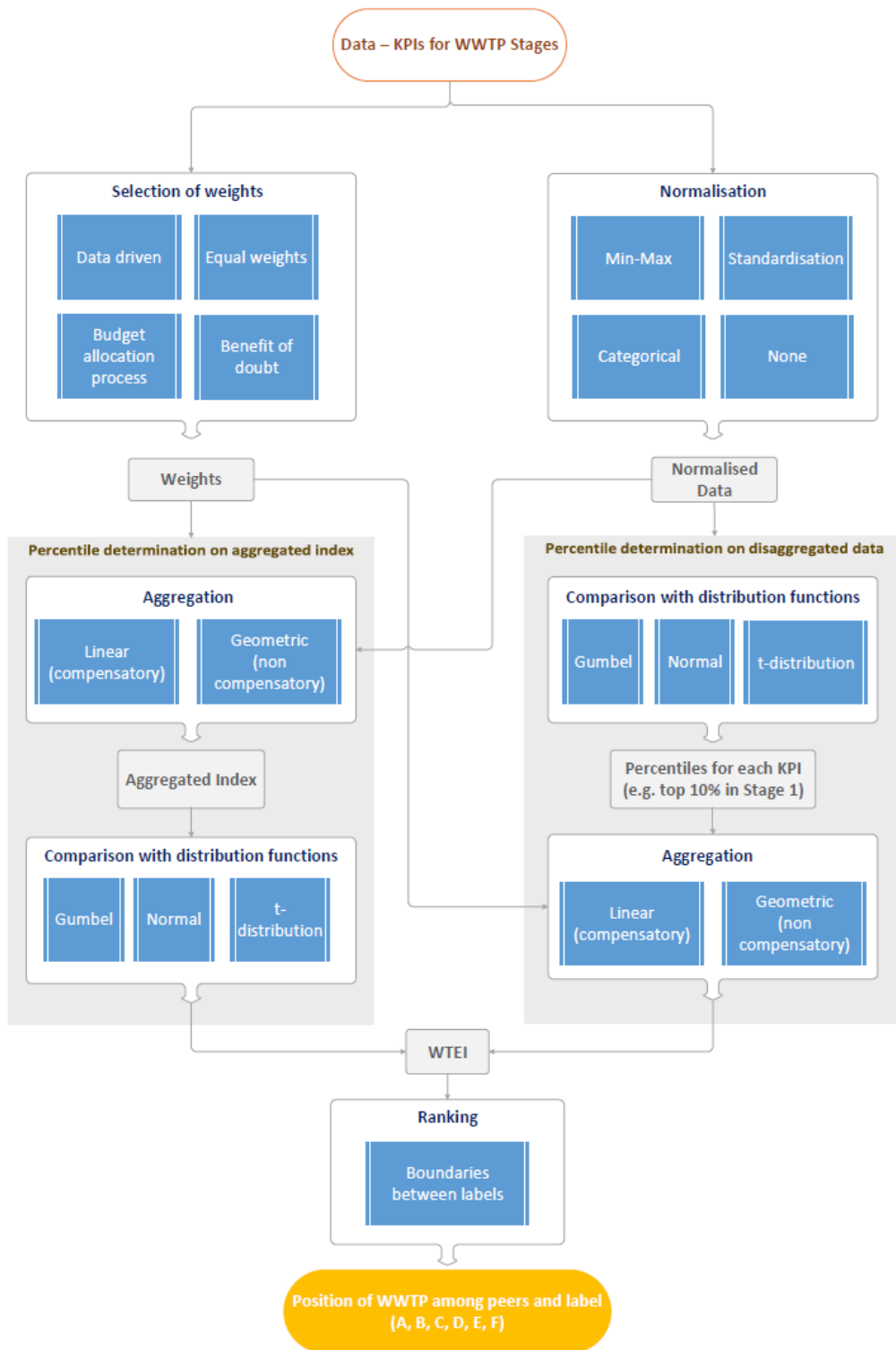


Figure 1. Flowchart for transforming KPIs into the composite index (WTEI) and posterior label allocation by Algorithm A showing the successive the steps and the possible methods in each one.

An interesting feature from Fig. 1 is that the comparison with the distribution function can be carried out with the data prior to aggregation or with the aggregated index. Hence, a fifth input is defined as “*Percentile determination*” which can take two values, whether the percentile of the distribution function is determined on the aggregated data (and hence only one distribution function is needed) or, the percentile is determined on the disaggregated data (and hence as many distribution functions as KPIs are needed). Designing the algorithm from the flowchart is effectively equivalent to select each of the steps along the transformations of the KPIs in the composite index. This can be formalised as selecting an entry in each of the rows of table 3.

Table 3. Summary of inputs and the value assigned to each of the governing choices in index building

Input description /Value	1	2	3	4
Normalisation (U_1)	<i>Min-Max</i>	<i>Standardisation</i>	<i>Categorical</i>	<i>None</i>
Weights (U_2)	<i>Data driven</i>	<i>Equal weights</i>	<i>BAP</i>	<i>BOD</i>
Aggregation (U_3)	<i>Linear additive</i>	<i>Geometric</i>		
Distribution Function (U_4)	<i>Normal</i>	<i>Gumbel</i>		
Percentile determination (U_5)	<i>On aggregated data</i>	<i>On disaggregated data</i>		

On the basis of previously studied composite indexes (Oree and Sayed Hassen, 2016) and internal discussion, a base algorithm was defined by making selecting the following sequence of inputs $U = [3\ 3\ 1\ 2\ 1]$, giving place to hereinafter called Algorithm A:

- Normalisation. Categorical scale ($U_1 = 3$), normalising all the KPIs into five categories that can be interpreted as percentiles with values 0, 25, 50, 75 and 100.
- Weights. Budget allocation process ($U_2 = 3$), whereby each KPI is given a weight equal to the contribution of the function to the overall energy consumption, as agreed by the experts in the consortium. These weights were determined as the average share of each function in the whole ENERWATER database (399 WWTPs), equal to [0.119, 0.535, 0.121, 0.225], which can be interpreted as pumping (KPI_1) accounting for 11.9% of the plant energy consumption, COD and nutrient removal (KPI_2) accounting for the 53.5%, and so forth. Note that, in contrast with the data driven method, these weights are fixed and do not depend on the actual sample of WWTPs.
- Aggregation, ($U_3 = 1$). Linear additive, as energy efficiency is considered as a desirable feature in a WWTP but not a requirement; hence, it is reasonable that energy inefficiency in a given function can be compensated to a certain extent by being efficiency in another function.
- Distribution Function, ($U_4 = 2$) Gumbel's function was selected as i) it reproduces well commonly observed efficiency distributions (i.e. large fraction of agents close to the average and long right tail) and ii) it can accommodate new observations out of the limits of the original database.

- Percentile determination, ($U_5 = 1$). Following the most common energy efficiency labelling methods, the percentiles are calculated on the aggregated data, leading to the WTEI that can be considered a percentile itself.

4.1 Evaluation of the proposed algorithm

The composite index must have a number of characteristics to provide robust representations of the measured characteristic while being understandable by the users. Ideally, the composite index should not depend on arbitrary choices and, if any should be taken, it must be duly justified so that it can be accepted by the community of users. The following items were identified as requirements for a composite index in this case:

1. Transparency. The composite index must be obtained in a transparent manner to ensure reproducibility and applicability (Wiréhn et al., 2015). It is important that the stakeholders understand the role of each step or mistrust will make the index unused.
2. Robustness. Composite indexes are determined after a long sequence of steps but, as they reflect a property that can be evaluated, they must not vary when changing the steps or methods, in particular:
 - a. The weights selected to aggregate the individual indexes must ideally have a negligible influence on the composite index. Otherwise, the weights must be based on common practice or widespread expert knowledge.
 - b. The composite index should not depend on the methods selected for its construction, being robust to other choices.
 - c. The index should not vary for a database representing the same population of WWTPs. In effect, as it is likely that data of all the WWTPs of a given area are not

available, it is important to ensure that the database for establishing the index is large enough so that minimal differences are expected if new WWTPs are added to the database.

The first item depends on the construction of the methodology, for which the rationale is explained here in detail (Longo et al., 2019). The second item refers to the robustness of the composite index with respect to choices in the selection of each method in the successive steps (i.e. which element per row is selected from table 2) and the fact that the database used for comparison is a sample of the population of WWTPs. Uncertainty and sensitivity analysis is used in this work, where the uncertain inputs are the choices in table 2, i.e. variables $U_1 \dots U_5$. Item 2c, which refers to the robustness of the database to the addition of new WWTPs can be inferred from an analysis of the uncertainty of subsampling, i.e. taking subsets of the database of WWTPs and evaluating the potential changes in the overall index.

4.2 Uncertainty and sensitivity analysis

The ranking uncertainty for each plant and the sensitivity to each of the choices in the construction of the overall index were estimated by variance-based sensitivity analysis.

Let us consider a function F that corresponds to building the overall index and a function G whose outcome is a ranking of the plants according to the overall index:

$$\text{Ranking} \equiv Y = F(KPI, U) \quad (8)$$

$$WTEI \equiv Z = G(KPI, U) \quad (9)$$

where U defines the sequence of steps. The different choices for the construction of the overall index are then represented as inputs to the building method defined as F e.g., a sequence of inputs $U = [1\ 2\ 2\ 2\ 1]$ would mean that the overall index is built by standardizing the KPIs, using equal weights, geometric aggregation done on normalized data and distribution percentiles based on Gumbel's distribution (Table 2).

To determine the uncertainty of the outcomes and the sensitivity of each of the choices, we followed the procedure proposed by Saltelli et al. (2010) and briefly described here for completeness. The impact of an input on the overall index is approximated by its first order sensitivity coefficient:

$$S_j = \frac{V_{U_j} \left(E_{U_{\sim j}}(Y|U_j) \right)}{V(Y)} \quad (10)$$

and the total effect of an input, which includes also higher order and interactive effects, is given by

$$S_{Tj} = \frac{E_{U_{\sim j}} \left(V_{U_j}(Y|U_{\sim j}) \right)}{V(Y)} \quad (11)$$

To determine both S_j and S_{Tj} , two independent sampling matrices A and B are built randomly sampling from the input space, hence with dimensions $N \times J$, where N is the number of iterations in the uncertainty analysis and J is the number of inputs (in this case 5). A new matrix A_B is built so that it has dimensions $NJ \times J$ and composed of J submatrices A_B^j . Each A_B^j is identical to A except the column j -th which is from B . Taking each row of matrices A , B and A_B as inputs U and determining $(2+J)N$ times the overall index, the sensitivity indexes can be determined by replacing in (eq. 12-13) the following formulas:

$$V_{U_j} \left(E_{U_{\sim j}}(Y|U_j) \right) = \frac{1}{N} \sum_{k=1}^N F(B)(F(A_B^j) - F(A)) \quad (12)$$

$$E_{U_{\sim j}} \left(V_{U_j} (Y | U_{\sim j}) \right) = \frac{1}{N} \sum_{k=1}^N F(A) \left(F(A) - F(A_B^j) \right) \quad (13)$$

Latin Hypercube Sampling (Helton and Davis, 2003) was used to provide maximizing the coverage of the input space and the number of iterations, N , was taken as 10 000 which was enough to achieve convergence in the sensitivity indexes estimations.

4.3 Evaluation of ranking consistency

As robustness metric in the capability to establish a ranking by the composite index, we used Spearman's rank order correlation coefficient.

$$d_{ki} = Y_i(U_k) - Y_i(U_{nominal}) \quad (14)$$

$$\rho_k = 1 - \frac{6 \sum_{i=1}^{N_{plants}} d_{ik}^2}{N_{plants}(N_{plants}^2 - 1)} \quad (15)$$

where eq. (14) defines the difference between paired ranks which allows calculating ρ_k , Spearman's coefficient, for a given selection k of the input space. A correlation equal to 1 indicates that two methods rank WWTPs identically whereas a value of 0 indicates a completely different ranking. We defined two averaged metrics of the previous coefficients, to account for the average impact in the whole set of WWTPs. These are defined as:

$$\bar{\rho} = \frac{1}{N} \sum_{k=1}^N \rho_k \quad (16)$$

$$\bar{d} = \frac{1}{NN_{plants}} \sum_{k=1}^N \sum_{i=1}^{N_{plants}} |d_{ik}| \quad (17)$$

4.4 Evaluation of subsampling error

To assess the uncertainty associated with having a finite sample, much smaller than the actual WWTP population, we have carried out the determination of the WTEI considering that only a randomly selected subset of 11 WWTPs (25% of the total set) is known. These 11 WWTPs are used to estimate data driven aggregation weights (which vary as the sample of 11 WWTPs change) and the parameters of the distribution functions. Then, the WTEI is determined for the whole set of WWTPs, in practice extrapolating to unknown 33 WWTPs. This procedure was repeated 10 000 times, each time randomly selecting the 11 WWTPs.

5. Evaluation results for algorithm A

Uncertainty analysis was carried out as described previously to determine the impact of the choices in each of the steps on the WWTPs ranking. The ranking, rather than the WTEI, was used as the uncertainty analysis output given that the absolute value of the WTEI is not relevant for benchmarking. On the contrary, the value of a plant's WTEI with respect to the rest of the plant population is the key for comparison, which is closely equivalent to a given plant's ranking.

Complying with the requirement of transparency and reproducibility is at odds with considering the BOD as a possible manner for weight selection. As BOD leads to different weights for each unit, it can be seen as non-transparent. Furthermore, BOD weights are very sensitive to outliers putting in doubt the reproducibility requirement; in effect, a WWTP with an extremely good performance in one dimension would greatly reduce the weights of the remaining WWTPs in that dimension

(Longo et al. 2018) , having a significant impact on the final index. Therefore, the BOD option will not be considered in the rest of this paper. For the sake of completeness, we have included the uncertainty and sensitivity analysis including BOD in Supplementary Material.

The results of the uncertainty analysis (figure 2, numerical data in Supp. Mat) show that the ranking can vary, within a 95% interval, up to 24 positions (plant ID 39). As can be seen, most of the plants (59%) can be ranked confidently within up to four positions but for 25% the ranking uncertainty is higher than 6 positions. All the plants can vary at least one position depending on the method. As a consequence, taking different choices in the determination of the WTEI would lead to poor decision-making as the individual steps selected would have a large influence on plant ranking.

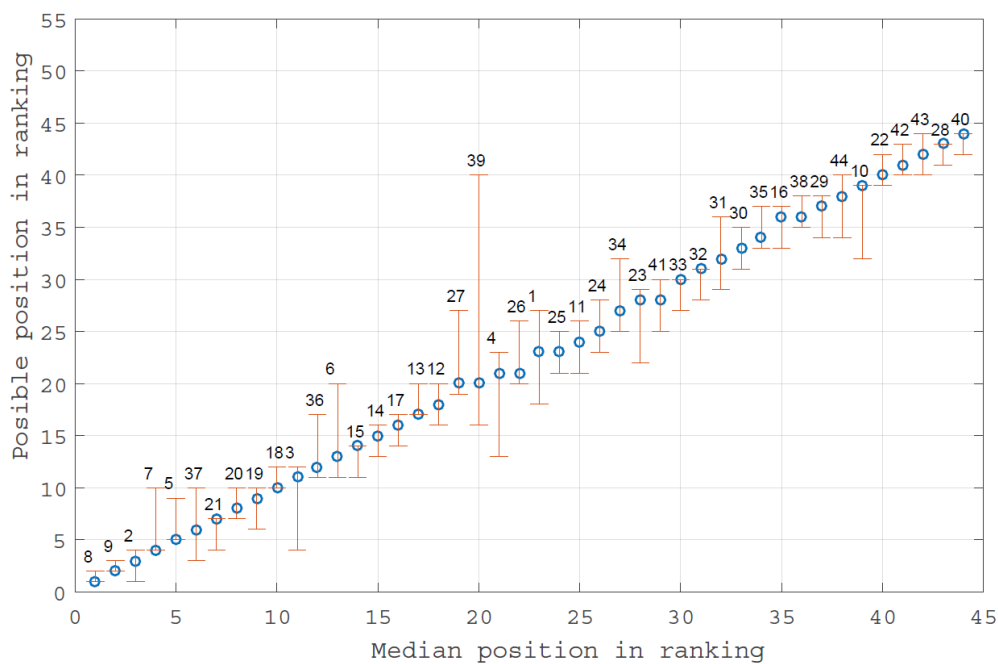


Figure 2. Uncertainty in plant ranking represented as the potential positions in ranking (2.5-97.5% percentile interval) vs the median position in the ranking. Numbers on top of the error bars are the plant ID

The root-cause for the changes in the ranking can be traced back to the different steps through sensitivity analysis. The sensitivity of each step in the WTEI construction, as defined in eq. 10, is represented in Figure 3. It can be seen that choices in every steps influence the ranking, in particular, normalisation and the non-additive contribution. This non-additive contribution refers to second-order effects which appear caused by the interactions between the choices in two or more steps, which therefore cannot be attributed to a single step alone. The importance of second-order effects is particularly relevant as they introduce uncertainty which is not justified by a rational choice. For example, it can be expected that linear and geometric aggregation lead to different results, as the first is a compensatory approach while the latter is mostly non-compensatory; hence choosing one aggregation method may respond to a policy decision. But second-order effects cannot be justified by policy decisions as they arise from interactions between choices in each of the stages and therefore, are to be suppressed.

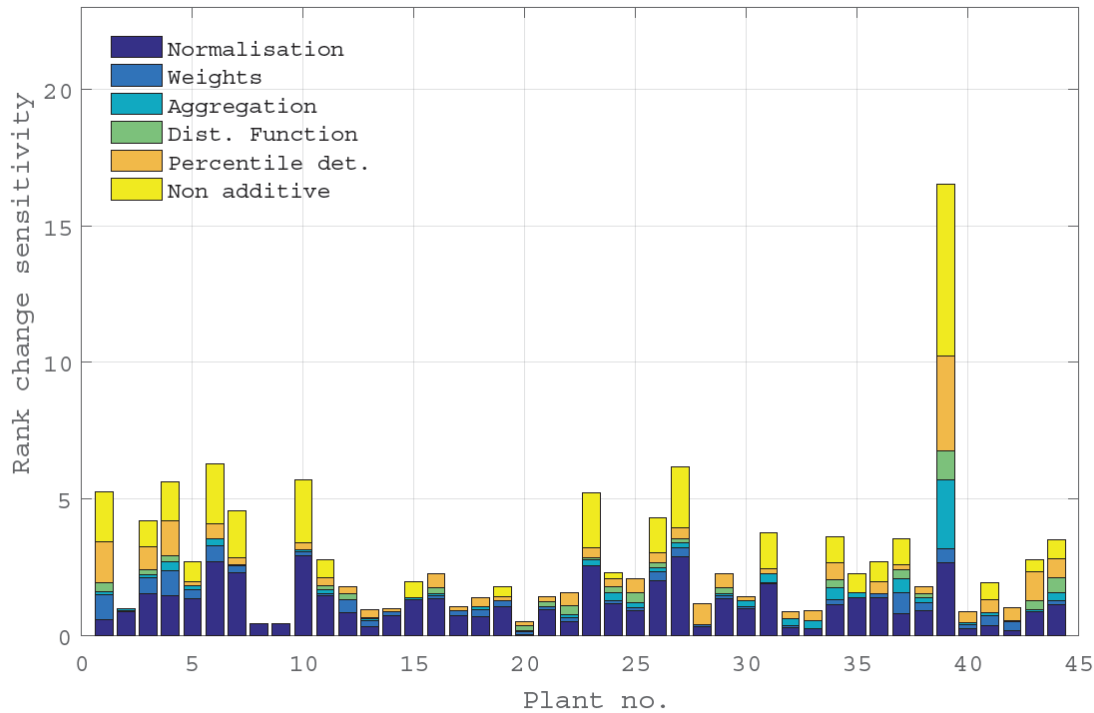


Figure 3. Sensitivity indexes (as defined in eq. 10). The non-additive component refers to second order effects coming from interactions between the different steps.

6. Proposal and evaluation of a new algorithm

From this root-cause analysis it is worth wondering whether it is possible to build the composite index in a way that it is more robust to choices. Taking into account that the second-order effects have a large influence we propose a different algorithm, with a different structure from those presented by the OECD Handbook (2008). Determining the percentiles of the normalized data (or of the raw data) is actually equivalent to merging the steps “*Normalisation*” and “*Distribution Function*”. Actually, comparing data with a distribution function is equivalent to using the “*Categorical*” normalisation method but with infinite categories, hence preventing undesired edge

effects. This new sequence (Algorithm B) is represented in Figure 4 and can be used when designing new robust composite indexes.

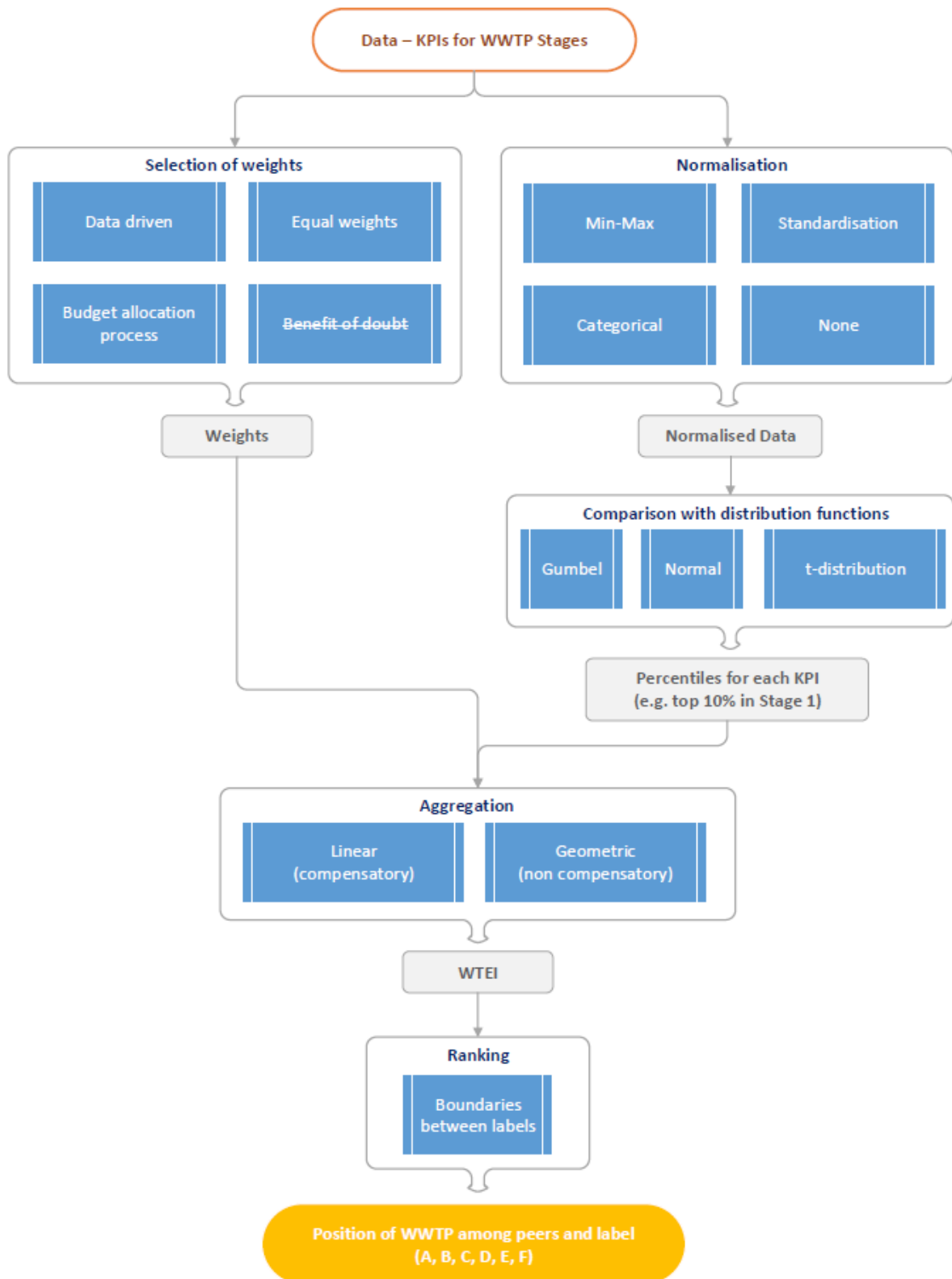


Figure 4. Flowchart for Algorithm B transforming KPIs into the composite index (WTEI) and posterior label allocation, showing the successive the steps and the possible methods in each one.

The uncertainty analysis on this new sequence of steps shows a more robust outcome (numerical data in Supp. Mat). Taking the 95% interval, 90% of the plants are confined within up to 4 positions whereas exceptions such as plant ID 39 presents a variation of up to 12 positions. For 20% of the plants, their positions are fixed regardless of the methods selected. The lower uncertainty related to this algorithm can be clearly observed in figure 5 where, except for plants ID 39, 37, 1 and 4, the rest are confidently ranked.

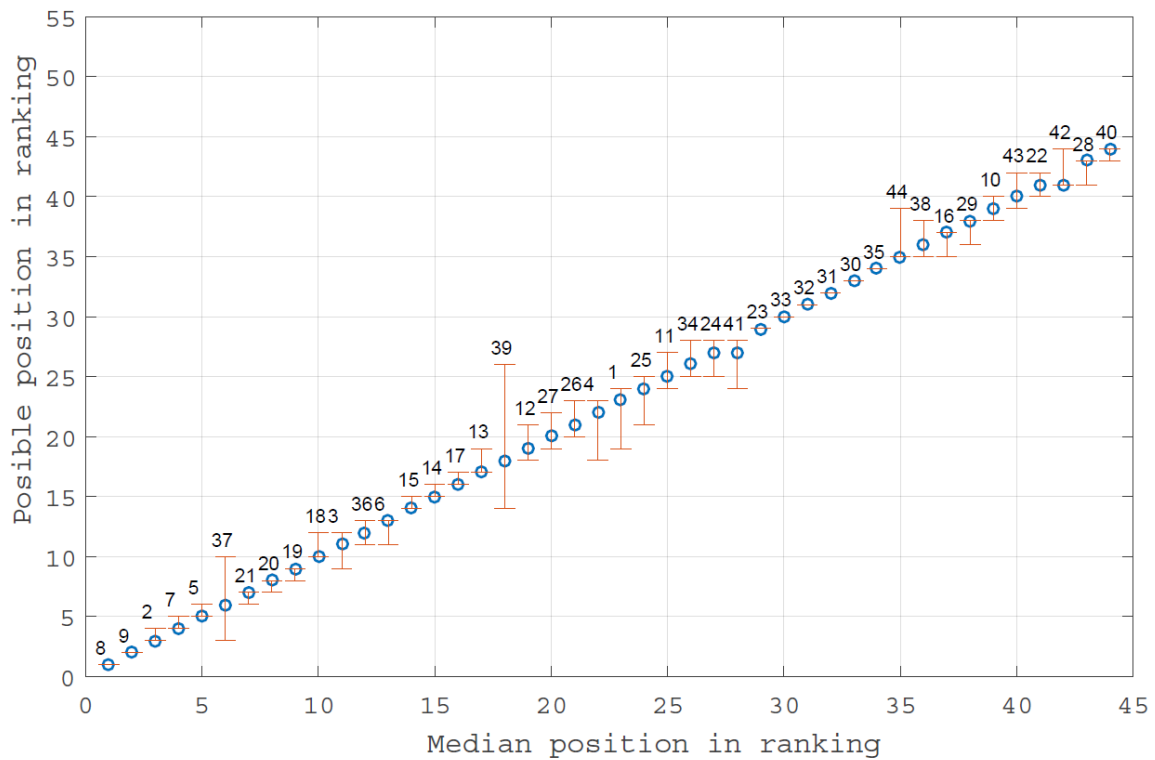


Figure 5. Uncertainty in plant ranking represented as the potential positions in ranking (2.5-97.5% percentile interval) vs the median position in the ranking for Algorithm B. Numbers on top of the error bars are the plant ID

The decomposition of the rank uncertainty (Fig. 6) shows that the selection of weights is now the most important factor for the large majority of plants, which is reasonable. It is not only impossible to make a composite index that is completely insensitive to the way it is calculated, it is also *undesirable*. Indeed, the goal when building a composite index is to obtain the same result regardless of arbitrary choices (such as using a given distribution function) but, on the other hand, the index should reflect the common practice in the sector. The use of weights selected by BAP representing the actual contribution to energy consumption in WWTP tries to reflect the different importance of each WWTP function, i.e., giving priority to energy efficiency in aeration rather than in tertiary treatment, as it is common practice in the wastewater treatment sector (Balmér and Hellström, 2012; Boiocchi et al., 2015).

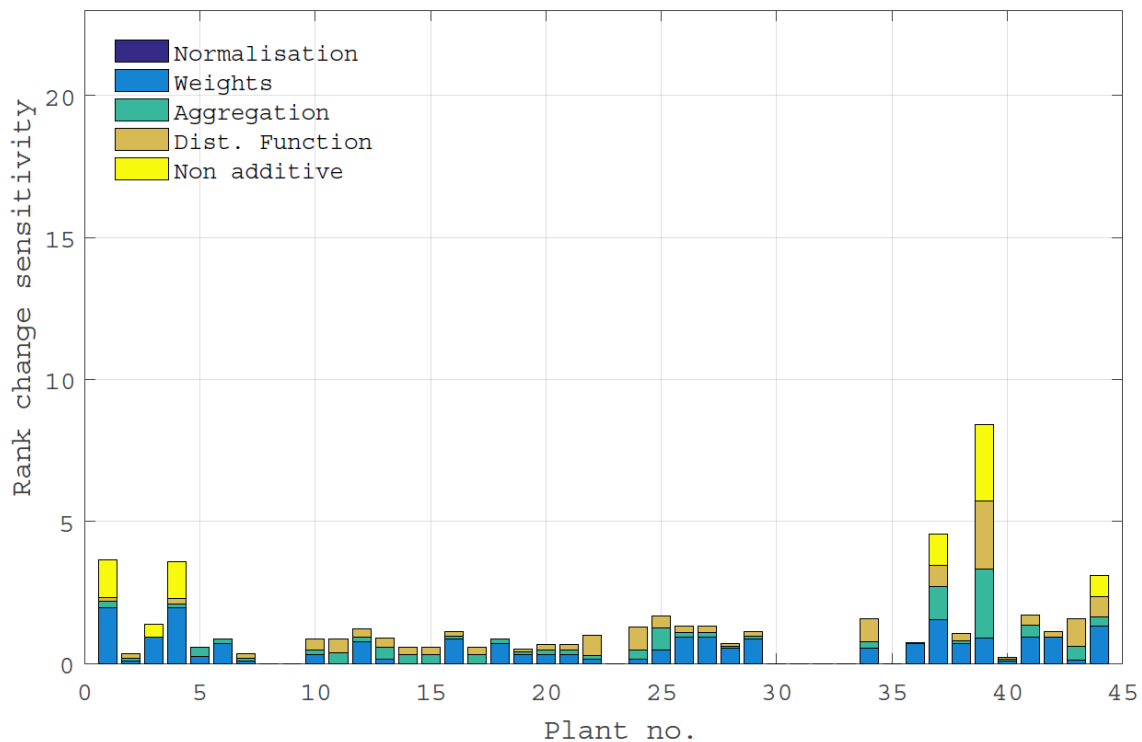


Figure 6. Sensitivity indexes (as defined in eq. 10). The non-additive component refers to second order effects coming from interactions between the different steps.

6.1 Comparison of ranking consistency

A common tool to assess the consistency of a ranking is by well-known Spearman's rank correlation coefficient. Here, Spearman's ρ (defined in eq. 15) was calculated between the nominal ranking method and each of the uncertainty analysis iterations. Although, both algorithms have a large average Spearman's coefficient (table 3), and close to 1, which means that all the variations are suitable for ranking the different WWTPs, it is seen that algorithm B is superior to algorithm A. This fact is confirmed when estimating the average ranking change per plant for each of the methods. It is expected that plants calculated with algorithm A would undergo a change of 2.21 positions if different methods are selected whereas this figure is reduced by a 70% to only 0.63 positions if following algorithm B. These indicators quantify the lower uncertainty associated with algorithm B as graphically presented in figures 2 and 5. As a conclusion, algorithm B is demonstrated as a more consistent method for building the WTEI.

Table 4. Ranking consistency indicators, as defined in equations 16-17, for different algorithms

Case	$\bar{\rho}$	\bar{d}
Algorithm A	0.970	2.21
Algorithm B	0.995	0.63
Algorithm B (subsampling uncertainty)	0.997	0.49

6.2 Subsampling uncertainty

It is important to acknowledge that the number of WWTPs in the database used for building the indicator is somewhat limited. As an important part of the WTEI estimation deals with the comparison with other WWTPs, this small sample number could lead to a large sensitivity to changes in the database. To assess the uncertainty associated with having a finite sample, much smaller than the actual WWTP population, we have carried out the determination of the WTEI considering that only a randomly selected subset of 11 WWTPs (25% of the total set) is known iterated 10 000 times.

The ranking consistency indicators show that even if only 25% of the WWTPs are taken as representative of their peers, the ranking would remain essentially unchanged. The representation of the potential positions in the ranking (fig. 7) confirms that, apart from plants ID 37, 39 and 43, the rest can be confidently ranked. In conclusion, the WTEI proposed can be in principle extended to other similar WWTPs out of the ENERWATER database.

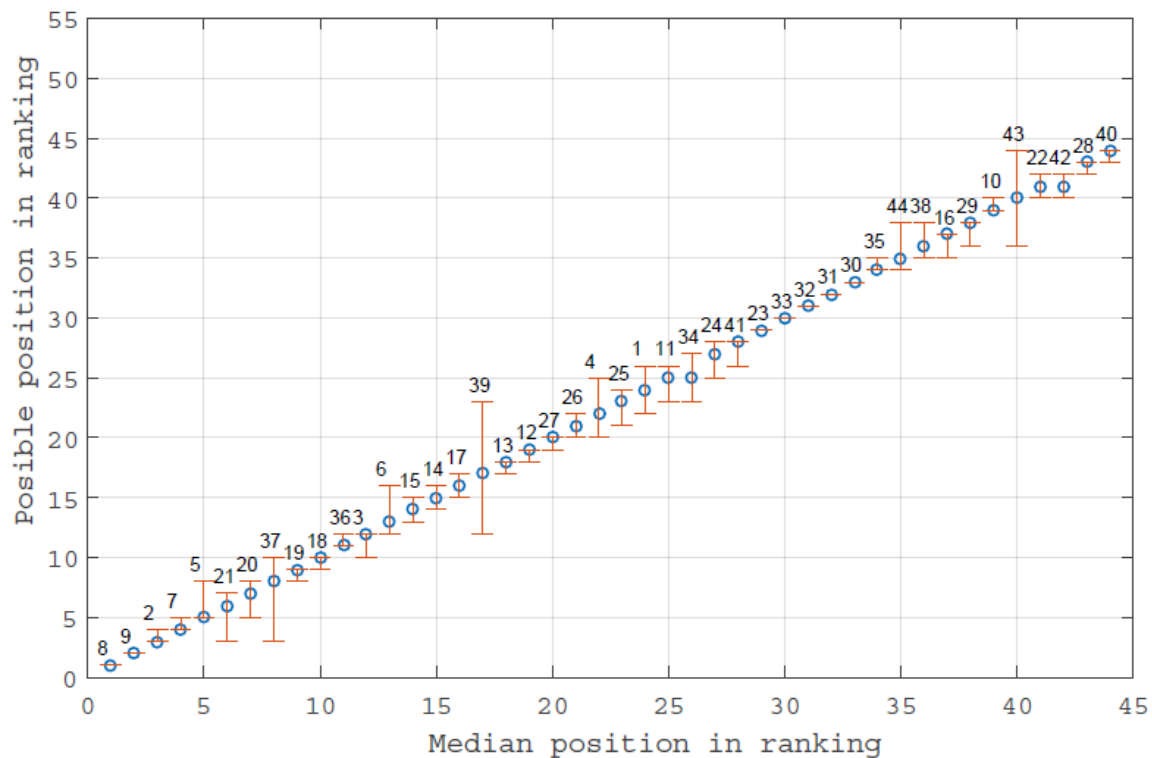


Figure 7. Uncertainty in plant ranking (2.5-97.5% percentile interval) caused by subsampling, selecting a subset of 25% of the WWTPs to determine the distribution function and the aggregation weights.

6.3 Implications as a policy support instrument

Since its approval, the Energy Efficiency Directive (2012/27/EU) is becoming an important driver for the targets on energy efficiency in Member States. Carrying out energy audit has evolved from convenient to an obligation for many water utilities, i.e. those with more than 250 employees and with annual trading volume greater than €50 million or whose annual balance sheet exceeds €43 million. For example, in Italy in 2017, 890 audits related to 302 water supplies, wastewater and waste management enterprises were carried out (European Commission, 2017). The Directive also recommends that, in provisions of public services with significant energy demand, tendering contracts should encourage contractors that provide long-term energy savings. A robust and

transparent index as the WTEI proposed in this contribution can facilitate the audits mandated by the Directive by making them more comparable. Besides, as the WTEI can be incorporated as a criterion in public tenders together with other green criteria, in so-called Green Public Procurement (Parikka-Alhola, 2008), it can be instrumental in helping the public administration apply energy saving actions in wastewater treatment.

7. Conclusions

Composite indexes are useful and increasingly used as means to communicate complex information in a summarised manner. We covered here the design of an index to represent the energy efficiency of a wastewater treatment plant, the WTEI, which is part of the broader ENERWATER methodology. The goal of the design of the index was to ensure that it properly reflected the energy efficiency of a wastewater treatment plant which is a multidimensional concept. Starting from commonly used steps in composite index construction, uncertainty analysis was used to achieve a robust algorithm that can be easily understood by users, a process that can be replicated in the conception of other composite indexes. It was also demonstrated that the WTEI is little sensitive to additions in the database of peers, making it suitable for extrapolation with other WWTPs and opening the door to the widespread use of the WTEI.

8. Acknowledgements

This work has been financially supported by ‘ENERWATER’ Coordination Support Action (www.enerwater.eu), as part of European Union's Horizon 2020 research and innovation programme under grant agreement No 649819. Authors belong to the Galician Competitive Research Group ED431C2017/029 and to the CRETUS Strategic Partnership (ED431E 2018/01),

both programmes co-funded by FEDER (EU). Authors would like to acknowledge the contribution of ENERWATER partners, in particular B. Mirko d'Antoni, Francesco Fatone and Ana Soares.

7. References

- ARWU, 2019. ARWU World University Rankings 2019 [WWW Document]. URL <http://www.shanghairanking.com> (accessed 4.25.19).
- Balmér, P., Hellström, D., 2012. Performance indicators for wastewater treatment plants. *Water Sci. Technol.* 65, 1304–1310. <https://doi.org/10.2166/wst.2012.014>
- Bandura, R., 2008. A Survey of Composite Indices Measuring Country Performance: 2008 Update. *Off. Dev. Stud. New York United ...* 1–95.
- Beirlant, J., Caeiro, F., Ivette Gomes, M., 2012. An overview and open research topics in statistics of univariate extremes. *Revstat Stat. J.*
- Benedetti, L., Dirckx, G., Bixio, D., Thoeye, C., Vanrolleghem, P.A., 2008. Environmental and economic performance assessment of the integrated urban wastewater system. *J. Environ. Manage.* 88, 1262–1272. <https://doi.org/10.1016/J.JENVMAN.2007.06.020>
- Boiocchi, R., Mauricio-Iglesias, M., Vangsgaard, A.K., Gernaey, K.V., Sin, G., 2015. Aeration control by monitoring the microbiological activity using fuzzy logic diagnosis and control. Application to a complete autotrophic nitrogen removal reactor. *J. Process Control* 30. <https://doi.org/10.1016/j.jprocont.2014.10.011>
- Borzooei, S., Campo, G., Cerutti, A., Meucci, L., Panepinto, D., Ravina, M., Riggio, V., Ruffino,

- B., Scibilia, G., Zanetti, M., 2019. Optimization of the wastewater treatment plant: From energy saving to environmental impact mitigation. *Sci. Total Environ.* 691, 1182–1189.
<https://doi.org/10.1016/j.scitotenv.2019.07.241>
- Cai, B., Guo, H., Ma, Z., Wang, Z., Dhakal, S., Cao, L., 2019. Benchmarking carbon emissions efficiency in Chinese cities: A comparative study based on high-resolution gridded data. *Appl. Energy* 242, 994–1009. <https://doi.org/10.1016/j.apenergy.2019.03.146>
- Corominas, L., Foley, J., Guest, J.S., Hospido, A., Larsen, H.F., Morera, S., Shaw, A., 2013. Life cycle assessment applied to wastewater treatment: State of the art. *Water Res.* 47, 5480–5492.
<https://doi.org/10.1016/j.watres.2013.06.049>
- DWA - German Association for Water Wastewater and Waste, 2015. Standard DWA-A 216E. Energy check and energy analysis - Instruments to Optimise the Energy Usage of Wastewater Systems.
- European Commission, 2017. Italy Energy Efficiency Action Plan.
- European Commission, 2012. Directive 2012/27/EU of the European Parliament and of the Council of 25 October 2012 on energy efficiency.
- European Commission, 2010. Directive 2010/30/EU of the European Parliament and of the Council of 19 May 2010 on the indication by labelling and standard product information of the consumption of energy and other resources by energy-related products. *Off. J. Eur. Union* 1–12.
- Fédération Internationale de Football Association, 2019. The FIFA/Coca-Cola World Ranking [WWW Document]. URL <https://www.fifa.com/fifa-world-ranking/index.html> (accessed 4.12.19).
- Ganora, D., Hospido, A., Husemann, J., Krampe, J., Loderer, C., Longo, S., Moragas Bouyat, L., Obermaier, N., Piraccini, E., Stanev, S., Váci, L., Pistocchi, A., 2019. Opportunities to

- improve energy use in urban wastewater treatment: A European-scale analysis. *Environ. Res. Lett.* 14, 44028. <https://doi.org/10.1088/1748-9326/ab0b54>
- Greco, S., Ishizaka, A., Tasiou, M., Torrisi, G., 2018. On the Methodological Framework of Composite Indices: A Review of the Issues of Weighting, Aggregation, and Robustness. *Soc. Indic. Res.* 1–34. <https://doi.org/10.1007/s11205-017-1832-9>
- Guinee, J.B., 2002. Handbook on life cycle assessment operational guide to the ISO standards. *Int. J. Life Cycle Assess.* <https://doi.org/10.1007/bf02978897>
- Haslinger, J., Krampe, J., Lindtner, S., 2016. Operating costs and energy demand of wastewater treatment plants in Austria: Benchmarking results of the last 10 years. *Water Sci. Technol.* <https://doi.org/10.2166/wst.2016.390>
- Helton, J.C., Davis, F.J., 2003. Latin hypercube sampling and the propagation of uncertainty in analyses of complex systems. *Reliab. Eng. Syst. Saf.* 81, 23–69. [https://doi.org/10.1016/S0951-8320\(03\)00058-9](https://doi.org/10.1016/S0951-8320(03)00058-9)
- Lange, K.L., Little, R.J.A., Taylor, J.M.G., 1989. Robust statistical modeling using the t distribution. *J. Am. Stat. Assoc.* <https://doi.org/10.1080/01621459.1989.10478852>
- Li, Z., Zou, Z., Wang, X., 2018. Energy efficiency evaluation for wastewater treatment plant. *Desalin. Water Treat.* 119, 276–281. <https://doi.org/10.5004/dwt.2018.22912>
- Liu, Y., Zhu, Y., Ren, L., Yong, B., Singh, V.P., Yuan, F., Jiang, S., Yang, X., 2019. On the mechanisms of two composite methods for construction of multivariate drought indices. *Sci. Total Environ.* 647, 981–991. <https://doi.org/10.1016/j.scitotenv.2018.07.273>
- Longo, S., D'Antoni, B.M., Bongards, M., Chaparro, A., Cronrath, A., Fatone, F., Lema, J.M., Mauricio-Iglesias, M., Soares, A., Hospido, A., 2016. Monitoring and diagnosis of energy consumption in wastewater treatment plants. A state of the art and proposals for improvement. *Appl. Energy* 179, 1251–1268. <https://doi.org/10.1016/j.apenergy.2016.07.043>

- Longo, S., Hospido, A., Lema, J.M.M., Mauricio-Iglesias, M., 2018. A systematic methodology for the robust quantification of energy efficiency at wastewater treatment plants featuring Data Envelopment Analysis. *Water Res.* 141, 317–328. <https://doi.org/10.1016/j.watres.2018.04.067>
- Longo, S., Mauricio-Iglesias, M., Soares, A., Campo, P., Fatone, F., Eusebi, A.L., Akkersdijk, E.A., Stefani, L., Hospido, A., 2019. ENERWATER – A standard method for assessing and improving the energy efficiency of wastewater treatment plants. *Appl. Energy* 242, 897–910. <https://doi.org/10.1016/j.apenergy.2019.03.130>
- Lorenzo-Toja, Y., Vazquez-Rowe, I., Chenel, S., Marín-Navarro, D., Moreira, M.T., Feijoo, G., 2015. Eco-efficiency analysis of Spanish WWTPs using the LCA + DEA method 68, 651–666.
- Luo, L., Dzakpasu, M., Yang, B., Zhang, W., Yang, Y., Wang, X.C., 2019. A novel index of total oxygen demand for the comprehensive evaluation of energy consumption for urban wastewater treatment. *Appl. Energy*. <https://doi.org/10.1016/j.apenergy.2018.11.101>
- Mahdi, S., 2005. Estimating Parameters of Gumbel Distribution Using the Methods of Moments, Probability Weighted Moments and Maximum Likelihood. *Rev.Mate.Teor.Aplic* 12, 151–156.
- Molinos-Senante, M., Gómez, T., Garrido-Baserba, M., 2014. Assessing the sustainability of small wastewater treatment systems : A composite indicator approach. *Sci. Total Environ.* 497–498, 607–617. <https://doi.org/10.1016/j.scitotenv.2014.08.026>
- OECD, 2008. Handbook on constructing composite indicators. Methodology and user guide, OECD. <https://doi.org/10.1787/533411815016>
- Oree, V., Sayed Hassen, S.Z., 2016. A composite metric for assessing flexibility available in conventional generators of power systems. *Appl. Energy* 177, 683–691. <https://doi.org/10.1016/j.apenergy.2016.05.138>
- Panepinto, D., Fiore, S., Zappone, M., Genon, G., Meucci, L., 2016. Evaluation of the energy efficiency of a large wastewater treatment plant in Italy. *Appl. Energy*.

<https://doi.org/10.1016/j.apenergy.2015.10.027>

Parikka-Alhola, K., 2008. Promoting environmentally sound furniture by green public procurement.

Ecol. Econ. <https://doi.org/10.1016/j.ecolecon.2008.05.004>

Saltelli, A., Annoni, P., Azzini, I., Campolongo, F., Ratto, M., Tarantola, S., 2010. Variance based sensitivity analysis of model output. Design and estimator for the total sensitivity index.

Comput. Phys. Commun. 181, 259–270. <https://doi.org/10.1016/j.cpc.2009.09.018>

United Nations Development Programme, 2016. Human Development Reports [WWW Document].

URL <http://hdr.undp.org/en/2016-report> (accessed 6.29.18).

Venkatesh, G., Brattebø, H., 2011. Energy consumption, costs and environmental impacts for urban water cycle services: Case study of Oslo (Norway). *Energy*.

<https://doi.org/10.1016/j.energy.2010.12.040>

Wiréhn, L., Danielsson, Å., Neset, T.S.S., 2015. Assessment of composite index methods for agricultural vulnerability to climate change. *J. Environ. Manage.* 156, 70–80.

<https://doi.org/10.1016/j.jenvman.2015.03.020>

**Supplementary Material to: Designing a robust index for WWTP energy efficiency: the
ENERWATER Water Treatment Energy Index**

Miguel Mauricio-Iglesias^{*}, Stefano Longo[‡], Almudena Hospido

CRETUS Institute, Department of Chemical Engineering, Universidade de Santiago de
Compostela, Spain.

[‡] Currently at Currently at HERA, Central Innovation Department, via Cesare Diana 40, 44121
Ferrara, Italy

* Corresponding author

A.1. Distribution functions defined for percentile assignment

The normal distribution has a cumulative distribution function defined as:

$$F(x) = \frac{1}{2} \left[1 + \operatorname{erf} \left(\frac{x - \mu}{\sigma\sqrt{2}} \right) \right] \quad (\text{A-1})$$

whose parameters μ and σ are estimated respectively as the sample mean and the sample standard deviation, i.e.

$$\bar{x} = \frac{1}{n} \sum_{i=1}^n x_i \quad (\text{A-2})$$

$$s^2 = \frac{1}{n-1} \sum_{i=1}^n (x_i - \bar{x})^2 \quad (\text{A-3})$$

Gumbel's probability distribution has a cumulative distribution function defined as:

$$F(x) = \exp\left(-\exp\left(-\frac{x-\alpha}{\beta}\right)\right) \quad (\text{A-4})$$

whose parameters α and β are estimated respectively following a maximum likelihood estimation method leading to the following equations (Mahdi, 2005). Eq. (A-5) is solved implicitly to obtain parameter β estimation and, by substitution, the value of α can be determined.

$$\bar{x} = \hat{\beta} + \frac{\sum_{i=1}^n x_i \exp\left(-\frac{x_i}{\hat{\beta}}\right)}{\sum_{i=1}^n \exp\left(-\frac{x_i}{\hat{\beta}}\right)} \quad (\text{A-5})$$

$$\hat{\alpha} = \hat{\beta} \left[\ln n - \ln \sum_{i=1}^n \exp\left(-\frac{x_i}{\hat{\beta}}\right) \right] \quad (\text{A-6})$$

The t-Student cumulative distribution is defined as

$$F(x) = \frac{1}{2} + z \Gamma\left(\frac{\nu+1}{2}\right) \times \frac{F\left(\frac{1}{2}, \frac{\nu+1}{2}, \frac{3}{2}, -\frac{z^2}{\nu}\right)}{\sqrt{\pi\nu} \Gamma\left(\frac{1}{2}\right)} \quad (\text{A-7})$$

$$\text{with } z = \frac{x - \hat{\mu}}{\hat{\sigma}}$$

where Γ is the gamma function, F is the hypergeometric function and ν is the number of degrees of freedom estimated for the distribution. An obvious issue of this distribution is that is not easily tractable analytically but, on the other hand, t-Student is preprogramed in many hand calculators, spreadsheets, etc.

A.2. Uncertainty analysis including benefit of doubt (BOD) as a potential way to assign weights.

Weights can be assigned following a benefit-of-doubt approach, i.e. allocating the highest weight to the best KPI for each WWTP. This approach considers that each WWTP is managed in order to optimise the objectives, which may not be common to other WWTPs. Hence, the array of weights changes from plant to plant, which is a drawback affecting the predictability of the index.

This section shows, through uncertainty analysis, the impact of including BOD as an index building alternative for weight selection. The matrix of input possibilities is shown in table 3 in the manuscript.

The results of the uncertainty analysis (figure A1) show that the ranking can vary widely. Plant ID 39 could vary up to 28 positions in a 95% interval and only 22% of the plants can be ranked within four positions. The mean position uncertainty is equal to 2.42 positions

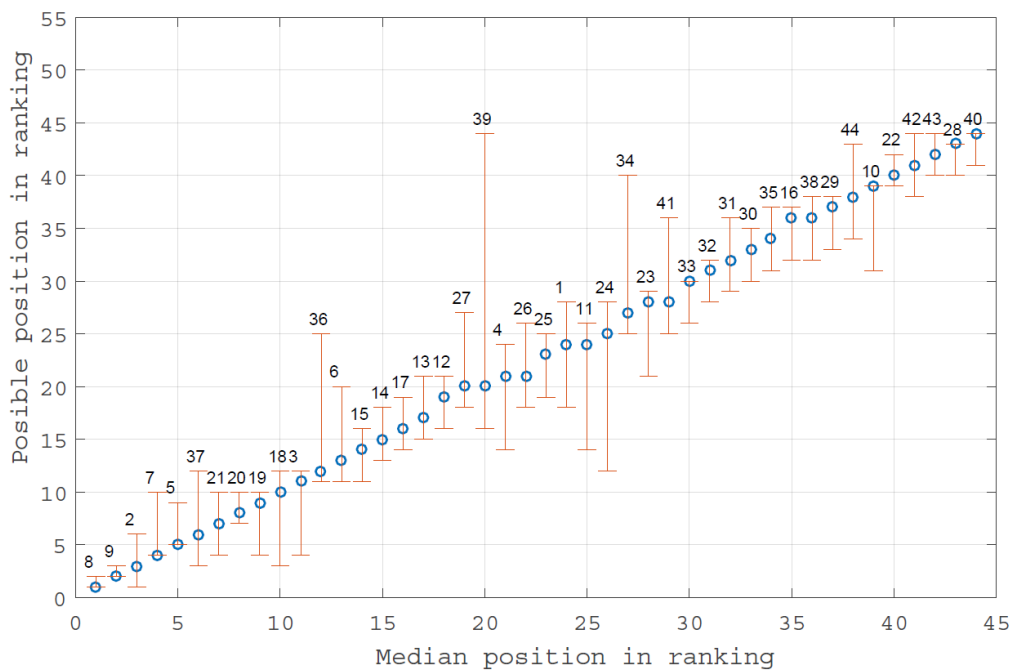


Figure A1. Uncertainty in plant ranking represented as the potential positions in ranking (2.5-97.5% percentile interval) vs the median position in the ranking. Numbers on top of the error bars are the plant ID

Examining the factors leading to this higher uncertainty, it is seen that the impact of adding the BOD does not lie on the selection of weights but on how these combine with the rest of methods. Indeed, the non-additive term is the one with the largest increase (Fig. A2) when compared with the results without BOD.

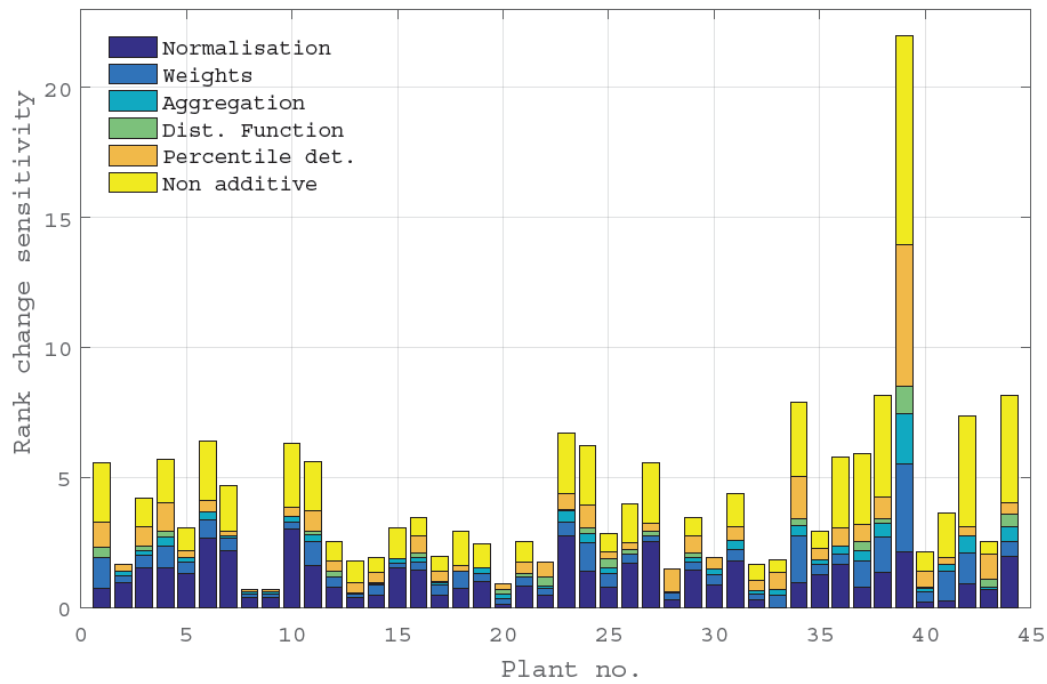


Figure A2. Sensitivity indexes (as defined in eq. 10).

A.3 Numerical results of uncertainty and sensitivity analysis

Table A1. Uncertainty in position (values for fig. 2) and index sensitivity (values for fig. 3) for algorithm A

Uncertainty in position results				Index sensitivity results					
Plant ID	Median position	Lower Errorbar	Upper Errorbar	Normalisation	Weights	Aggregation	Dist. Function	Agg. Order	Non additive
1	23	-5	4	0.57	0.94	0.10	0.33	1.48	1.84
2	3	-2	1	0.89	0.04	0.05	0.00	0.00	0.00
3	11	-7	1	1.52	0.59	0.13	0.17	0.84	0.96
4	21	-8	2	1.45	0.93	0.32	0.22	1.27	1.43
5	5	0	4	1.36	0.33	0.11	0.00	0.18	0.72
6	13	-2	7	2.70	0.59	0.27	0.00	0.55	2.19
7	4	0	6	2.31	0.24	0.06	0.00	0.24	1.73
8	1	0	1	0.43	0.00	0.00	0.00	0.00	0.00
9	2	0	1	0.43	0.00	0.00	0.00	0.00	0.00
10	39	-7	0	2.93	0.16	0.06	0.00	0.27	2.30
11	25	-3	2	1.45	0.10	0.13	0.14	0.29	0.67
12	18	-2	2	0.86	0.46	0.00	0.20	0.27	0.00
13	17	0	3	0.33	0.22	0.07	0.05	0.27	0.00
14	15	-2	1	0.72	0.15	0.00	0.00	0.12	0.00
15	14	-3	0	1.31	0.01	0.06	0.00	0.00	0.60
16	35	-3	1	1.34	0.14	0.07	0.20	0.52	0.00
17	16	-2	1	0.73	0.18	0.01	0.00	0.12	0.00
18	10	0	2	0.70	0.26	0.09	0.00	0.33	0.00
19	9	-3	1	1.07	0.20	0.00	0.00	0.16	0.38
20	8	-1	2	0.04	0.12	0.04	0.17	0.15	0.00
21	7	-3	0	0.95	0.10	0.01	0.17	0.18	0.00
22	40	-1	2	0.50	0.14	0.13	0.34	0.46	0.00
23	28	-6	1	2.56	0.00	0.20	0.07	0.40	2.00
24	26	-2	3	1.18	0.11	0.29	0.22	0.28	0.25
25	24	-2	2	0.92	0.10	0.20	0.34	0.51	0.00

26	22	-1	5	2.03	0.33	0.14	0.18	0.36	1.27
27	19	-1	7	2.89	0.33	0.16	0.17	0.40	2.22
28	43	-2	0	0.33	0.05	0.03	0.00	0.74	0.00
29	37	-3	1	1.34	0.14	0.07	0.20	0.52	0.00
30	33	-2	2	1.00	0.05	0.21	0.00	0.18	0.00
31	32	-3	4	1.91	0.04	0.30	0.00	0.19	1.32
32	31	-3	0	0.28	0.07	0.26	0.00	0.28	0.00
33	30	-3	0	0.26	0.00	0.28	0.00	0.38	0.00
34	27	-2	5	1.12	0.20	0.43	0.30	0.61	0.95
35	34	-1	3	1.38	0.00	0.18	0.00	0.00	0.72
36	12	-1	5	1.40	0.12	0.00	0.00	0.45	0.73
37	6	-3	4	0.79	0.77	0.53	0.33	0.17	0.96
38	36	-1	2	0.93	0.27	0.20	0.14	0.25	0.00
39	20	-4	20	2.65	0.52	2.53	1.05	3.48	6.31
40	44	-2	0	0.25	0.14	0.08	0.00	0.42	0.00
41	29	-3	2	0.37	0.36	0.12	0.00	0.45	0.63
42	41	-1	2	0.19	0.32	0.01	0.00	0.48	0.00
43	42	-2	2	0.88	0.00	0.06	0.34	1.04	0.44
44	38	-4	2	1.13	0.14	0.29	0.55	0.69	0.70

Table A2. Uncertainty in position (values for fig. 4) and index sensitivity (values for fig. 5) for algorithm B

Uncertainty in position results				Index sensitivity results			
Plant ID	Median position	Lower Errorbar	Upper Errorbar	Weights	Aggregation	Dist. Function	Non additive
1	23	-4	1	1.96	0.25	0.11	1.33
2	3	0	1	0.08	0.11	0.16	0.00
3	11	-2	1	0.93	0.00	0.00	0.45
4	22	-4	1	1.96	0.15	0.20	1.28
5	5	0	1	0.24	0.34	0.00	0.00
6	13	-2	0	0.70	0.17	0.00	0.00
7	4	0	1	0.08	0.11	0.16	0.00
8	1	0	0	0.00	0.00	0.00	0.00
9	2	0	0	0.00	0.00	0.00	0.00
10	39	-1	1	0.32	0.16	0.39	0.00
11	25	-1	2	0.00	0.37	0.48	0.00
12	19	-1	2	0.77	0.16	0.28	0.00
13	17	0	2	0.14	0.44	0.30	0.00
14	15	0	1	0.00	0.33	0.23	0.00
15	14	0	1	0.00	0.33	0.23	0.00
16	37	-2	0	0.86	0.11	0.17	0.00
17	16	0	1	0.00	0.33	0.23	0.00
18	10	0	2	0.70	0.17	0.00	0.00
19	9	-1	0	0.31	0.11	0.08	0.00
20	8	-1	0	0.31	0.17	0.21	0.00
21	7	-1	0	0.31	0.17	0.21	0.00
22	41	-1	1	0.15	0.14	0.70	0.00
23	29	0	0	0.00	0.00	0.00	0.00
24	27	-2	1	0.16	0.32	0.79	0.00
25	24	-3	1	0.48	0.78	0.41	0.00

26	21	-1	2	0.93	0.17	0.23	0.00
27	20	-1	2	0.93	0.17	0.23	0.00
28	43	-2	0	0.55	0.04	0.10	0.00
29	38	-2	0	0.86	0.11	0.17	0.00
30	33	0	0	0.00	0.00	0.00	0.00
31	32	0	0	0.00	0.00	0.00	0.00
32	31	0	0	0.00	0.00	0.00	0.00
33	30	0	0	0.00	0.00	0.00	0.00
34	26	-1	2	0.55	0.22	0.80	0.00
35	34	0	0	0.00	0.00	0.00	0.00
36	12	-1	1	0.69	0.05	0.00	0.00
37	6	-3	4	1.55	1.17	0.73	1.10
38	36	-1	2	0.70	0.10	0.27	0.00
39	18	-4	8	0.91	2.41	2.40	2.68
40	44	-1	0	0.09	0.05	0.09	0.00
41	28	-3	1	0.93	0.43	0.35	0.00
42	42	0	3	0.94	0.00	0.18	0.00
43	40	-1	2	0.13	0.48	0.96	0.00
44	35	0	4	1.33	0.31	0.70	0.76

1 **References**

2

3 Mahdi, S., 2005. Estimating Parameters of Gumbel Distribution Using the Methods of Moments,
4 Probability Weighted Moments and Maximum Likelihood. *Rev.Mate.Teor.Aplic* 12, 151–156.

5