UNIVERSITY OF MACEDONIA

DEPARTMENT OF BUSINESS ADMINISTRATION

A novel Multi-level Data Envelopment Analysis variation for the construction of composite indicators and the measurement of sustainability

Georgios Tsaples

A thesis submitted to the University of Macedonia for the degree of Doctor of Philosophy

Supervisor

Dr. Jason Papathanasiou (Associate Professor, Department of Business Administration)

Members of Advisory Committee

Professor Andreas C. Georgiou

Professor Nikolaos Samaras

Thessaloniki, 2022

UNIVERSITY OF MACEDONIA

DEPARTMENT OF BUSINESS ADMINISTRATION

A novel Multi-level Data Envelopment Analysis variation for the construction of composite indicators and the measurement of sustainability

PhD Dissertation by

Georgios Tsaples

Τριμελής Συμβουλευτική Επιτροπή:	Επταμελής Εξεταστική Επιτροπη	Υπογραφή
1. Ιάσων Παπαθανασίου, Αναπλ. Καθ. ΟΔΕ (Επιβλέπων)	Ιάσων Παπαθανασίου, Αναπλ. Καθ. ΟΔΕ, Πανεπιστήμιο Μακεδονίας (Επιβλέπων)	__________
2. Ανδρέας Γεωργίου, Καθ. ΟΔΕ	Ανδρέας Γεωργίου, Καθ. ΟΔΕ, Πανεπιστήμιο Μακεδονίας	__________
3. Νικόλαος Σαμαράς, Καθ. Εφαρμ. Πληροφορική	Νικόλαος Σαμαράς, Καθ. Εφαρμ. Πληροφορική, Πανεπιστήμιο Μακεδονίας	__________
	Γεώργιος Τσιότρας, Καθ. ΟΔΕ, Πανεπιστήμιο Μακεδονίας	__________
	Alexis Tsoukiàs, Prof. Université Paris Dauphine - PSL	__________
	Γεώργιος Αρετούλης, Αναπ. Καθ., Πολιτ. Μηχ. ΑΠΘ	__________
	Κωνσταντίνος Καπάρης, Επικ. Καθ. ΟΔΕ, Πανεπιστήμιο Μακεδονίας	__________

Θεσσαλονίκη, 2022

This page intentionally left blank

Acknowledgments

The completion of the present dissertation would not be possible without the assistance of a lot of persons. First and foremost, I am deeply indebted to my supervisor, Associate Professor Jason Papathanasiou, who gave me the opportunity and has supported me throughout my PhD with his enthusiasm for the research, patience, encouragement and knowledge. I would like to express my deepest appreciation to Professor Andreas C. Georgiou who took the time to contribute so thoroughly through his comments and suggestions. Moreover, I am grateful to Professor Nikolaos Samaras for his trust in allowing me to pursue the research wherever it led.

I could not have undertaken this journey without the love and guidance of my parents and my sister who support me unconditionally in whatever I pursue.

Lastly, I would be remiss in not mentioning my wife Eleni, who has stood by me through all my travails, absences, fits of pique and impatience. Her belief in me has kept my motivation and focus during this process. Her name should be on this thesis as much as mine.

Γεώργιος Τσαπλές

Θεσσαλονίκη, 2022

Publications

- Tsaples, G., Papathanasiou, J., Georgiou, A. C., & Samaras, N. (2019, May). Assessing multidimensional sustainability of European countries with a novel, two-stage DEA. *Lecture Notes in Business Information Processing*, 348, pp. 111-122. Springer, Cham (Q3)
- Tsaples, G., & Papathanasiou, J. (2020, May). Using multi-level DEA to go beyond the three dimensions of sustainability. *Lecture Notes in Business Information Processing, 384*, pp. 15-29. Springer, Cham (Q3)
- Tsaples, G., & Papathanasiou, J. (2020). Multi-level DEA for the construction of multi-dimensional indices. *MethodsX*, *7*, 101169 (CiteScore: 1.8, Q2)
- Tsaples, G., & Papathanasiou, J. (2021). Data envelopment analysis and the concept of sustainability: A review and analysis of the literature. Renewable and Sustainable Energy Reviews, 138, 110664 (IF: 14,982, Q1)
- Tsaples, G., Papathanasiou, J., Digkoglou, P.(2021). Decision-making in the context of Sustainability. In Urban Sustainability: A game-based approach. (Eds.) Papathanasiou, J., Tsaples, G., Blouchoutzi, A., Springer.
- Tsaples, G., & Papathanasiou, J. (2021). Measuring agricultural sustainability of European countries with a focus on transportation. International Journal of Sustainable Agricultural Management and Informatics, 7(4), 304-320. (CiteScore: 0.8, Q2)
- Tsaples, G., Papathanasiou, J., & Georgiou, A. C. (2022). An Exploratory DEA and Machine Learning Framework for the Evaluation and Analysis of Sustainability Composite Indicators in the EU. Mathematics, 10(13), 2277. (IF: 2.592, CiteScore: 2.9, Q2)

Abstract

In public decision-making factors such as personal values, cultural background and different individual perspectives play a central role in the policy cycle of design, test, implementation and review. To assist policy makers, analysts have used an array of qualitative and quantitative methods to all steps of the cycle.

However, the increasing use of sophisticated methods seems not to be always accompanied by an improvement in the quality of policy making; on the contrary, it seems to attract criticism that is focused on their disadvantages. Furthermore, the rise of Artificial Intelligence (AI) and its expanding use in decision- and/or policy-making, has brought forth the issue of interpretability of algorithms and whether their output can be trusted. Questions such as “which specific feature made the model/algorithm reach the specific decision”, hence issues of transparency and interpretability of the methods, are becoming central issues of the critique on quantitative methods and algorithms.

This criticism is not without its merits. The complexity of contemporary problems means that there are issues about which an analyst can only make assumptions due to the existence of deep uncertainty. Moreover, in such complexity, the perception of the analyst may limit the view of the policy cycle under study. As a result, the success of a quantitative method relies on all of the above choices to be exactly “correct”.

Sustainable development perfectly encapsulates these issues and in order to achieve it, public policies should have economic, social and environmental dimensions, while taking into account the current technological developments, the cultural context and the value system in which they are applied. Thus, sustainable development is a multi-dimensional concept and from early on arose the need to find an appropriate proxy to measure it; that measure was sustainability.

In addition to their multi-dimensional nature, both sustainable development and sustainability have been characterized by different perceptions on how to explicitly define them. Complementary to the lack of a unified definition is also the absence of an official and unified methodological framework. Composite indicators have emerged as a suitable means that allow the proper measurement of sustainability.

One of the methods that has been proposed and used in the literature for both the measurement of sustainability and the construction of composite indicators is Data Envelopment Analysis. A literature review was performed in the context of the current thesis for the years 2016-2020 and it was discovered that researchers have made an effort to include parameters that represent the social dimension of sustainability (a feature that was missing the previous years), while more dimensions of sustainability are included in recent studies such as technological innovation and advancement, despite the fact that the three-dimensional construct (economy, environment, society) seems to be the preferred one. Moreover, it was discovered that the choice of inputs and outputs (and intermediate measures) despite commonalities is unique to each research work. In addition, the choice of DEA variation and/or combination with other methodologies implies that the perception of each analyst affects the final result of their work.

Apart from those identified gaps, the methodology of DEA itself does not come without its own limitations. First, in its traditional form the efficiency of Decision Making Units is calculated with weights that are most favorable to themselves; i.e. each DMU is evaluated under the most favorable weighting scheme with the purpose of maximizing its own efficiency. As a result, the weights that are chosen for one DMU may be completely different from those selected for another. Moreover, DEA needs to be used in the appropriate context, which means that there is the requirement to decide which parameters will best explain different dimensions of sustainability. This is especially important since, the number of inputs and outputs that can be used is limited by the number of DMUs under evaluation for the measurement to be meaningful, otherwise there would be an increased number of efficient DMUs that would result in inconsistencies.

As a result, the purpose of the current thesis is to address all the above gaps and more specifically:

1) to propose an alternative version of two-stage Data Envelopment Analysis with a different optimization metric that attempts to intervene on the weights of the inputs, intermediate measures and outputs to better reflect their importance for the DMUs by considering positive and negative deviations in the calculations and limiting the distance of these deviations from the maximum and minimum values.

2) to propose a computational framework that will attempt to incorporate different perceptions (meaning different combinations of inputs and outputs) and apply it in the measurement of sustainability of the EU 28 countries.

To achieve this objective, the framework will rely on Exploratory Modeling and Analysis (EMA). EMA is a school of thought developed at RAND corporation and promotes the exploratory use of quantitative methods despite methodological limitations, uncertainties and different perceptions (meaning different combinations of inputs, intermediate measures and outputs). Employing an exploratory approach to sustainability measurement could reveal unanticipated implications of the initial assumptions regarding inputs and outputs.

The thesis is developed in a series of consecutive steps. First, an alternative two-stage DEA model is introduced that employs positive and negative deviational variables both in the objective function (thus altering the optimization metric) and in the constraints. The model attempts to find the best possible weights for the inputs, intermediate measures and outputs, by minimizing the deviations of both the first and second stage of the model. By minimizing simultaneously the deviations of each stage, the efficiencies of both stages are maximized at the same time and no priority is given into which stage should take precedence. Three lemmas and one theorem are proved, and it is proven that the alternative model has a feasible solution that is optimal.

The alternative optimization metric, two-stage proposed DEA model is applied in two case studies: one that calculates the environmental performance of European countries and a second to calculate the agricultural sustainability of European countries.

Following the definition of the new model, a new computational framework is defined for the construction of composite indicators. The proposed model is used for the calculation of each sub-indicator that the final indicator will consist of. The calculated sub-indicators are then used as parameters in a Benefit-of-the-Doubt (BoD) model that generates the value of the final index. The computational framework is tested two times in the measurement of sustainability of European countries: once with the proposed, alternative, two-stage DEA model and once with the typical two-stage model of Chen et al. (2012).

The above calculation of sustainability however is limited by the same notion that was identified in the beginning: since there is no unique, “correct” definition of sustainability, the same indicator can be calculated by using different variations of DEA and/or different combinations of inputs, intermediate measures and outputs.

Consequently, there is the need to have an indicator of sustainability that will incorporate all these different perceptions that may arise, where perceptions mean different DEA variation and/or different combination of inputs, intermediate outputs and outputs. The proposed computational framework is based on this principle, and it consists of the following steps:

Step 1: Define different perceptions of sustainability and for each perception:

a) define how many sub-indicators will be entailed in this perception’s sustainability index

b) define the inputs, intermediate measures and outputs that each sub-indicator will entail

c) Repeat for all perceptions

Step 2: Define the variation of DEA that will calculate the value of the sub-indicators

a) calculate the sub-indicators

b) calculate the perception’s sustainability index using DEA models

c) Once all sustainability indices for all perceptions are calculated, calculate the mean value for each country/DMU

Step 3: Use machine learning to gain insights into the sustainability of each country under different perceptions

The proposed computational framework is used with four different versions of two-stage DEA models, and different combinations of inputs, intermediate measures and outputs to the calculation of sustainability of European countries.

The final step of the proposed computational framework is to use Machine Learning techniques in the results of the generated computations with the purpose of revealing insights into how the sustainability of countries behaves under different perceptions.

Following the logic of EMA, several techniques will be employed in an effort to mitigate intrinsic methodological limitations and find the common, emergent elements that remain robust despite the different methods.

The first insights will be revealed by using clustering techniques and more specifically K-Means and Density based clustering (DBSCAN). For the clustering algorithms, the values of the sub-indicators along with those of the sustainability indices under all the computational regimes were used.

For the current thesis, three additional techniques were used: Classification and Regression Decision Trees (CART), Random Forests and Boosting Regression.

Classification and Regression Decision Trees (CART) since they are not computationally costly, they can be used as communication tools to non-experts and offer deep interpretational capabilities. However, CART trees tend to overfit the data to their training set and are considered weak learners and for that reason two additional ML techniques will be used: Random Forests and boosting regression.

Random forests train trees independently using random samples of the available data and the sampling happens with bootstrapping both the sample and the features at every repetition. As a result, they tend to be slower than CART trees, but the generated results are more robust and tend to avoid the pitfalls of overfitting. More specifically, with the random forests 80% of the data will be used for training and the remaining will be used for prediction. Furthermore, for each data row (point) of the remaining data, the contribution of the individual features to the predicted value will be calculated. The average of all the contributions will be plotted in a boxplot to reveal insights on how individual sub-indicators affect the value of the sustainability index.

Similarly, boosting regression is also considered a slow learner, but compared to random forests, each tree is generated using information from previous ones. Moreover, the technique will also reveal the relative influence of the individual sub-indicator to the index of sustainability, which could provide further insights into the analysis of the results. Both random forests and boosting regression are more robust than CART trees, but this robustness comes at the detriment of intuitive communication capabilities that are the main characteristic of CART trees. Consequently, the use of all three Machine Learning techniques will limit the methodological weaknesses of each method, while providing results and insights that are robust and independent of the used technique.

The final results illustrated that a balance among the performance of various dimensions can be a good policy to achieve sustainable development and when the inclusion of all DEA variations does not alter significantly the mean value of sustainability then the trust in the results increases, thus making them robust.

Finally, the blend of DEA with machine learning (applied on the results of DEA for the various scenarios) revealed insights on the areas that policy makers could direct investments to increase sustainability. In addition, the ML applications contributed in the identification of the most important features of sustainability for the various countries something that could have direct implications in the area of EU policy making: for example, countries that share similar features that drive the behavior of sustainability could be grouped together in clusters and policies, laws, regulations etc. could be adapted to those clusters in order to boost the particular features that would increase their sustainability. As a result, policy making has the potential to become customized (adapted to the specifics of each group) without missing its overall and principal theme of pursuing sustainable development. This adaptive and adaptable policy making could greatly be of assistance especially when new countries are negotiating their entry to the Union; based on the features that affect the sustainability of the new countries, they could follow the regulations and laws of the appropriate cluster. Finally, the inclusion of new layers and perceptions renders the algorithms more inclusive and participatory, increasing their transparency, thus improving the trust to the final results.

Περίληψη

Κατά τη λήψη δημόσιων αποφάσεων, παράγοντες όπως οι προσωπικές αξίες, το πολιτισμικό υπόβαθρο και οι διαφορετικές ατομικές προοπτικές διαδραματίζουν κεντρικό ρόλο στον κύκλο σχεδιασμού, δοκιμής, εφαρμογής και αναθεώρησης της πολιτικής. Για να βοηθήσουν τους υπεύθυνους χάραξης πολιτικής, οι αναλυτές έχουν χρησιμοποιήσει μια σειρά ποιοτικών και ποσοτικών μεθόδων σε όλα τα στάδια του κύκλου αποφάσεων.

Ωστόσο, η αυξανόμενη χρήση εξελιγμένων μεθόδων δεν φαίνεται να συνοδεύεται πάντα από βελτίωση της ποιότητας της χάραξης πολιτικής - αντίθετα, φαίνεται να προσελκύει κριτική που επικεντρώνεται στα μειονεκτήματά τους. Επιπλέον, η ραγδαία άνοδος της χρήσης τεχνικών μηχανικής μάθησης και τεχνητής νοημοσύνης και η ολοένα διερευνώμενη χρήση μεθοδολογιών βαθιάς μάθησης στη λήψη αποφάσεων ή/και στη χάραξη πολιτικής, έφερε στο προσκήνιο το ζήτημα της ερμηνευσιμότητας των αποτελεσμάτων των αλγορίθμων αυτών και κατά πόσον τα αυτά τους μπορούν να θεωρηθούν αξιόπιστα και έγκυρα (verified and validated). Ερωτήματα όπως "ποιο συγκεκριμένο χαρακτηριστικό (feature/variable/attribute) ή μετρική (metric) έκανε το μοντέλο/αλγόριθμο να καταλήξει στη συγκεκριμένη απόφαση", άρα ζητήματα διαφάνειας (transparency), ερμηνευσιμότητας (interpretability) και εμπιστοσύνης (trustworthiness) των μεθόδων, καθίστανται κεντρικά ζητήματα της κριτικής θεώρησης στις ποσοτικές μεθόδους και στους αλγορίθμους.

Η κριτική αυτή δεν στερείται θεωρητικής και πρακτικής βάσης. Η εγγενής πολυπλοκότητα των σύγχρονων προβλημάτων προκαλεί ερωτήματα για τα οποία ένας αναλυτής μπορεί να κάνει μόνο υποθέσεις λόγω της ύπαρξης πλήρους αβεβαιότητας ή περιβάλλοντος υψηλού ρίσκου. Επιπλέον, σε ένα τέτοιο περιβάλλον πολυπλοκότητας, οι αντιλήψεις και τα στερεότυπα του αναλυτή (cognitive biases) μπορούν να περιορίσουν το οπτικό του πεδίο και τη δυνατότητα ερμηνείας του υπό μελέτη κύκλου πολιτικής. Ως αποτέλεσμα, η επιτυχία μιας ποσοτικής μεθόδου βασίζεται στο ότι όλες οι παραπάνω επιλογές που αφορούν την ανάλυση του προβλήματος θα είναι κατάλληλες.

Η έννοια της αειφόρου ανάπτυξης περικλείει αυτά τα ζητήματα και για να επιτευχθεί οι δημόσιες πολιτικές θα πρέπει να έχουν οικονομικές, κοινωνικές και

περιβαλλοντικές διαστάσεις, λαμβάνοντας παράλληλα υπόψη τις τρέχουσες τεχνολογικές εξελίξεις, το πολιτισμικό πλαίσιο και το σύστημα αξιών στο οποίο εφαρμόζονται.

Εκτός από τον πολυδιάστατο χαρακτήρα τους, τόσο η αειφόρος ανάπτυξη όσο και η αειφορία χαρακτηρίζονται από διαφορετικές αντιλήψεις σχετικά με τον τρόπο που πρέπει να οριστούν. Συμπληρωματικά με την έλλειψη ενός ενιαίου ορισμού είναι και η απουσία ενός επίσημου και ενιαίου μεθοδολογικού πλαισίου. Οι σύνθετοι δείκτες έχουν αναδειχθεί ως ένα κατάλληλο μέσο που επιτρέπει τη σωστή μέτρηση της αειφορίας.

Μια από τις μεθόδους που έχουν προταθεί και χρησιμοποιηθεί στη βιβλιογραφία τόσο για τη μέτρηση της αειφορίας όσο και για την κατασκευή σύνθετων δεικτών είναι η Περιβάλλουσα Ανάλυση Δεδομένων (Data Envelopment Analysis, DEA). Στο πλαίσιο της παρούσας διατριβής πραγματοποιήθηκε βιβλιογραφική ανασκόπηση για τα έτη 2016-2020 και διαπιστώθηκε ότι οι ερευνητές έχουν καταβάλει προσπάθεια να συμπεριλάβουν παραμέτρους που αντιπροσωπεύουν την κοινωνική διάσταση της αειφορίας (χαρακτηριστικό που έλειπε τα προηγούμενα χρόνια), ενώ σε πρόσφατες μελέτες περιλαμβάνονται περισσότερες διαστάσεις της αειφορίας, όπως η τεχνολογική καινοτομία και πρόοδος, παρά το γεγονός ότι η τρισδιάστατη προσέγγιση (οικονομία, περιβάλλον, κοινωνία) φαίνεται να είναι η προτιμώμενη. Επιπλέον, διαπιστώθηκε ότι η επιλογή των εισροών και εκροών (και των ενδιάμεσων μέτρων) παρά τις σχετικές ομοιότητες είναι μοναδική για κάθε ερευνητική εργασία. Επιπλέον, η επιλογή της παραλλαγής της DEA και/ή ο συνδυασμός με άλλες μεθοδολογίες συνεπάγεται ότι η αντίληψη κάθε αναλυτή επηρεάζει το τελικό αποτέλεσμα της μελέτης του.

Εκτός από τα κενά που διαπιστώθηκαν παραπάνω, η ίδια η μεθοδολογία της DEA δεν είναι απαλλαγμένη από τους δικούς της περιορισμούς. Πρώτον, στην παραδοσιακή της μορφή, η αποδοτικότητα των Μονάδων Λήψης Αποφάσεων (Decision Making Units-DMUs) υπολογίζεται με τα πιο ευνοϊκά για τις ίδιες βάρη, δηλαδή κάθε DMU αξιολογείται με το πιο ευνοϊκό σχήμα στάθμισης με σκοπό τη μεγιστοποίηση της δικής της αποδοτικότητας, κάτι που είναι ως ένα βαθμό αποδεκτό θεωρώντας ότι κάθε DMU δικαιούται το ελαφρυντικό της αμφιβολίας ως προς τους λόγους της μειωμένης αποδοτικότητάς της (benefit of the doubt principle). Ως αποτέλεσμα, τα βάρη που επιλέγονται για μία DMU μπορεί να είναι εντελώς διαφορετικά από εκείνα που επιλέγονται για μια άλλη υπό την έννοια ότι οι αντίστοιχες μονάδες με τις οποίες συγκρίνεται κάθε μία από αυτές μεταβάλλονται ανάλογα με την άριστη λύση του σχετικού μοντέλου ώστε να εντοπιστεί η καλύτερη δυνατή απόδοση αλλά και οι ιδανικοί ομότιμοι προς τους οποίους θα πρέπει να κοιτάξει για να βελτιωθεί (efficient peers). Επιπλέον, η DEA πρέπει να χρησιμοποιείται στο κατάλληλο πλαίσιο, πράγμα που σημαίνει ότι υπάρχει η απαίτηση να αποφασιστεί ποιες παράμετροι θα εξηγήσουν καλύτερα τις διάφορες διαστάσεις της αειφορίας. Αυτό είναι ιδιαίτερα σημαντικό δεδομένου ότι, ο αριθμός των εισροών και εκροών που μπορούν να χρησιμοποιηθούν περιορίζεται από τον αριθμό των DMUs που αξιολογούνται για να έχει νόημα η μέτρηση, διαφορετικά θα υπήρχε αυξημένος αριθμός αποδοτικών DMUs που θα οδηγούσε σε ασυνέπειες.

Ως εκ τούτου, σκοπός της παρούσας διατριβής είναι να αντιμετωπίσει όλα τα παραπάνω κενά και πιο συγκεκριμένα:

1) να προτείνει μια εναλλακτική εκδοχή της Περιβάλλουσας Ανάλυσης Δεδομένων δύο σταδίων με μια διαφορετική αντικειμενική συνάρτηση που επιχειρεί να παρέμβει στα βάρη των εισροών, των ενδιάμεσων μέτρων και των εκροών ώστε να αντικατοπτρίζει καλύτερα τη σημασία τους για τα DMUs, λαμβάνοντας υπόψη τις θετικές και αρνητικές αποκλίσεις στους υπολογισμούς και περιορίζοντας την απόσταση αυτών των αποκλίσεων από τις μέγιστες και ελάχιστες τιμές.

2) να προτείνει ένα υπολογιστικό πλαίσιο που θα επιχειρήσει να ενσωματώσει διαφορετικές αντιλήψεις (δηλαδή διαφορετικούς συνδυασμούς εισροών και εκροών) και να το εφαρμόσει στη μέτρηση της αειφορίας των χωρών της ΕΕ των 28, δηλαδή προτείνεται ένα νέο υπολογιστικό πλαίσιο που στηρίζεται σε τεχνικές μηχανικής μάθησης μέσω του οποίου κατασκευάζονται σύνθετοι δείκτες απόδοσης αειφορίας για κάθε χώρα.

Για την επίτευξη αυτού του στόχου, το πλαίσιο θα βασίζεται στη διερευνητική μοντελοποίηση και ανάλυση (Exploratory Modeling and Analysis- EMA). Η EMA είναι μια σχολή σκέψης που αναπτύχθηκε στο ερευνητικό κέντρο RAND και προωθεί τη διερευνητική χρήση ποσοτικών μεθόδων παρά τους μεθοδολογικούς περιορισμούς τους, τις αβεβαιότητες και τις διαφορετικές αντιλήψεις (δηλαδή διαφορετικούς συνδυασμούς εισροών, ενδιάμεσων μέτρων και εκροών). Η χρήση μιας διερευνητικής προσέγγισης για τη μέτρηση της αειφορίας θα μπορούσε να αποκαλύψει απρόβλεπτες επιπτώσεις των αρχικών υποθέσεων σχετικά με τις εισροές και τις εκροές.

Η διατριβή αναπτύσσεται σε μια σειρά διαδοχικών βημάτων. Πρώτον, εισάγεται ένα εναλλακτικό μοντέλο DEA δύο σταδίων που χρησιμοποιεί θετικές και αρνητικές αποκλίνουσες μεταβλητές τόσο στην αντικειμενική συνάρτηση (μεταβάλλοντας έτσι το μέτρο της βελτιστοποίησης), όσο και στους περιορισμούς. Το μοντέλο επιχειρεί να βρει τα καλύτερα δυνατά βάρη για τις εισροές, τα ενδιάμεσα μέτρα και τις εκροές, ελαχιστοποιώντας τις αποκλίσεις τόσο του πρώτου όσο και του δεύτερου σταδίου του μοντέλου. Με την ταυτόχρονη ελαχιστοποίηση των αποκλίσεων κάθε σταδίου, μεγιστοποιούνται ταυτόχρονα οι αποδοτικότητες και των δύο σταδίων και δεν δίνεται προτεραιότητα στο ποιο στάδιο θα πρέπει να υπερισχύσει. Το προτεινόμενο πλαίσιο/μοντέλο στηρίζεται θεωρητικά σε τρία λήμματα και ένα θεώρημα όπου αποδεικνύεται ότι έχει τουλάχιστον μια εφικτή λύση που είναι βέλτιστη.

Η προτεινόμενη παραλλαγή της μεθόδου εφαρμόζεται σε δύο μελέτες περίπτωσης: η μία υπολογίζει τις περιβαλλοντικές επιδόσεις των ευρωπαϊκών χωρών και η δεύτερη ελέγχει τη γεωργική βιωσιμότητα των ευρωπαϊκών χωρών.

Μετά τον ορισμό του νέου μοντέλου, ορίζεται ένα νέο υπολογιστικό πλαίσιο για την κατασκευή σύνθετων δεικτών. Το προτεινόμενο μοντέλο χρησιμοποιείται για τον υπολογισμό κάθε επιμέρους δείκτη από τον οποίο θα αποτελείται ο τελικός δείκτης. Οι υπολογισμένοι υποδείκτες χρησιμοποιούνται στη συνέχεια ως παράμετροι σε ένα μοντέλο Benefit-of-the-Doubt (BoD) που παράγει την τιμή του τελικού δείκτη. Το υπολογιστικό πλαίσιο δοκιμάζεται δύο φορές στη μέτρηση της αειφορίας των ευρωπαϊκών χωρών: μία φορά με το προτεινόμενο, εναλλακτικό, μοντέλο DEA δύο σταδίων και μία φορά με το τυπικό μοντέλο δύο σταδίων των Chen et al. (2012).

Ωστόσο, ο παραπάνω υπολογισμός της βιωσιμότητας περιορίζεται από την ίδια έννοια που εντοπίστηκε στην αρχή: δεδομένου ότι δεν υπάρχει μοναδικός, "σωστός" ορισμός της βιωσιμότητας, ο ίδιος δείκτης μπορεί να υπολογιστεί χρησιμοποιώντας διαφορετικές παραλλαγές της DEA ή/και διαφορετικούς συνδυασμούς εισροών, ενδιάμεσων μέτρων και εκροών.

Κατά συνέπεια, γεννιέται η ανάγκη να υπάρξει ένας δείκτης αειφορίας που θα ενσωματώνει όλες αυτές τις διαφορετικές αντιλήψεις που μπορεί να προκύψουν, όπου αντιλήψεις σημαίνει διαφορετική παραλλαγή της DEA ή/και διαφορετικός συνδυασμός εισροών, ενδιάμεσων μέτρων και εκροών. Το προτεινόμενο υπολογιστικό πλαίσιο βασίζεται σε αυτή την αρχή και αποτελείται από τα ακόλουθα βήματα:

Βήμα 1: Καθορισμός διαφορετικών αντιλήψεων για τη βιωσιμότητα και για κάθε αντίληψη:

- α) καθορισμός των επιμέρους δεικτών που θα περιλαμβάνονται στον δείκτη βιωσιμότητας αυτής της αντίληψης
- β) καθορισμός των εισροών, τις ενδιάμεσων μέτρων και εκροών που θα περιλαμβάνει κάθε υποδείκτης
- γ) Επανάληψη για όλες τις αντιλήψεις

Βήμα 2: Καθορισμός της παραλλαγής της DEA που θα υπολογίζει την τιμή των υποδεικτών

- α) Υπολογισμός των υποδεικτών
- β) υπολογισμός του δείκτη αειφορίας της κάθε αντίληψης χρησιμοποιώντας τα μοντέλα DEA
- γ) Αφού υπολογιστούν όλοι οι δείκτες αειφορίας για όλες τις αντιλήψεις, υπολογισμός της μέσης τιμής για κάθε χώρα/DMU

Βήμα 3: Χρήση μηχανικής μάθησης για ανάδειξη πληροφοριών σχετικά με τη βιωσιμότητα κάθε χώρας υπό διαφορετικές αντιλήψεις

Το προτεινόμενο υπολογιστικό πλαίσιο χρησιμοποιείται με τέσσερις διαφορετικές παραλλαγές μοντέλων DEA δύο σταδίων και διαφορετικούς συνδυασμούς εισροών, ενδιάμεσων μέτρων και εκροών για τον υπολογισμό της αειφορίας των ευρωπαϊκών χωρών.

Το τελικό βήμα του προτεινόμενου υπολογιστικού πλαισίου είναι η χρήση τεχνικών μηχανικής μάθησης στα αποτελέσματα των παραγόμενων υπολογισμών με σκοπό την αποκάλυψη γνώσεων σχετικά με το πώς συμπεριφέρεται η αειφορία των χωρών υπό διαφορετικές αντιλήψεις.

Ακολουθώντας τη λογική της EMA, χρησιμοποιούνται διάφορες τεχνικές σε μια προσπάθεια να αμβλυνθούν οι εγγενείς μεθοδολογικοί περιορισμοί και να βρεθούν τα κοινά, αναδυόμενα στοιχεία που παραμένουν ισχυρά παρά τις διαφορετικές μεθόδους.

Στο πρώτο στάδιο του βήματος 3 της μεθοδολογίας που προτείνεται, χρησιμοποιούνται δύο τυπικές τεχνικές συσταδοποίησης μία η οποία στηρίζεται στα centroids και μία η οποία στηρίζεται στην χωρική πυκνότητα. Πιο συγκεκριμένα χρησιμοποιούνται η μέθοδος K-Means και η μέθοδος DBSCAN. Για τους αλγορίθμους αυτούς χρησιμοποιήθηκαν ως δεδομένα οι τιμές των επιμέρους δεικτών μαζί με εκείνες των δεικτών αειφορίας σε όλα τα υπολογιστικά καθεστώτα.

Για την παρούσα διατριβή χρησιμοποιήθηκαν επίσης και τεχνικές που πηγάζουν καταρχάς από τα κλασικά Classification and Regression Decision Trees και επεκτείνονται σε Random Forests και Boosting Regression.

Τα δέντρα αποφάσεων ταξινόμησης και παλινδρόμησης (CART), δεδομένου ότι δεν έχουν υπολογιστικό κόστος, μπορούν να χρησιμοποιηθούν ως εργαλεία επικοινωνίας σε μη ειδικούς και προσφέρουν βαθιές ερμηνευτικές δυνατότητες. Ωστόσο, τα CART τείνουν να προσαρμόζουν υπερβολικά τα δεδομένα και θεωρούνται αδύναμοι learners και για το λόγο αυτό θα χρησιμοποιηθούν δύο πρόσθετες τεχνικές μηχανικής μάθησης: Random Forests και boosting regression.

Τα τυχαία δάση εκπαιδεύουν τα δέντρα ανεξάρτητα, χρησιμοποιώντας τυχαία δείγματα των διαθέσιμων δεδομένων και η δειγματοληψία γίνεται με bootstrapping τόσο του δείγματος όσο και των χαρακτηριστικών σε κάθε επανάληψη. Ως αποτέλεσμα, τείνουν να είναι πιο αργά από τα απλά CART, αλλά τα παραγόμενα αποτελέσματα είναι πιο ισχυρά και τείνουν να αποφεύγουν τις παγίδες της υπερπροσαρμογής. Συνήθως στα τυχαία δάση το 80% των δεδομένων θα χρησιμοποιηθεί για εκπαίδευση (training) και το υπόλοιπο θα χρησιμοποιηθεί για την πρόβλεψη (prediction). Ο μέσος όρος όλων των συνεισφορών θα απεικονιστεί σε ένα θηκόγραμμα (boxplot) για να αποκαλυφθούν γνώσεις σχετικά με τον τρόπο με τον οποίο οι επιμέρους υποδείκτες επηρεάζουν την τιμή του δείκτη βιωσιμότητας.

Παρομοίως, η παλινδρόμηση boosting θεωρείται επίσης αργός learner, αλλά σε σύγκριση με τα τυχαία δάση (Random Forests), κάθε δέντρο δημιουργείται χρησιμοποιώντας πληροφορίες από τα προηγούμενα. Επιπλέον, η τεχνική μπορεί να αποκαλύψει τη σχετική επιρροή του κάθε επιμέρους υποδείκτη στον δείκτη βιωσιμότητας, γεγονός που θα μπορούσε να προσφέρει περαιτέρω πληροφορίες για την ανάλυση των αποτελεσμάτων. Τόσο τα Random Forests όσο και η boosting regression είναι πιο εύρωστα από τα δέντρα CART, αλλά αυτή η ευστάθεια αποβαίνει εις βάρος των διαισθητικών δυνατοτήτων επικοινωνίας και εμηνευσιμότητας που είναι τα κύρια χαρακτηριστικά των δέντρων CART. Κατά συνέπεια, η χρήση και των τριών τεχνικών Μηχανικής Μάθησης βοήθησε στο να περιορίσει τις μεθοδολογικές αδυναμίες της κάθε μεθόδου, παρέχοντας παράλληλα αποτελέσματα και γνώσεις που είναι εύρωστες και ανεξάρτητες από τη χρησιμοποιούμενη τεχνική.

Τα τελικά αποτελέσματα κατέδειξαν ότι η ισορροπία μεταξύ των επιδόσεων των διαφόρων διαστάσεων μπορεί να είναι μια καλή πολιτική για την επίτευξη της βιώσιμης ανάπτυξης και όταν η συμπερίληψη όλων των παραλλαγών της DEA δεν μεταβάλλει σημαντικά τη μέση τιμή της αειφορίας, τότε η εμπιστοσύνη στα αποτελέσματα αυξάνεται, καθιστώντας τα έτσι ισχυρά.

Τέλος, ο συνδυασμός της DEA με τη μηχανική μάθηση αποκάλυψε πληροφορίες σχετικά με τους τομείς στους οποίους οι υπεύθυνοι χάραξης πολιτικής θα μπορούσαν να κατευθύνουν τις επενδύσεις για την αύξηση της αειφορίας. Επιπλέον, η εφαρμογή τεχνικών μηχανικής μάθησης συνέβαλε στον προσδιορισμό των πιο σημαντικών χαρακτηριστικών της αειφορίας για τις διάφορες χώρες, κάτι που θα μπορούσε να έχει άμεσες επιπτώσεις στον τομέα της χάραξης πολιτικής της ΕΕ: για παράδειγμα, οι χώρες που μοιράζονται παρόμοια χαρακτηριστικά θα μπορούσαν να ομαδοποιηθούν και οι πολιτικές, οι νόμοι, οι κανονισμοί κ.λπ. θα μπορούσαν να προσαρμοστούν σε αυτές τις ομάδες, προκειμένου να ενισχυθούν τα συγκεκριμένα χαρακτηριστικά που θα αύξαναν την αειφορία τους. Ως αποτέλεσμα, η χάραξη πολιτικής έχει τη δυνατότητα να γίνει εξατομικευμένη (προσαρμοσμένη στις ιδιαιτερότητες κάθε ομάδας) χωρίς να χάνει το γενικό και κύριο θέμα της, δηλαδή την επιδίωξη της βιώσιμης ανάπτυξης. Αυτή η προσαρμοστική και ευπροσάρμοστη χάραξη πολιτικής θα μπορούσε να βοηθήσει σημαντικά, ιδίως όταν νέες χώρες διαπραγματεύονται την είσοδό τους στην Ένωση- με βάση τα χαρακτηριστικά που επηρεάζουν την αειφορία των νέων χωρών, θα μπορούσαν να ακολουθήσουν τους κανονισμούς και τους νόμους της κατάλληλης ομάδας. Τέλος, η συμπερίληψη νέων διαστάσεων και αντιλήψεων καθιστά τους αλγορίθμους πιο περιεκτικούς και συμμετοχικούς, αυξάνοντας τη διαφάνειά τους, βελτιώνοντας έτσι την εμπιστοσύνη στα τελικά αποτελέσματα.

List of Figures

List of Tables

Abbreviations

AI	Artificial Intelligence
DEA	Data Envelopment Analysis
DMU	Decision Making Unit
ML	Machine Learning
LP	Linear Programming
CRS	Constant Returns to Scale
VRS	Variable Returns to Scale
PPS	Production Possibility Set
DEA CCR	Data Envelopment Analysis under Constant Returns to Scale
DEA BCC	Data Envelopment Analysis under Variable Returns to Scale
MCDA	Multi-Criteria Decision Aid
CART	Classification and Regression Tree
ML	Machine Learning

Table of Contents

1. Introduction[1]

1.1 General Background

In public decision-making factors such as personal values, cultural background and different individual perspectives play a central role in the policy cycle of design, test, implementation and review (Tsoukias, Montibeller, Lucertini, & Belton, 2013). To assist policy makers, analysts have used an array of qualitative and quantitative methods to all steps of the cycle.

In particular, quantitative methods have seen a growth that started in the 1970s (Longo & McNutt, 2018), which continues today with the steep increase in computational power and availability of data (Bankes, 1993). However, the increasing use of sophisticated methods seems not to be always accompanied by an improvement in the quality of policy making; on the contrary, it seems to attract criticism that is focused on their disadvantages (Bankes, 1993). Furthermore, the rise of Artificial Intelligence (AI) and its expanding use in decision- and/or policy-making, has brought forth the issue of interpretability of algorithms and whether their output can be trusted. Questions such as "which specific feature made the model/algorithm reach the specific decision" (Moraffah, Karami, Guo, Raglin, & Liu, 2020), hence issues of transparency and interpretability of the methods (Lewis, Li, & Sycara, 2021), are becoming central issues of the critique on quantitative methods and algorithms.

This criticism is not without its merits. The complexity of contemporary problems means that there are issues about which an analyst can only make assumptions due to the existence of deep uncertainty (Kwakkel & Pruyt, 2013). For example, there is no easy way to quantify human behavior or have knowledge and data that illustrate how the components of a system precisely interact. Moreover, in such complexity, the perception of the analyst may limit the view of the policy cycle under study. As a result, the success of a quantitative method relies on all of the above choices to be exactly appropriate (Bankes, 1993).

Sustainable development perfectly encapsulates these issues. It entered the sphere of public policy-making and analysis in the 1980s, when the Brundtland report defined

[1] The Introduction appeared on: Tsaples, G., Papathanasiou, J., & Georgiou, A. C. (2022). An Exploratory DEA and Machine Learning Framework for the Evaluation and Analysis of Sustainability Composite Indicators in the EU. Mathematics, 10(13), 2277

sustainable development as: “the ability to meet the needs of the present without compromising the ability of future generations to meet their own needs” (Brundtland, Khalid, Agnelli, Al-Athel, & Chidzero, 1987). This definition implies the existence of limits on both the human side (how can humans with the present state of technology use environmental resources) and nature (how can it absorb the effects of human activities) (Kates, et al., 2001). Furthermore, the definition implies a (multi-generational) fairness in distributing the resources and effects of human activities (Hassanzadeh, Yousefi, Saen, & Hosseininia, 2018). Consequently, the Brundtland report brought to the limelight of policy-making the idea that policy makers cannot solely seek to promote social development in parallel to economic prosperity (Lin & Chiu, 2018), but they need to always consider the balance between economic growth and its environmental consequences (Munda & Saisana, 2011).

These implications, omissions and conflicts imply that in order to achieve sustainable development, public policies should have economic, social and environmental dimensions, while taking into account the current technological developments (Robinson, 2004), the cultural context and the value system in which they are applied (Santana, Mariano, Camioto, & Rebelatto, 2015). Thus, sustainable development is a multi-dimensional concept and from early on arose the need to find an appropriate proxy to measure it (Tyteca, 1998).

Sustainability, a notion stemming from ecology, has been used in that aspect. At its basic form, it is an indication of a natural system’s endurance, its ability to retain its essential properties and naturally replenish its population (Zhou, Yang, Chen, & Zhu, 2018). In human systems, sustainability is regarded as the ability to live without environmental degradation (Robinson, 2004), while encompassing all dimensions of human systems and processes (Sneddon, Howarth, & Norgaard, 2006).

In addition to their multi-dimensional nature, both sustainable development and sustainability have been characterized by different perceptions on how to explicitly define them (Drucker, 2014; Robinson, 2004). So far, all interpretations of both sustainable development and sustainability fall into two categories: there is the three-dimensional approach that seeks to integrate an economic, social and environmental dimension and the dualistic approach that emphasizes the interlinked relationship between humans and nature (Robinson, 2004). Lately however, another category has emerged, one that focuses on technology and innovation as the means to achieve sustainable development (Drucker, 2014).

Complementary to the lack of a unified definition is also the absence of an official and unified methodological framework (Munda & Saisana, 2011). The existence of such a framework could be of great assistance, since the increasing complexity of policy-making is increasingly characterizing the effort to achieve sustainable development (He, Wan, Feng, Ai, & Wang, 2016). To achieve its objective, such a framework should entail certain properties. First, the multi-dimensional nature of sustainability dictates that any quantitative method cannot rely only on terms of costs and benefits (Adler, 2012). Moreover, any such method should have integrating properties, since sustainability seeks to combine different dimensions into a single measure (Ramanathan, 2002) and finally it should be transparent, easy to communicate to non-experts and subjected to the review of experts (Robinson, 2004).

These characteristics facilitated the use of composite indicators as a suitable means that allow the proper measurement of sustainability (Coli, Nissi, & Rapposelli, 2011). A composite indicator is a mathematical construction that can integrate multi-dimensional concepts with different units of measurement (Zhou, Yang, Chen, & Zhu, 2018). Furthermore, it can be used as a communication tool while offering meaningful policy monitoring (Zhang, Kong, & Choi, 2014). Consequently, indicators have been used extensively in the measurement of multi-dimensional concepts such as energy efficiency (Ervural, Zaim, & Delen, 2018), human development (Despotis D. K., 2005) and regional sustainability (Zhou, Yang, Chen, & Zhu, 2018).

Nonetheless, the use of composite indicators did not come without some criticism. At their early stages, they mainly focused on the environmental aspect of sustainability (Zofío & Prieto, 2001). Moreover, they relied on weighted linear aggregation implying compensability among the parameters, which is not always a realistic assumption (Munda & Nardo, 2009). Finally, this type of weighted aggregation reflects the personal values of the policy maker and/or the analyst that might be different from the perceptions and values of other analysts, the general public etc. (Kuosmanen & Kortelainen, 2005).

This criticism has stirred the research towards efforts to increase the robustness of composite indicators. Firstly, the focus has shifted from the environmental aspect of sustainability and more (sub-) indicators have been integrated. For example, the social dimension (Winfield, Gibson, Markvart, Gaudreau, & Taylor, 2010), innovation as a force of change (Drucker, 2014) and the capacity to produce sustainable technological products (Santana, Mariano, Camioto, & Rebelatto, 2015).

Regarding the methodological limitations of linear aggregation, other methods have been suggested in the literature and one that is being increasingly used is Data Envelopment Analysis (DEA). DEA is a non-parametric, mathematical programming technique that is used for the assessment of the technical efficiency of Decision Making Units (DMUs) relative to one another (Førsund & Sarafoglou, 2002), where technical efficiency can be viewed as the ability of a DMU to transform its inputs to outputs and is defined as the ratio of the sum of its weighted outputs over the sum of its weighted inputs (Thanassoulis, 2001) as indicated in expression (1):

$$technical\ efficiency = \frac{\sum w_{output} * y}{\sum w_{input} * x}, where\ x = input\ level\ and\ y = output\ level \qquad (1)$$

The foundations of DEA can be traced in the works of Debreu (1951), Farrell (1957) and Diewert (1973), while the method was established in the seminal papers of Charnes, Cooper and Rhodes (1978) and Banker, Charnes and Cooper (1984). The method does not require the knowledge of price information (Kuosmanen & Kortelainen, 2005), it requires knowledge neither of the relationship between inputs and outputs nor of the statistical distribution of the data that are used (Hajiagha, Hashemi, & Mahdiraji, 2016; He, Wan, Feng, Ai, & Wang, 2016). Moreover, DEA is flexible enough to be combined with other methods (Amiri, Zandieh, Vahdani, Soltani, & Roshanaei, 2010; Georgiou, Thanassoulis, & Papadopoulou, 2021; Kamvysi, Gotzamani, Georgiou, & Andronikidis, 2010), thus increasing its methodological robustness. These advantages were crucial in recognizing that DEA can be a suitable tool for assessing sustainable development (Callens & Tyteca, 1999) and as a result it has been increasingly used in sustainability policy making (Zhou, Yang, Chen, & Zhu, 2018).

Zhou et al. (2018) performed a literature review on the use of Data Envelopment Analysis in regional sustainability studies and their study covers the years until 2016. In their paper, Zhou et al. (2018) identified the trend of using DEA to measure sustainability, however they also noted several gaps in the literature. Firstly, it appeared that the main focus of the studies has been on the economic and environmental dimensions of sustainability, while the inclusion of the societal aspect was not equally extensive. Secondly, the authors observed a trend of combining DEA with other methodologies such as Tobit regression (Ervural, Zaim, & Delen, 2018) or lifecycle assessment (Gonzalez-Garcia, Manteiga, Moreira, & Feijoo, 2018) in order to increase the robustness of the measurement by mitigating the methodological limitations of DEA. Moreover, the authors identified that while early studies tend to employ classic DEA models, in later years more sophisticated versions are used. Nonetheless, Zhou et al. (2018) also identify that there is still the need to decide which parameters will be used in the model that best describe the multi-dimensional concept of sustainability.

Tsaples and Papathanasiou (2021) performed a literature review on DEA and sustainability for the years 2016-2020 and discovered that since 2016, the studies have made an effort to include parameters that represent the social dimension of sustainability. Moreover, there are efforts to include other aspects that represent technological advancement and innovation, despite the fact that the three-dimensional construct appears to be the preferred one. However, they also revealed the lack of a unified context in which sustainability is measured in two forms: Firstly, the choice of inputs and outputs (and intermediate measures) despite commonalities is unique to each research work. Secondly, the choice of DEA variation and/or combination with other methodologies implies that the perception of each analyst affects the final result of their work. In their work, Tsaples and Papathanasiou (2020a) calculate four different versions of sustainability using different combinations of inputs and outputs. However, they still use one variation of DEA for all versions of sustainability and they offer no explanation of how the choice of parameters affects the final results.

Consequently, DEA does not come without limitations. First, in its traditional form the efficiency of Decision Making Units is calculated with weights that are most favorable to themselves; i.e. each DMU is evaluated under the most favorable weighting scheme with the purpose of maximizing its own efficiency (Sun, Wu, & Guo, 2013). As a result, the weights that are chosen for one DMU may be completely different from those selected for another (Pedraja-Chaparro, Salinas-Jimenez, & Smith, 1997). In other words, one DMU might place more importance on one of the inputs that it uses while another might do so on another.

Furthermore, Zhou et al. (2018) identified that there is the need to use DEA in the appropriate context, which means that there is the requirement to decide which parameters will best explain different dimensions of sustainability. This particular methodological limitation was not unknown; Moutinho, Madaleno and Robaina (2017; 2018) identified that DEA is sensitive to the choice of inputs and outputs, meaning that the calculated efficiency depends on what inputs and outputs will be chosen. Finally, the number of inputs and outputs that can be used is limited by the number of DMUs under evaluation for the measurement to be meaningful, otherwise there would be an increased number of efficient DMUs that would result in inconsistencies (Hassanzadeh, Yousefi, Saen, & Hosseininia, 2018). Hence, the problem of context cannot be solved simply by adding more parameters to the model.

Using appropriate inputs and outputs is an item of ongoing research in the DEA literature, with researchers attempting to utilize different techniques to increase the robustness of the method. For example, Benítez-Peña et al. (2020) propose the use of Mixed Integer Programming in choosing the appropriate inputs and outputs, while Lee and Cai (2020) propose a least absolute shrinkage and selection operator (LASSO) variable selection that not only addresses the issue of which inputs and outputs to use but also can be used with small datasets.

Moreover, researchers understood that the robustness of a DEA model increases if the DMU under study is not considered a “black box” and for that reason, the intermediate steps of DEA were increased (Färe & Grosskopf, 2009). Hence, two-stage DEA model were considered as more appropriate to capture the appropriate weights that are used in the calculation of efficiency. However, little attention has been paid to these types of models with regard to weight distribution and their discriminatory power (Mahdiloo, Jafarzadeh, Saen, Tatham, & Fisher, 2016).

Consequently, the power of DEA as a monitoring tool for sustainability is diminished by the same issue that was identified in the beginning of this section; the choice of the appropriate parameters must be “correct” in order for the measurement to be meaningful. The lack of a unified definition, however, means that there is no singular, “correct” choice. Different people (policy makers, analysts, the public etc.) have different values and perceptions of what sustainable development means and what should be used to measure sustainability. Thus, there is the need to increase the robustness of DEA by incorporating as many perceptions as possible in the measurement of sustainability without losing the value of its advantages.

1.2 Thesis Objective

The purpose of the current thesis is to propose a computational framework with a twofold functionality:

1) to propose an alternative two-stage Data Envelopment Analysis model with an alternative optimization metric that attempts to intervene on the weights of the inputs, intermediate measures and outputs to better reflect their importance for the DMUs by considering positive and negative deviations in the calculations and limiting the distance of these deviations from the maximum and minimum values.

2) to propose a computational framework that will attempt to incorporate different perceptions (meaning different combinations of inputs and outputs) and apply it in the measurement of sustainability of the EU 28 countries.

To achieve this objective, the framework will rely on Exploratory Modeling and Analysis (EMA). EMA is a school of thought developed at RAND corporation (Bankes, 1992) and promotes the exploratory use of quantitative methods despite methodological limitations, uncertainties and different perceptions. Employing an exploratory approach to sustainability measurement could reveal unanticipated implications of the initial assumptions regarding inputs and outputs. The use of computational experimentations to explore conjectures, models and datasets is not new. It has been applied to mathematics (for example (Bailey, Borwein, & Bradley, 2006)), in simulation models (for example (Kwakkel & Pruyt, 2013)), and of course in various disciplines with the emergence of big data (for example (Fraedrich, Schneider, & Westermann, 2009)). The approach requires computational power, development of new algorithms and techniques to analyze the data that will be generated.

Thus, EMA relies on Machine Learning (ML) techniques, even though the developed models may not be possible to be validated. However, even when it is not possible to validate a model, exploration could lead to insights on how the different perceptions on sustainability give rise to unexpected results. Moreover, the use of computational explorations could facilitate the explanation of known facts and the discovery of commonalities among different perceptions of sustainability, hence leading towards the development of a composite definition of sustainable development.

For that reason, the combination of DEA and ML has been gaining traction in the literature: Samoilenko and Osei-Bryson (2008) used cluster analysis to separate the original set of DMUs, calculated the efficiencies in each cluster and finally employed decision tree induction to gain insights into the specifics of each cluster. Wu (2009a) proposed a hybrid model that consists of two modules: in the first module suppliers are evaluated under DEA and are separated into efficient and inefficient ones. In the second module, neural networks are trained with the generated efficiencies to predict the performance of new suppliers. Hu et al. (2012) used DEA to calculate the efficiency of IT investments in Chinese companies and then employ Classification and Regression Trees (CART) to identify the main factors that affect the performance. De Nicola et al. (2012) combined DEA with CART to evaluate the Italian health system. and Salehi, Veitch and Musharhag (2020) used DEA to analyze the influence of resilience engineering and then employ multilayer perceptron to estimate the level of adaptive capacity. Nandy and Singh (2020a; 2020b) used DEA to evaluate the efficiency of farms in India and employ machine learning to gain insights into which variables are crucial in predicting performance. Aydin and Yurdakul (2020) separated countries in groups via clustering and then calculate the efficiency of how countries responded to Covid-19 in each cluster with DEA. Finally, Thaker et al. (2021) employed DEA to evaluate the efficiency of Indian banks and then use Random Forest Regression to analyze the impact of corporate governance (and other bank characteristics) on the calculated efficiencies. Consequently, combining DEA with ML offers an alternative approach to the issue of inputs and outputs selection.

However, all the above combinations of DEA with ML are limited by the repeating theme of this introduction, that they do not consider different perceptions into the calculations. Furthermore, all the above attempts, in essence worked towards reducing the size of the available data with the introduction of ML (e.g using clustering). In the current thesis the opposite occurs; the variety of calculations under different perceptions can be seen as new data generator that are used as inputs for the ML stage of the model. The new data add new layers of insights; hence, the integration of DEA with ML under an exploratory, multi-perspective (similar to a full factorial experimental design pattern) will not only calculate the performance of EU countries on sustainability but at the same time provide insights relevant to policy makers and the general public.

1.3 Contributions

The contributions of the current thesis are summarized below:

- A literature review on Data Envelopment Analysis and sustainability covering the years 2016-2020 (Section 2)
- The proposal and development of a new two-stage DEA variation with a different optimization metric and the inclusion of deviations to handle the weight distribution and the proof of lemmas and a theorem (Section 3.3.1)
- The application of the proposed model to the assessment of the environmental performance (Section 3.3.2)
- The application of the proposed model to the assessment of the agricultural sustainability (Section 3.3.3)
- The design of a framework to construct composite indicators based on DEA (Section 4.1)
- The design and development of an Exploratory, multi-dimensional DEA-ML framework (Section 4.2) and its application in the calculation of the sustainability of EU countries under different perspectives and assumptions (Section 4.3).

1.4 Thesis Structure

The rest of the thesis is structured as follows:

Section 2 is focused on the literature and especially how Data Envelopment Analysis has been employed to calculate sustainability of countries and regions.

In **Section 3**, the methodology of Data Envelopment Analysis is presented mathematically. Furthermore, the new two-stage DEA variation is proposed, formalized mathematically, applied in two case studies and finally, its sensitivity on the rank reversal phenomenon is tested.

In **Section 4**, the Exploratory, Multi-dimensional Data Envelopment Analysis – Machine Learning computational framework is proposed, designed and applied in the calculation of sustainability of EU countries' sustainability under different perceptions and methodological assumptions.

Conclusions, lessons learnt from the research and literature as well as future research directions are presented and discussed in **section 5**.

2. Literature Review[2]

2.1 Sustainable Development and Sustainability

The term of sustainable development became widely known in the 1980s with the Brundtland report (Brundtland, Khalid, Agnelli, Al-Athel, & Chidzero, 1987). The report served as a reminder that all human activities had damaged the natural integrity and had caused unbalances to ecosystems that could seriously threaten the security of human societies (Coli, Nissi, & Rapposelli, 2011).

The report brought to the limelight of public debate the notion that the objectives of policy makers and governments cannot be solely to promote social development while facilitating economic prosperity (Lin & Chiu, 2018), but there needs to be a constant reminder in the decision-making process that conflicts can exist between economic growth and the environment (Munda & Saisana, 2011)**.**

Hence, to achieve sustainable development governmental policies should have economic, social and environmental dimensions. Their consequences should contribute neither to overexploitation of the natural resources nor to widen the gap in distribution of social services (Hassanzadeh, Yousefi, Saen, & Hosseininia, 2018). Finally, sustainable development should always reconcile technological development and efficiency (Robinson, 2004), while considering the cultural context and the values system in which it is applied (Santana, Mariano, Camioto, & Rebelatto, 2015).

Despite of the immense complexity associated with the concept of sustainable development - even from its beginning - efforts to achieve that state have become common practice in all levels of public policy, from government laws to regional and private decision-making (Li, et al., 2009). The reason behind that effort is best captured by the implications if sustainable development is not achieved: the ability of the natural environment to provide critical resources will be severely diminished (Daniels & Moore, 2001) followed by dire consequences for human societies.

[2] The Literature Review appeared on: Tsaples, G., & Papathanasiou, J. (2021). Data envelopment analysis and the concept of sustainability: A review and analysis of the literature. *Renewable and Sustainable Energy Reviews*, *138*, 110664. (IF: 14.982, Q1)

Thus, the achievement of sustainable development is an enormous, complex and ongoing effort, but the first step should always be to address the widely recognized need of how best to measure it (Tyteca, 1998).

The notion that has been used to measure the extent to which sustainable development has been achieved is the one of *sustainability*. It originates from the field of ecology and in its most basic form it signals the ability of a natural system to retain its essential properties and naturally replenish its population. Hence, sustainability is a measure of endurance of natural systems (Zhou, Yang, Chen, & Zhu, 2018), while in terms of human systems and processes, sustainability focuses on the ability to live without environmental degradation (Robinson, 2004; Sneddon, Howarth, & Norgaard, 2006).

Despite their importance, both sustainable development and sustainability are characterized by a plethora of definitions and meanings for people and organizations (Drucker, 2014; Robinson, 2004). In general, however, all the definitions fall under two categories: there is the three-dimension approach (integration of economic, social and environmental dimensions) and the dualistic topology that emphasizes the relationship between human and nature (Robinson, 2004). Lately, an even more contested term has entered the arguments of the opposing categories, with some claiming that the road to sustainable development can be achieved through *technology and innovation* while the rest claiming that this road could only lead to further environmental degradation (Drucker, 2014).

Consequently, to continue to be an uncontested governmental activity (Coli, Nissi, & Rapposelli, 2011), sustainable development needs to be defined in such a way that it does not exclude any views; whether one perceives sustainability as a three-dimensional construct or as a measure of balance between humans and nature, there is the need to develop a measure or index that incorporates both (or even more) perceptions.

The lack of a unified definition notwithstanding, policy makers understood the importance of trying to achieve sustainable development and a series of international treaties and policy frameworks have been reached. The most important examples of recent years are the Sustainable Development Goals of the United Nations that attempt to synchronize the effort across countries (Adler, 2012) and the United

Nations Conference of the Parties 2015 Agreement (or else known as the COP-21/Paris Climate Agreement), the focus of which has been to underline the connection between sustainable patterns of consumption and the fight against climate change (DiMaria, 2019).

What these attempts do not seem to offer however, is methodological guidance and a unified framework on how to measure sustainability in practice and therefore achieve sustainable development (Munda & Saisana, 2011). The use of such an appropriate framework and/or guidance could immensely help policy makers reaching effective decisions related to sustainable policies (He, Wan, Feng, Ai, & Wang, 2016), especially since decision-making is increasingly characterized by multi-dimensional complexity.

Composite indicators emerged as a tool for the proper measurement of sustainability (Coli, Nissi, & Rapposelli, 2011). A composite indicator can be considered as a mathematical construction that can measure multidimensional concepts, derived from individual indicators that usually have no common units of measurement (Zhou, Yang, Chen, & Zhu, 2018). Their advantages include the fact that they can be easily communicated and act as justification tools for policy makers, while - if properly constructed - they can lead to meaningful comparisons, policy monitoring and benchmarking (Zhang, Kong, & Choi, 2014).

These indicators have been considered essential for regional sustainability measurement (Ervural, Zaim, & Delen, 2018), but in the beginning, they mainly focused on the environmental aspect of sustainable development (Zofío & Prieto, 2001), covering the effects of economic activities to the environment. The criticism over this focus, has lead the research towards integrating more aspects of the sustainability structure such as the social dimension (Gibson, 2006; Pope, Annandale, & Morrison-Saunders, 2004; Winfield, Gibson, Markvart, Gaudreau, & Taylor, 2010), innovation as a force for socioeconomic change (Drucker, 2014), and the capacity of a country to produce a steady stream of sustainable technological products (Santana, Mariano, Camioto, & Rebelatto, 2015).

Apart from the criticism on what these indices should include and measure, there have also been voices of concern on how they are constructed. The main objection is that they are usually developed by using a framework of weighted linear aggregation.

Linear aggregation however implies compensability among the parameters (or sub-indicators) that construct the overall indicator; disadvantages of one sub-indicator could be offset by a sufficiently large advantage of another sub-indicator (Munda & Saisana, 2011). In the case of sustainability for example this could mean that the loss of potable water or the diminished levels of clean air could be substituted by economic growth (Munda & Nardo, 2009).

Such an assumption is not realistic and even goes against the very notion of sustainability. For that reason, a robust methodological framework is necessary to mitigate the methodological limitations and assist in constructing effective and appropriate sustainability indices.

Furthermore, the linear aggregation function demands the determination of weights from the analyst/policy-maker that builds the function. However, the problem with this type of weighting is that the resulting indicator will represent the values of the analyst/policy-maker, which may differ even within the same society/environment etc. (Kuosmanen & Kortelainen, 2005).

In conclusion, the studying of sustainable development and sustainability led to the identification of the following gaps: First, the lack of a unified definition on what sustainable development is, resulted in different and even conflicting interpretations, which may have a negative effect on the communication of why sustainable development is necessary. Second, the lack of a unified methodological framework for measuring sustainability led to the employment of methods that may not be suitable to capture its multi-dimensional nature, which may have resulted in policies that are not sufficient and effective. Finally, these two gaps are interrelated: misguided assumptions about what sustainable development is accompanied by misguided methodological assumptions on how to measure it, lead to wrong estimations, hence the individual consequences of each one is amplified, increasing the overall complexity of the endeavor.

The next section is focused on how Data Envelopment Analysis has attempted to mitigate the methodological limitations of measuring sustainability, while in parallel offering examples within DEA of how the problem of different definitions still persists affecting the results.

2.2 DEA and Sustainability

As it was mentioned in the introduction, Data Envelopment Analysis emerged as a suitable method to measure sustainability. It is a non-parametric method that is used for the assessment of the *technical efficiency of Decision Making Units (DMUs)* relative to one another (Adler, 2012; Førsund & Sarafoglou, 2002), where technical efficiency can be defined as a measure of how well a DMU can transform inputs into outputs.

The definition of efficiency for DEA originates in engineering and (as was already indicated in the Introduction) is defined as the ratio of the sum of its weighted outputs over the sum of its weighted inputs.

$$technical\ efficiency = \frac{\sum w_{output} * y}{\sum w_{input} * x}, where\ x = input\ level\ and\ y = output\ level \qquad \text{(1 revisited)}$$

The method was established in the seminal papers of Charnes, Cooper and Rhodes (1978) and Banker, Charnes and Cooper (1984). In its most basic form, it is assumed that there are N DMUs that use *m* inputs to produce *s* outputs. The notation includes the variables of x_{ij} (i=1…m, j=1…N) the level of the i^{th} input of DMU *j,* and y_{rj} (r= 1…s, j = 1…N) the level of rth output of DMU *j*.

Then the calculation for the *technical efficiency for the input-oriented model* can be found by solving the Linear Program (LP):

$$min\ \Theta_0 - e\left(\sum_{i=1}^{m} S_i^- + \sum_{r=1}^{s} S_r^+\right) \qquad (2)$$

$$subject\ to\ constraints:$$

$$\sum_{j=1}^{N} \lambda_j * x_{ij} = \Theta_0 * x_{ij_o} - S_i^-, i = 1 \dots m \qquad (3)$$

$$\sum_{j=1}^{N} \lambda_j * y_{rj} = y_{rj_o} + S_r^+, r = 1 \dots s \qquad (4)$$

$$\lambda_j \geq 0, j = 1 \dots N, S_r^+, S_i^- \geq 0 \qquad (5)$$

The variable λ_j is the weight calculated by DEA with the equations (2)-(5) for DMU$_j$ while the variables S_r^+, S_i^- are the slack variables that are used in Linear Programming. They represent any additional output increase or input decrease that is feasible to be achieved by the DMU.

The technical efficiency of the above problem for DMUj_o is the variable Θ_0 and it takes values between 0 and 1 (or 0 and 100%). The mathematical program represented with equations (2)-(5) is solved separately for each DMU and there are three options for the results after the solution:

1. DMUj_o is Pareto-efficient if and only if $\Theta_0 = 1$ at the optimal solution and $S_r^+, S_i^- = 0$ for all inputs and outputs
2. If the value of one of the slack variables S_r^+, S_i^- is positive at the optimal solution, the corresponding input (or output) of DMUj_o can be further improved
3. If none of the above applies, then DMUj_o has technical efficiency ${\Theta_0}^*$. In the particular case, the technical efficiency at the optimal solution $\Theta_0 < 1$ reflects the maximum radial contraction of the input levels, without worsening the output levels, in order for DMUj_o to be considered efficient.

This simple model and the notion of efficiency has proven to be appropriate for the measurement of sustainability. Zhou et al. (2018) performed an extensive literature review of how Data Envelopment Analysis has been used in the context of sustainable development and sustainability. Their paper covers research efforts until 2016, and the authors consider their work an extension of the review by Dakpo, Jeanneaux and Latruffe (2016). Their focus is not only the environmental dimension of sustainability, but they attempt to include and search for the social factors that can contribute to sustainable development.

For the current thesis, a search was performed in bibliographic databases (Scopus and Google Scholar) for the years after 2016 to investigate the extent to which the method has been used for sustainability (using as keywords the terms "composite indicators" and "data envelopment analysis" or "DEA"). From the initial sample of articles, a further screening was performed by reading the abstracts (and where necessary the main texts) to check for measurement of sustainability (or similar notions). Finally, several articles from the original review are included in the final sample because they were used to draw alternate interpretations of the results. Tables 1, 2 and 3 below summarize the new search, grouped per region of application.

Table 1 Summary of the new research on the literature- Applications in Europe

Work	**Input**	**Intermediate**	**Output**	**Index**	**DEA variation**	**Combination with other method**	**Area of Application**
(Moutinho, Madaleno, & Robaina, 2017)	Labor productivity, capital productivity, the weight of fossil energy and the share of renewable energy in GDP	-	GDP/GHG	Efficiency	Classic DEA	Quantile regression	EU countries
(Masternak-Janus & Rybaczewska-Błażejowska, 2017)	consumption of electricity, consumption of heat, consumption of fuel, consumption of sawn wood and particle boards, consumption of fiberboard, consumption of sheets of float glass, consumption of paper and cardboard, consumption of cement, consumption of basic chemicals and plastics, consumption of metallurgical products, water consumption, wastewater discharged in waters, emissions of air pollutants, waste production	-	GDP, gross value added	Eco-efficiency	Classic DEA	-	Polish regions

(Gonzalez-Garcia, Manteiga, Moreira, & Feijoo, 2018)	AROPE rate, unemployment rate, LCA result, Public school vacancies, number of crimes, inhabitants with higher education	-	Net disposable income	Efficiency	Classic DEA	Material Flow Analysis+ Life Cycle Assessment	Spanish cities
(Moutinho, Madaleno, Robaina, & Villar, 2018)	GDP, population density, labor productivity, total resource productivity, patent applications per 10000 inhabitants	-	GDP per capita, CO_2emissions	Eco-efficiency	Classic DEA	Malmquist index	German and English cities
(Cucchiella, D'Adamo, Gastaldi, & Miliacca, 2018)	Greenhouse gases, Gross final energy consumption, renewable energy consumption	-	GDP, population	Efficiency	Classic DEA	Zero Sum Gains DEA	EU countries
(Biresselioglu, Demir, & Turan, 2018)	mathematical programming scores and scores from the energy trilemma	-	energy consumption, GHG generations, share of renewable energy in gross final energy consumption	Efficiency	Classic DEA	Mathematical Programming	EU countries

(Carboni & Russu, 2018)	infrastructure, efficiency of the legal system, tourists, high school qualifications, unauthorized buildings	-	Environmental index, GDP per capita	Eco-efficiency	Classic DEA	Malmquist index	Italian regions
(Pozo, et al., 2019)	Percentage of people with low income, Carbon emissions, Traffic flow, House Price, Anxiety	-	Happiness, Life Satisfaction, Income of tax Payers	Efficiency	Non radial DEA	Temporal analysis	London boroughs
(Tsaples G. , Papathanasiou, Georgiou, & Samaras, 2019)	Gross Fixed Capital in PPS, Total Labor Force	GDP per capita in PPS	Share of renewable energy in gross final energy consumption, Greenhouse gas emissions (in CO2 equivalent), Overall life satisfaction, Satisfaction with living environment, Satisfaction with financial situation, Intramural R&D expenditure for all sectors of the economy	Sustainability index	Multi-stage DEA	-	EU countries
(Tsaples & Papathanasiou, 2020b)	Fixed Capital in Purchasing Power Standards (PPS), Total Labor Force,	GDP per capita in PPS	Share of renewable energy in gross final energy consumption, Greenhouse gas emissions (in CO2 equivalent), Overall life	4 sustainability indices using combinations of inputs and outputs	Multi-stage DEA applied 4 times		EU countries

			satisfaction, Satisfaction with living environment, Satisfaction with financial situation, Intramural R&D expenditure for all sectors of the economy, Mean equivalized net income, ability to face unexpected financial expenses as percentage of the population				

Table 2 Summary of the new research on the literature- Applications in Asia

Work	**Input**	**Intermediate**	**Output**	**Index**	**DEA variation**	**Combination with other method**	**Area of Application**
(Sueyoshi & Yuan, 2017)	Capital, Labor, Energy	-	Gross Regional Product, CO_2 emissions, SO_2 emissions, soot, wastewater, Chemical Oxygen Demand, NO	Efficiency under natural and managerial disposability	Intermediate DEA	-	Chinese regions
(Sueyoshi, Yuan, Li, & Wang, 2017)	Capital, Labor, Energy	-	Gross Regional Product, CO_2 emissions, SO_2 emissions, soot, waste water, Chemical Oxygen Demand, NO emissions	Efficiency under natural and managerial disposability	Intermediate DEA	-	Chinese regions
(Lin & Chiu, 2018)	population, investment in energy industry	coal consumption, oil consumption, electricity consumption, natural gas consumption	CO_2 emissions, GDP	Efficiency	Two-stage DEA	-	Chinese regions
(Sueyoshi & Yuan, 2018)	electricity consumption, total primary energy consumption	-	GDP, GDP per capita, total CO_2 emissions, CO_2/total primary energy	Efficiency and natural and managerial disposability	Intermediate DEA	-	Asian countries

(Song, Peng, Wang, & Zhao, 2018)	Employment, Total Energy Consumption, Fixed capital input	-	Total discharge of industrial wastewater, Discharge of industrial waste gas, amount of industrial solid waste	Efficiency	Ray slack-based DEA	-	Chinese regions
(Ervural, Zaim, & Delen, 2018)	Total renewable energy potential, network length, total installed power of renewable energy, transformer capacity	-	Gross energy generation from renewable energy, number of consumers, total exports, GDP per capita, HDI, Total energy production, Population, area	Super efficiency	Super efficiency DEA	Tobit regression analysis	Turkish regions
(Zhang, Li, & Gao, 2018)	Capital, Labor, Energy	-	Gross regional product (GRP), CO_2emissions, SO_2 emissions, soot and dust, wastewater, COD, Ammonia nitrogen	Efficiency	Intermediate DEA		Chinese regions
(Zhao, Zha, Zhuang, & Liang, 2019)	Capital, Labor, Energy, RFE %	GDP	Wastewater, waste gas, Solid waste, SHC, SBE, SSSE	Efficiency	Parallel DEA models	-	Chinese regions

Table 3 Summary of the new research on the literature- Applications in various countries

Work	**Input**	**Intermediate**	**Output**	**Index**	**DEA variation**	**Combination with other method**	**Area of Application**
(Hassanzadeh, Yousefi, Saen, & Hosseininia, 2018)	total material consumption, labor unemployment	-	GDP per capita, CO_2 emissions, employment protection index	Efficiency	SORM DEA	Inverse SORM DEA	OECD countries
(DiMaria, 2019)	Labor, capital	-	GDP, ecological reserve deficit	Aggregation of efficiency and anti-efficiency	RAM DEA	-	Various countries
(Tajbakhsh & Shamsi, 2019)	Imports of goods and services in current US$, total annual freshwater withdrawals in percentage of internal resources, public expenditure per capita in current US$, duration of compulsory education	-	exports of goods and services in current US$, GNI per capita in current US$, total life expectancy at birth in years, total employment, proportion of seats held by women in national parliaments in percentage, CO2 emissions, total refugees leaving the country	Efficiency	Classic DEA	-	Various countries

The tables above indicate papers that were published until 2020. The first aspect that can be noticed is that apart from the indicators that were identified as inputs and outputs by Zhou et al.(2018), an effort has been made to diversify their types with the aim of including social sub-indicators. Furthermore, their variety has increased with authors trying to diversify the types of pollutant emissions, waste, consumption etc. Interestingly, of all the papers that were studied only five diversify completely from the norm.

Gonzalez-Garcia et al. (2018) who used as inputs several social indicators that were not used before like level of higher education, crimes etc. and as output the net disposable income to study the efficiency of Spanish cities with regards to sustainability. Furthermore, Carboni and Russu (2018) included the quality of the legal system along with notions of corruption and quality of life to investigate the eco-efficiency of Italian regions.

However, only in 2019 notions like “happiness”, “proportion of seats held by women in national parliaments in percentage” and “total refugees leaving the country” have started to be included explicitly as equally important inputs and outputs in analyses (Pozo, et al., 2019; Tajbakhsh & Shamsi, 2019).

The inclusion of diverse social indicators continued in the work of Tsaples et al. (2019) who included notions such as “overall life satisfaction” and “satisfaction with the living environment” in their calculation of sustainability. Furthermore, Tsaples and Papathanasiou (2020a) further continued this trend with variables such as the “ability to face unexpected financial expenses as percentage of the population”.

Despite the inclusion of more societal dimensions, the literature is lagging in including technology and innovation in the measurement of sustainability. In the work by Santana et al. (2015), the authors used as input the gross domestic expenditure on Research and Development along with the total number of applications, as outputs they used GDP per capita, means of schooling years, life expectance, CO_2 emissions to measure the efficiency of sustainable development of BRICS and G7 countries. Finally, Tsaples et al. (2019) and Tsaples and Papathanasiou (2020a) included the variable of “Intramural R&D expenditure for all sectors of the economy” in their calculations.

Consequently, in the debate of how to define and measure sustainable development, the research on DEA illustrates that the concept of the three-dimensional sustainability appears to be predominant. It must however be noted that in that way, other views and definitions are usually excluded from the analysis. Only recently, individual efforts have started to look how sustainability could be defined in an alternate way.

Regarding the actual notion of sustainability, Figure 1 below illustrates the type of index that has been employed.

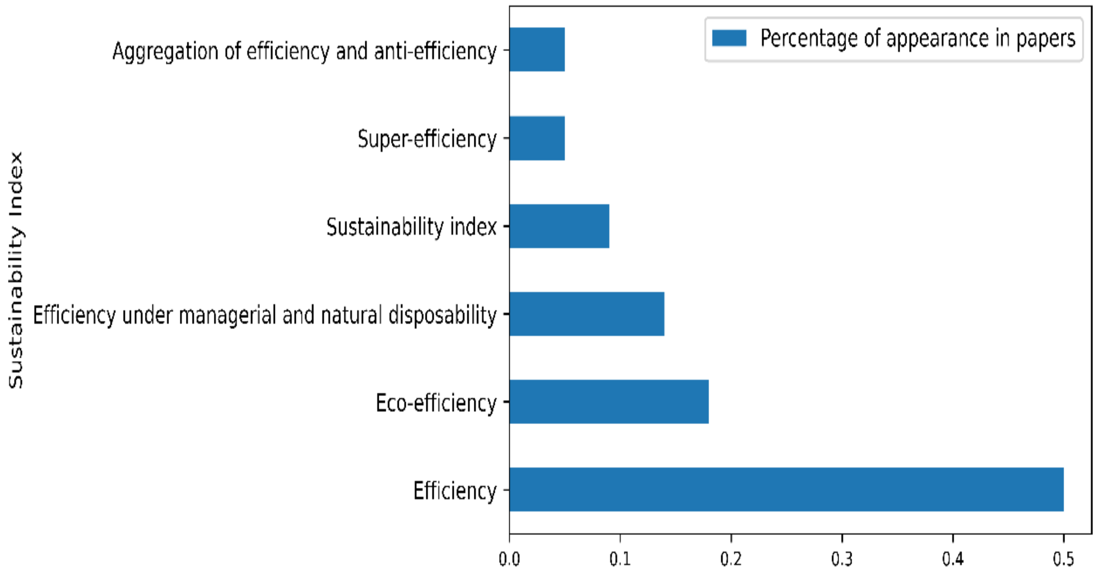


Figure 1 Frequency of appearance of sustainability index

As it can be observed, the majority of the papers used Efficiency (or some variation) as a proxy for sustainability and only in two papers there is an explicit mention of calculating a sustainability indicator. Furthermore, the second most-used term is that of eco-efficiency.

Eco-efficiency is one of the most widely used indicators that is related to the more encompassing notion of sustainability (Kuosmanen & Kortelainen, 2005). The aim of an eco-efficient system is the maximization of the production while keeping the environmental consequences to a minimum (Moutinho, Madaleno, & Robaina, 2017; WBCSD, 2000). OECD (1998) defined it more formally as “the efficiency with which ecological resources are used to meet human needs”.

Its concept can be traced in the decade of 1970s when it was linked to the efforts to achieve environmental efficiency (Carboni & Russu, 2018; Freeman III, Haveman, &

Kneese, 1973);. According to Huppes and Ishikawa (2005), the notion of eco-efficiency can be measured in real life with 4 ways:

- As the ratio of economic output to environmental pollution (named as environmental productivity).
- As the ratio of environmental pollution to economic activity (named as environmental intensity).
- As the ratio of improvement cost to environmental improvement (named as environmental improvement cost).
- As the ratio of environmental improvement to improvement cost (named as environmental cost effectiveness).

Similar to the discussion on which inputs and outputs should be used with DEA to measure sustainability, the above notions of economic output, environmental pollution etc. are perceived differently by different authors and different combinations of inputs and outputs are used to measure eco-efficiency.

The notion of eco-efficiency is considered critical since it provides a pathway to the design of policies that could reduce environmental pressure (Kuosmanen, 2005). Furthermore, the ratio of economic value over environmental damage (or its inverse) is considered intuitive and clear, thus making eco-efficiency a measure of sustainability that is easy to communicate (Kuosmanen & Kortelainen, 2005).

For those reasons, eco-efficiency has been extensively used with DEA as a measure of sustainability of regions and countries (Xing, Wang, & Zhang, 2018).

Masternak-Janus and Rybaczewska-Błażejowska (2017) used a classic DEA model to measure eco-efficiency of Polish regions, while in Moutinho et al. (2018) the authors did the same for German and English cities. Finally, a stream of research utilized the notion of eco-efficiency to measure sustainable development of various regions. This stream includes the works of Carboni and Russu (2018) for Italy and Lin and Chiu (2018) for Chinese regions.

Despite its popularity, eco-efficiency has attracted a lot of criticism. Ehrenfeld (2004) sees eco-efficiency as a symptomatic solution where technological innovation is solving the problems that were created by technological innovation. Furthermore, attempting to achieve adequate levels of eco-efficiency does not guarantee a state of sustainable development (Kuosmanen & Kortelainen, 2005). A recurring criticism that was observed in all the literature, is which inputs and outputs should be used in the DEA model with the aim of measuring eco-efficiency.

As a result, the lack of a unified definition has an impact of how sustainability is perceived and how it should be measured. In the DEA literature, efficiency is used as equivalent to sustainability, while other notions like eco-efficiency are regarded as sufficient proxies. Furthermore, the diverse definitions have an effect on which inputs and outputs should be used to measure sustainability. In the DEA literature, the typical three-dimensional construct is prevalent, but recently efforts have been made to include technological aspects. Finally, only one recent paper attempted to integrate different definitions of sustainability within the same measurement.

Apart from the effects of different definitions of sustainability, differences are also observed in the methodology used even within the same DEA framework. Figure 2 below illustrates the frequency of the method that has been used in the literature.

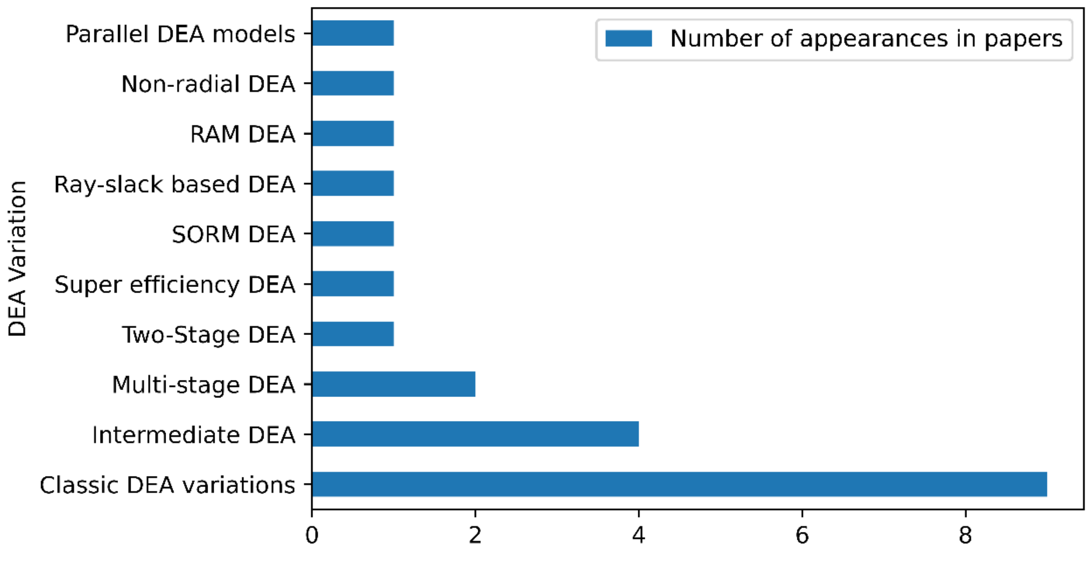


Figure 2 Frequency for DEA variations in the papers under study

Thus far, it appears that classic variations of DEA are the preferred option for researchers, with non-radial and multi-stage approaches gaining traction in the last years.

Moreover, a trend is observed where authors enrich the results generated by DEA with another method to gain another layer of knowledge. For example, a combination of a classic CRS DEA model with another method is the one proposed by Cucchiella et al. (2018). After analyzing the DEA model, the authors perform a second analysis to identify the input values that make the system under study globally efficient.

A similar idea but with a different approach was executed by Hassanzadeh et al. (2018). The authors combined a DEA model with its inverse; its purpose was to determine the most desirable inputs and outputs that keep the levels of efficiency unchanged (Lertworasirikul, Charnsethikul, & Fang, 2011).

Another stream of work observed in the literature is the use of the Intermediate DEA. The method was proposed by Sueyoshi et al. (2017) and a typical example of its use is the work on (Sueyoshi & Yuan, 2018). The authors measured efficiency but they did so under the concepts of natural and managerial disposability.

The examples that were described thus far use the typical, one-stage version of DEA. However, in recent years, researchers understood that the robustness of a model increases (whereby robustness it is meant to increase the validity of the results by mitigating some of the limitations of DEA) if the region/country under study is not considered a black box; for this reason, a network-version of DEA could be used. Furthermore, in any DEA analysis, the number of inputs and outputs depends on the number of DMUs under study for the results to be meaningful i.e. the number of DMUs, must be no less than three times of the total number of inputs and outputs (for example if one uses 2 inputs and 2 outputs to measure sustainability then the number of DMUs or regions under study should not be less than 12) (Wu D. , 2009b).

As a result, increasing the number of intermediate stages increases the discriminatory power of the DEA model and the analysis. For example, Zhao et al. (2019) used parallel settings of DEA to explicitly model the three dimensions of sustainability and their potential interactions. In a similar direction, the work by Tsaples et al. (2019) and Tsaples and Papathanasiou (2020a) use multi-stage DEA models to calculate the different sub-indicators that sustainability entails and integrate them in a final sustainability index.

Consequently, even within the DEA framework it appears that there are differences on the variation that is being used to measure sustainability. As a result, the lack of a unified methodological framework from the birth of the notion of sustainable development persists to date.

Finally, by examining the area of application, it is revealed that for the last years the scope of research is tailored towards urban environments with a focus on Chinese regions. The measurement of sustainability of European countries is steadily increasing, but it appears that the papers that investigated the sustainability among the EU countries was still lower than those measuring sustainability of Chinese regions. One possible explanation could be that the rapid economic development that was observed in China the previous years, made the research community to reflect on what the impact of this development could be in the environment. Figure 3 below illustrates the frequency of appearance for the various areas of application.

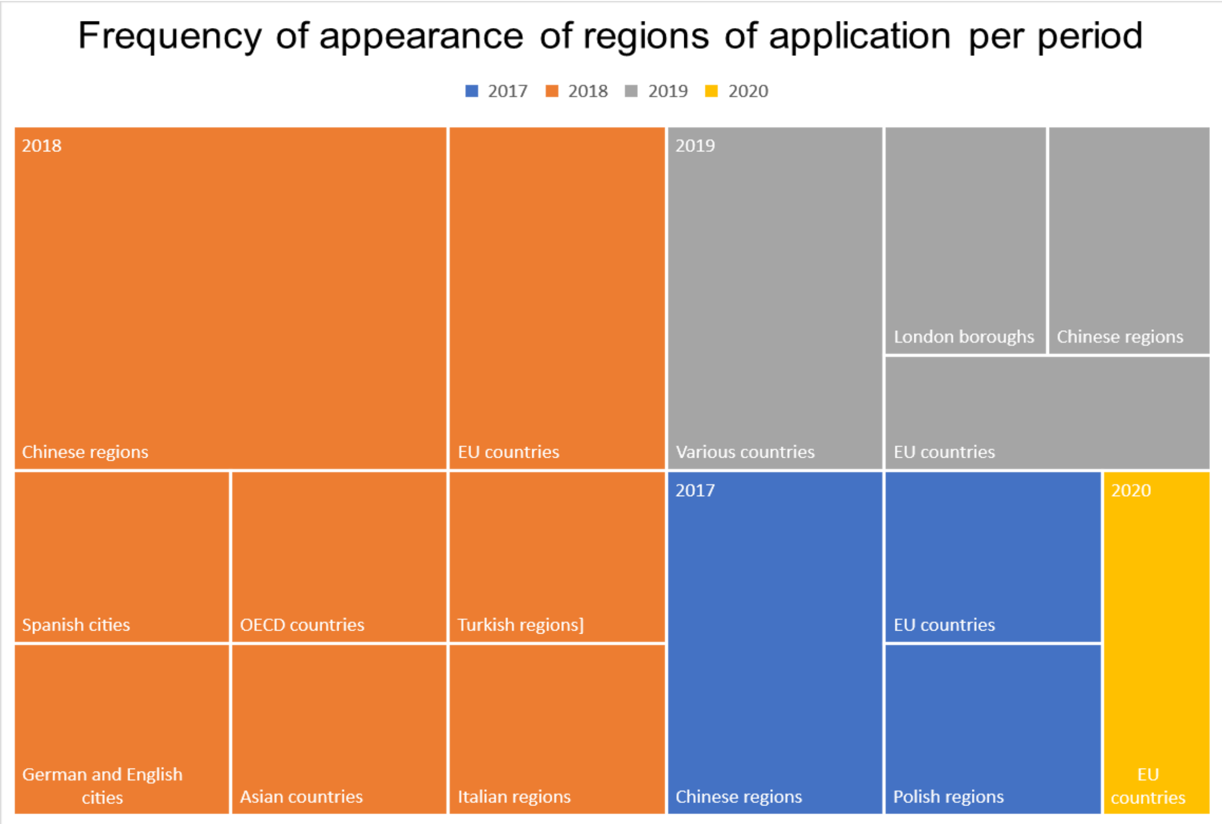


Figure 3 Frequency of appearance for the various regions of application of DEA in sustainability measurement per year

2.3 Lessons learned from the literature

The studying of the literature on Data Envelopment Analysis and Sustainability has highlighted that the ambiguity of the definition of sustainable development has permeated to the research. The three-dimensional structure of sustainability appears to be the preferred option, however there are approaches where there is an effort to integrate different dimensions, like technology and innovation, in the measurement of sustainability.

Moreover, even sustainability as a measurement of sustainable development appears to have different definitions. In the DEA literature, efficiency is used as equivalent to sustainability, while other notions like eco-efficiency are regarded as appropriate proxies. These diverse definitions have an effect on which inputs, outputs and data should be used to measure sustainability, thus impacting the final results.

In addition to the lack of a common definition of sustainability, differences are observed on the variation of DEA that is being used. All these differences result in different measurements of sustainability, which may cloud the robustness of the research efforts and ultimately affect the policy-making that depends on those measurements.

Finally, by examining the area of application, it is revealed that for the last years the scope of research is tailored towards urban environments with a focus on Chinese regions. One possible explanation could be that the rapid economic development that was observed in China the previous years, forced policy makers and the research community to reflect on the impact of this development.

3. Methodology

3.1 Data Envelopment Analysis[3]

Data Envelopment Analysis (Førsund & Sarafoglou, 2002; Thanassoulis, 2001) is a non-parametric, mathematical programming technique that is used for the evaluation of the Technical Efficiency of a group of (generally) homogeneous units under assessment. These units are referred to as Decision Making Units (DMUs) and the technical efficiency measures their performance relative to one another.

The definition of efficiency for DEA originates in engineering and is defined as the ratio of the sum of its weighted outputs over the sum of its weighted inputs (Ishizaka & Nemery, 2013) as expressed in:

$$technical\ efficiency = \frac{\sum w_{output} * y}{\sum w_{input} * x}, where\ x = input\ level\ and\ y = output\ level \quad \text{(1 revisited)}$$

By outputs it is meant what the DMU produces, while inputs are the factors that the DMU consumes. All the input-output correspondences that can be achieved by a DMU (regardless if these correspondences are observed in practice) form the Production Possibility Set (PPS). Finally, the weights give a measure of importance of the inputs and outputs that are used.

Hence, technical efficiency is a measure of how well a Decision Making Unit can transform inputs into outputs, without considering how the transformation can be achieved. Moreover, Decision Making Units represent any type of entity and are considered homogeneous in terms of identical inputs and outputs. Finally, the technical efficiency that is measured cannot be considered "absolute" because the performance measurement is done with reference to the set of the other (homogeneous) DMUs (Thanassoulis, 2001).

As a result, a DMU is considered efficient compared to the rest, when there cannot be an improvement in one of the levels of inputs or outputs without a simultaneous

[3] This section appeared on the book chapter: Tsaples, G., Papathanasiou, J., Digkoglou, P.(2021). Decision-making in the context of Sustainability. In Urban Sustainability: A game-based approach. (Eds.) Papathanasiou, J., Tsaples, G., Blouchoutzi, A., Springer.

deterioration in one of the others. This efficiency is defined as Pareto-Koopmans efficiency (Cooper, Seiford, & Zhu, 2011).

DEA offers a series of advantages compared to other parametric methods of efficiency assessment:

1. As mentioned before, it is not necessary to identify the relationships/processes that transform the input(s) of a DMU to output(s)
2. It requires less information that traditional methods, since it can allow comparative assessment of DMUs in situations where price information is not available
3. DEA can provide insights into the reasons for which a DMUS is not efficient and propose directions towards its improvement (He, Wan, Feng, Ai, & Wang, 2016; Shi, Bi, & Wang, 2010)
4. Inputs and outputs are not required to have common units of measurement.

However, the method does not come without limitations:

1. The results obtained by DEA cannot be generalized and concern only the specific DMUs set under evaluation
2. Noise and outliers can cause errors in the results
3. Efficient DMUs cannot be ranked, since all take the maximum value of efficiency
4. The general method cannot aggregate different dimensions of efficiency.

The seminal papers that fully established the two basic models of DEA are by Charnes, Cooper and Rhodes (1978) which assume that the DMUs operate under Constant Returns to Scale (CRS) and in a perfectly competitive environment, and by Banker, Charnes and Cooper (1984) that relaxe these assumptions and assume that the DMUs operate under Variable Returns to Scale (VRS).

The notions of Returns to Scale are borrowed from economics and describe what happens in a production when inputs are increased by a factor α, where $\alpha > 0$. Formally, assuming that a production uses x inputs to produce y outputs, and the inputs are increased by a factor α, where α > 0, then the outputs could vary by the factor β. Furthermore, assuming that:

$$p = \lim_{a \to \infty} \left(\frac{b-1}{a-1} \right) \tag{6}$$

The following cases are possible:

- p > 1: In this situation an increase in inputs leads to a more than proportional increase in output. It is noted as Increasing Returns to Scale
- p=1: In this situation output increases by the same proportional increase as the inputs. It is noted as Constant Returns to Scale
- p<1: In this situation, an increase in inputs leads to a less than proportional increase in outputs. It is noted as Decreasing Returns to Scale (Thanassoulis, 2001).

Furthermore, a DEA model can be input- or output-oriented. An input-oriented DEA model minimizes input for a given output. Thus, it indicates how much a DMU should reduce its levels of input(s) to achieve the given level of output. On the contrary, an output-oriented DEA model maximizes output for a given input. Thus, it indicates how much a DMU should increase its levels of output(s) given the provided level of input(s).

In both cases, DEA identifies a frontier of "best practices" or efficient frontier that envelopes all the DMUs; those that lie on the frontier are considered the most efficient or "best-practice units" and those that lie beneath the frontier are considered inefficient. In the inefficient units, a number is attributed that represents their radial distance from the efficient frontier. Its difference from the maximum value (reserved for the DMUs on the efficient frontier) indicates:

- The level of output(s) decrease that the DMU should proceed to for its given inputs (input-oriented model) to be considered efficient
- The level of output(s) increase that the DMU should proceed to for its given inputs (output-oriented model) to be considered efficient (Coelli, Rao, O'Donnell, & Battese, 2005).

3.1.1 DEA under Constant Returns to Scale

The DEA models either under Constant or Variable Returns to Scale are approached as Linear Programming models (Cooper, Seiford, & Zhu, 2011; Thanassoulis, 2001; Zhu J. , 2014). For their calculation, there is the need to define the Production Possibility Set which contains all input-output correspondences that can be achieved by a Decision Making Unit, regardless if these correspondences are observed in practice or not.

The measurement of the relative efficiency of the DMU is achieved in two steps:

(1) The construction of the Production Possibility Set; meaning the definition of inputs and outputs of the Decision Making Units

(2) The estimation of the degree to which the outputs can be expanded or the inputs can be contracted within the limits defined in the Production Possibility Set.

The DEA model under Constant Returns to Scale is known as DEA-CCR model and was first introduced by Charnes, Cooper and Rhodes (1978). Under this assumption if there is an input scaling (either upwards or downwards) of a feasible input-output correspondence then another feasible correspondence is obtained, where the output levels have the same factorial scaling as the input levels.

For its formulation, we assume that there are *N* DMUs that use *m* inputs to produce *s* outputs. We denote denote x_{ij} (i=1…m, j=1…N) the level of the *i*th input of DMU *j*, and y_{rj} (r= 1…s, j = 1…N) the level of DMU .

Then the calculation for the *technical efficiency for the input-oriented model* can be found by solving the LP:

$$min\ \Theta_0 - e\left(\sum_{i=1}^{m} S_i^- + \sum_{r=1}^{s} S_r^+\right) \quad (7)$$

Subject to Constraints:

$$\sum_{j=1}^{N} \lambda_j * x_{ij} = \Theta_0 * x_{ij_o} - S_i^-, i = 1 \dots m \quad (8)$$

$$\sum_{j=1}^{N} \lambda_j * y_{rj} = y_{rj_o} + S_r^+, r = 1 \dots s \quad (9)$$

$$\lambda_j \geq 0, j = 1 \dots N, S_r^+, S_i^- \geq 0 \quad (10)$$

As it is already mentioned, the variable λ_j is the weight calculated by DEA with the equations (7)-(10) for DMU_j while the variables S_r^+, S_i^- are the slack variables that are used in Linear Programming. They represent any additional output increase or input decrease that is feasible to be achieved by the DMU.

The technical efficiency of the above problem for DMUj_o is the variable Θ_0 and it takes values between 0 and 1 (or 0 and 100%). The variable *e* takes a very small value. The problem is solved in two stages:

1. The problem is solved without the slack variables S_r^+, S_i^- and the model provides the technical efficiency ${\Theta_0}^*$.

2. The technical efficiency ${\Theta_0}^*$ is substituted in the LP problem, and it is solved again with objective of maximizing $\sum_{i=1}^{m} S_i^- + \sum_{r=1}^{s} S_r^+$. The result of the second stage is the optimal values of the slack variables, when it is ensured that the optimal technical efficiency is included in the calculation.

The technical input efficiency of a DMU reflects the extent to which the inputs of the DMU can be reduced/contracted without detriment to its output levels (Thanassoulis, 2001). Regarding the solution of the mathematical problem represented with equations (7)- (10), there are three options for the results after the solution:

1. DMUj_o is Pareto-efficient[4] if and only if ${\Theta_0}^* = 1$ and $S_r^+, S_i^- = 0$ for all inputs and outputs

2. If the value of one of the slack variables is positive at the optimal solution, the corresponding input (or output) of DMUj_o can be further improved

3. If none of the above applies, then DMUj_o has technical efficiency ${\Theta_0}^*$ (Thanassoulis, 2001). In the particular case, the technical efficiency ${\Theta_0}^* < 1$ reflects the maximum radial contraction of the input levels, without worsening the output levels, in order for DMUj_o to be considered efficient.

[4] A DMU is Pareto-efficient when any effort to reduce any of its inputs or expand any of its outputs will adversely affect other inputs or outputs. In DEA, Pareto efficiency and efficiency equal to 1 are equivalent (Charnes, Cooper and Rhodes, 1978).

Since the DEA model formulated above are based on Linear Programming, there is the possibility to formulate its dual; it measures efficiency in a value context. These types of models are referred to as value-based DEA models (Thanassoulis, 2001).

The equivalent value-based DEA model to the one represented by equations (7)-(10) is:

$$max \sum_{r=1}^{s} u_r * y_{rj_0} \tag{11}$$

Subject to Constraints

$$\sum_{i=1}^{m} v_i * x_{ij_0} = 1 \tag{12}$$

$$\sum_{r=1}^{s} u_r * y_{rj} - \sum_{i=1}^{m} v_i * x_{ij} \leq 0, j = 1 \dots j_0 \dots N \tag{13}$$

$$u_r \geq \varepsilon, r = 1 \dots s \tag{14}$$

$$v_i \geq \varepsilon, i = 1 \dots m \tag{15}$$

The variables u_r and v_i are the dual variables of the *r*th and *i*th constraints of the problem formulation (7)-(10) respectively and the weight coefficients assigned inputs and outputs respectively by DEA (Thanassoulis, 2001). Furthermore, DMUj_o's technical efficiency is calculated by the $\sum_{r=1}^{s} u_r^* * y_{rj_0}$, where u_r^* is the optimal solution of the problem above. If $\sum_{r=1}^{s} u_r^* * y_{rj_0} = 1$, then DMU$j_o$ is considered efficient.

Thus far, the DEA models described were formulated under input orientation. However, there are the corresponding models for output orientation. For its formulation, we assume that there are *N* DMUs that use *n* inputs to produce *s* outputs. We denote x_{ij} (*i*=1…m, *j*=1…N) the level of the *i*th input of DMU *j*, and y_{rj} (r= 1…s, *j* = 1…N) the level of DMU .

Then the calculation for the *technical efficiency for the output-oriented model* can be found by solving the LP:

$$max\ h_0 + e\left(\sum_{i=1}^{m} S_i^- + \sum_{r=1}^{s} S_r^+\right) \tag{16}$$

Subject to Constraints

$$\sum_{j=1}^{N} \lambda_j * x_{ij} = x_{ij_o} - S_i^-, i = 1 \dots m \tag{17}$$

$$\sum_{j=1}^{N} \lambda_j * y_{rj} = h_0 * y_{rj_o} + S_r^+, r = 1 \dots s \tag{18}$$

$$\lambda_j \geq 0, j = 1 \dots N, S_r^+, S_i^- \geq 0 \tag{19}$$

The technical efficiency of the above problem for DMUj_o is the variable $\frac{1}{h_{j_0}^*}$ and it takes values between 0 and 1 (or 0 and 100%). The variable *e* takes a very small value. The slack variables S_r^+, S_i^- represent any additional output augmentation and/or input reduction that may be necessary in order for DMUj_o to be considered efficient.

The technical output efficiency represents the expansion of the output levels that could be achieved in the DMU under increased performance, without any effect on its input levels. Similar to the DEA CCR model under input orientation, the above model is solved in two stages and the potential solutions to the problem could be:

1. DMUj_o is Pareto-efficient if and only if ${h_0}^* = 1$ and $S_r^+, S_i^- = 0$ for all inputs and outputs
2. DMUj_o has technical output efficiency $\frac{1}{h_{j_0}^*}$ (Thanassoulis, 2001).

It should be stated that the technical efficiency under input orientation and that under output orientation are equal in the case of Constant Returns to Scale (Cooper, Seiford, & Zhu, 2011).

Finally, the equivalent value-based DEA model under output orientation can be solved by the following equations:

$$min \sum_{i=1}^{m} v_i * x_{ij_0} \tag{20}$$

Subject to Constraints

$$\sum_{r=1}^{s} u_r * y_{rj_0} = 1 \tag{21}$$

$$\sum_{r=1}^{s} u_r * y_{rj} - \sum_{i=1}^{m} v_i * x_{ij} \leq 0, j = 1 \dots j_0 \dots N \quad (22)$$

$$u_r \geq \varepsilon, r = 1 \dots s \quad (23)$$

$$v_i \geq \varepsilon, i = 1 \dots m \quad (24)$$

DMUj_o's technical efficiency is calculated by the equation $\frac{1}{\sum_{i=1}^{m} v_i^* * x_{ij_0}}$, where v_i^* is the optimal solution of the problem above. If $\sum_{i=1}^{m} v_i^* * x_{ij_0} = 1$, then DMU$j_o$ is considered efficient (Thanassoulis, 2001).

In conclusion, it should be stated that both the general and the value-based DEA models (under input and output orientations) offer more insights apart from the comparative technical efficiency of the DMUs. In detail, the results from the LP models indicate the efficient DMUs that the inefficient ones should emulate with the purpose of improving their performance. For the general models, the *peers* of (inefficient) DMUj_o are those that have the same mix of input-output levels as DMUj_o but operate at more efficient levels. These peers correspond to the DMUs that exhibit non-zeros λ^* at the optimal solution. For the value-based DEA models, the *peers* of (inefficient) DMUj_o are those that correspond to the binding constraints of the optimal solution (Thanassoulis, 2001).

The finding of peers in the solution of DEA (either CRS or VRS) is a useful feature that can provide insights regarding the performance of a particular DMU. By the notion of peers it is meant identifying efficient DMUs whose practices an inefficient DMU can emulate to enhance its performance (Thanassoulis, 2001). For example, the frequency that a DMU_j is considered a peer to inefficient DMUs can increase the robustness of the result of the particular unit, thus increasing the confidence to the conclusion that the particular DMU_j is efficient.

3.1.2 DEA under Variable Returns to Scale

The DEA model under Variable Returns to Scale is known as DEA-BCC model is an update of the DEA-CCR model and was first introduced by Banker, Charnes and Cooper (1984). In its core the BCC model relaxes the assumption of constant returns to scale.

For its formulation, we assume that there are *N* DMUs that use *m* inputs to produce *s* outputs. We denote x_{ij} (i=1…m, j=1…N) the level of the ith input of DMU j, and y_{rj}

(r= 1…s, *j* = 1…N) the level of DMU. Then the calculation for the *technical efficiency for the input-oriented model (*defined as *pure technical efficiency)* can be found by solving the LP:

$$min\ \Theta_0 - e\left(\sum_{i=1}^{m} S_i^- + \sum_{r=1}^{s} S_r^+\right) \tag{25}$$

Subject to Constraints

$$\sum_{j=1}^{N} \lambda_j * x_{ij} = \Theta_0 * x_{ij_o} - S_i^-, i = 1 \dots m \tag{26}$$

$$\sum_{j=1}^{N} \lambda_j * y_{rj} = y_{rj_o} + S_r^+, r = 1 \dots s \tag{27}$$

$$\sum_{j=1}^{N} \lambda_j = 1 \tag{28}$$

$$\lambda_j \geq 0, j = 1 \dots N, S_r^+, S_i^- \geq 0 \tag{29}$$

The potential results of the above problem are:

1. DMUj_o is Pareto-efficient if and only if ${\Theta_0}^* = 1$ and $S_r^+, S_i^- = 0$ for all inputs and outputs
2. If the value of one of the slack variables is positive at the optimal solution, the corresponding input (or output) of DMUj_o can be further improved (Cooper, Seiford, & Zhu, 2011).

Finally, it should be stated that the *pure technical efficiency* cannot be less than its technical input efficiency as calculated by equations (7)-(10) (Thanassoulis, 2001).

The equivalent value-based DEA model to the one represented by equations (25)-(29) is:

$$max \sum_{r=1}^{s} u_r * y_{rj_0} + w \tag{30}$$

Subject to Constraints

$$\sum_{i=1}^{m} v_i * x_{ij_0} = 1 \tag{31}$$

$$\sum_{r=1}^{s} u_r * y_{rj} - \sum_{i=1}^{m} v_i * x_{ij} - w \ \leq 0, j = 1 \dots j_0 \dots N \tag{32}$$

$$u_r \geq \varepsilon, r = 1 \dots s \tag{33}$$

$$v_i \geq \varepsilon, i = 1 \dots m \text{ and w free} \tag{34}$$

For the value-based model, DMUj_o is considered efficient if $\sum_{r=1}^{s} u_r^* * y_{rj_0} + \text{w} = 1$

Similarly, the DEA-BCC model under output orientation is solved by:

$$max\ h_0 + e\left(\sum_{i=1}^{m} S_i^- + \sum_{r=1}^{s} S_r^+\right) \tag{35}$$

Subject to Constraints

$$\sum_{j=1}^{N} \lambda_j * x_{ij} = x_{ij_o} - S_i^-, i = 1 \dots m \tag{37}$$

$$\sum_{j=1}^{N} \lambda_j * y_{rj} = h_0 * y_{rj_o} + S_r^+, r = 1 \dots s \tag{38}$$

$$\sum_{j=1}^{N} \lambda_j = 1 \tag{39}$$

$$\lambda_j \geq 0, j = 1 \dots N, S_r^+, S_i^- \geq 0 \tag{40}$$

The potential results of the above problem are:

1. DMUj_o is Pareto-efficient if and only if ${h_0}^* = 1$ and $S_r^+, S_i^- = 0$ for all inputs and outputs
2. DMUj_o has technical output efficiency $\frac{1}{h_{j_0}^*}$ (Thanassoulis, 2001).

It should be stated that the *pure technical efficiency* $\frac{1}{h_{j_0}^*}$ cannot be less than its technical efficiency as calculated by equations (16)-(19) (Thanassoulis, 2001).

Finally, the equivalent value-based DEA-BCC model is:

$$min \sum_{r=1}^{s} u_r * y_{rj_0} + w \tag{41}$$

Subject to Constraints

$$\sum_{i=1}^{m} v_i * x_{ij_0} = 1 \tag{42}$$

$$\sum_{r=1}^{s} u_r * y_{rj} - \sum_{i=1}^{m} v_i * x_{ij} + w \leq 0, j = 1 \dots j_0 \dots N \tag{43}$$

$$u_r \geq \varepsilon, r = 1 \dots s \tag{44}$$

$$v_i \geq \varepsilon, i = 1 \dots m \text{ and w free} \tag{45}$$

For the value-based model, DMUj_o is considered efficient if $\sum_{r=1}^{s} u_r^* * y_{rj_0} - \mathrm{w} = 1$ (Cooper, Seiford, & Zhu, 2011; Thanassoulis, 2001).

In general, the efficiency as calculated under CRS is noted as *technical efficiency*. It can be analysed into two dimensions: *pure technical efficiency* (as calculated by the DEA-BCC models) and *scale efficiency*. Thus, if:

$$Technical\ Efficiency = TE$$

$$Pure\ Technical\ Efficiency = PTE$$

$$Scale\ Efficiency = SE$$

The different efficiencies are related by the equation:

$$TE = PTE * SE \tag{46}$$

3.2 Two-stage DEA models[5]

One of the advantages of Data Envelopment Analysis is that there is no need for knowledge on how inputs are transformed to outputs. As a result, a performance analysis can occur without having to resort to complex functions that represent the relationship between inputs and outputs.

However, this advantage can be seen also as a limitation, since the method is considered a “black box”; the analyst has knowledge only of the inputs and outputs but not what happens among them. Färe and Grashhopf (2009) studied extensively the “black box” property of DEA and suggested the concept of two-stage models to mitigate its effects. These models can be considered a special category of network-DEA models (Seiford & Zhu, 1999; Wang, Gopal, & Zionts, 1997). Figure 4 illustrates the structure of a typical two-stage DEA model.

[5] Part of this section appeared on: Tsaples, G., Papathanasiou, J., & Georgiou, A. C. (2022). An Exploratory DEA and Machine Learning Framework for the Evaluation and Analysis of Sustainability Composite Indicators in the EU. Mathematics, 10(13), 2277

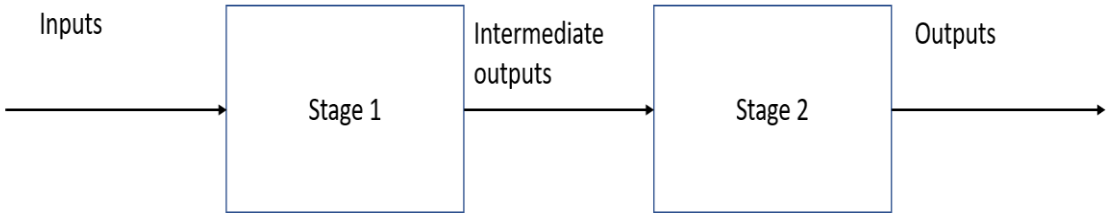


Figure 4 Structure of two-stage DEA

Each stage is regarded as a separate decision center with the overall process managed by one decision maker. The goal is the simultaneous improvement of efficiency both of the individual stages (efficiency of the Stage 1 named E1 and efficiency of Stage 2 named E2) and the overall efficiency (named E0) (Ross A. D., 2000). Ross and Drage (2002) analyzed what internal and external improvement of efficiency means and proposed that at the internal level, each decision center aims at succeeding the best allocation of resources while accounting for its individual preferences and needs. The optimal allocation refers to higher efficiency at the individual stage. On the other hand, at the external level the decision center aims at achieving a bigger market share, which reflects the contribution of that individual stage to the overall process. Hence, two-stage DEA models increase the dimensions of performance measurement compared to the classic one-stage DEA models, thus offering greater insights, which in turn reduces the characterization of DEA as a "black-box" method.

Halkos, Tzeremes and Kourtzidis (2014) studied in detail the two-stage DEA models and classified them into four categories:

- Independent two-stage models. In this category, the classic DEA variation is applied to each stage without accounting for interactions between the stages (Cook, Liang, & Zhu, 2010)
- Connected two-stage models, where the interaction between the stages are accounted for
- Relational two-stage DEA which assumes a formal mathematical relationship between the overall and individual efficiencies
- Finally, there are the two-stage models that are based on game theory approaches.

Chen et al. (2009);(2012) and Halkos, Tzeremes and Kourtzidis (2015) proposed a model of efficiency decomposition, whose mathematical formulation is:

$$E_0 = \xi_1 \frac{\sum_{d=1}^{D} w_d z_{d0}}{\sum_{i=1}^{m} v_i x_{i0}} + \xi_2 \frac{\sum_{r=1}^{s} u_r y_{r0}}{\sum_{d=1}^{D} w_d z_{d0}} \tag{47}$$

$$E_0 = max \sum_{d=1}^{D} \mu_d z_{d0} + \sum_{r=1}^{s} \gamma_r y_{r0} \tag{48}$$

Subject to Constraints

$$\sum_{i=1}^{m} \omega_i x_{i0} + \sum_{d=1}^{D} \mu_d z_{d0} = 1 \tag{49}$$

$$\sum_{d=1}^{D} \mu_d z_{dj} - \sum_{i=1}^{m} \omega_i x_{ij} \leq 0 \tag{50}$$

$$\sum_{r=1}^{s} \gamma_r y_{rj} - \sum_{d=1}^{D} \mu_d z_{dj} \leq 0 \tag{51}$$

$$\gamma_r, \omega_i, \mu_d \geq 0 \tag{52}$$

The above model can produce results that may not be optimal. To solve the issue Kao and Hwang (2008) proposed the maximization of one of f the E_0^1, E_0^2 while maintaining the overall efficiency at E_0. For example, maximizing the individual efficiency E_0^2:

$$E_0^2 = max \sum_{r=1}^{s} \gamma_r y_{r0} \tag{53}$$

Subject to Constraints

$$\sum_{d=1}^{D} \mu_d z_{d0} = 1 \tag{54}$$

$$\sum_{r=1}^{s} \gamma_r y_{r0} - E_0 \sum_{i=1}^{m} \omega_i x_{i0} \leq 0 \tag{55}$$

$$\sum_{r=1}^{s} \gamma_r y_{rj} - \sum_{i=1}^{m} \omega_i x_{ij} \leq 0 \tag{56}$$

$$\sum_{d=1}^{D} \mu_d z_{dj} - \sum_{i=1}^{m} \omega_i x_{ij} \leq 0 \tag{57}$$

$$\sum_{r=1}^{s} \gamma_r y_{rj} - \sum_{d=1}^{D} \mu_d z_{dj} \leq 0 \tag{58}$$

$$\gamma_r, \omega_i, \mu_d \geq 0 \tag{59}$$

Consequently, the other efficiency will be calculated by:

$$E_0 = \xi_1 E_o^1 + \xi_2 E_o^2 \tag{60}$$

Where ξ_1 *and* ξ_2 are the weights assigned to the individual efficiency of each stage.

Liang et al. (2006; 2008), Kao and Hwang (2008), Chen et al. (2009; 2012) and Cook et al. (2010) provided alternative integrated models that attempted to simultaneously optimize the efficiencies of the two stages. For indicative extensive reviews of two-stage and network DEA models, the reader is referred to the works by Castelli et al. (2010), Halkos et al. (2014), Kao (2014) and Despotis et al. (2016).

The two-stage model that is used as a basis in the context of the current thesis is the one by Chen et al. (2012) and is presented below.

$$\max E_0 = \sum_{r=1}^{s} \gamma_r y_{r0} \tag{61}$$

Subject to Constraints

$$\sum_{d=1}^{D} \mu_d z_{dj} - \sum_{i=1}^{m} \omega_i x_{ij} \leq 0, j = 1, \dots, N \tag{62}$$

$$\sum_{r=1}^{s} \gamma_r y_{rj} - \sum_{d=1}^{D} \mu_d z_{dj} \leq 0, j = 1, \dots, N \tag{63}$$

$$\sum_{i=1}^{m} \omega_i x_{i0} = 1 \tag{64}$$

$$\omega_i \geq 0, i = 1, \dots, m \tag{65}$$

$$\mu_d \geq 0, d = 1, \dots, D \tag{66}$$

$$\gamma_r \geq 0, r = 1, \dots, s \tag{67}$$

The optimal value calculated by the model described in the equations (61)-(67) represents the overall efficiency of DMU_0, a number between 0 (inefficient) and 1 (efficient). Furthermore, the optimal values of the weights $\omega_i, \mu_d \; and \; \gamma_r$ are used for the decomposition of the overall efficiency and the calculation of the efficiencies of the individual stages according to the equations:

$$E_0^1 = \frac{\sum_{d=1}^{D} \mu_d^* z_{d0}}{\sum_{i=1}^{m} \omega_i^* x_{i0}} \tag{68}$$

$$E_0^2 = \frac{\sum_{r=1}^{s} \gamma_r^* y_{r0}}{\sum_{d=1}^{D} \mu_d^* z_{d0}} \tag{69}$$

The values of $\omega_i^*, \; \mu_d^* \; and \; \gamma_r^*$ are the optimal values of the weights of the inputs, intermediate outputs, and outputs respectively that were calculated by the equations (61)-(67).

3.3 Weight Flexibility in DEA[6]

Data Envelopment Analysis (either in its classic or two-stage form) does not come without limitations. In its traditional forms the efficiency of Decision Making Units is calculated with weights that are most favorable to themselves; i.e. each DMU is evaluated under the most favorable weighting scheme with the purpose of maximizing its own efficiency (Sun, Wu, & Guo, 2013). As a result, the weights that are chosen for one DMU may be completely different from those selected for another (Pedraja-Chaparro, Salinas-Jimenez, & Smith, 1997). In other words, one DMU might place more importance on one of the inputs that it uses while another might do so on another.

[6] This section appeared on: Tsaples, G., Papathanasiou, J., & Georgiou, A. C. (2022). An Exploratory DEA and Machine Learning Framework for the Evaluation and Analysis of Sustainability Composite Indicators in the EU. Mathematics, 10(13), 2277

Hence, this weighting flexibility has been met with some criticism since a DMU's efficiency assessment might be dominated by secondary activities, thus concealing inefficiencies of important factors in the Production Possibility Set. Furthermore, by being assessed on different weighting schemes, DMUs cannot be compared at the same basis (Wang, Luo, & Lan, 2011).

To solve these limitations several approaches have been proposed. Sexton et al. (1986) introduced the notion of cross efficiency, where each DMU is assessed in a peer evaluation mode with the optimal weights of all the other DMUs instead of its own (Örkcü, Özsoy, Örkcü, & Bal, 2019). However, Kao and Hung (2005) argued that cross efficiency limits the information contained in the weights of the DMU's inputs and outputs.

In another approach several authors suggested modifications to the classic DEA models in order to obtain a Common Set of Weights on which the efficiency of all DMUs is calculated (Ganley & Cubbin, 1992; Roll, Cook, & Golany, 1991; Roll & Golany, 1993). One stream of research in this approach is the use of virtual DMUs in the set that act as reference points for the real DMUs. Such virtual DMUs are hypothetical units that can act as a reference for the existing ones either by assuming that they use inputs in the most economical ways to produce the maximum level of outputs (ideal DMUs), thus incentivizing the existing DMUs to imitate them or by assuming that they use inputs in the most expensive way to produce the minimum level of outputs (anti-ideal DMUs), thus incentivizing the existing DMUs to deviate from the behavior as much as possible.

Jahanshahloo et al. (2010) used an ideal line to determine a common set of weights in order to calculate a new efficiency and rank efficient DMUs. Lotfi et al. (2011) introduced a virtual DMU based on the aggregate units to obtain a full ranking of DMUs. Barzegarinegad et al. (2014) proposed a variation of DEA with the purpose of fully ranking DMUs based on ideal and anti-ideal points in the production possibility set. Similarly, Khalili-Damghani and Fadaei (2018) used both an ideal and an anti-ideal DMU to increase the discrimination power of DEA. Finally, Azadi et al. (2020) proposed models to calculate the efficiency of DMUs based on the distances to two virtual DMUs, considering both the pessimistic and optimistic approach of DEA.

Moreover, there are approaches to increase the discriminatory power of DEA by combining the classic model with another method. For example, Kritikos (2017) combined DEA with TOPSIS in order to fully rank DMUs, Simuany-Stern and Friedman (1998) used the classic DEA models with non linear discriminant analysis, while Yang et al. (2018) proposed a new approach inspired by the Z score (the distance of an observation from the mean expressed in standard deviation units) combined with DEA. Another common approach is the combination of DEA with AHP, for example in Thanassoulis et al. (2017) the authors combine the two methods to evaluate higher education teaching performance. Finally, weight restrictions as part of the DEA model have been proposed as a solution to their flexibility in the literature. Weight restrictions can be seen as value judgments (Podinovski, 2016) that not only limit their flexibility, but also act positively on the discriminatory power of the model (Thanassoulis, et al., 2008). Examples of weight restriction methods include the work by Alirezaee and Afsharian (2010) who used the trade-offs approach with an expanded Malmquist index to increase the discrimination of DMUs or the Cone-Ratio models (Angulo-Meza & Lins, 2002) and Assurance Regions (Allen, et al., 1997).

However, Jain et al. (2015) point out that these weight restriction methods might also have some limitations such as increased subjectivity since the models incorporate a priori information, lack of guarantee for feasibility or the assumption of a single policy maker.

Moreover, all the approaches that were mentioned above are solutions for classic, one-stage Data Envelopment Analysis. Similarly, in two-stage DEA the weight distribution for the calculation of both individual stages and overall efficiency are still free, thus the discriminatory power of two-stage DEA can be further improved. Following the works of researchers for typical DEA models, the research followed different streams on how to better handle the weight distribution of two-stage DEA models.

Mahdiloo et al. (2016) argued that little attention has been paid to the weight distribution and weak discriminatory power of network models. In their paper, they propose a multi-criteria DEA model, which is tested by assessing the sustainable design performances of car products. Gharakhami et al. (2018) proposed a DEA variation that is based on goal programming, while Mavi et al. (2019) use a similar approach to analyze the joint effects of eco-efficiency and eco-innovation. Finally, Kiaei and Matin (2020) suggest a method based on separation vector to change a Multiple Objective problem into a single objective linear programming problem in two-stage DEA. For a complete overview of the ways that researchers have attempted to address the weight distribution in DEA, Contreras (2020) performed a comprehensive literature review. Consequently, for two-stage DEA models more efforts are necessary to address the limitations of weight distribution especially when it comes to measuring the efficiency of complex processes like sustainable development; in such cases, incorrect weight distribution might conceal inefficiencies of very important factors. For example, GHG emissions might not get equal importance compared to increased economic activity, despite the fact that they are crucial for the measurement of sustainability, thus providing an erroneous picture of sustainable development for the country/DMU under study. The present thesis contributes to that aim.

3.3.1 Proposed model

To overcome this limitation the following model is proposed (the model is based on the works by Mahdiloo et al. (2016) and Sun et al. (2013):

Assume a two-stage process for N DMUs. Each $DMU_j (j = 1 \dots N)$ in the first stage uses m inputs, the level of which is notated as $x_{ij}, (i = 1,2 \dots m)$, to produce D intermediate outputs, the level of which is notated as $z_{dj}, (d = 1,2 \dots D)$. The intermediate outputs are used as inputs in the second stage to produce s outputs, notated as $y_{rj}, (r = 1,2 \dots s)$.

$$min\, d_0 + n_0 + d_0' + n_0' \quad (70)$$

Subject to Constraints

$$\sum_{i=1}^{m} \omega_i x_{ij} - \sum_{d=1}^{D} \mu_d z_{dj} \geq 0, j = 1, \dots, N \quad (71)$$

$$\sum_{i=1}^{m} \omega_i x_{i0} - \sum_{d=1}^{D} \mu_d z_{d0} - d_0 + n_0 = 0 \quad (72)$$

$$\sum_{d=1}^{D} \mu_d z_{dj} - \sum_{r=1}^{s} \gamma_r y_{rj} \geq 0, j = 1, \dots, N \quad (73)$$

$$\sum_{d=1}^{D} \mu_d z_{d0} - \sum_{r=1}^{s} \gamma_r y_{r0} - d'_0 + n'_0 = 0 \tag{74}$$

$$\sum_{d=1}^{D} \mu_d z_{d0} = 1 \tag{75}$$

$$\omega_i \geq 0, i = 1, \ldots, m \tag{76}$$

$$\mu_d \geq 0, d = 1, \ldots, D \tag{77}$$

$$\gamma_r \geq 0, r = 1, \ldots, s \tag{78}$$

$$d_0, n_0, d'_0, n'_0 \geq 0 \tag{79}$$

The main property of the model is the introduction of the variables d_0, d'_0, n_0, n'_0. The variable d_0 represents the positive deviation and variable n_0 represents the negative deviation of the efficiency of stage 1 from reaching the maximum value. Variable d'_0 represents the positive deviation while the variable n'_0 represents the negative deviation of the efficiency of stage 2 from reaching its maximum value.

Consequently, since in two-stage DEA it is assumed that the ratio of the sum of the weighed intermediate outputs to the sum of the weighted inputs (stage 1) and the sum of weighted outputs to the sum of weighted intermediate outputs should be smaller or equal to 1 (constraints (71), (73)), the introduction of the variables d_0, d'_0, n_0, n'_0 occurs in constraints (71) and (73) in order to make them equal to 0 in the following manner:

$$\frac{\sum_{d=1}^{D} \mu_d z_{dj}}{\sum_{i=1}^{m} \omega_i x_{ij}} \leq 1 \;\Rightarrow \sum_{i=1}^{m} \omega_i x_{ij} - \sum_{d=1}^{D} \mu_d z_{dj} \geq 0$$
$$\Rightarrow \sum_{i=1}^{m} \omega_i x_{ij} - \sum_{d=1}^{D} \mu_d z_{dj} - d_0 + n_0 = 0 \tag{80}$$

$$\frac{\sum_{r=1}^{S} \gamma_r y_{r0}}{\sum_{d=1}^{D} \mu_d z_{dj}} \leq 1 \ \Rightarrow \sum_{d=1}^{D} \mu_d z_{dj} - \sum_{r=1}^{S} \gamma_r y_{r0} \geq 0$$
$$\Rightarrow \sum_{d=1}^{D} \mu_d z_{dj} - \sum_{r=1}^{S} \gamma_r y_{r0} - d_0' + n_0' = 0 \qquad (81)$$

Thus, the model represented by equations (70)-(79) attempts to find the best possible values for ω_i, μ_d and γ_r by minimizing the deviations of both the first and second stage of the DEA model. By minimizing simultaneously the deviations of each stage (i.e. $d_0 + n_0$ for stage 1 and $d_0' + n_0'$ for stage 2), the efficiencies of both stages are maximized at the same time and no priority is given into which stage should take precedence. Moreover, since we wish to attain the specific value of 1 for both stage efficiencies, both groups of deviational variables are included in the objective function to be minimized. The introduction of positive and negative deviations in such models is typical in the Multi-Criteria Decision Analysis literature and although the model might work with fewer deviations, in the context of the current thesis the author believes that the inclusion of both positive and negative deviations, although it adds a level of complexity to the model, it also offers a layer of rigor, which might increase the computational time but at the same time provides a more nuanced distribution of the weights without altering the core elements of the DEA methodology. In addition, the significance of each deviational variable in the objective function can be further fine-tuned, either within a level or between levels with appropriate weights thus offering a trade-off vehicle of trading between the efficiencies.

Equations (71) and (73) ensure that the efficiency scores of the first and second stage are smaller than 1. Equations (72) and (74) indicate that the efficiency score for the first and second stage respectively when their respective deviations are added should be 1 and to achieve that it is assumed that $\sum_{d=1}^{D} \mu_d z_{d0} = 1$ (equation (75)).

Finally, the efficiencies of the individual stages are calculated according to equations (68) and (69) while the overall efficiency equals the average of the individual efficiencies similar to the work by Mahdiloo et al. (2016). The author of the current thesis recognizes that the proposed alternative, two-stage DEA with a different optimization metric described by equations (70)-(79) can be one of many approaches that use deviational variables; in order to examine whether the proposed approach was valid, three lemmas and one theorem were proved about the model, which are presented below. Finally, the proposed alternative approach to two-stage DEA might not offer a unique solution and one approach to mitigate the potential effects of that fact will be presented in a next section of the current thesis.

The following lemmas and theorem about the model described by the equations (70)-(79) are proved based on the work by Sun et al. (2013) and Khalili-Damghani and Fadaei (2018).

Lemma 1. The constraints of the model described by the equations (70)-(79) form a non-empty, convex set.

Proof. The constraints (70)-(79) constitute a non-empty set named S. If $(\omega_i^*,\ \mu_d^*\ and\ \gamma_r^*)\ and\ (\omega_i^{*\prime},\ \mu_d^{*\prime}\ and\ \gamma_r^{*\prime})\ \in S$ then $\forall \alpha \in [0,1]$: $\left(\alpha\omega_i^* + (1-\alpha)\omega_i^{*'}, \alpha\mu_d^* + (1-\alpha)\mu_d^{*'}, \alpha\gamma_r^* + (1-\alpha)\gamma_r^{*'}\right) \in S$. Hence, S is a convex set. ■

Lemma 2. The objective function (70) is convex in the defined domain.

Proof 2. The Hessian matrix of the objective function is:

$$\begin{pmatrix} \frac{\partial_{obj}^2}{\partial^2\omega_1} & \frac{\partial_{obj}^2}{\partial\omega_1\partial\omega_2} & \cdots & \frac{\partial_{obj}^2}{\partial\omega_1\partial\gamma_s} \\ \vdots & \frac{\partial_{obj}^2}{\partial^2\omega_2} & \ddots & \vdots \\ \frac{\partial_{obj}^2}{\partial\gamma_s\partial\omega_1} & \frac{\partial_{obj}^2}{\partial\gamma_s\partial\omega_2} & \cdots & \frac{\partial_{obj}^2}{\partial^2\gamma_s} \end{pmatrix}$$

The matrix is always zero, thus, objective function (11) is a strictly convex function. ■

Lemma 3. The model described by the equations (70)-(79) is always feasible.

Proof 3. Suppose that an arbitrary solution of the model of the form:

$$\omega_i = \frac{1}{mx_{i0}} \tag{82}$$

$$\mu_d = \frac{1}{Dz_{d0}} \tag{83}$$

$$\gamma_r = \frac{1}{sy_{r0}} \tag{84}$$

$$d_0 = \left(\frac{1}{D}\sum_{d=1}^{D}\frac{z_{dj}}{z_{d0}} - \frac{1}{m}\sum_{i=1}^{m}\frac{x_{ij}}{x_{i0}}\right), \forall \mathrm{j} \tag{85}$$

$$d_0' = \left(\frac{1}{s}\sum_{r=1}^{s}\frac{y_{rj}}{y_{r0}} - \frac{1}{D}\sum_{d=1}^{D}\frac{z_{dj}}{z_{d0}}\right), \forall \mathrm{j} \tag{86}$$

$$n_0 = \left(\frac{1}{D}\sum_{d=1}^{D}\frac{z_{dj}}{z_{d0}} - \frac{1}{s}\sum_{r=1}^{s}\frac{y_{rj}}{y_{r0}}\right), \forall \mathrm{j} \tag{87}$$

$$n_0' = \left(\frac{1}{D}\sum_{d=1}^{D}\frac{z_{dj}}{z_{d0}} - \frac{1}{s}\sum_{r=1}^{s}\frac{y_{rj}}{y_{r0}}\right), \forall \mathrm{j} \tag{88}$$

Substituting equations (82), (83) in equation (71):

$$\sum_{i=1}^{m}\frac{1}{mx_{i0}}x_{ij} - \sum_{d=1}^{D}\frac{1}{Dz_{d0}}z_{dj} \geq 0, j = 1, \dots, n$$

Substituting equations (82), (83), (85) and (87) in equation (72):

$$\sum_{i=1}^{m}\frac{1}{mx_{i0}}x_{i0} - \sum_{d=1}^{D}\frac{1}{Dz_{d0}}z_{d0} - \left(\frac{1}{D}\sum_{d=1}^{D}\frac{z_{dj}}{z_{d0}} - \frac{1}{m}\sum_{i=1}^{m}\frac{x_{ij}}{x_{i0}}\right) + \left(\frac{1}{m}\sum_{i=1}^{m}\frac{x_{ij}}{x_{i0}} - \frac{1}{D}\sum_{d=1}^{D}\frac{z_{dj}}{z_{d0}}\right) = 0$$

Substituting equations (83), (84) in equation (73):

$$\sum_{d=1}^{D}\frac{1}{Dz_{d0}}z_{dj} - \sum_{r=1}^{s}\frac{1}{sy_{r0}}y_{rj} \geq 0, j = 1, \dots, n$$

Substituting equations (83), (84), (86) and (88) in equation (74):

$$\sum_{d=1}^{D}\frac{1}{Dz_{d0}}z_{d0}-\sum_{r=1}^{s}\frac{1}{sy_{r0}}y_{r0}-\left(\frac{1}{s}\sum_{r=1}^{s}\frac{y_{rj}}{y_{r0}}-\frac{1}{D}\sum_{d=1}^{D}\frac{z_{dj}}{z_{d0}}\right)+\left(\frac{1}{D}\sum_{d=1}^{D}\frac{z_{dj}}{z_{d0}}-\frac{1}{s}\sum_{r=1}^{s}\frac{y_{rj}}{y_{r0}}\right)$$
$$=0$$

Finally, substituting equation (82) in equation (75):

$$\sum_{d=1}^{D}\frac{1}{Dz_{d0}}z_{d0}=1$$

All constraints are satisfied with the arbitrary solution, thus the model (70)-(79) is feasible.■

Theorem 1. The model described by the equations (70)-(79) has an optimal solution.

Proof 1. S is a non-empty, convex set (Lemma 1), the objective function is strictly convex (Lemma 2) and the model has a feasible solution (Lemma 3). Consequently, the solution obtained by the model is optimal. ■

3.3.2 Case Study 1

To test the proposed variation, the environmental performance of European countries is calculated.

The environmental performance can provide concise information to decision-makers especially when dealing with interactions between energy and the environment (Esty, et al., 2006). Furthermore, environmental performance has been and continues to be used as a measure of sustainable development. Finally, an environmental performance index can be an effective communication tool to convey complex notions to the general public and non-experts.

Table 4 below illustrates some of the papers that were found in scientific databases when searching for environmental performance and DEA.

Table 4 Review of the literature on DEA and environmental performance

Paper	**Desirable outputs**	**Undesirable outputs**	**Production inputs**
(Zhou, Poh, & Ang, 2007)	GDP	CO_2 emissions, SO_2 emissions, NOx emissions	Labor, primary energy consumption

(Zhou, Ang, & Poh, 2008)	GDP	CO_2 emissions	energy consumption
(Bian & Yang, 2010)	GDP	SO_2 emissions, COD, Nitrogen emissions	labor, capital, energy, water
(Guo, Zhu, Fan, & Xie, 2011)	GDP	CO_2 emissions	labor, capital, energy
(Li & Hu, 2012)	GDP	CO_2 emissions, SO_2 emissions	labor, capital, energy
(Wang, Wei, & Zhang, 2012)	GDP	CO_2 emissions, SO_2 emissions	labor, capital, coal, crude oil, natural gas
(Wu, Fan, Zhou, & Zhou, 2012)	Industrial added value	CO_2 emissions	labor, capital, energy
(Wei, Ni, & Du, 2012)	GDP	-	labor, capital, energy
(Wang, Zhao, Zhou, & Zhou, 2013)	GDP		labor, capital, energy
(Wang & Wei, 2014)	Industrial added value	Waste water, waste gas, solid waste	labor, capital, energy
(Wang, Feng, & Zhang, 2014)	Industrial added value	CO_2 emissions, SO_2 emissions	labor, capital, energy
(Wu, An, Yao, & Wang, 2014)	Industrial added value	NO_2 emissions	capital, electricity
(Zhou, Sun, & Zhou, 2014)	GDP	CO_2 emissions	labor, capital, energy
(Li & Lin, 2015)	GDP	CO_2 emissions, SO_2 emissions	labor, capital, coal, electricity

(Yao, Zhou, Zhang, & Li, 2015)	GDP	CO_2 emissions	labor, capital, energy
(Du, Matisoff, Wang, & Liu, 2016)	GDP	CO_2 emissions	labor, capital, energy
(Long, Wang, & Chen, 2016)	GDP	Solid waste	labor, capital, coal
(Sueyoshi & Yuan, 2016)	GDP, primary secondary and tertiary industry	PM10, SO_2 emissions, NO_2 emissions	coal, oil, gas, electricity, energy investment
(Wu, Yin, Sun, Chu, & Liang, 2016)	Industrial added value	Waste water, solid waste	labor, capital, coal
(Zhang, Hao, & Song, 2016)	Industrial added value	CO_2 emissions	labor, capital, energy
(Chen & Jia, 2017)	GDP	SO_2 emissions, solid waste	labor, capital, energy
(Chen, Wang, Lai, & Feng, 2017)	GDP	CO_2 emissions, SO_2 emissions, COD	labor, capital, energy
(Feng, Zhang, & Huang, 2017)	GDP	CO_2 emissions	labor, capital, energy
(Li, Zhang, Zhou, & Yao, 2017)	GDP	CO_2 emissions	labor, capital, energy
(Li, Peng, Wang, & Yao, 2017)	GDP	CO_2 emissions, SO_2 emissions	labor, capital, energy
(Sueyoshi & Yuan, 2017)	Gross regional product (GRP)	CO_2 emissions, SO_2 emissions, soot, waste water, Chemical Oxygen Demand, NO	labor, capital, energy
(Sueyoshi, Yuan, & Goto, 2017)	Gross regional product (GRP)	CO_2 emissions, SO_2 emissions, Smoke and dust, Waste	labor, capital, energy

		water, chemical oxygen demand, NO	
(Sueyoshi, Yuan, Li, & Wang, 2017)	GDP	CO_2 emissions	labor, capital, energy
(Zhu, Wu, Li, & Xiong, 2017)	GDP		coal, oil, gas, fixed investment
(Song, Peng, Wang, & Zhao, 2018)	GDP	SO_2 emissions, waste water, solid waste	labor, capital, energy
(Du, Chen, & Huang, 2018)	GDP	CO_2 emissions, SO_2 emissions, solid waste, industrial dust	labor, capital, energy
(Zhang, Li, & Gao, 2018)	Gross regional product (GRP)	CO_2 emissions, SO_2 emissions, soot and dust, waste water, COD, Ammonia nitrogen	labor, capital, energy
(Biresselioglu, Demir, & Turan, 2018)	share of renewable energy in gross final energy consumption	energy consumption, GHG generations	mathematical programming scores and scores from the energy trilemma
(Ervural, Zaim, & Delen, 2018)	Gross energy generation from renewable energy, number of consumers, total exports, GDP per capita, HDI, Total energy production, Population, area		Total renewable energy potential, network length, total installed power of renewable energy, transformer capacity

The works that are displayed on Table 4 have used the classic DEA model to calculate the environmental performance of regions. Thus, another contribution of the current thesis is the calculation of the environmental performance of European countries with the two-stage DEA model described by the equations (70)-(79). For the determination of the inputs, intermediate outputs and outputs, apart from the works summarized on table 1, the efforts of Tsaples and Papathanasiou (2021) were also used. The authors study the concept of sustainability and how DEA has tackled it and consider environmental performance as one of the dimensions of sustainability. Furthermore, Tsaples et al. (2019) and Tsaples and Papathanasiou (2020a) use different combinations of inputs, intermediate outputs and outputs to calculate the environmental performance of EU countries.

For the case study, the following measures are used:

- Inputs: Population, Gross electricity production [Thousand tonnes of oil equivalent (TOE)]
- Intermediate measures: Final energy consumption [Terajoule]
- Outputs: Terrestrial protected area (km2), Share of renewable energy in gross final energy consumption, Greenhouse gas emissions (in CO2 equivalent)

The data source is Eurostat[7] for the year 2018, which was the latest common year for which data was available for all countries. Moreover, the output Greenhouse gas emissions is considered undesirable which contradicts the nature of outputs in Data Envelopment Analysis, which should always be maximized. Thus, the undesirable output of the case study is rendered into a "desirable" one with a linear monotonic transformation (Seiford & Zhu, 1999).

Table 5 below illustrates the results obtained from the proposed two-stage DEA variation and the corresponding results obtained from the model of Chen et al. (2012).

[7] https://ec.europa.eu/eurostat (accessed January 2021)

Table 5 Results from the proposed variation and the ones obtained by the models of Chen et al. (2009)

	Chen et al. (2012)			**Proposed, alternative metric two-stage DEA**				**Chen et al. (2012)**			**Proposed, alternative metric two-stage DEA**		
Country	E0	E1	E2	E0	E1	E2	**Country**	E0	E1	E2	E0	E1	E2
Belgium	0.032 (26)	0.733	0.044	0.388 (21)	0.733	0.044	**Luxemburg**	0.255 (10)	1	0.255	0.628 (9)	1	0.255
Bulgaria	0.296 (9)	0.296	1	0.648 (7)	0.296	1	**Hungary**	0.160 (16)	0.715	0.224	0.469 (14)	0.715	0.224
Czech Republic	0.077 (22)	0.664	0.116	0.390 (20)	0.664	0.116	**Malta**	0.223 (11)	0.223	1	0.612 (10)	0.223	1
Denmark	0.145 (18)	0.818	0.178	0.498 (12)	0.818	0.178	**Netherlands**	0.023 (27)	0.570	0.041	0.306 (28)	0.570	0.041
Germany	0.038 (25)	0.657	0.058	0.358 (23)	0.657	0.058	**Austria**	0.153 (17)	0.756	0.202	0.479 (13)	0.756	0.202
Estonia	0.680 (2)	0.680	1	0.840 (2)	0.680	1	**Poland**	0.093 (21)	0.512	0.183	0.347 (24)	0.512	0.183
Ireland	0.132 (20)	0.599	0.220	0.410 (19)	0.599	0.221	**Portugal**	0.136 (19)	0.272	0.502	0.387 (22)	0.272	0.502

Greece	0.188 (13)	0.348	0.541	0.444 (15)	0.348	0.541	**Romania**	0.191 (12)	0.469	0.408	0.438 (16)	0.469	0.408
Spain	0.171 (14)	0.320	0.534	0.427 (17)	0.320	0.534	**Slovenia**	0.405 (6)	0.506	0.800	0.653 (6)	0.506	0.800
France	0.061 (23)	0.583	0.105	0.344 (25)	0.583	0.105	**Slovakia**	0.169 (15)	0.378	0.449	0.414 (18)	0.378	0.449
Croatia	0.457 (5)	0.659	0.693	0.676 (5)	0.659	0.693	**Finland**	0.495 (4)	1	0.495	0.748 (4)	1	0.495
Italy	0.056 (24)	0.538	0.104	0.321 (26)	0.538	0.104	**Sweden**	0.380 (7)	0.753	0.505	0.629 (8)	0.753	0.505
Cyprus	0.316 (8)	0.398	0.793	0.596 (11)	0.398	0.793	**United Kingdom**	0.020 (28)	0.604	0.033	0.318 (27)	0.604	0.033
Latvia	0.697 (1)	0.714	0.977	0.845 (1)	0.714	0.977							
Lithuania	0.589 (3)	1	0.589	0.794 (3)	1	0.589							

The column named as E0 indicates the overall environmental performance of the country, E1 the performance of the first stage and E2 the performance of the second stage. The first aspect to observe is that the results of the individual stages for the proposed alternative are almost similar to the ones calculated with the model of Chen et al. (2012), nonetheless with notable differences. These differences, it is argued, are the result of the introduction of the deviational variables that apart from restricting the values of the optimal weights, in essence, they induce the optimization of a different metric in the objective function and the following calculation of the overall efficiency by the arithmetic mean of the individual efficiencies. Furthermore, the two models differ in the average overall efficiency, which is larger in the proposed variation and the range of the values of the overall efficiency, where in the proposed model is smaller than that of Chen et al. (2012). Notably, the model of Chen et al. (2012) assumes that the overall efficiency is derived as the product of the divisional efficiency scores whereas in the proposed model the overall efficiency is calculated as the weighted average of the divisional efficiency scores. Therefore, it is true that comparing the overall efficiency scores wouldn't make much of a sense. It is worth mentioning though that the overall efficiency in the proposed model is defined through a compensatory approach whereas the model of Chen et al. (2012) employs a non-compensatory approach. Consequently, it is mathematically expected that the proposed model would certainly be allowed to attain higher (or equal) overall efficiency scores. In Table 5, as both models estimate almost identical divisional efficiency scores, they will also provide almost the same overall efficiency scores under any common definition of the overall efficiency.

The accumulated small differences result in different rankings of the countries according to their overall efficiency. Countries that seemed to perform better under the variation of Chen et al. (2012) like Poland, Portugal and Greece, move down the ranking in the proposed variation of the current thesis. Finally, it is observed that despite the above notable observations, the values of the efficiencies of the individual stages do not differ much and the fact that the countries that are in the efficient frontiers of the first and the second stages respectively remain the same under both variations, increases the confidence in the results. In combination with the results from the paper of Mahdiloo et al. (2016), where the authors observed similar behavior when they compared their model with that of Chen et al. (2012), the robustness of both the results and the models increases.

To get a better understanding of how the alternative variation compares to that of Chen et al. (2012), an additional illustration was performed by adding: to the inputs the Gross fixed capital at current prices (PPS) and the Total Labor force (x1000 persons), to the intermediate measures the Domestic material consumption [Thousand tons] and to the outputs the Total expenditure [Euro per inhabitant]. The results are tabulated in Table 6 below.

Table 6 Another set of results from the two variations with additional parameters

	Chen et al. (2012)			Proposed, alternative metric two-stage DEA				Chen et al. (2012)			Proposed, alternative metric two-stage DEA		
Country	E0	E1	E2	E0	E1	E2	**Country**	E0	E1	E2	E0	E1	E2
Belgium	0.066 (26)	0.746	0.088	0.475 (22)	0.891	0.059	**Luxemburg**	1 (1)	1	1	1 (1)	1	1
Bulgaria	1 (1)	1	1	1 (1)	1	1	**Hungary**	0.321 (17)	0.965	0.333	0.652 (14)	0.986	0.317
Czech Republic	0.117 (23)	0.716	0.164	0.456 (23)	0.795	0.116	**Malta**	0.647 (10)	0.647	1	0.822 (9)	0.645	1
Denmark	0.228 (20)	1	0.228	0.614 (17)	1	0.228	**Netherlands**	0.052 (27)	0.527	0.098	0.364 (28)	0.665	0.062
Germany	0.069 (25)	0.656	0.105	0.410 (26)	0.758	0.062	**Austria**	0.204 (21)	0.765	0.267	0.511 (21)	0.815	0.206
Estonia	0.843 (5)	0.843	1	0.917 (5)	0.907	0.927	**Poland**	0.252 (19)	1	0.252	0.626 (16)	1	0.252
Ireland	0.267 (18)	0.857	0.311	0.588 (20)	0.896	0.279	**Portugal**	0.350 (16)	0.740	0.473	0.605 (19)	0.763	0.446

Greece	0.723 (6)	0.864	0.836	0.847 (6)	0.894	0.801	**Romania**	0.429 (13)	0.985	0.435	0.710 (12)	1	0.420
Spain	0.366 (15)	0.429	0.854	0.641 (15)	0.429	0.854	**Slovenia**	0.656 (9)	0.701	0.934	0.818 (10)	0.701	0.934
France	0.109 (24)	0.551	0.198	0.393 (27)	0.676	0.111	**Slovakia**	0.376 (14)	0.607	0.620	0.613 (18)	0.607	0.620
Croatia	0.935 (4)	0.935	1	0.968 (4)	0.935	1	**Finland**	0.670 (8)	1	0.670	0.835 (8)	1	0.670
Italy	0.130 (22)	0.613	0.212	0.452 (24)	0.795	0.108	**Sweden**	0.497 (12)	0.755	0.658	0.706 (13)	0.755	0.658
Cyprus	0.717 (7)	0.882	0.813	0.847 (6)	0.898	0.797	**United Kingdom**	0.044 (28)	0.634	0.070	0.447 (25)	0.858	0.037
Latvia	0.998 (3)	0.998	0.999	0.999 (3)	0.998	1							
Lithuania	0.632 (11)	1	0.632	0.816 (11)	1	0.632							

The inclusion of more parameters in the two-stage models alters the results which is not unexpected. However, the conclusions from the previous table also hold for the results on Table 6. Nonetheless, the differences in the overall efficiency between the proposed method and that of Chen et al. (2012) are larger.

In conclusion, the introduction of both negative and positive deviational variables distinguishes the proposed model from that of Chen et al. (2012) (and the model by Mahdiloo et al. (2016)) in a qualitative manner: the deviational variables in the first stage push the stage efficiency to increase in detriment of the efficiency of the second stage. At the same time, the deviational variables of the second stage, push the stage efficiency to increase in detriment to the efficiency of the first stage. Hence, a trade-off occurs between the two stage efficiencies that ultimately drives the overall efficiency upwards.

3.3.3 Case Study 2[8]

Agriculture is one of the most important sectors in the economy and it can have effects (negative or positive) in environmental conservation and economic development (Pang, Chen, Zhang, & Li, 2016). Furthermore, agriculture plays an important role in social support since it provides nutrition to an increased global population.

However, current nutrition choices and consequently current production practices are considered unsustainable and one of the main drivers of climate change (Poore & Nemecek, 2018; Pradhan, et al., 2020). Moreover, an increased urbanization and globalization of supply chains means that food demands are met only after transportation over long distances (Kissinger, 2012; Weber & Matthews, 2008).

Nonetheless, the increased transportation of goods (and people) is considered one of the main sources for Greenhouse Gas emissions (Fuglestvedt, Berntsen, Myhre, Rypdal, & Skeie, 2008), which contributes further to climate change. For example, the dairy sector emitted 4% of the total greenhouse gas emissions (Aggestam & Buick, 2017).

[8] The case study appeared in the paper of Tsaples, G., & Papathanasiou, J. (2021). Measuring agricultural sustainability of European countries with a focus on transportation. International Journal of Sustainable Agricultural Management and Informatics, 7(4), 304-320. (CiteScore: 0.8, Q2)

As a result, efficient and cost-effective transportation can be a driver for sustainability (Gao, Erokhin, & Arskiy, 2019); in the opposite case, climate change will hinder sustainable agriculture, which due to globalization will cause disruptions with cascading, global effects. Furthermore, the effects will not be only global and even in the medium-term. An inefficient transportation has a severe impact on the small (and medium) agricultural enterprises (Han, Pervez, Wu, Shen, & Zhang, 2020).

Since the primary objective of farmers is to increase their profits, an increased transportation cost might lead them to employ unsustainable production practices that cause further damage to the environment (Hoang & Alauddin, 2012). Hence, farmers are trapped in a vicious cycle where in order to survive they must be detriment to the environment which further hinders any effort for sustainability that results in more dire consequences for farmers.

In conclusion, sustainable agriculture is a key goal for all. Therefore measuring the performance of the agricultural sector of countries can provide essential information to policy makers in order for them to design appropriate policies that could lead to sustainable development (Picazo-Tadeo, Gómez-Limón, & Reig-Martínez, 2011).

DEA has been used to measure performance in various echelons of the agricultural sector. Dhungana et al. (Dhungana, Nuthall, & Nartea, 2004) used DEA to 76 Nepalese rice farms to reveal significant variations in the levels of inefficiency that were attributed to the manner that the inputs were used; Rebolledo-Leiva et al. (Rebolledo-Leiva, Angulo-Meza, Iriarte, & González-Araya, 2017) combined carbon footprint assessment with DEA to measure the eco-efficiency of organic blueberry orchards, while Atici and Podinovski (2015) analyzed the efficiency of wheat production in various farms in Turkey.

The efficiency measurement was not limited to the farm level. Toma et al. (2015) applied DEA to measure the sustainability of agriculture in regions of Romania, while Li et al. (Li, Jiang, Mu, & Yu, 2018) calculated the relative efficiency of 30 regions in China for several years. Furthermore, Arnade (1994) measured the efficiency of agricultural sectors in 77 countries; Hoang and Rao (2010) evaluated the agricultural efficiency of 29 OECD countries. Moreover, DEA has been used to measure agricultural efficiency in Europe (Bojnec, Fertő, Jámbor, & Tóth, 2014; Kočišová, 2015; Toma, Miglietta, Zurlini, Valente, & Petrosillo, 2017).

However, several gaps were identified in the literature. Fistly, Toma et al. (2017) identified that in general there is a lack of studies to evaluate agricultural efficiency at national level. Secondly, to the best of our knowledge, transportation of agricultural goods has not been considered in the studies associated with agricultural efficiency. Finally, the classic, one-stage DEA models were used before. However, evaluating agricultural efficiency is a complex effort and one-stage DEA models can be considered “black boxes” since no knowledge is required about how inputs are transformed into outputs (Färe & Grosskopf, 2009). Thus, two-stage DEA models might be more suitable, since they can provide a decomposing of the overall efficiency into two different elements and reveal more insights to the policy maker.

The objective of the current case study is to provide a calculation of the sustainability of agriculture of European countries taking also into account the transportation of goods by employing the two-stage Data Envelopment Analysis (DEA) model described by the equations (70)-(79). Thus, other contributions of this thesis are: (1) to the best of our knowledge transportation of agricultural goods has not been considered in previous studies, (2) the employment of two-stage DEA models for the evaluation of agricultural sustainability of countries and the comparison among the different models contributes to the relevant literature to the relevant.

For the objective of the current thesis, both models will be used to calculate the agricultural efficiency of European countries. As inputs, the following measures were used:

- Utilized Agricultural Area
- Total labor force in agriculture
- Transport infrastructure investment and maintenance spending

As intermediates, the following measures were used:

- Output of agricultural industry – Production value at basic price
- Greenhouse gas emissions from agriculture
- Total transported goods

Finally, as outputs the following measures were used:

- Agricultural income
- Gross value added at basic prices

The data were obtained by Eurostat[9] and OECD[10] for the year 2018. The choice of the year depended on the data availability and 2018 was the most recent, common year for which there was availability for the chosen countries.

The results are illustrated on table 7 below.

Table 7 Results of the agricultural sustainability of European countries according to the models by Chen et al. (2009) and the proposed variation

	Chen et al (2012)			**Proposed, alternative metric two-stage DEA**		
Country	**Overall agricultural sustainability**	**Efficiency of stage 1**	**Efficiency of stage 2**	**Overall agricultural sustainability**	**Efficiency of stage 1**	**Efficiency of stage 2**
Belgium	0,37	0,82	0,44	0,64	0,95	0,33
Bulgaria	0,15	0,21	0,73	0,32	0,38	0,26
Czech	0,28	0,37	0,75	0,62	1	0,24
Denmark	0,87	0,87	1	0,93	0,86	1
Germany	0,58	1	0,58	0,79	1	0,58
Estonia	0,15	0,48	0,32	0,41	0,57	0,24
Ireland	0,24	0,37	0,64	0,50	0,37	0,64
Greece	0,18	0,20	0,87	0,41	0,27	0,56
Spain	0,48	0,50	0,95	0,64	0,61	0,66
France	0,77	1	0,77	0,88	1	0,77
Croatia	0,20	0,24	0,83	0,39	0,29	0,49
Italy	1	1	1	1	1	1
Cyprus	0,40	0,47	0,84	0,55	0,82	0,27

[9] https://ec.europa.eu/eurostat/data/database (accessed in April 2021)
[10] https://stats.oecd.org/ (accessed in April 2021)

Latvia	0,17	0,32	0,54	0,37	0,57	0,16
Lithuania	0,33	1	0,33	0,66	1	0,33
Luxembourg	1	1	1	0,97	1	0,94
Hungary	0,78	1	0,78	0,89	1	0,78
Malta	1	1	1	1	1	1
Netherlands	0,66	1	0,66	0,83	1	0,66
Austria	0,30	0,40	0,74	0,55	0,77	0,33
Poland	0,14	0,23	0,60	0,35	0,45	0,24
Portugal	0,55	0,82	0,67	0,75	0,82	0,67
Romania	0,45	0,58	0,78	0,68	0,58	0,78
Slovenia	0,28	0,28	1	0,50	0,70	0,30
Slovakia	0,27	0,43	0,63	0,52	0,73	0,31
Finland	0,20	0,48	0,42	0,51	0,85	0,16
Sweden	0,28	0,68	0,42	0,63	1	0,37

The first column of the table contains the countries that form the DMU set. The next three columns have the overall agricultural sustainability, the efficiency of the first stage and the efficiency of the second stage as they were calculated with the model by Chen et al. (2012). The countries/DMUs can be separated into three general groups: the first group has the countries that have an overall agricultural efficiency of 1 and these are Italy, Luxembourg and Malta. The inclusion of countries like Italy (considered as agriculturally developed) with that of Malta, highlights that agricultural output should not be the only measure of development. Malta, a small country, utilizes a small percentage of its available land for agriculture and has a very small agricultural sector in general. However, the available resources are used in an efficient way and although the country relies on imports for such products, the small population and the relatively effective spending on the infrastructure along with the reduced greenhouse emissions, means that Malta is as agriculturally efficient as Italy.

The second group contains the countries that have an overall agricultural sustainability below 1 and over 0.5. These countries are: Denmark, Germany, France, Hungary, Netherlands and Portugal. These countries are considered among the most developed in the European Union and it appears that the agricultural sector follows (or affects) the overall development of the country.

Finally, the third group contains the rest of the countries that have an overall agricultural performance below 0.5. The majority of these countries are either ones that are considered new in the union (like Poland) or have been affected heavily by the economic crisis of the previous decade (Greece).

The last three columns of Table 7 illustrate the same results but calculated with the two-stage DEA variation proposed in the current thesis. The first fact to observe is that all the values of the overall agricultural sustainability are larger than in the first case, but the values themselves have a smaller variation. Similar to the results as calculated with the DEA variation of Chen et al. (2012), three groups can be recognized. The first entails Italy and Malta that have the largest values of overall agricultural sustainability. The second group consists of the majority of the countries and the final group with Bulgaria, Estonia, Croatia, Latvia and Poland.

To get a better understanding of the countries' performance, their rank is displayed on Table 8. The results illustrate that differences are observed on the two DEA variations; there are countries that perform better in their ranking but not by a lot like the Czech Republic and there are others where the two rankings are significantly different, like Slovenia.

The differences in the values of agricultural efficiency and the overall ranking can be explained in a similar way as those of the previous case study.

Table 8 Ranking of the countries for the overall agricultural sustainability according to the two variations

	Chen et al (2012)	Proposed, alternative metric two-stage DEA
Belgium	13	12
Bulgaria	26	27
Czech	18	15
Denmark	4	4
Germany	8	8
Estonia	25	23
Ireland	20	20
Greece	23	22
Spain	10	13
France	6	6
Croatia	21	24
Italy	1	1
Cyprus	12	17
Latvia	24	25
Lithuania	14	11
Luxembourg	1	3
Hungary	5	5
Malta	1	1
Netherlands	7	7
Austria	15	16
Poland	27	26
Portugal	9	9
Romania	11	10
Slovenia	17	21
Slovakia	19	18
Finland	22	19
Sweden	16	14

4. Exploratory, Multi-dimensional Data Envelopment Analysis[11]

As it was mentioned in Case Study 1 of the previous Chapter, environmental performance is considered only one of the three (or more) dimensions of sustainability. Consequently, moving in the direction of adding more dimensions to measure sustainability, the need arises to move from a two-stage DEA model to a multi-level or multi-dimensional model that will allow the incorporation of these dimensions without succumbing to the methodological limitations of DEA. In the following sub-sections, the new framework is proposed for the incorporation of multiple dimensions.

4.1 Multi-Dimensional DEA for the construction of composite indicators

The typical calculation of sustainability involves three dimensions: economic, environmental and social. Thus, the calculation of the environmental performance in the previous section can be considered as part of sustainability, despite the fact that many of the inputs, intermediate measures and outputs that have been used by the various authors resemble those that are used in the DEA literature for the calculation of sustainability.

However, for a more inclusive calculation of sustainability that is not limited by the number of inputs and outputs that can be used, the proposed variation that was described by equations (70)-(79) can be incorporated in the framework proposed by Tsaples and Papathanasiou (2020b) that is depicted in Figure 5 below.

[11] This section appeared on:
1. Tsaples, G., & Papathanasiou, J. (2020). Multi-level DEA for the construction of multi-dimensional indices. *MethodsX*, *7*, 101169.(CiteScore 1.8)
2. Tsaples, G., Papathanasiou, J., & Georgiou, A. C. (2022). An Exploratory DEA and Machine Learning Framework for the Evaluation and Analysis of Sustainability Composite Indicators in the EU. Mathematics, 10(13), 2277

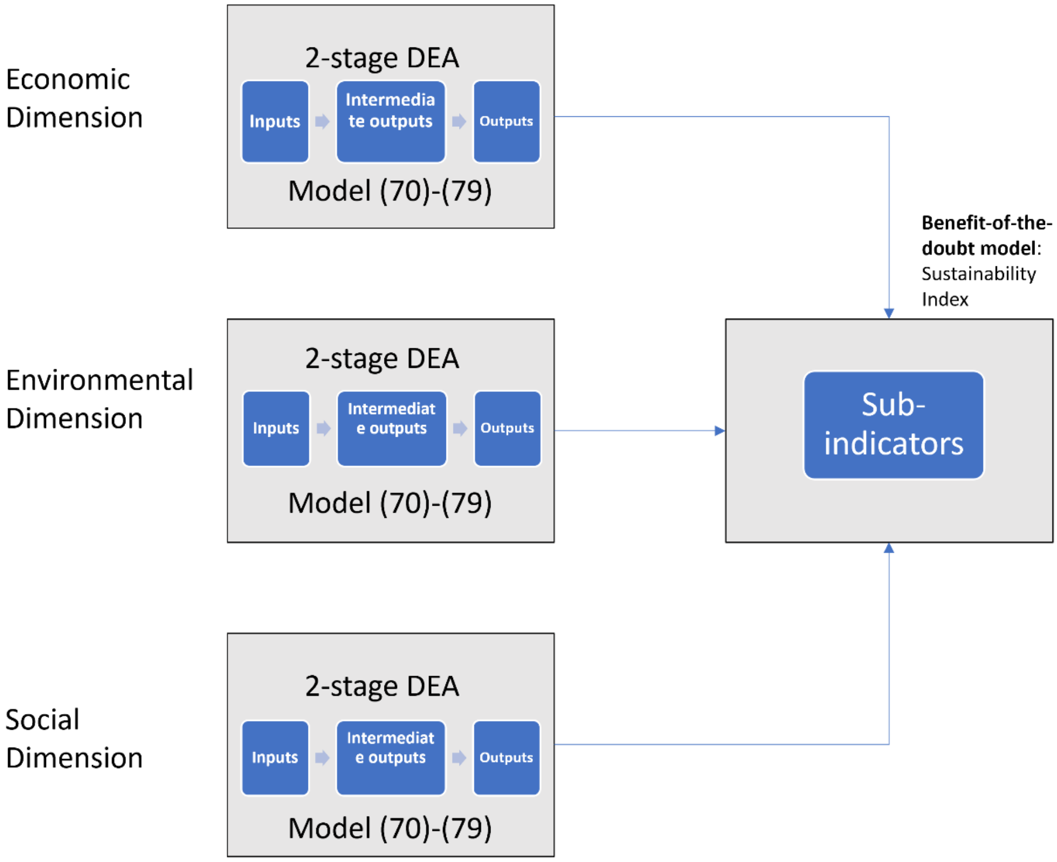


Figure 5 Framework for the construction of composite indicators with two-stage DEA models

Based on the figure, each sub-indicator is calculated using the equations (70)-(79) and the overall performance of each sub-indicator is used in a Benefit-of-the-Doubt (BoD) model to calculate the overall sustainability index. The BoD model is described by the equations (92)-(94) below (Cherchye, Moesen, Rogge, & Van Puyenbroeck, 2007):

$$\max \sum_{r=1}^{s} w_{ri} y_{ri} \tag{92}$$

Subject to:

$$\sum_{r=1}^{s} w_{ri} y_{rj} \leq 1 \ (N \ constraints, one \ for \ each \ DMU \ j = 1 \ldots N) \tag{93}$$

$$w_{rj} \geq 0, (r = 1 \ldots s, s \ constraints \ one \ for \ each \ sub - indicator) \tag{94}$$

The BoD model described by the equations (92)-(94) is a typical DEA model with the inputs designated as 1. As a result, the model calculates the optimal weights allowing maximum flexibility. In contrast to the proposed two-stage variation of equations (70)-(79), the BoD model does not include any restrictions to the weights because the dimensions that are included are typical of sustainability (despite the differences in the underlying measures that are used to calculate those indicators) and limited in number. Moreover, the simplicity of the BoD model, the opportunities that it allows to account for different (countries') backgrounds (Rogge, et al., 2017), the fact that it has been used by numerous studies (see (Karagiannis & Karagiannis, 2018) for an inclusive account) and has been proposed by OECD for the construction of composite indicators (OECD, 2008) mean that it can be used without any intervention in the weights. As mentioned in the above paper, its main advantage is that "it results in idiosyncratic weights to aggregate sub-indicators that vary both across sub-indicators and evaluated decision-making units (DMUs)". In other words, "each evaluated DMU is allowed to choose a set of weights that maximizes its performance in terms of the resulting value of the composite indicator under the restriction that if the same set of weights is used by any other evaluated DMU it will not result in a value of the composite indicator that is greater than one" (Karagiannis & Karagiannis, 2018, p. 1). The use of a BoD model is not unique and alternative methods can be used equally successfully and efficiently. However, in the context of the current thesis, the BoD approach is preferred because the overall proposed model continues to be two-stage DEA, in which the first stage consists of two-stage DEA models that calculate the (more refined) dimensions that will be used in the BoD model that brings the above desired properties for the construction of a final scalar index for each country. Thus, the framework is characterized by an esoteric, elegant consistency.

4.1.1 Case Study 1[12]

To test the proposed framework, the measurement of sustainability of European countries is performed. In the current thesis, the following measures are used for each dimension:

- Economic
  - Inputs: Gross fixed capital at current prices (PPS), Total Labour force (x1000 persons)
  - Intermediate measures: GDP per capita in PPS-Index (EU28 = 100)
  - Outputs: Median equivalised net income [Purchasing power standard (PPS)], Final consumption expenditure of households [Current prices, million euro]
- Environmental
  - Inputs: Population, Gross electricity production [Thousand tonnes of oil equivalent (TOE)]
  - Intermediate measures: Final energy consumption (Terajoule)
  - Outputs: Terrestrial protected area (km2), Share of renewable energy in gross final energy consumption (%), Greenhouse gas emissions (in CO_2 equivalent)
- Social
  - Inputs: Gross fixed capital at current prices (PPS), GDP per capita in PPS-Index (EU28 = 100)
  - Intermediate measures: Total expenditure (Euro per inhabitant)
  - Outputs: Patent applications to the European patent office (EPO) by priority year, Overall life satisfaction, Satisfaction with living

[12] The case study appeared on: Tsaples, G., Papathanasiou, J., & Georgiou, A. C. (2022). An Exploratory DEA and Machine Learning Framework for the Evaluation and Analysis of Sustainability Composite Indicators in the EU. Mathematics, 10(13), 2277

environment, Percentage of females in total labor population (Tsaples & Papathanasiou, 2020a)

Table 9 Results for the construction of a sustainability composite indicator

Country	Sustainability	Economic sub-indicator	Environmental sub-indicator	Social sub-indicator
Belgium	0,68	0,44	0,38	0,47
Bulgaria	0,94	0,4	0,75	0,6
Czech Republic	0,62	0,4	0,39	0,40
Denmark	0,82	0,44	0,5	0,57
Germany	1	0,49	0,35	0,87
Estonia	1	0,53	0,84	0,46
Ireland	0,62	0,40	0,41	0,38
Greece	0,61	0,34	0,44	0,37
Spain	0,68	0,47	0,42	0,44
France	0,85	0,50	0,34	0,70
Croatia	0,82	0,43	0,67	0,43
Italy	0,70	0,46	0,32	0,54
Cyprus	0,90	0,80	0,59	0,47
Latvia	1	0,46	0,84	0,53
Lithuania	0,93	0,41	0,79	0,49
Luxemburg	0,93	0,79	0,62	0,53
Hungary	0,68	0,32	0,46	0,44
Malta	1	1	0,61	0,58
Netherlands	0,65	0,40	0,30	0,50

Austria	0,74	0,45	0,47	0,47
Poland	0,68	0,43	0,34	0,50
Portugal	0,60	0,36	0,38	0,39
Romania	0,80	0,29	0,43	0,60
Slovenia	0,85	0,54	0,65	0,46
Slovakia	0,61	0,33	0,41	0,39
Finland	0,93	0,46	0,74	0,52
Sweden	0,87	0,41	0,62	0,54
United Kingdom	0,75	0,50	0,31	0,58

As it can be observed on table 9, there are 4 countries that are considered sustainable compared to the rest of the set: Germany, Estonia, Latvia and Malta. The rest of the countries can be grouped in two broad categories; those that have a sustainability index above 0.7 compared to the other countries and those that have a sustainability index below 0.7. which are Belgium, Czech Republic, Ireland, Greece, Spain, Hungary, Netherlands, Portugal and Slovakia. Furthermore, the Spearman Correlation Coefficient was calculated for:

- Sustainability-Economic sub-indicator: 0.635
- Sustainability-Environmental sub-indicator: 0.627
- Sustainability-Social sub-indicator: 0.616

The coefficients illustrate that the sustainability of each country depends almost equally on each sub-indicator, with the economic-sub-indicator however, having a slightly larger coefficient.

4.1.2 Case Study 2[13]

To get a better sense of the versatility of the computational framework, the same measurement of sustainability occurred but this time using the typical two-stage DEA variations and not the proposed one. Thus, the equations for the measurement of each dimension are:

$$\max E_0 = \sum_{r=1}^{s} \gamma_r y_{r0} \tag{95}$$

Subject to Constraints

$$\sum_{d=1}^{D} \mu_d z_{dj} - \sum_{i=1}^{m} \omega_i x_{ij} \leq 0, j = 1, \dots, N \tag{96}$$

$$\sum_{r=1}^{s} \gamma_r y_{rj} - \sum_{r=1}^{s} \mu_d z_{dj} \leq 0, j = 1, \dots, N \tag{97}$$

$$\sum_{r=1}^{s} \gamma_r y_{rj} - \sum_{i=1}^{m} \omega_i x_{ij} \leq 0, j = 1, \dots, N \tag{98}$$

$$\sum_{i=1}^{m} \omega_i x_{i0} = 1 \tag{99}$$

$$\omega_i \geq 0, i = 1, \dots, m \tag{100}$$

$$\mu_d \geq 0, d = 1, \dots, D \tag{101}$$

$$\gamma_r \geq 0, r = 1, \dots, s \tag{102}$$

Furthermore, the BoD model is altered, and the new equations are:

$$\max \sum_{r=1}^{s} w_{ri} y_{ri} \tag{103}$$

Subject to:

[13] The case study appeared on the paper Tsaples, G., & Papathanasiou, J. (2020). Multi-level DEA for the construction of multi-dimensional indices. *MethodsX*, *7*, 101169 (CiteScore 1.8)

$$\sum_{r=1}^{s} w_{ri} y_{rj} \leq 1 \ (N \ constraints, one \ for \ each \ DMU \ j = 1 \dots N) \quad (104)$$

$$w_{rj} \geq a, (r = 1 \dots s, s \ constraints \ one \ for \ each \ sub - indicator) \quad (105)$$

$$\sum_{r=1}^{s} w_r = 1 \quad (106)$$

Constraint (104) indicates that the constructed indicator will be less than or equal to 1. Constrain (105) ensures that all sub-indicators will participate in the construction of the overall composite index with the parameter a defined by the analyst. Constraint (106) indicates that the sum of the calculated weights will be equal to 1.

As it was mentioned, the new equations are used for the calculation of sustainability of European countries with the same dimensions (inputs, intermediates, outputs) as before:

- Economic
    - Inputs: Gross fixed capital at current prices (PPS), Total Labour force (x1000 persons)
    - Intermediate measures: GDP per capita in PPS-Index (EU28 = 100)
    - Outputs: Median equivalised net income [Purchasing power standard (PPS)], Final consumption expenditure of households [Current prices, million euro]
- Environmental
    - Inputs: Population, Gross electricity production [Thousand tonnes of oil equivalent (TOE)]
    - Intermediate measures: Final energy consumption (Terajoule)
    - Outputs: Terrestrial protected area (km2), Share of renewable energy in gross final energy consumption (%), Greenhouse gas emissions (in CO_2 equivalent)

- Social
    - Inputs: Gross fixed capital at current prices (PPS), GDP per capita in PPS-Index (EU28 = 100)
    - Intermediate measures: Total expenditure (Euro per inhabitant)
    - Outputs: Patent applications to the European patent office (EPO) by priority year, Overall life satisfaction, Satisfaction with living environment, Percentage of females in total labor population

The results are summarized on table 10 below.

Table 10 Results for the sustainability indicator by using the typical two-stage DEA variation for the calculation of the individual dimensions

Country	**Sustainability index**	**Economic sub-indicator**	**Environmental sub-indicator**	**Social sub-indicator**
Malta	0.778	1	0.223	0.296
Latvia	0.544	0.122	0.697	0.249
Estonia	0.536	0.186	0.681	0.214
Germany	0.524	0.004	0.038	0.739
Cyprus	0.509	0.615	0.316	0.205
Lithuania	0.459	0.095	0.589	0.214
Luxemburg	0.458	0.584	0.255	0.0719
Finland	0.381	0.067	0.495	0.157
Croatia	0.358	0.07	0.457	0.186
France	0.343	0.006	0.061	0.475
Slovenia	0.338	0.168	0.405	0.198
Sweden	0.300	0.032	0.38	0.198
Bulgaria	0.244	0.036	0.296	0.213
United Kingdom	0.235	0.003	0.02	0.331

Italy	0.206	0.007	0.056	0.281
Romania	0.17	0.008	0.191	0.2
Spain	0.162	0.008	0.171	0.194
Poland	0.159	0.008	0.093	0.206
Greece	0.157	0.035	0.188	0.131
Netherlands	0.146	0.021	0.023	0.2
Slovakia	0.146	0.039	0.169	0.145
Hungary	0.144	0.022	0.16	0.167
Austria	0.139	0.039	0.153	0.157
Denmark	0.132	0.064	0.145	0.142
Portugal	0.13	0.033	0.136	0.149
Czech Republic	0.123	0.023	0.077	0.155
Ireland	0.119	0.072	0.132	0.107
Belgium	0.115	0.032	0.032	0.151

The countries are also ranked from the lowest to the highest sustainability value. For example, Malta has the highest sustainability index compared to the other countries with a value of 0.778. The individual sub-indicators that are entailed in sustainability have a value of 1 for the economic sub-indicator, 0.223 for the environmental sub-indicator and 0.296 for the social sub-indicator. These results indicate that Malta has a high sustainability index due to its great economic performance (compared to the other countries). However, the economic prosperity is not accompanied by the same environmental or social performance. As a result, policy makers in Malta might consider placing greater importance in the environment and the social processes of the country, should they wish to achieve sustainable development.

Furthermore, the rest of the countries can be roughly separate into 2 groups: those countries like Latvia, Germany and Slovenia that have a sustainability index between 0.55 and 0.3. The last group that performs the best includes countries like Greece, Netherlands and Belgium.

Hence, the use of the typical two-stage DEA model with the altered BoD version produces different results. These results can be attributed to the more strict constraints that are present in the altered BoD version. Furthermore, the typical two-stage model has wider values than those of the proposed variation, since it does not attempt to minimize the calculated values from the minimum ones.

In conclusion, the proposed computational framework for the construction of composite indicators can be used with any variation of DEA (either two-stage or simple-stage) according to the requirements of the problem at hand. This method allows the calculation of sub-indicators and the final indicator by limiting the bias that can be inserted in the models by the analyst and/or the policy-maker.

4.2 Proposed DEA-ML computational framework[14]

Nonetheless, the above calculated sustainability indices suffer from the same limitation that was identified in the Introduction: since there is no unique, "correct" definition of sustainability, the same indicator can be calculated by using different variations of DEA and/or different combinations of inputs, intermediate measures and outputs.

Furthermore, the proposed two-stage DEA model with the alternative optimization metric might not offer a unique solution that could alter the final results of the calculated index. Finally, one could argue that the BoD model that was used to aggregate the individual dimensions into one sustainability index does not pose any restrictions to the weights, similar to those proposed in the initial two-stage DEA model. Thus, methodological limitations might limit the value of the final results.

Consequently, there is the need to have an indicator of sustainability that will incorporate all these different perceptions that may arise, where perceptions mean different DEA and BoD variations and/or different combinations of inputs, intermediate measures and outputs, and at the same time limit the impact of methodological limitations. Such an indicator would be useful in policy design (and policy making in general) because as Foster and Sen (1997) proposed, uniqueness is

[14] This sub-section appeared on: Tsaples, G., Papathanasiou, J., & Georgiou, A. C. (2022). An Exploratory DEA and Machine Learning Framework for the Evaluation and Analysis of Sustainability Composite Indicators in the EU. Mathematics, 10(13), 2277

not a prerequisite to make agreed judgments. Hence, the proposed computational framework is based on this principle, and it consists of the following steps:

Step 1: Define different perceptions of sustainability and for each perception:

a) define how many sub-indicators will be entailed in this perception's sustainability index
b) define the inputs, intermediate outputs and outputs that each sub-indicator will entail
c) Repeat for all perceptions

Step 2: Define the variation of DEA that will calculate the value of the sub-indicators

d) calculate the sub-indicators

e) calculate the perception's sustainability index using model (92)-(94)

f) Once all sustainability indices for all perceptions are calculated, calculate the mean value for each country/DMU

Step 3: Use machine learning to gain insights into the sustainability of each country under different perceptions

Figure 6 below illustrates the proposed computational framework.

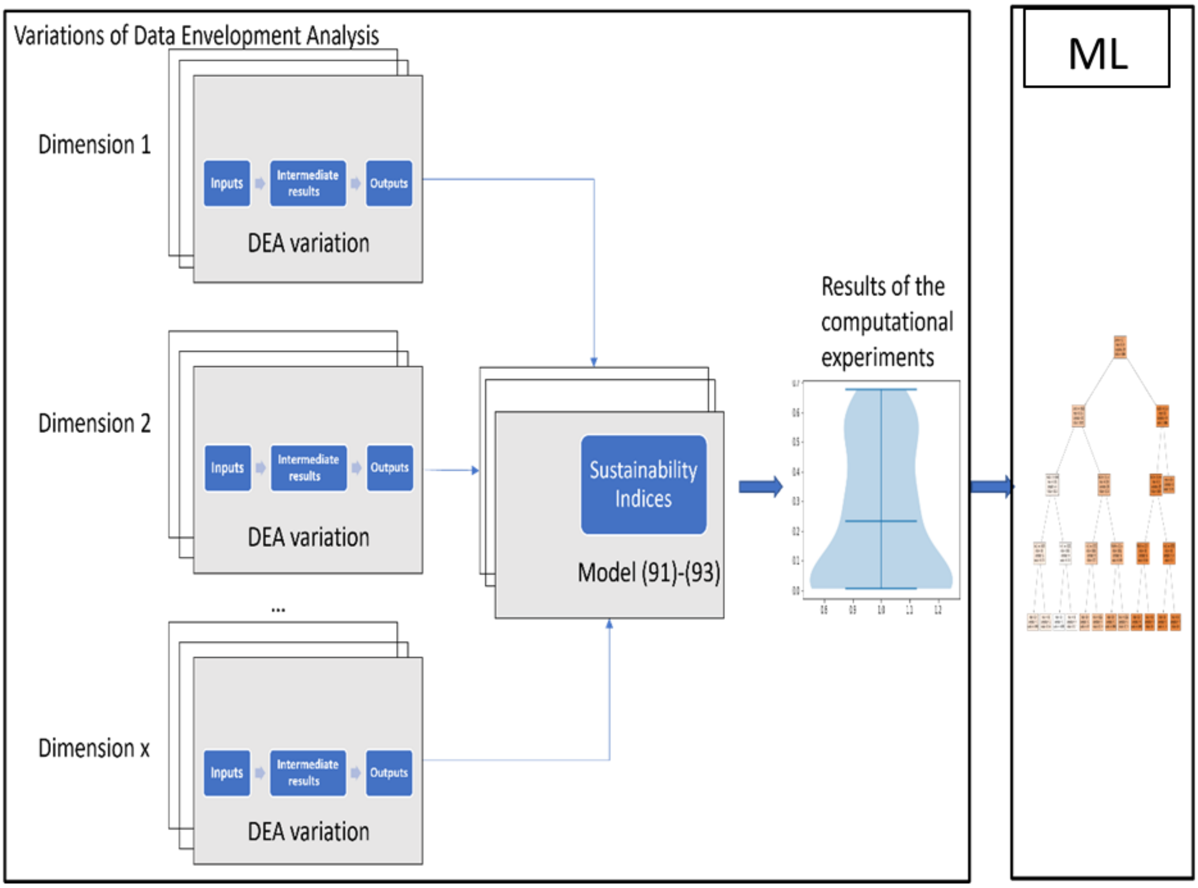


Figure 6 Exploratory, Multi-Dimensional Data Envelopment Analysis

Consequently, by blending DEA with ML the available data and analyses are expanded which contributes to investigating the topic under study (thus implicitly adding new layers to the initial problem) meaning that greater insights are revealed. Furthermore, the absence of a unique solution by the proposed variation of equations (70)-(79) can be considered a methodological limitation, however, the issue becomes not of central importance per se, since in the context of the current thesis the variation will be used repeatedly and with different data to generate different results, in accordance to the philosophy of Exploratory Modeling and Analysis, where methodological limitations lose their impact from the generation of numerous results under different assumptions. Hence, the exploratory framework offers not only a slight deviation to the typical way that DEA is used, but also a complementary research avenue on the issues of interpretability and transparency of algorithms and/or quantitative methods: by blending methodologies under a multi-perspective design, algorithms become more inclusive and democratic (in the sense that the Benefit-of-the-Doubt notion inherent in the aforementioned DEA formulations, is further enriched). Hence, decision support can take a step towards the generation of collective knowledge that includes different values, perceptions and dimensions.

Regarding the proposed computational framework, the author recognizes that it deviates slightly from the traditional process of using Data Envelopment Analysis. In its most basic form, the process starts with a question/problem to be solved, it moves to researching the area of research and gathering the appropriate data, followed by the employment of the method, the analysis of the results and the reporting of the conclusions. Figure 7 below illustrates the typical process of using DEA.

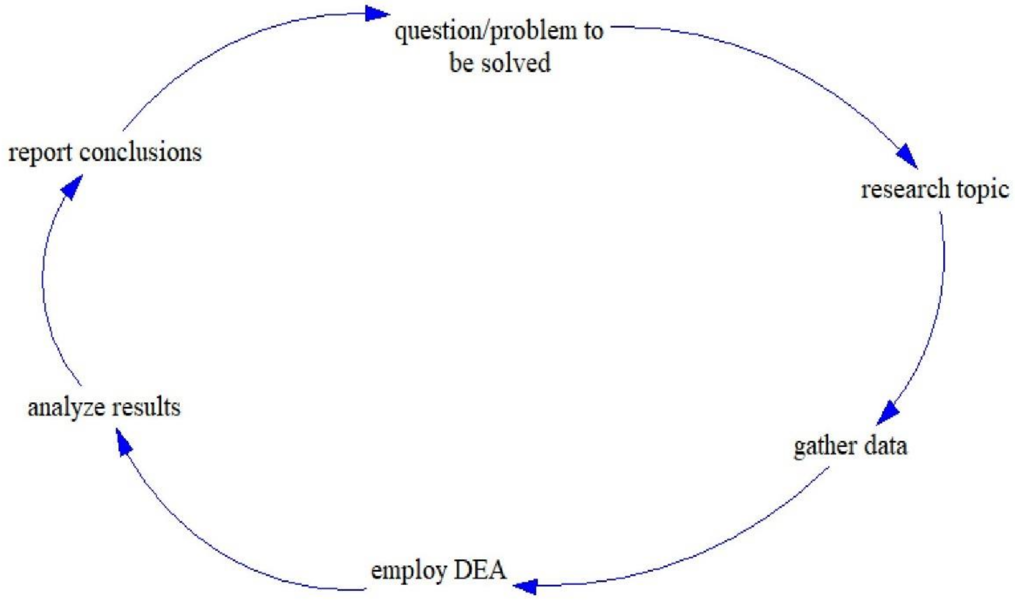


Figure 7 Typical DEA process

However, in the proposed DEA-ML computational framework, additional loops are added to the above elements: new perceptions are added to the research of the topic, and every time new data are gathered followed by the choice and implementation of a DEA variation. Thus, new data are generated which are then fed into an ML methodology to either provide new avenues of research and perception additions or intervene between the use of DEA and analysis of the results (Figure 8).

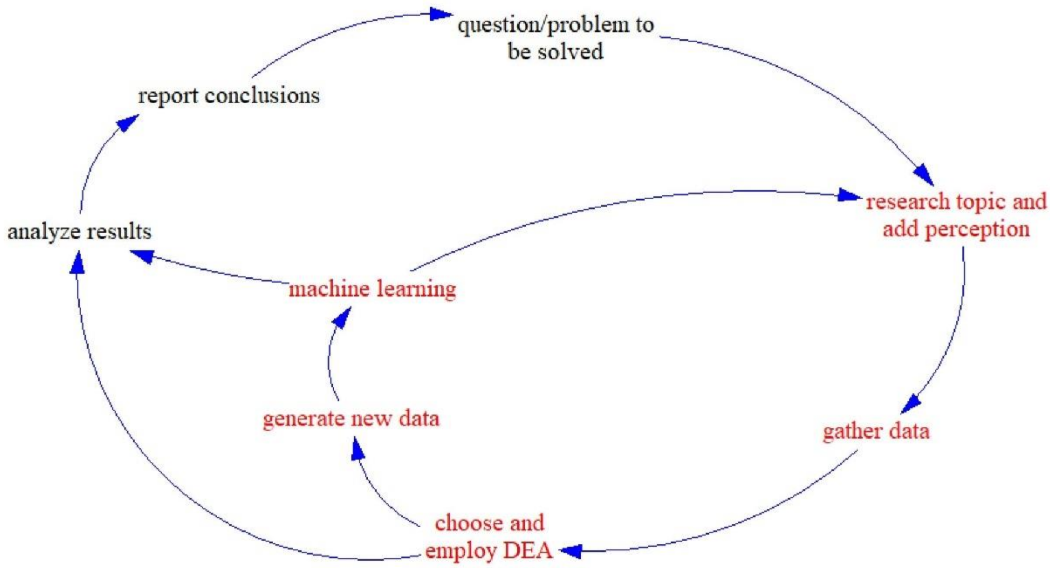


Figure 8 The process in the proposed computational framework

Consequently, by blending DEA with ML the available data and analyses are expanded which contribute to investigating the topic under study (thus implicitly adding new layers to the initial problem) meaning that greater insights are revealed. Moreover, the framework offers a complementary research avenue on the issues of interpretability and transparency of algorithms and/or quantitative methods: by blending methodologies under a multi-perspective design, algorithms become more inclusive and democratic (in the sense that the Benefit-of-the-Doubt notion inherent in the aforementioned DEA formulations, is further enriched). Hence, decision support can take a step towards the generation of collective knowledge that includes different values, perceptions and dimensions.

4.3 Illustration of the proposed DEA-ML computational framework[15]

Following the steps that were proposed in the previous section:

Step 1: In the context of the current thesis, three types of Economic, three types of Environmental, three types of Social and two types of Research and Development (R&D) dimensions were defined and are displayed in the list below:

- Economic sub-indicator 1

[15] This sub-section appeared on: Tsaples, G., Papathanasiou, J., & Georgiou, A. C. (2022). An Exploratory DEA and Machine Learning Framework for the Evaluation and Analysis of Sustainability Composite Indicators in the EU. Mathematics, 10(13), 2277

    - Inputs: Total Labour force (x1000 persons), Gross fixed capital at current prices (PPS), Final energy consumption [Million tonnes of oil equivalent (TOE)]
    - Intermediate: GDP per capita in PPS-Index (EU28 = 100)
    - Outputs: Median equivalised net income [Purchasing power standard (PPS)], Final consumption expenditure of households [Current prices, million euro], People at risk of poverty or social exclusion [thousand persons] [modified to be maximized]
- Economic sub-indicator 2
    - Inputs: Total Labour force (x1000 persons), Gross fixed capital at current prices (PPS)
    - Intermediate: GDP per capita in PPS-Index (EU28 = 100)
    - Outputs: Median equivalised net income [Purchasing power standard (PPS)], Final consumption expenditure of households [Current prices, million euro]
- Economic sub-indicator 3
    - Inputs: Total Labour force (x1000 persons), Gross fixed capital at current prices (PPS)
    - Intermediate: GDP per capita in PPS-Index (EU28 = 100)
    - Outputs: Satisfaction with financial situation, People at risk of poverty or social exclusion [thousand persons] [modified to be maximized]
- Environmental sub-indicator 1
    - Inputs: Gross fixed capital at current prices (PPS), Total Labour force (x1000 persons)
    - Intermediate: GDP per capita in PPS-Index (EU28 = 100)
    - Outputs: Share of renewable energy in gross final energy consumption, Greenhouse gas emissions (in CO_2 equivalent) [modified to be maximized]

- Environmental sub-indicator 2
  - Inputs: Population-, Gross electricity production [Thousand tonnes of oil equivalent (TOE)]-
  - Intermediate: Final energy consumption [Terajoule]
  - Outputs: Terrestrial protected area (km2), Share of renewable energy in gross final energy consumption, Greenhouse gas emissions (in CO2 equivalent) [modified to be maximized]
- Environmental sub-indicator 3
  - Inputs: Gross fixed capital at current prices (PPS), Population, Gross electricity production [Thousand tonnes of oil equivalent (TOE)]-
  - Intermediate: Final energy consumption [Terajoule], GDP per capita in PPS-Index (EU28 = 100)
  - Outputs: Share of renewable energy in gross final energy consumption, Greenhouse gas emissions (in CO2 equivalent) [modified to be maximized], Carbon dioxide [thousand tonnes] [modified to be maximized]
- Social sub-indicator 1
  - Inputs: Total Labour force (x1000 persons), Gross fixed capital at current prices (PPS)
  - Intermediate: GDP per capita in PPS-Index (EU28 = 100)
  - Outputs: Overall life satisfaction, Satisfaction with living environment, Satisfaction with financial situation
- Social sub-indicator 2
  - Inputs: Total Labour force (x1000 persons), Gross fixed capital at current prices (PPS)
  - Intermediate: Total expenditure [Euro per inhabitant]
  - Outputs: Overall life satisfaction, Satisfaction with living environment, Percentage of females in total labor population-2018

- Social sub-indicator 3
    - Inputs: Gross fixed capital at current prices (PPS
    - Intermediate: Total expenditure [Euro per inhabitant], Mean consumption expenditure of private households on cultural goods and services by COICOP consumption purpose [Purchasing power standard (PPS)]
    - Outputs: Percentage of females in total labor population, Life expectancy at birth, People at risk of poverty or social exclusion [thousand persons] [modified to be maximized]
- Research and Development sub-indicator 1
    - Inputs: Total Labour force (x1000 persons), Gross fixed capital at current prices (PPS)
    - Intermediate: GDP per capita in PPS-Index (EU28 = 100)
    - Outputs: Intramural R&D expenditure (GERD) by sectors of performance [Euro per inhabitant], Pupils and students enrolled All ISCED 2011 levels excluding early childhood educational development, Participation rate in education and training (last 4 weeks) by sex and age From 25 to 64 years Percentage
- Research and Development sub-indicator 2
    - Inputs: Final energy consumption [Million tonnes of oil equivalent (TOE)], Population, Gross fixed capital at current prices (PPS
    - Intermediate: GDP per capita in PPS-Index (EU28 = 100)
    - Outputs: Patent applications to the European patent office (EPO) by priority year, Intramural R&D expenditure (GERD) by sectors of performance [Euro per inhabitant]

These 11 different types of dimensions are combined in all the possible combinations of three and four dimensions, resulting in 135 different perceptions of sustainability. Consequently, in the context of the current thesis, the choice of parameters for the models becomes secondary in importance with the purposing of reducing the bias of the analyst or decision maker and the methodological limitations of DEA. All the parameters/variables that are used in the calculations along with summary statistics are presented in Appendix B.

Step 2: Each of these 135 perceptions are used with the proposed DEA variation that is described by the equations (70)-(79) and (92)-(94). The mean sustainability of the countries with the proposed DEA variation is displayed in table 11 below:

Table 11 Mean sustainability calculated with the proposed variation (equations (70)-(79) and ((91)-(93))

Country	Mean Sustainability (proposed two-stage DEA model)
Belgium	0,50
Bulgaria	0,83
Czech Republic	0,47
Denmark	0,73
Germany	0,75
Estonia	0,87
Ireland	0,47
Greece	0,57
Spain	0,52
France	0,65
Croatia	0,77
Italy	0,48
Cyprus	0,86
Latvia	0,91
Lithuania	0,79

Luxemburg	0,93
Hungary	0,57
Malta	1
Netherlands	0,45
Austria	0,60
Poland	0,53
Portugal	0,51
Romania	0,73
Slovenia	0,73
Slovakia	0,51
Finland	0,82
Sweden	0,81
United Kingdom	0,58

Hence, the inclusion of different perceptions alters the results that were illustrated in table 11. Under multiple perceptions, Malta, Latvia and Luxemburg have the highest sustainability compared to the rest of countries. Moreover, with all the different variations of sub-indicators, there are countries for which the mean sustainability falls below 0.5 like the Czech Republic, Ireland, Italy and the Netherlands. Finally, there are no countries for which the mean sustainability increased with the inclusion of different parameters; only Malta managed to keep the sustainability at the value of 1 in both cases.

Figure 9 below illustrates the results of the 135 calculations of sustainability on violin plots.

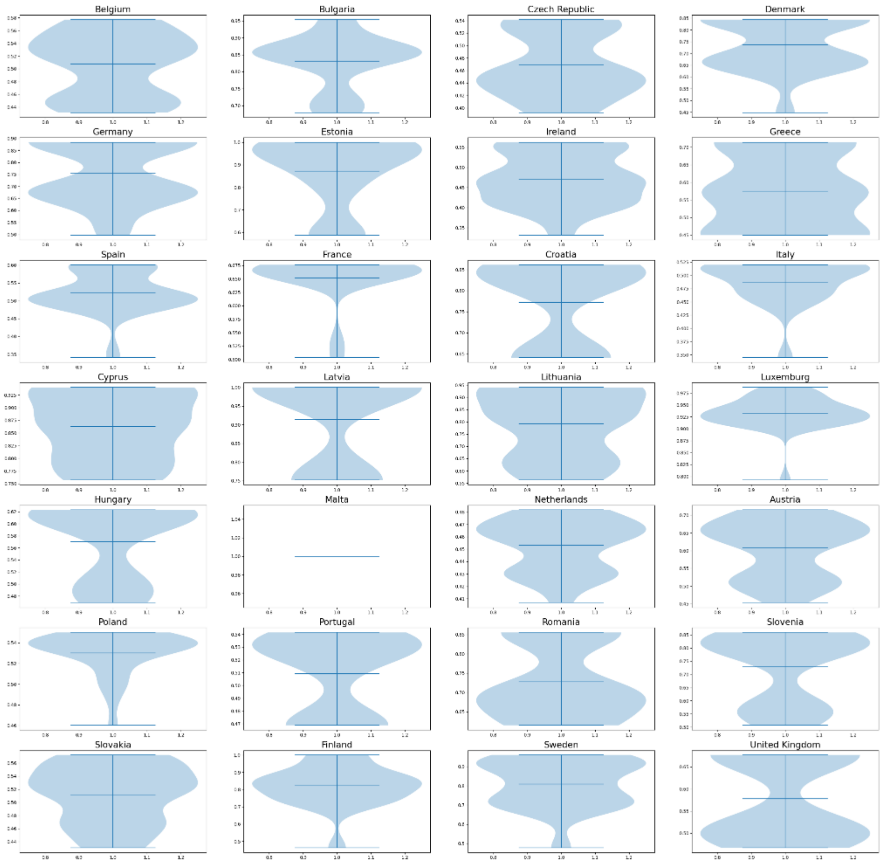


Figure 9 Violin plots of the sustainability of 135 different perceptions calculated with the proposed two-stage DEA variation

The y axis indicates the measurement of sustainability, while the x axis indicates the distribution of the sustainability indices that were calculated; wider sections of the density indicate that there is a higher probability that data points will take the given value, while narrow sections indicate lower probability.

The first aspect to observe is that Latvia concentrates the majority of their values on the upper side of the violin plot, while Malta has a constant value of 1 for all calculations, indicating that compared to the rest, the two countries have a high sustainability no matter the perception, thus the robustness of the conclusion increases. For the rest of the countries, the different perceptions create different situations for their sustainability. For example, Greece has a mean sustainability of 0.57, however its values can change depending to the perception from 0.45 to 0.75, with the majority of the values concentrated between 0.5 and 0.6. Thus, the sustainability of Greece changes with different perceptions in a significant way, weakening the declaration of any robust conclusions.

Apart from the calculation of the sustainability indices under different perceptions with the proposed variation, Step 2 of the computational framework includes the use of different variations of DEA. In the present application, the classic two-stage model of Chen et al. (2012), and an adaptation of the Constant Returns to Scale (CRS) DEA (Charnes, et al., 1978) and of the Variable Returns to Scale (VRS) (Banker, et al., 1984) are used (along with the proposed variation). For the last two DEA variations, the classic DEA models are used in chained way to accommodate the 2-stage nature of the models. More specifically, each combination of inputs, intermediate measures and outputs was used in the chained way of the classic one stage models as follows: the efficiency of the first stage is calculated with the inputs and the intermediate measures as outputs using either CRS or VRS DEA. The efficiency of the second stage is calculated with the intermediate measures as inputs and the outputs using (similar to the first stage) either CRS or VRS DEA. The sustainability index is calculated by multiplying the efficiencies of the two stages. Finally, the sustainability index of each perception is calculated using the BoD model (Cherchye, et al., 2007) The inclusion of different variations of DEA (which can be chosen by the analyst and/or the policy maker) with different combinations of inputs and outputs increases the robustness of the results since many sustainability indices will be calculated that can capture different perceptions both methodological (which method is the more "correct") and context-wise (which combination of inputs, outputs and intermediates is the more "correct").

Table 12 below summarizes the results when all DEA variations are used.

Table 12 Mean sustainability calculated with all DEA variations

Country	**Mean Sustainability (all DEA variations)**	**Mean Sustainability (proposed DEA variation)**
Belgium	0.56	0,50
Bulgaria	0.51	0,83
Czech Republic	0.35	0,47
Denmark	0.69	0,73
Germany	0.60	0,75
Estonia	0.82	0,87
Ireland	0.55	0,47
Greece	0.49	0,57
Spain	0.49	0,52
France	0.55	0,65
Croatia	0.60	0,77
Italy	0.47	0,48
Cyprus	0.85	0,86
Latvia	0.80	0,91
Lithuania	0.67	0,79
Luxemburg	0.95	0,93
Hungary	0.41	0,57

Malta	0.99	1
Netherlands	0.55	0,45
Austria	0.57	0,60
Poland	0.42	0,53
Portugal	0.49	0,51
Romania	0.44	0,73
Slovenia	0.67	0,73
Slovakia	0.40	0,51
Finland	0.75	0,82
Sweden	0.71	0,81
United Kingdom	0.55	0,58

The inclusion of different variations of DEA (which can be chosen by the analyst and/or the policy maker) with different combinations of inputs and outputs increases the robustness of the results since many sustainability indices will be calculated that can capture different perceptions both methodological (which method is the more "correct") and context-wise (which combination of inputs, outputs and intermediates is the more "correct").

As it can be observed on Table 12, the mean sustainability changes again indicating that the methodological framework that is used matters in the calculation of the final index. In the current illustration, there are countries where the mean sustainability increases with the inclusion of other DEA variations (like Belgium, Luxemburg and the Netherlands), others where it is almost the same (like Malta) and those for which the mean sustainability decreases (rest of the countries). Moreover, the Spearman correlation coefficient for the two mean sustainability indices was calculated and found to be equal to 0.752, which indicates a strong positive correlation between the two indices.

Finally, for the majority of the countries, the mean sustainability index is similar under the two DEA variations, which indicates an increased confidence in the results. Hence, those sustainability indices can be considered relatively robust, since they do not have major differences when methods and parameters change. However, there is a set of countries (Bulgaria, Czech Republic, Germany, Latvia, Hungary, Poland, Romania and Slovakia) for which the sustainability index changes significantly under the different DEA variations. Thus, the results seem sensitive in the choice of method which decreases the relative robustness of the results.

These changes are mirrored also in the violin plots of the sustainability, displayed on Figure on figure 10.

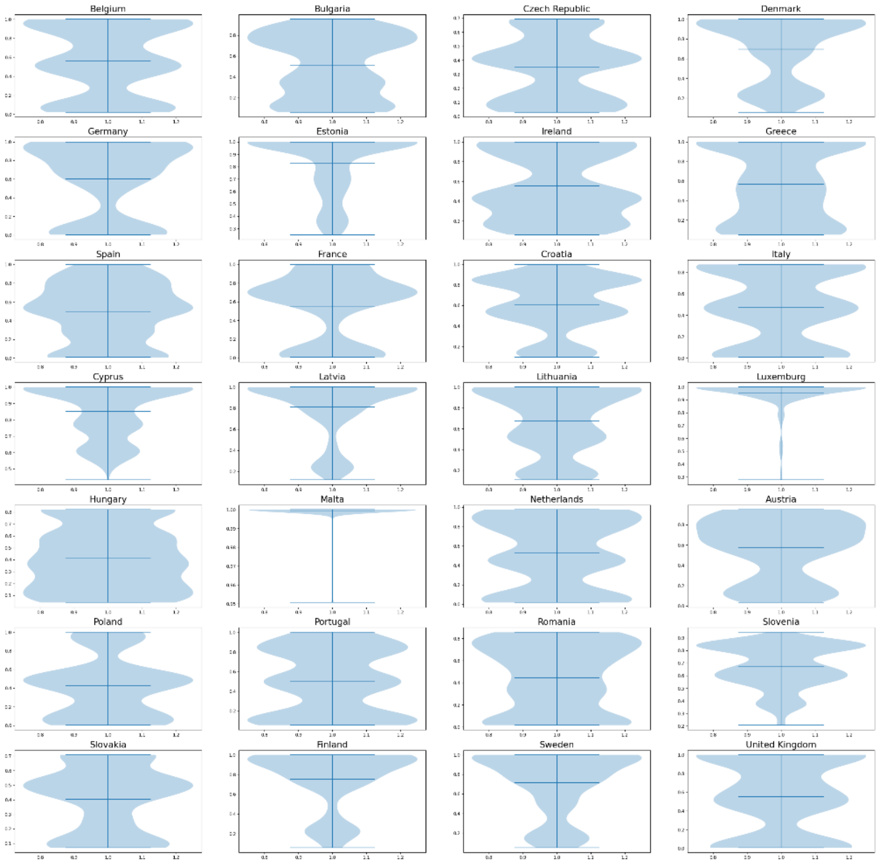


Figure 10 Violin plots of the sustainability of 135 different perceptions calculated with all DEA variations

Step 3: The final step of the proposed computational framework is to use Machine Learning techniques in the results of the generated computations with the purpose of revealing insights into how the sustainability of countries behaves under different perceptions.

Following the logic of EMA, several techniques will be employed in an effort to mitigate intrinsic methodological limitations and find the common, emergent elements that remain robust despite the different methods. All the techniques were deployed with the Python package sci-kit learn (Pedregosa, et al., 2011).

The first insights will be revealed by using clustering techniques and more specifically K-Means (Krishna & Murty, 1999; Likas, Vlassis, & Verbeek, 2003) and Density based clustering (DBSCAN) (Ester, Kriegel, Sander, Xu, & others, 1996). For the clustering algorithms, the values of the sub-indicators along with those of the sustainability indices under all the computational regimes were used. For the K-Means algorithms 4 clusters were defined and Figure 11 below illustrates the results.

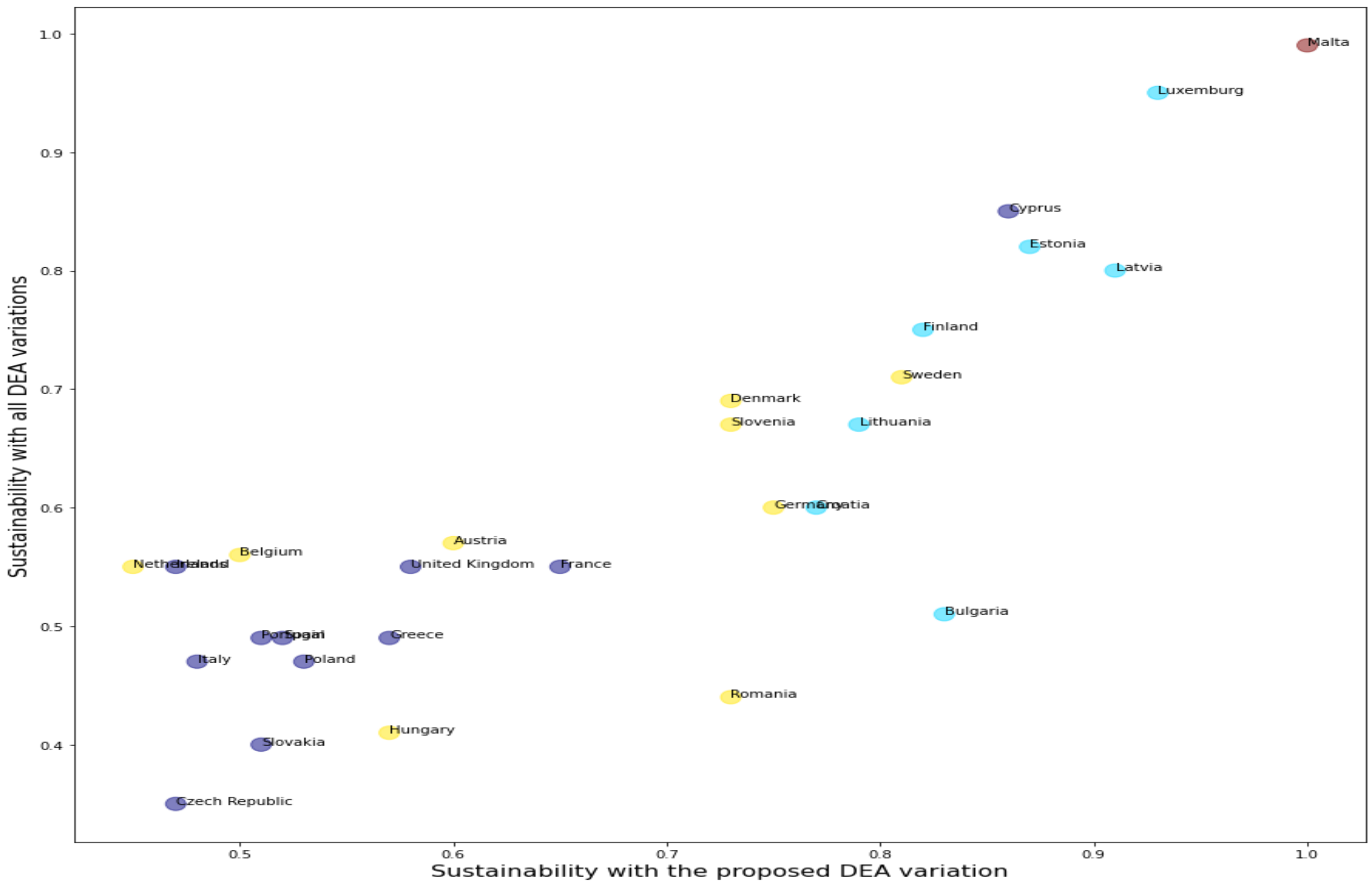


Figure 11 K-Means Clusters of the EU countries

The x-axis illustrates the mean sustainability of the countries under all the perceptions when only the proposed DEA variation is used, while the y-axis entails the mean sustainability of the countries under all perceptions when all DEA variations are employed. As it can be observed, the 4 clusters have different colors. Malta, having the highest mean sustainability values in all calculations, belongs to one cluster (Cluster 0) by itself. The second cluster (light blue color) contains the countries of Luxemburg, Estonia, Latvia, Finland, Lithuania, Croatia and Bulgaria. As it can be observed, these countries have a mean sustainability of 0.5 to 0.9 for all calculations. In this cluster, only Croatia has the lowest mean values.

The third cluster (yellow color) contains the countries Sweden, Denmark, Slovenia, Hungary, Romania, Austria, Germany, Belgium and the Netherlands. The final cluster (purple color) contains the rest of the countries. In the last cluster, an anomaly is detected since Cyprus, despite the fact that belongs to a cluster in which the majority of the countries has a mean sustainability of 0.6 or lower, is the country with the third highest mean sustainability under all perceptions. Thus, it appears thus despite the large mean sustainability, Cyprus shares similar characteristics in the dimensions as the other countries in the cluster.

Apart from the K-Means clustering, the Density Based clustering algorithm was used and the results are illustrated on Figure 12 below.

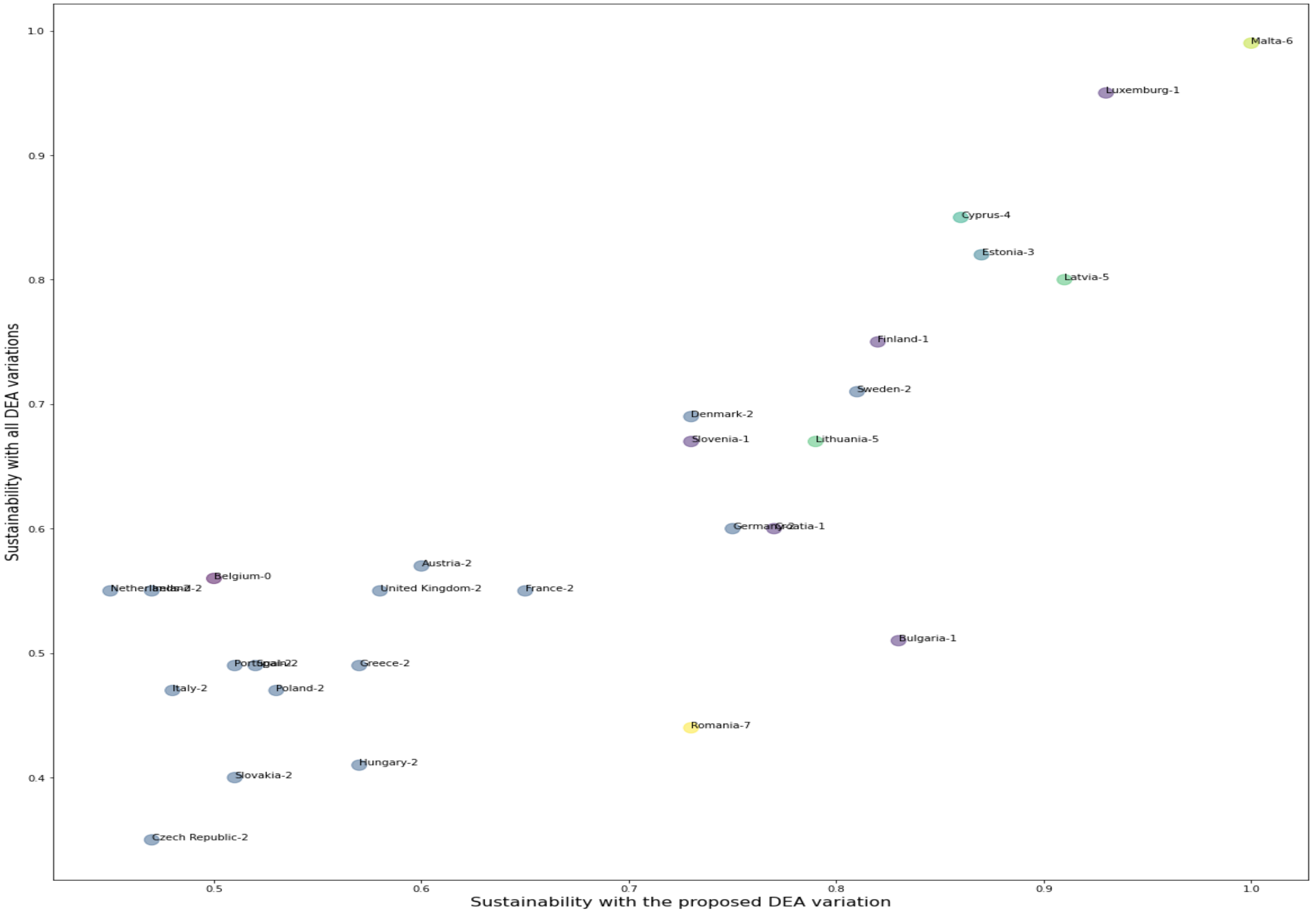


Figure 12 DBSCAN clusters of the EU countries

The DBSCAN algorithm produces 8 clusters and the largest clusters are 1 and 2, with the cluster 2 containing the majority of the countries. Table 13 below summarizes the mean sustainability and the clusters with the two methods for each country.

Table 13 Summary of the clustering process

Country	**Mean Sustainability (all DEA variations)**	**Mean Sustainability (proposed DEA variation)**	**K- Means Cluster**	**DBSCAN Cluster**
Belgium	0.56	0,50	2	0
Bulgaria	0.51	0,83	1	1
Czech Republic	0.35	0,47	0	2
Denmark	0.69	0,73	2	2
Germany	0.60	0,75	2	2
Estonia	0.82	0,87	1	3
Ireland	0.55	0,47	0	2
Greece	0.49	0,57	0	2
Spain	0.49	0,52	0	2
France	0.55	0,65	0	2
Croatia	0.60	0,77	1	1
Italy	0.47	0,48	0	2
Cyprus	0.85	0,86	0	4
Latvia	0.80	0,91	1	5
Lithuania	0.67	0,79	1	5
Luxemburg	0.95	0,93	1	1
Hungary	0.41	0,57	2	2

Malta	0.99	1	3	6
Netherlands	0.55	0,45	2	2
Austria	0.57	0,60	2	2
Poland	0.42	0,53	0	2
Portugal	0.49	0,51	0	2
Romania	0.44	0,73	2	7
Slovenia	0.67	0,73	2	1
Slovakia	0.40	0,51	0	2
Finland	0.75	0,82	1	1
Sweden	0.71	0,81	2	2
United Kingdom	0.55	0,58	0	2

The two clustering techniques illustrate that Malta forms a cluster on its own, since it is the country that has the highest mean sustainability compared with the other countries. Furthermore, there are countries that behave differently than the others. Belgium, Estonia, Cyprus and Romania are countries that in the K-Means clustering belong to clusters with other countries, but with the DBSCAN algorithm they are in a cluster on their own. That is also the case with Latvia, where with the density-based algorithm, it belongs to a cluster with Lithuania.

As a result, these countries will be further analysed in the following paragraphs with other machine learning techniques.

For the current thesis, three additional techniques were used: Classification and Regression Decision Trees (CART) since they are not computationally costly, they can be used as communication tools to non-experts and offer deep interpretational capabilities (Vayssières, Plant, & Allen-Diaz, 2000). However, CART trees tend to overfit the data to their training set and are considered weak learners (Friedman, 2017) and for that reason two additional ML techniques will be used: Random Forests (Kam & others, 1995) and boosting regression (Bühlmann & Yu, 2003).

Random forests train trees independently using random samples of the available data and the sampling happens with bootstrapping both the sample and the features at every repetition. As a result, they tend to be slower than CART trees, but the generated results are more robust and tend to avoid the pitfalls of overfitting. More specifically, with the random forests 80% of the data will be used for training and the remaining will be used for prediction. Furthermore, for each data row (point) of the remaining data, the contribution of the individual features to the predicted value will be calculated. The average of all the contributions will be plotted in a boxplot to reveal insights on how individual sub-indicators affect the value of the sustainability index.

Similarly, boosting regression is also considered a slow learner, but compared to random forests, each tree is generated using information from previous ones (James, Witten, Hastie, & Tibshirani, 2013). Moreover, the technique will also reveal the relative influence of the individual sub-indicator to the index of sustainability, which could provide further insights into the analysis of the results. Both random forests and boosting regression are more robust than CART trees, but this robustness comes at the detriment of intuitive communication capabilities that are the main characteristic of CART trees. Consequently, the use of all three Machine Learning techniques will limit the methodological weaknesses of each method, while providing results and insights that are robust and independent of the used technique.

Finally, following the logic of the previous steps, each country will be studied individually first when only the proposed two-stage DEA variation is used and later with all the DEA variations included in the calculations of sustainability.

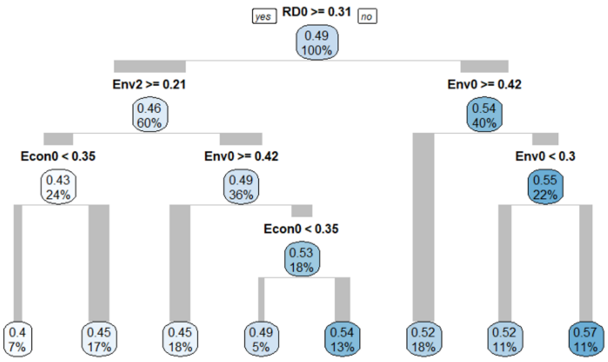
Belgium

Figure 13 below illustrates the results from the employment of the three ML techniques in the generated sustainability indices of Belgium under the proposed variation.

As it can be observed, the overall indicator of Research and Development has the largest influence on the sustainability indices (under the proposed DEA variation), followed by the overall Environmental performance and the efficiency of the second stage of the Environmental sub-indicator.

By studying the CART tree it can be observed, that when the overall Research and Development sub-indicator is lower than 0.31 and the overall Environmental sub-indicator is lower than 0.42, the sustainability index takes its largest values. Thus, it can be concluded that the increase Research and Development and Environmental performance of the country cannot drive the sustainability of the country in high values, which implies that the Economic and Social sub-indicators push the sustainability index to its current levels.

Belgium



Relative Influence

	Relative Influence
RD0	50.160
Env0	31.566
Env2	14.211
Econ0	4.063
Econ1	0.000
Econ2	0.000
Env1	0.000
Soc0	0.000
SOc1	0.000
Soc2	0.000
RD1	0.000
RD2	0.000

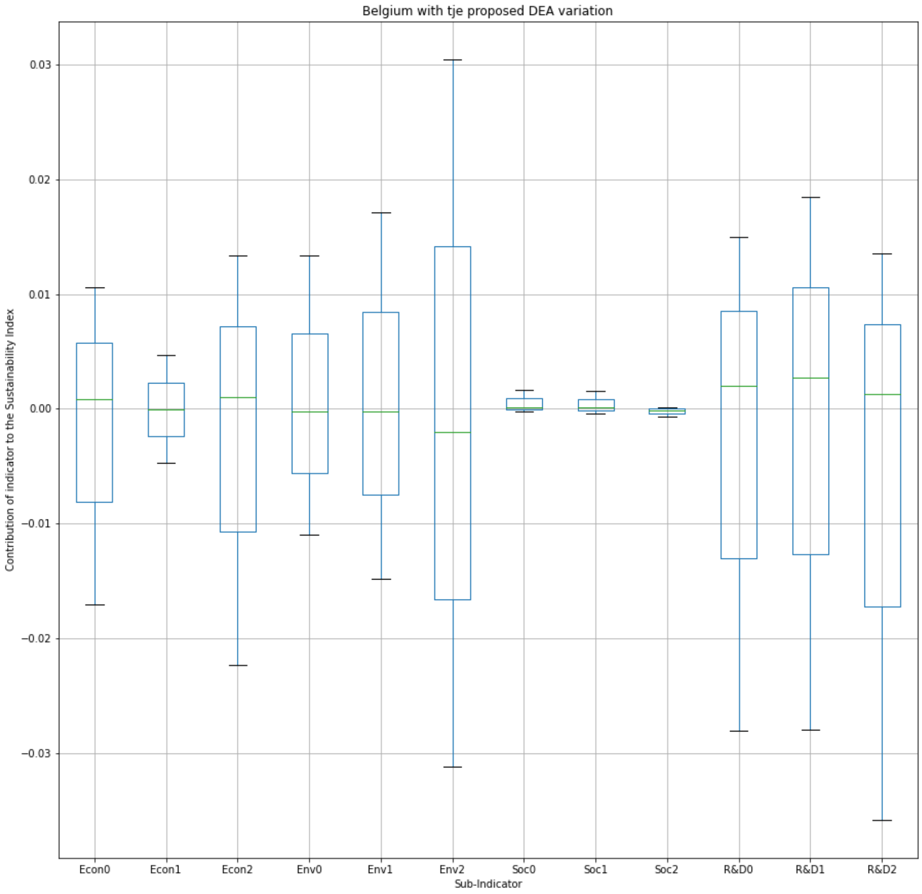


Figure 13 CART tree (upper left), relative influence of sub-indicators from boosting regression (upper right) and average contribution of the sub-indicators to the index from Random forests (bottom) for Belgium under the proposed computational framework with equations (70)-(79)

The results are corroborated by the boxplot, which illustrates that the Research and Development and Environmental sub-indicators (along with the efficiencies of their individual stages) have the widest range of contribution in the sustainability index depending on the choice of the parameters that are use in the framework. Another interesting result that is revealed by the boxplot is that the Social dimension has almost no influence in the values of the final sustainability index.

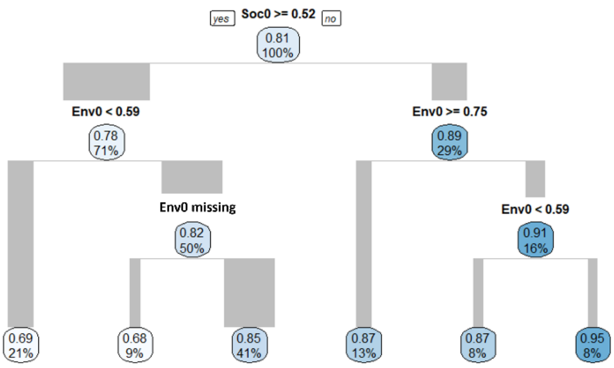
Consequently, the implications for the policy-makers are that Belgium should focus more on the social dimension of sustainability and make an attempt to increase the positive influence of the economic dimension.

Bulgaria

Figure 14 below illustrates the results from the employment of the three ML techniques in the generated sustainability indices of Bulgaria under the proposed variation.

As it can be observed, the overall Environmental performance has the largest relative influence on the sustainability results followed by the overall Social performance. By studying the CART tree, it can be deduced that when the overall Social performance is not larger that 0.52 and the overall Environmental sub-indicator is not larger than 0.75, the sustainability of the country takes its largest values.

Bulgaria



Relative Influence

	Relative Influence
Env0	56.494
Soc0	43.506
Econ0	0.000
Econ1	0.000
Econ2	0.000
Env1	0.000
Env2	0.000
SOc1	0.000
Soc2	0.000
RD0	0.000
RD1	0.000
RD2	0.000

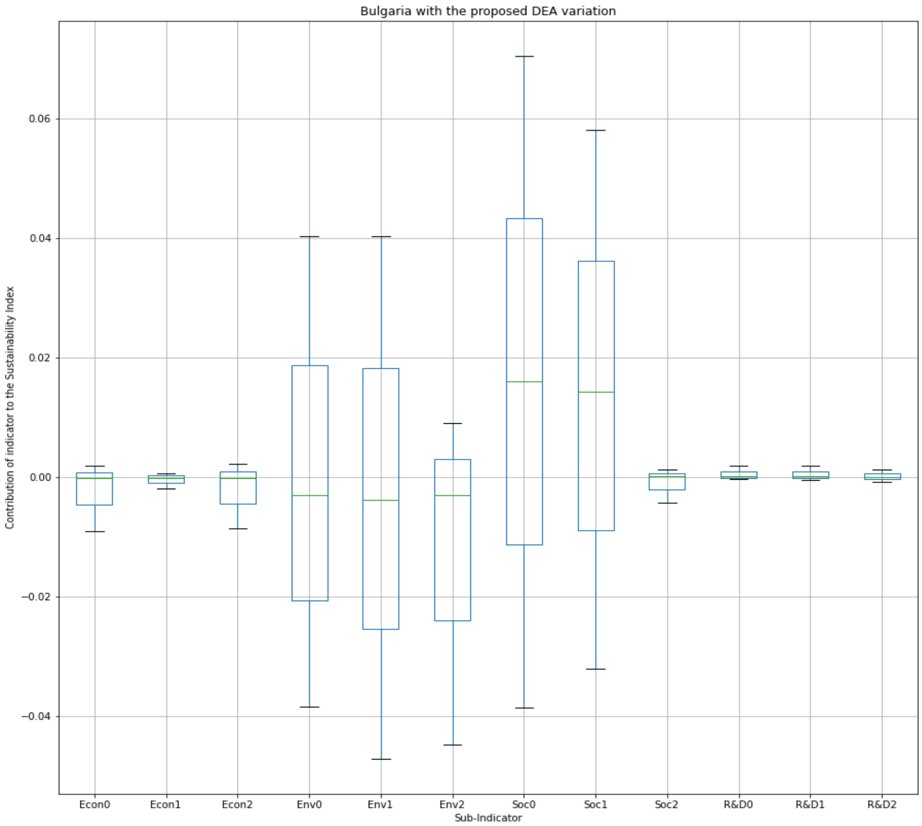


Figure 14 CART tree (upper left), relative influence of sub-indicators from boosting regression (upper right) and average contribution of the sub-indicators to the index from Random forests (bottom) for Bulgaria under the proposed DEA variation

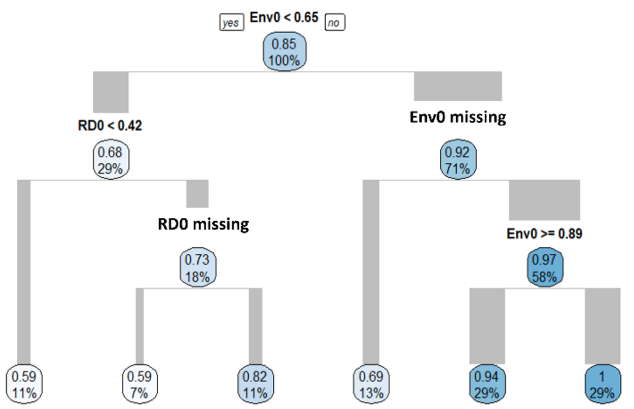
The boxplot illustrates that the Economic and Research and Development sub-indicators (and their corresponding stage efficiencies) have almost no impact in the final calculations of sustainability. Furthermore, the overall Social performance of the country can have a positive impact on the final index further supporting the results of both the boosting regression and CART tree.

Estonia

Figure 15 below illustrates the results from the employment of the three ML techniques in the generated sustainability indices of Estonia under the proposed variation.

The boosting regression algorithm indicates that the overall Environmental performance of the country has the largest influence on the country. Furthermore, the CART tree illustrates that when its value is not smaller than 0.65 (or not missing) then the sustainability index has its largest values that accounts for 58% of the generated results.

Estonia



Relative Influence

	Relative Influence
Env0	81.842
RD0	15.831
Econ0	1.284
Soc0	1.043
Econ1	0.000
Econ2	0.000
Env1	0.000
Env2	0.000
S0c1	0.000
Soc2	0.000
RD1	0.000
RD2	0.000

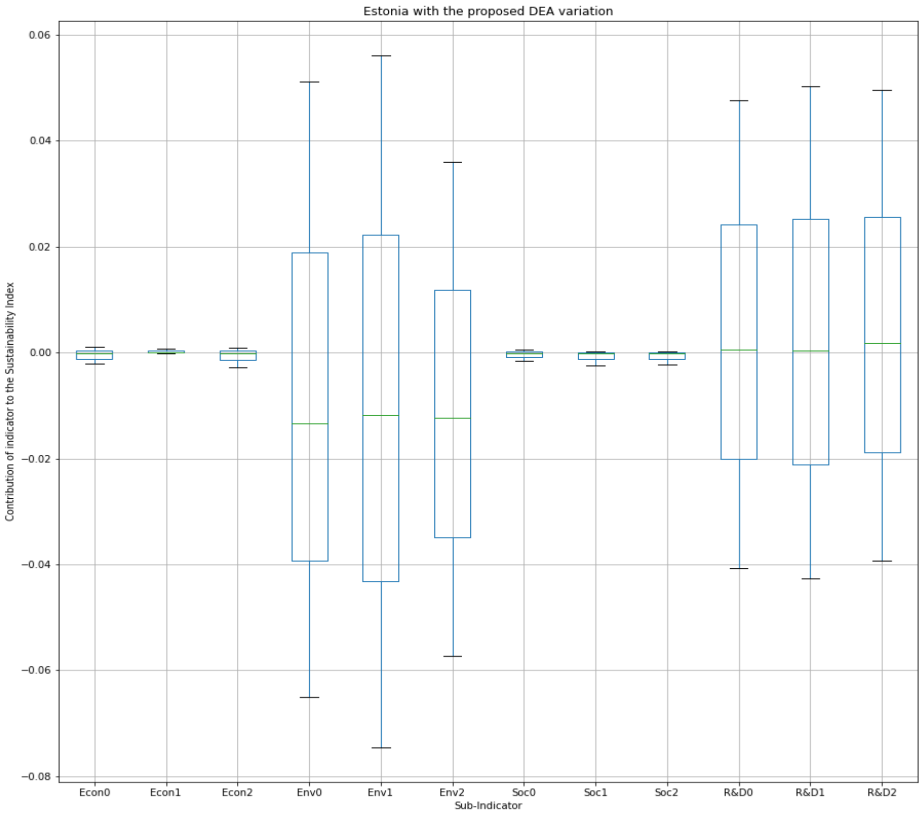


Figure 15 CART tree (upper left), relative influence of sub-indicators from boosting regression (upper right) and average contribution of the sub-indicators to the index from Random forests (bottom) for Estonia under the proposed DEA variation

Consequently, the Environmental performance of the country is the one that drives the sustainability to higher level for the majority of the calculations. In addition, the Random Forest algorithm indicates that the environmental efficiency of the first stage can have both the most positive and most negative contribution to the sustainability of the country. Moreover, the Economic and Social dimensions seem to have a negligent contribution to the final index.

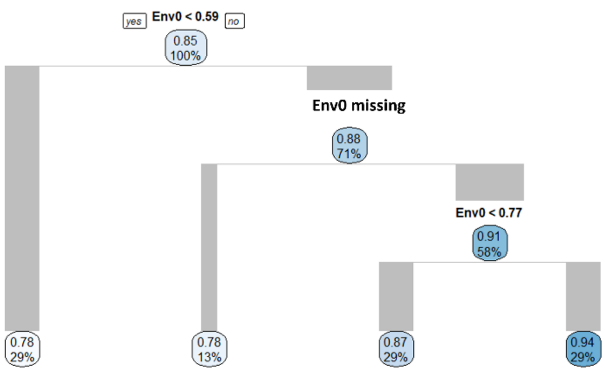
Thus, the implications for policy makers could be that the country should continue the efforts towards the Environment (and the Research and Development front), however more could be done in the aspects of the Economy and Society.

Cyprus

Figure 16 below illustrates the results from the employment of the three ML techniques in the generated sustainability indices of Cyprus under the proposed variation.

Similar to Estonia, for Cyprus the overall Environmental performance has the largest relative influence to the final calculations of sustainability, followed by the economic efficiency of the first stage. Moreover, the CART tree indicates that when the value of the Environmental performance is larger than 0.77, then the final sustainability index has its largest values which accounts for 58% of the calculations.

Cyprus



Relative Influence	
	Relative Influence
Env0	91.822
Econ1	7.225
SOc1	0.953
Econ0	0.000
Econ2	0.000
Env1	0.000
Env2	0.000
Soc0	0.000
Soc2	0.000
RD0	0.000
RD1	0.000
RD2	0.000

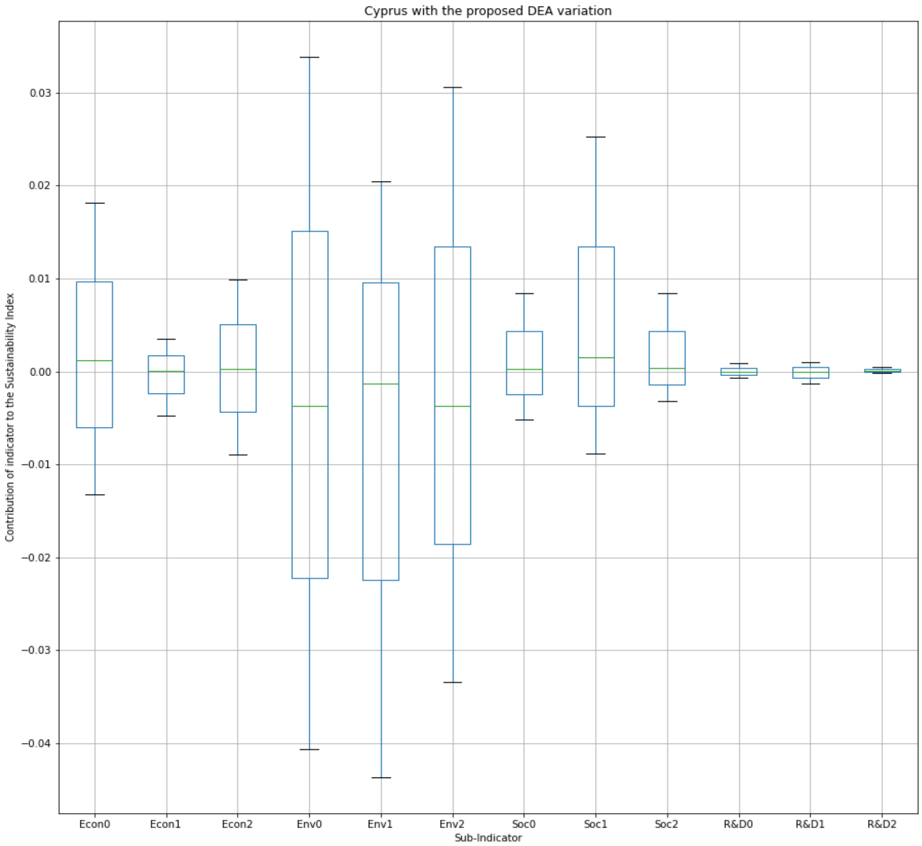


Figure 16 CART tree (upper left), relative influence of sub-indicators from boosting regression (upper right) and average contribution of the sub-indicators to the index from Random forests (bottom) for Cyprus under the proposed DEA variation

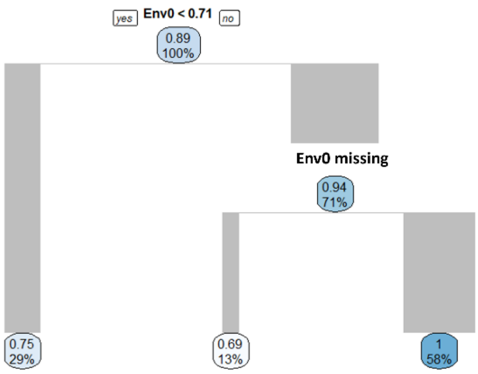
The results from the Boosting Regression and the CART tree are supported by the Random Forests, which indicate that the overall Environmental performance can have the most positive contribution to the final sustainability index. Hence, the sustainability of Cyprus shares many similarities with that of Estonia. The difference of the two countries however can be seen in the contribution of the Research and Development index to sustainability: for Cyprus it is almost negligent. Thus, the implications for policy makers could be that more efforts should be devoted to increasing the capacity of the country to innovate, while keeping the performance for environmental sustainability.

Latvia

Figure 17 below illustrates the results from the employment of the three ML techniques in the generated sustainability indices of Latvia under the proposed variation.

The boosting regression algorithm indicated that overall Environmental performance along with the environmental efficiency of the second stage for the particular sub-indicator are the ones with the largest relative influence to the final sustainability index. In addition, according to the CART tree, when the overall Environmental performance is not smaller than 0.71 then the final sustainability index has a mean value of 0.91, which accounts for 71% of the calculations.

Latvia



Relative Influence

	Relative Influence
Env0	56.463
Env2	31.315
Soc2	4.160
Soc0	2.476
Econ0	2.015
Econ1	1.843
RD1	1.055
RD0	0.673
Econ2	0.000
Env1	0.000
SOc1	0.000
RD2	0.000

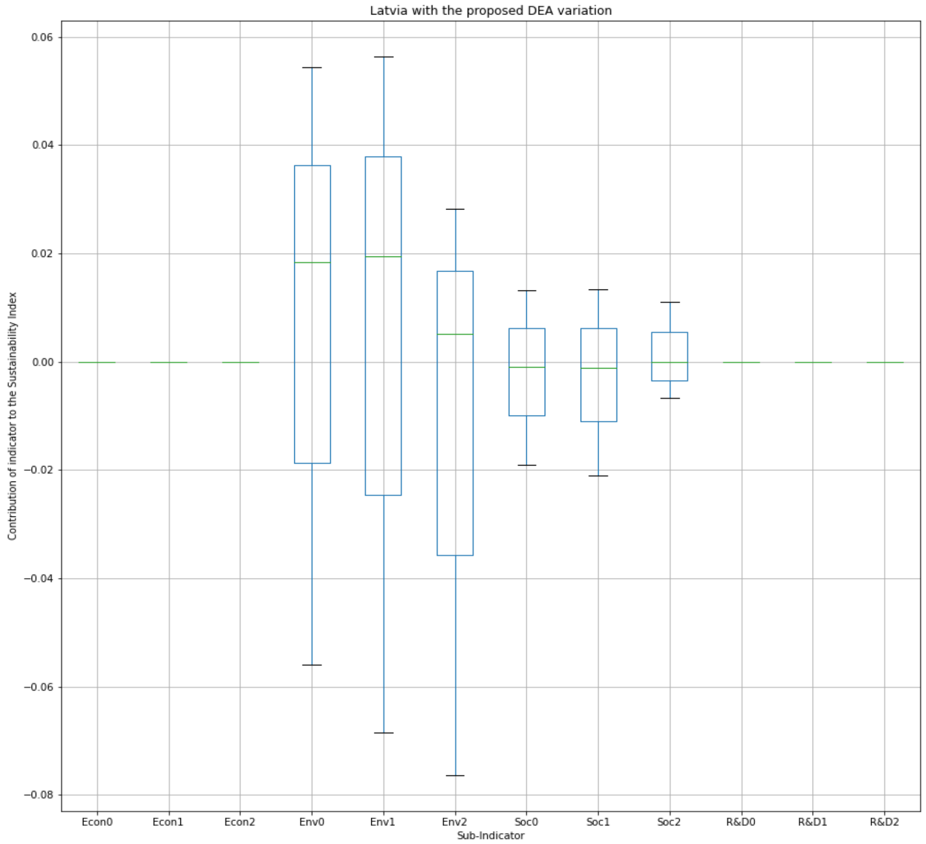


Figure 17 CART tree (upper left), relative influence of sub-indicators from boosting regression (upper right) and average contribution of the sub-indicators to the index from Random forests (bottom) for Latvia under the proposed DEA variation

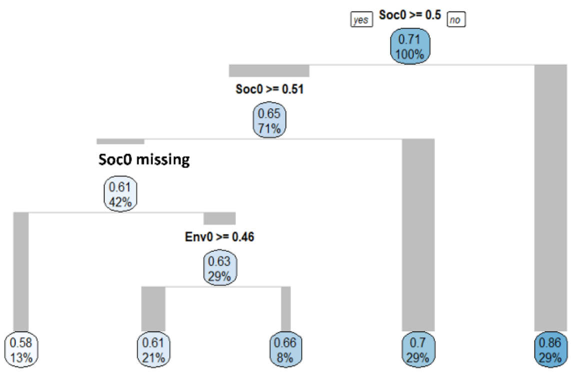
Thus, Latvia also shares similarities with Cyprus and Estonia, since the Environmental performance is the one that drives the final sustainability index to high values. However, contrary to the other two countries, Latvia has two sub-indicators with zero contribution to the calculations according to the Random Forest calculations: the Economic and Research and Development sub-indicators. Consequently, the implications for policy makers could be on the one hand, to retain the efforts on the environmental sustainability, while on the other hand to concentrate resources on the economy and research and development capacity of the country.

Romania

Figure 18 below illustrates the results from the employment of the three ML techniques in the generated sustainability indices of Romania under the proposed variation.

The boosting regression algorithm indicated that the overall Social performance is the one with the largest relative influence to the final sustainability calculations. This result is supported both by the CART tree and the Random Forest. Moreover, the latter also illustrates that the overall Environmental performance can have a positive impact on the country’s sustainability.

Romania



Relative Influence

	Relative Influence
Soc0	99.495
Econ0	0.194
RD0	0.140
Env0	0.137
Env1	0.035
Econ1	0.000
Econ2	0.000
Env2	0.000
SOc1	0.000
Soc2	0.000
RD1	0.000
RD2	0.000

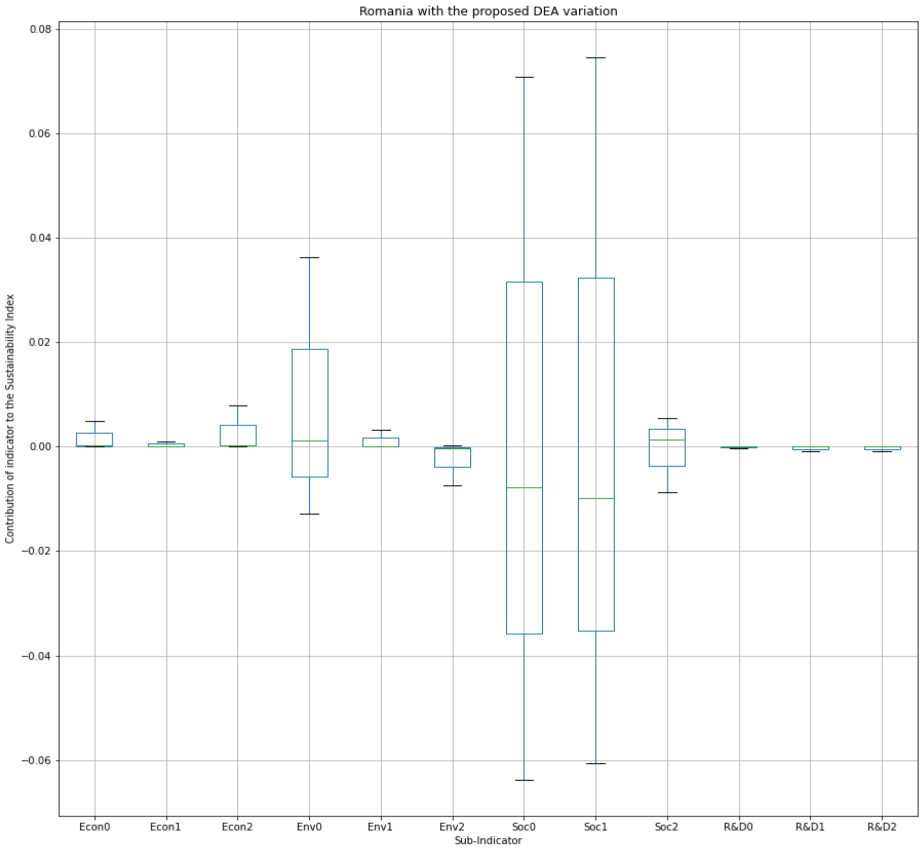


Figure 18 CART tree (upper left), relative influence of sub-indicators from boosting regression (upper right) and average contribution of the sub-indicators to the index from Random forests (bottom) for Romania under the proposed DEA variation

In conclusion, the societal dimension of sustainability is the one that drives the country's effort towards sustainable development. Thus, the implications for policy makers could be that the country needs resources and efforts for all the other dimensions in order for their sustainability to be balanced and with higher values.

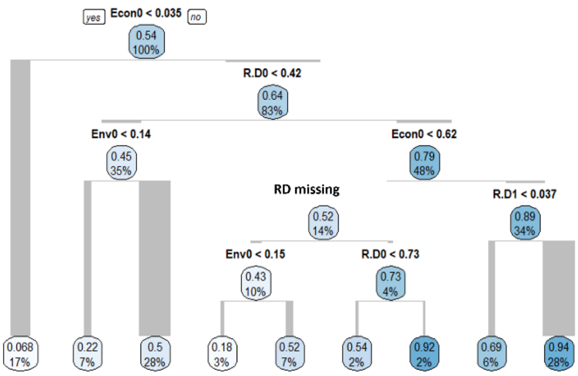
The final part of the analysis is to perform data mining on the countries when all DEA variations are used.

Belgium with all DEA variations

Figure 19 below illustrates the results from the employment of the three ML techniques in the generated sustainability indices of Belgium under all DEA variations.

Contrary to the previous results for the country, when all DEA variations are used, the sub-indicators that have the largest influence on the sustainability index are the overall Economic, Social and Research and Development performances. Moreover, the CART tree becomes more complex with more branches, however, the largest values of sustainability are observed when the individual sub-indicators have in general larger values.

Belgium all DEA variations



Relative Influence

	Relative Influence
Econ0	30.474
Soc0	23.965
R&D0	19.509
Econ2	10.227
R&D1	4.437
Env1	3.779
Env2	2.798
Econ1	1.331
Env0	1.264
Soc2	1.203
Soc1	1.012
R&D2	0.000

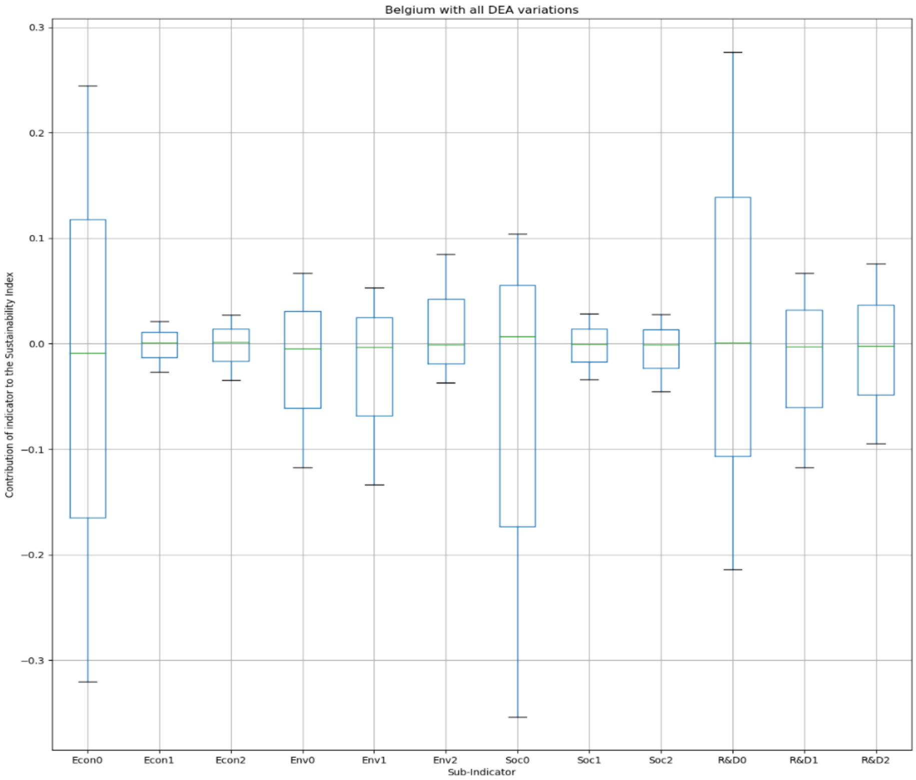


Figure 19 CART tree (upper left), relative influence of sub-indicators from boosting regression (upper right) and average contribution of the sub-indicators to the index from Random forests (bottom) for Belgium under all DEA variations

In addition, the boxplot of the Random Forest results illustrate that the overall Social performance can have the most negative impact on the final sustainability index. Contrary to the results of sustainability when only the proposed DEA variation was used, the Environmental sub-indicator has a diminished role in the sustainability of the country.

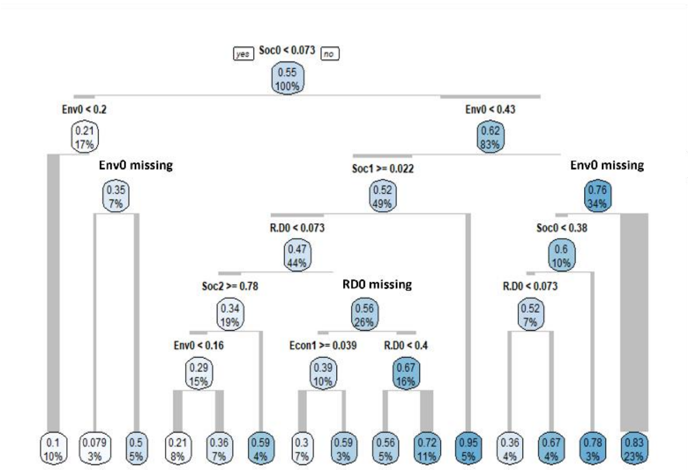
Thus, based on all calculations, the Research and Development sub-indicator of Belgium appears to drive the sustainability to high levels. However, the contribution and influence of the other three sub-indicators differ when the DEA variations change, thus the results become more sensitive and less robust.

Bulgaria with all DEA variations

Figure 20 below illustrates the results from the employment of the three ML techniques in the generated sustainability indices of Bulgaria under all DEA variations.

The results from the boosting regression algorithm illustrate that the overall Environmental, Social and Research and Development sub-indicators share the relative influence to the calculations of the sustainability index. Moreover, the CART tree (similar to the case of Belgium) becomes more complex, however it becomes clear that when the overall Environmental performance is not missing and is not smaller than 0.43 then the sustainability index has its largest values. In conclusion, results when all DEA variations are used support the importance of the Environmental dimension for Bulgaria, since it had similarly large influence when the proposed DEA variation was used.

Bulgaria all DEA variations



Relative Influence	
	Relative Influence
Env0	27.847
Soc0	25.973
R&D0	25.525
Econ0	7.035
Soc1	6.651
Env1	4.587
R&D2	1.155
R&D1	0.616
Soc2	0.610
Econ1	0.000
Econ2	0.000
Env2	0.000

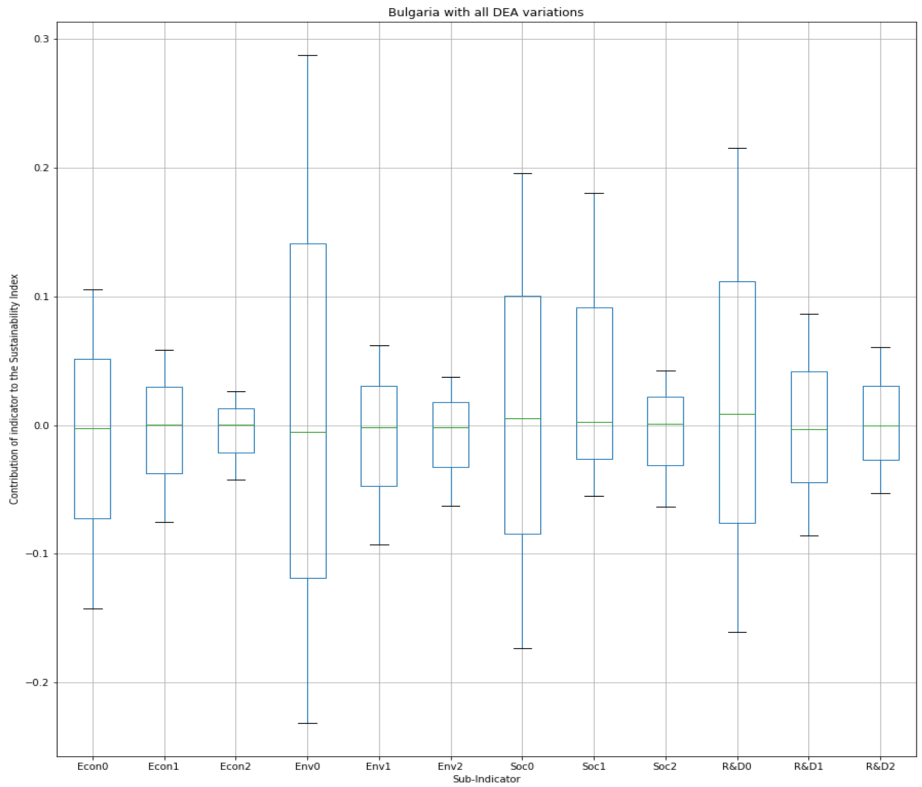


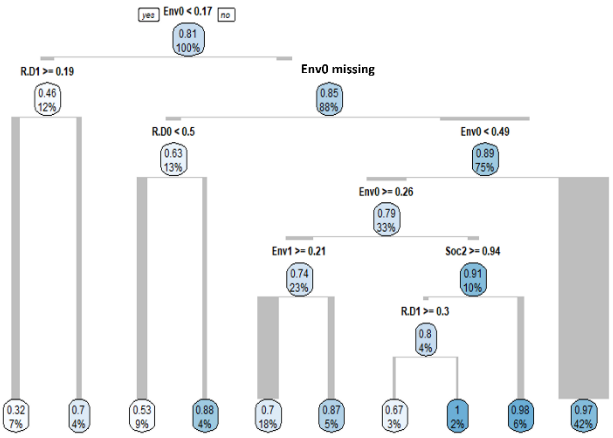
Figure 20 CART tree (upper left), relative influence of sub-indicators from boosting regression (upper right) and average contribution of the sub-indicators to the index from Random forests (bottom) for Bulgaria under all DEA variations

Estonia with all DEA variations

Figure 21 below illustrates the results from the employment of the three ML techniques in the generated sustainability indices of Estonia under all DEA variations.

The boosting regression algorithm indicates that when all DEA variations are used, the overall Environmental performance continues to have the largest relative influence in the calculations of sustainability. Moreover, the CART tree, despite its increase in complexity, illustrates that influence clearly: when the overall Environmental performance is not smaller than 0.49 then the sustainability index has a mean value of 0.97 which accounts for 42% of the calculations.

Estonia all DEA variations



Relative Influence	
	Relative Influence
Env0	51.504
Econ0	11.428
Env2	8.569
R&D2	7.079
R&D0	5.132
Econ1	4.241
Env1	4.061
Soc1	2.786
R&D1	2.565
Soc0	1.651
Econ2	0.499
Soc2	0.485

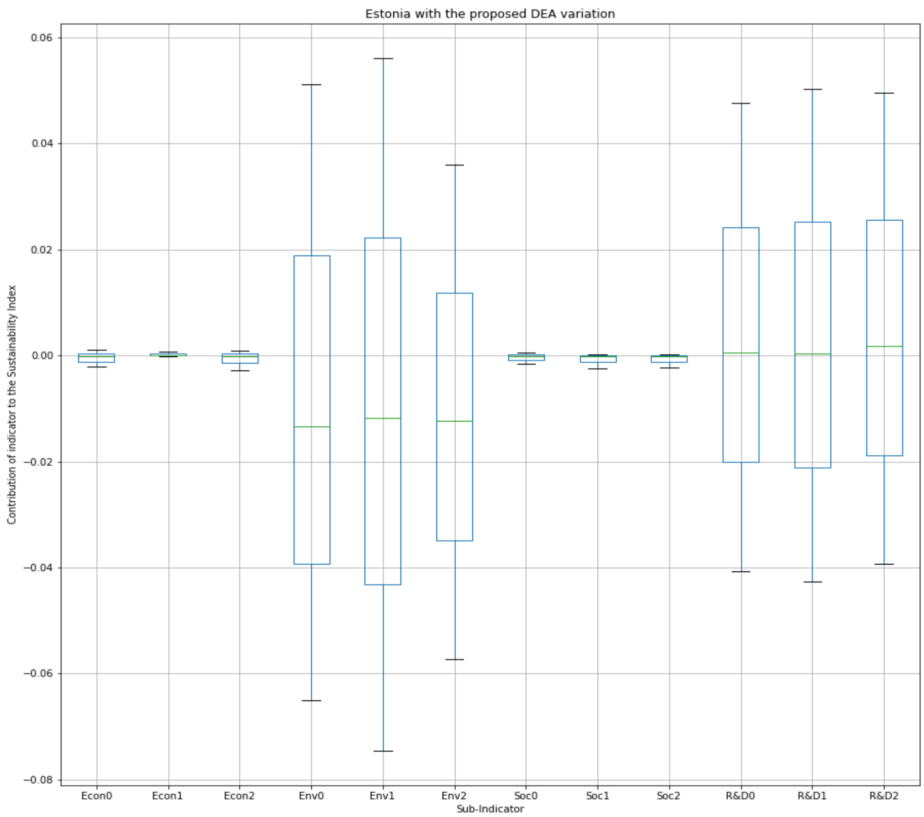


Figure 21 CART tree (upper left), relative influence of sub-indicators from boosting regression (upper right) and average contribution of the sub-indicators to the index from Random forests (bottom) for Estonia under all DEA variations

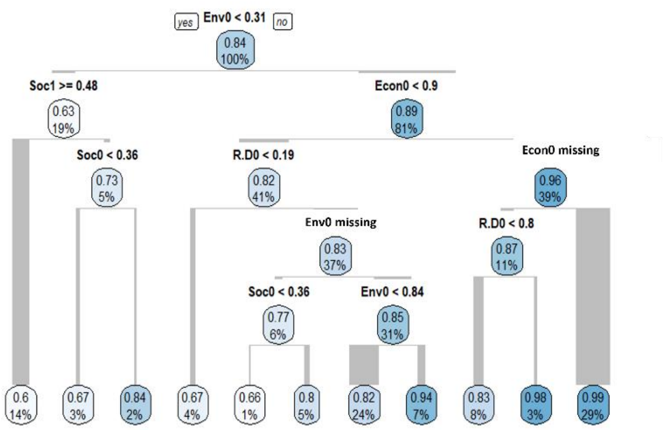
Moreover, the Random Forest algorithm produces a boxplot similar to the one when only the proposed DEA variation was used. Thus, it can be concluded that the results for Estonia are relatively robust since they show sensitivity neither to the choice of DEA variation nor to the parameters that are used in the calculations.

Cyprus with all DEA variations

Figure 22 below illustrates the results from the employment of the three ML techniques in the generated sustainability indices of Cyprus under all DEA variations.

The boosting regression algorithm indicates that the lead in the relative influence is taken by the overall Economic performance, despite the fact that the overall Environmental performance follows in the second place. The CART tree becomes also more complex compared to the CART of Cyprus when only the proposed DEA variation is used. In addition, the CART tree illustrates that an increased overall Environmental performance in combination with a very high (above 0.9) level of the overall Economic performance drives the value of the sustainability index to its highest levels.

Cyprus all DEA variations



Relative Influence

	Relative Influence
Econ0	37.063
Env0	22.530
R&D0	17.647
Soc0	11.165
Econ2	8.361
Env2	1.392
Env1	0.802
Soc1	0.525
Econ1	0.396
Soc2	0.119
R&D1	0.000
R&D2	0.000

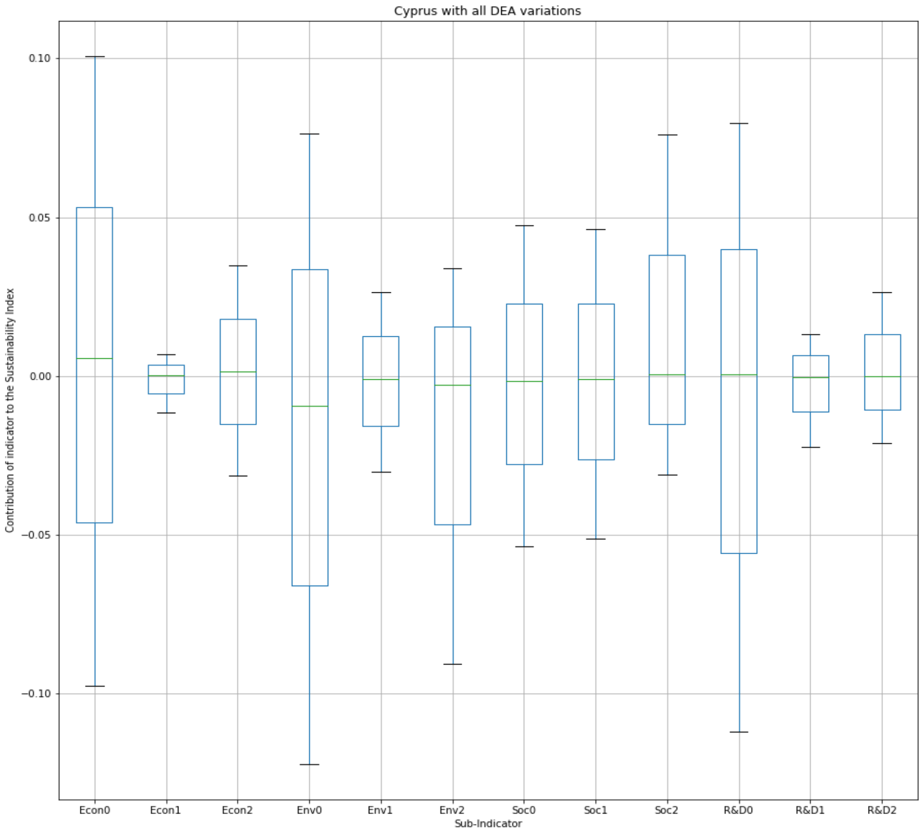


Figure 22 CART tree (upper left), relative influence of sub-indicators from boosting regression (upper right) and average contribution of the sub-indicators to the index from Random forests (bottom) for Cyprus under all DEA variations

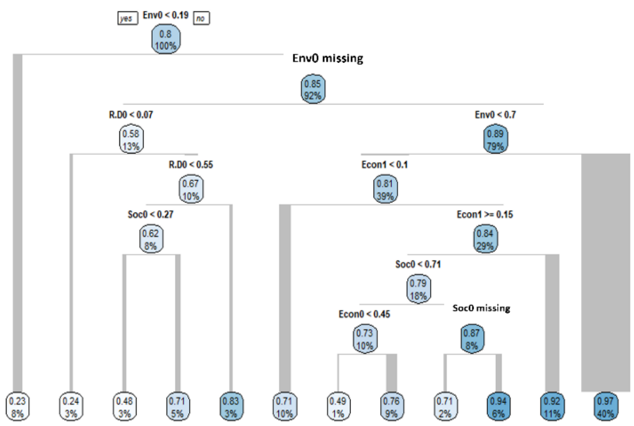
Finally, the boxplot from the Random Forest algorithm shows a more balanced contribution of the various sub-indicators. Compared to the one that was produced when only the proposed DEA variation was used, the results show an increased influence of the Research and Development sub-indicator to the calculated sustainability index. Thus for Cyprus, the results show robustness of the influence of the overall Environmental Performance. On the other hand, the Research and Development sub-indicator is more sensitive to the choice of DEA variation.

Latvia with all DEA variations

Figure 23 below illustrates the results from the employment of the three ML techniques in the generated sustainability indices of Latvia under all DEA variations.

Similar to the case when only the proposed DEA variation was used, for Latvia the importance of the overall Environmental performance remains, since it is the sub-indicator with the largest relative influence. Moreover, similar to all other cases, the CART tree becomes more complex, however, a combination of high values of the Environmental, Economic and Social sub-indicators lead to the largest values of sustainability.

Latvia all DEA variations



Relative Influence

	Relative Influence
Env0	65.222
Env1	17.999
Econ0	6.233
Soc0	5.760
Env2	3.395
R&D0	1.148
Soc2	0.242
Econ1	0.000
Econ2	0.000
Soc1	0.000
R&D1	0.000
R&D2	0.000

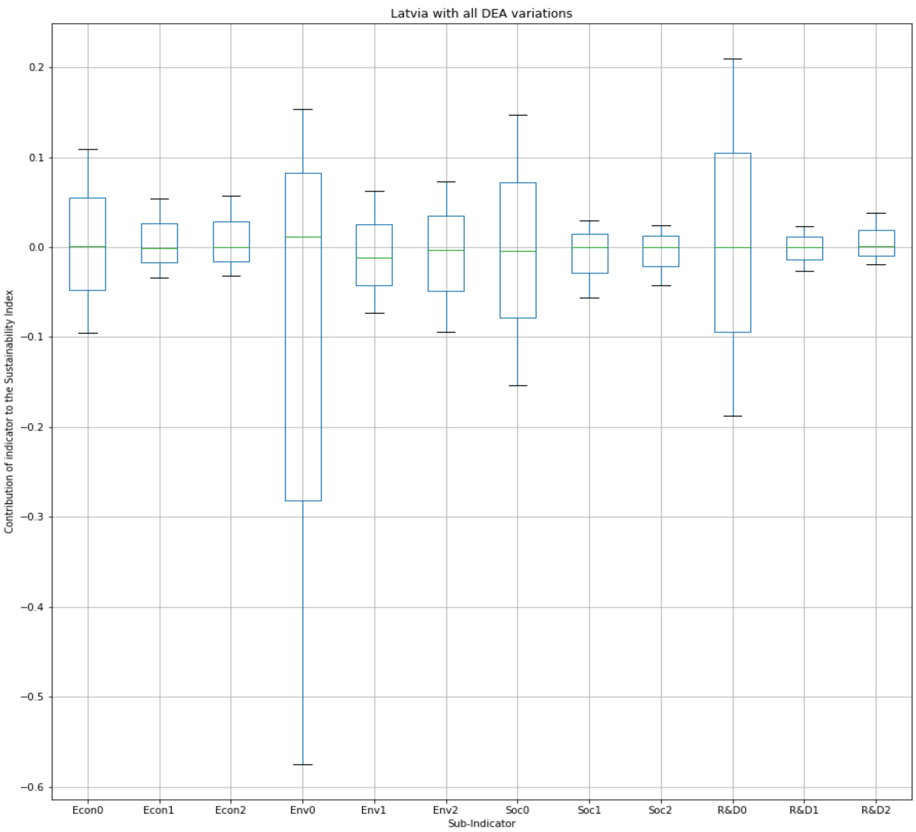


Figure 23 CART tree (upper left), relative influence of sub-indicators from boosting regression (upper right) and average contribution of the sub-indicators to the index from Random forests (bottom) for Latvia under all DEA variations

Finally, the produced boxplot illustrates that the overall Environmental sub-indicator can have the largest negative contribution to sustainability, while compared to the case when only the proposed DEA variation was used, the other dimensions participate in the value of the final sustainability index.

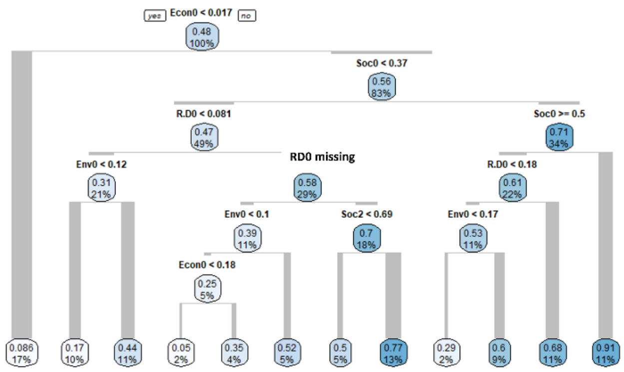
Hence, it can be concluded that the results for Latvia are relatively robust (based on the mean values of sustainability) and the importance of the Environmental sub-indicator (along with that of the Social sub-indicator) remains regardless of the choice of DEA variation. However, the Economic and Research and Development dimensions are more sensitive.

Romania with all DEA variations

Figure 24 below illustrates the results from the employment of the three ML techniques in the generated sustainability indices of Romania under all DEA variations.

The boosting regression algorithm introduces two more dimensions that share the largest influence to the final sustainability index compared to the case when only the proposed DEA variation was used. In this case, the overall Economic and Research and Development sub-indicators share the influence with the overall Social performance.

Romania all DEA variations



Relative Influence		
		Relative Influence
Econ0		26.198
R&D0		22.491
Soc0		21.896
Env0		15.318
Soc1		10.926
R&D1		2.072
R&D2		0.592
Soc2		0.346
Env2		0.163
Econ1		0.000
Econ2		0.000
Env1		0.000

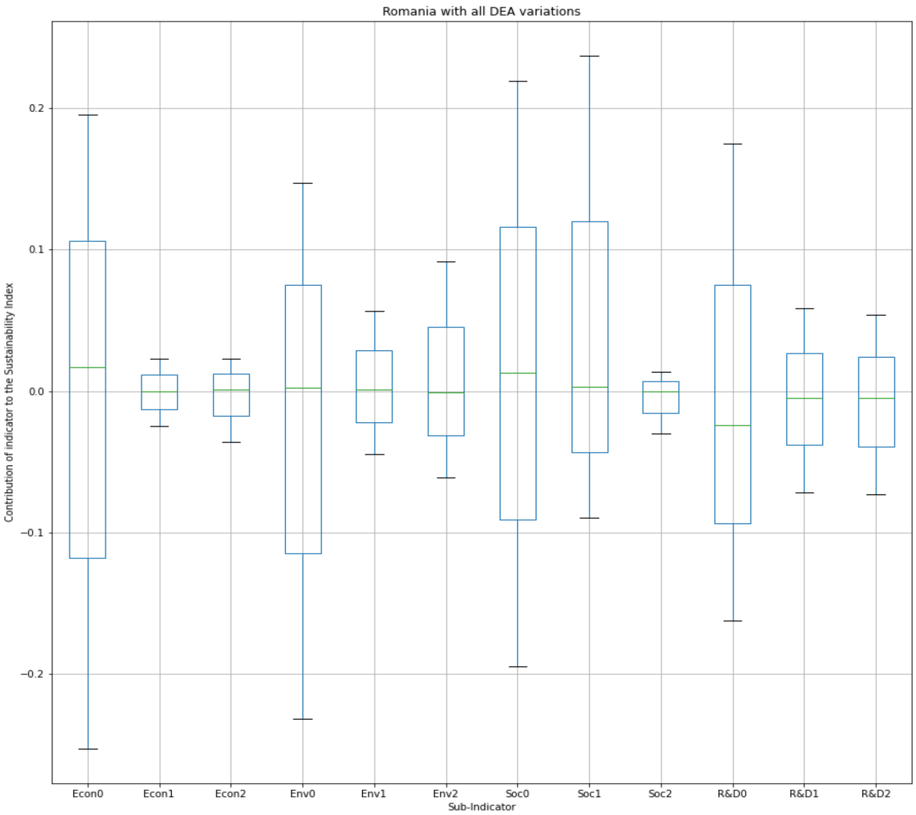


Figure 24 CART tree (upper left), relative influence of sub-indicators from boosting regression (upper right) and average contribution of the sub-indicators to the index from Random forests (bottom) for Romania under all DEA variations

Moreover, the resulted boxplot from the Random Forest algorithm supports the importance of the Social Dimension, but similarly to all other cases, the other dimensions contribute also to the calculated sustainability index.

In conclusion, when all DEA variations are used, the influence is shared among all or most of the dimensions and/or sub-indicators. This can be partly attributed to the availability of increased data. Thus, the difference with the previous process (when only the proposed DEA variation was used), a balance among the influence of the dimensions does not guarantee high values of sustainability. Finally, when the inclusion of all DEA variations does not alter significantly the mean value of sustainability and/or the most important sub-indicators (like Estonia and Latvia) then the trust in the results increases, thus making them robust.

5. Conclusions

The purpose of the current thesis was twofold:

1) to propose a new variation on two-stage Data Envelopment Analysis that attempts to intervene on the weights of the inputs, intermediates and outputs to better reflect their importance for the DMUs by considering positive and negative deviations in the calculations and limiting the distance of these deviations from the maximum and minimum values.

2) to propose a computational framework that will attempt to incorporate different perceptions (meaning different combinations of inputs and outputs) and apply it in the measurement of sustainability of the EU 28 countries, thus assisting in the transition of Data Envelopment Analysis to a more Exploratory and Investigative mode.

To achieve the objective, a series of steps was designed and developed. First, a literature review on the use of Data Envelopment Analysis in the context of sustainability was performed. The purpose of the review was to extend the literature review performed by Zhou et al. (2018) and investigate whether the lack of unified definition and methodological framework for the measurement of sustainable development has affected the research.

To do so, bibliographic databases were searched for research efforts concerning the years from 2017 until 2020. Several interesting insights are revealed in the literature. First, the vague definition of sustainability has led to different approaches on how to measure it. However, in the DEA literature the authors heavily use the 3-dimensional structure (economic, social and environmental), with individual efforts attempting to incorporate different dimensions such as technology and innovation.

Moreover, even when the 3-dimensional structure of sustainability is used, differences are observed on the combinations of inputs and outputs of the DEA model. This is to be expected since the social dimension, for example, might have a different meaning for different people. Nonetheless, each of these approaches, by using one combination of parameters excludes the other perceptions from the analysis.

Another issue that was also mentioned in the review by Zhou et al. (2018) is that the social dimension of sustainability has been underrepresented in the studies so far. In fact, a lot of environmental and energy studies use the same combinations of inputs and outputs as those that explicitly measure sustainability. Despite the recent approaches that seek to remedy the issue, there is still a lot of effort needed to fully capture the multi-dimensional notion of sustainability in a coherent and mathematically sound way.

Moreover, in the last few years the research on sustainability has shifted towards urban environments and within country regions. A large portion of the research activity concerned Chinese regions. One possible explanation could be that the research community is focused on investigating the possible visible and not visible consequences that the country's economic development could have on the environment and the society. Finally, from the papers that were reviewed, those that focus on the comparison of the EU countries with regards to sustainability appear to be lagging in numbers. Nonetheless, it is deemed important to address the specific gap especially since the Sustainable Development Goals are part of the European Policy Framework.

All these gaps result in different measurements of sustainability, which may have a negative impact on the robustness of the research efforts. Equally important, this fact could have negative implications in policy making. First, decisions based on those measurements may be rendered ineffective because the measurements cannot really capture the full scope of sustainability. Moreover, these decisions could produce undesired consequences in areas of public life that were not addressed in the analysis. Finally, these differences in measurements have an effect on communicating policy efforts to the general public. As a result, citizens may be less inclined to abide by policies if these appear to be based on contested measurements.

A second step was to propose and develop a new variation of a two-stage DEA model which takes advantage of deviational variables to handle the variations attributed to the weights distribution. The contribution to the literature is the inclusion of deviations (for the individual stages) in a two-stage DEA model. The deviational variables provide a vehicle of interventions on the weights distribution through the goal programming formulation inherent to DEA thus reducing the fluctuations of the weight distribution. In addition, several Lemmas and a Theorem regarding the model are provided.

The proposed variation was used to calculate the environmental performance of EU countries and a comparison was provided with the two-stage variations of Chen et al.

(2009). The results illustrate that the two variations do not share the same values for environmental performances and obviously there exist changes in the rankings. This is attributed to the fact that the proposed two-stage DEA variation uses an augmented set of constraints and additional variables that impose limitations to the distributions of the calculated weights (as desired).

Furthermore, to test the robustness of the new two-stage, DEA variation, a rank reversal analysis was performed to investigate how the results change if DMUs are added to or deleted from the original set. To have a better understanding of the sensitivity, a new was developed that analyses not only how many DMUs remain common in the rankings, but also how similar are the rankings themselves.

The final step of the thesis was to develop a DEA-ML computational framework for the evaluation of EU policies, especially in the case of multi-dimensional constructs like sustainability. The proposed framework relies (a) on the use of multi-level Data Envelopment Analysis in combination with classic DEA variations, (b) on the application of these models for different combinations of inputs, intermediates and outputs that represent different perceptions of what sustainability is and finally (c) on the exploratory analysis on the outcomes with the use of Machine Learning methodologies such as CART decision trees.

In this direction, it is worth pointing out that this framework follows the school of thought of Exploratory Modeling and Analysis that supports the use of models and quantitative methods in an exploratory way, not to predict or monitor policy cycles accurately (which can be considered impossible), but to gain insights by incorporating different perceptions and methodological approaches at the same problem, thus increasing the robustness of the results (Moallemi, Kwakkel, de Haan, & Bryan, 2020). In the current thesis, this approach was applied in the measurement of sustainability of EU 28 countries. Concretely, the computational experiments illustrated that the different perceptions of how sustainability is measured, and the use of different DEA variations (hence different methodological frameworks) affect the final results. More specifically, the results illustrated that a balance among the performance of various dimensions can be a good policy to achieve sustainable development and when the inclusion of all DEA variations does not alter significantly the mean value of sustainability then the trust in the results increases, thus making them robust.

Finally, the blend of DEA with machine learning (applied on the results of DEA for the various scenarios) revealed insights on the areas that policy makers could direct investments to increase sustainability. In addition, ML application contributed in the identification of the most important features of sustainability for the various countries something that could have direct implications in the area of EU policy making: for example, countries that share similar features that drive the behavior of sustainability could be grouped together in clusters and policies, laws, regulations etc. could be adapted to those clusters in order to boost the particular features that would increase their sustainability. As a result, policy making has the potential to become customized (adapted to the specifics of each group) without missing its overall and principal theme of pursuing sustainable development. This adaptive and adaptable policy making could greatly be of assistance especially when new countries are negotiating their entry to the Union; based on the features that affect the sustainability of the new countries, they could follow the regulations and laws of the appropriate cluster. Finally, the inclusion of new layers and perceptions renders the algorithms more inclusive and participatory, increasing their transparency, thus improving the trust to the final results.

However, the thesis is not without limitations. Regarding the definition and/or methodological framework for sustainability, a new approach could be taken, a bottom-up approach, where scientists propose a unified methodological and/or computational framework that attempts to mitigate the limitations of individual methods and integrates different and diverse definitions of sustainability into the same measurement.

Furthermore, the addition of new layers and perceptions means that the process becomes more computationally costly and new conceptual questions arise; for example, when is it valid to stop adding new perceptions and report the conclusions? How many new perceptions are necessary to get a clearer picture?

Such questions will drive future research efforts of the current study. Further directions of research include the development of a user interface that could be used by non-experts, and the inclusion of supplementary variations of DEA, the generation of additional sub-indicators along with various data sources. Finally, the framework could be enriched with methods other than DEA which would allow the Machine Learning techniques to identify not only the differences in the context (sub-indicators) but also in the method that was used.

Finally, the current thesis can be seen as an example for the use of an Exploratory approach to Data Envelopment Analysis (and Operational Research) and could be a useful source on future research efforts on sustainability and/or Data Envelopment Analysis.

Bibliography

Adler, M. (2012). *Well-being and fair distribution: beyond cost-benefit analysis.* Oxford University Press.

Aggestam, V., & Buick, J. (2017). A comparative analysis of vehicle-related greenhouse gas emissions between organic and conventional dairy production. *The Journal of dairy research, 84,* 360.

Amiri, M., Zandieh, M., Vahdani, B., Soltani, R., & Roshanaei, V. (2010). An integrated eigenvector–DEA–TOPSIS methodology for portfolio risk evaluation in the FOREX spot market. *Expert Systems with Applications, 37,* 509–516.

Arnade, C. A. (1994). *Using data envelopment analysis to measure international agricultural efficiency and productivity.* US Department of Agriculture, Economic Research Service.

Atici, K. B., & Podinovski, V. V. (2015). Using data envelopment analysis for the assessment of technical efficiency of units with different specialisations: An application to agriculture. *Omega, 54,* 72–83.

Aydin, N., & Yurdakul, G. (2020). Assessing countries' performances against COVID-19 via WSIDEA and machine learning algorithms. *Applied Soft Computing, 97,* 106792.

Azadi, M., Izadikhah, M., Ramezani, F., & Hussain, F. K. (2020). A mixed ideal and anti-ideal DEA model: an application to evaluate cloud service providers. *IMA Journal of Management Mathematics, 31,* 233–256.

Bailey, D. H., Borwein, J. M., & Bradley, D. M. (2006). Experimental Determination of Apéry-like Identities for ς (2n+ 2). *Experimental Mathematics, 15,* 281–289.

Bal, H., Örkcü, H. H., & Çelebioğlu, S. (2010). Improving the discrimination power and weights dispersion in the data envelopment analysis. *Computers & Operations Research, 37,* 99–107.

Banker, R. D., Charnes, A., & Cooper, W. W. (1984). Some models for estimating technical and scale inefficiencies in data envelopment analysis. *Management science, 30,* 1078–1092.

Bankes, S. C. (1992). *Exploratory modeling and the use of simulation for policy analysis.* Tech. rep., Rand Corp Santa Monica CA.

Bankes, S. C. (1993). Exploratory modeling for policy analysis. *Operations research, 41*, 435–449.

Barzegarinegad, A., Jahanshahloo, G., & Rostamy-Malkhalifeh, M. (2014). A full ranking for decision making units using ideal and anti-ideal points in DEA. *The Scientific World Journal, 2014*.

Belton, V., & Gear, T. (1983). On a short-coming of Saaty's method of analytic hierarchies. *Omega, 11*, 228–230.

Benítez-Peña, S., Bogetoft, P., & Morales, D. R. (2020). Feature selection in data envelopment analysis: A mathematical optimization approach. *Omega, 96*, 102068.

Bian, Y., & Yang, F. (2010). Resource and environment efficiency analysis of provinces in China: A DEA approach based on Shannon’s entropy. *Energy Policy, 38*, 1909–1917.

Biresselioglu, M. E., Demir, M. H., & Turan, U. (2018). Trinity on thin ice: Integrating three perspectives on the European Union's likelihood of achieving energy and climate targets. *Energy Research & Social Science, 42*, 247–257.

Blancard, S., & Hoarau, J.-F. (2013). A new sustainable human development indicator for small island developing states: A reappraisal from data envelopment analysis. *Economic Modelling, 30*, 623–635.

Bojnec, Š., Fertő, I., Jámbor, A., & Tóth, J. (2014). Determinants of technical efficiency in agriculture in new EU member states from Central and Eastern Europe. *Acta Oeconomica, 64*, 197–217.

Borucke, M., Moore, D., Cranston, G., Gracey, K., Iha, K., Larson, J., . . . Galli, A. (2013). Accounting for demand and supply of the biosphere's regenerative capacity: The National Footprint Accounts’ underlying methodology and framework. *Ecological indicators, 24*, 518–533.

Brundtland, G. H., Khalid, M., Agnelli, S., Al-Athel, S., & Chidzero, B. (1987). Our common future. *New York*, 8.

Bruni, M. E., Guerriero, F., & Patitucci, V. (2011). Benchmarking sustainable development via data envelopment analysis: an Italian case study.

Bühlmann, P., & Yu, B. (2003). Boosting with the L 2 loss: regression and classification. *Journal of the American Statistical Association, 98*, 324–339.

Callens, I., & Tyteca, D. (1999). Towards indicators of sustainable development for firms: a productive efficiency perspective. *Ecological Economics, 28*, 41–53.

Camarero, M., Castillo, J., Picazo-Tadeo, A. J., & Tamarit, C. (2013). Eco-efficiency and convergence in OECD countries. *Environmental and Resource Economics, 55*, 87–106.

Carboni, O. A., & Russu, P. (2018). Measuring and forecasting regional environmental and economic efficiency in Italy. *Applied Economics, 50*, 335–353.

Charnes, A., & Cooper, W. W. (1962). Programming with linear fractional functionals. *Naval Research logistics quarterly, 9*, 181–186.

Charnes, A., Cooper, W. W., & Rhodes, E. (1978). Measuring the efficiency of decision making units. *European journal of operational research, 2*, 429–444.

Chen, C.-M., Du, J., Huo, J., & Zhu, J. (2012). Undesirable factors in integer-valued DEA: Evaluating the operational efficiencies of city bus systems considering safety records. *Decision Support Systems, 54*, 330–335.

Chen, J., Song, M., & Xu, L. (2015). Evaluation of environmental efficiency in China using data envelopment analysis. *Ecological indicators, 52*, 577–583.

Chen, L., & Jia, G. (2017). Environmental efficiency analysis of China's regional industry: a data envelopment analysis (DEA) based approach. *Journal of Cleaner Production, 142*, 846–853.

Chen, L., Wang, Y., Lai, F., & Feng, F. (2017). An investment analysis for China's sustainable development based on inverse data envelopment analysis. *Journal of cleaner production, 142*, 1638–1649.

Chen, S.-J., & Hwang, C.-L. (1992). Fuzzy multiple attribute decision making methods. *Fuzzy multiple attribute decision making*, 289–486.

Chen, Y., & Zhu, J. (2004). Measuring information technology's indirect impact on firm performance. *Information Technology and Management, 5*, 9–22.

Chen, Y., Cook, W. D., Li, N., & Zhu, J. (2009). Additive efficiency decomposition in two-stage DEA. *European Journal of Operational Research, 196*, 1170–1176.

Cherchye, L., Moesen, W., Rogge, N., & Van Puyenbroeck, T. (2007). An introduction to 'benefit of the doubt'composite indicators. *Social indicators research, 82*, 111–145.

Coelli, T. J., Rao, D. S., O'Donnell, C. J., & Battese, G. E. (2005). *An introduction to efficiency and productivity analysis.* springer science & business media.

Coli, M., Nissi, E., & Rapposelli, A. (2011). Monitoring environmental efficiency: an application to Italian provinces. *Environmental Modelling & Software, 26*, 38–43.

Contreras, I. (2020). A review of the literature on DEA models under common set of weights. *Journal of Modelling in Management, 15*, 1277–1300.

Cook, W. D., Liang, L., & Zhu, J. (2010). Measuring performance of two-stage network structures by DEA: a review and future perspective. *Omega*, *38*, 423–430.

Cooper, W. W., Seiford, L. M., & Zhu, J. (2011). *Handbook on data envelopment analysis.* Springer Science & Business Media.

Cucchiella, F., D'Adamo, I., Gastaldi, M., & Miliacca, M. (2018). Efficiency and allocation of emission allowances and energy consumption over more sustainable European economies. *Journal of Cleaner Production, 182*, 805–817.

Dakpo, K. H., Jeanneaux, P., & Latruffe, L. (2016). Modelling pollution-generating technologies in performance benchmarking: Recent developments, limits and future prospects in the nonparametric framework. *European Journal of Operational Research*, *250*, 347–359.

Daniels, P. L., & Moore, S. (2001). Approaches for quantifying the metabolism of physical economies: Part I: Methodological overview. *Journal of Industrial Ecology, 5*, 69–93.

Debreu, G. (1951). The coefficient of resource utilization. *Econometrica: Journal of the Econometric Society*, 273–292.

Delias, P., Grigoroudis, E., & Matsatsinis, N. F. (2018). Group Decision Support for Crop Planning: A Case Study to Guide the Process of Preferences Elicitation. In *Elicitation* (pp. 529–542). Springer.

Despotis, D. K. (2005). A reassessment of the human development index via data envelopment analysis. *Journal of the operational research society, 56*, 969–980.

Despotis, D. K., & Koronakos, G. (2014). Efficiency assessment in two-stage processes: A novel network DEA approach. *Procedia Computer Science, 31*, 299–307.

Dhungana, B. R., Nuthall, P. L., & Nartea, G. V. (2004). Measuring the economic inefficiency of Nepalese rice farms using data envelopment analysis. *Australian Journal of Agricultural and Resource Economics, 48*, 347–369.

Diewert, W. E. (1973). Functional forms for profit and transformation functions. *Journal of Economic Theory, 6*, 284–316.

DiMaria, C.-H. (2019). An indicator for the economic performance and ecological sustainability of nations. *Environmental Modeling & Assessment*, *24*, 279–294.

Drucker, P. (2014). *Innovation and entrepreneurship.* Routledge.

Du, H., Matisoff, D. C., Wang, Y., & Liu, X. (2016). Understanding drivers of energy efficiency changes in China. *Applied Energy, 184*, 1196–1206.

Du, J., Chen, Y., & Huang, Y. (2018). A modified malmquist-luenberger productivity index: Assessing environmental productivity performance in China. *European journal of operational research, 269*, 171–187.

Dunning, I., Huchette, J., & Lubin, M. (2017). JuMP: A Modeling Language for Mathematical Optimization. *SIAM Review, 59*, 295-320. doi:10.1137/15M1020575

Dyer, J. S. (1990). Remarks on the analytic hierarchy process. *Management science, 36*, 249–258.

Ehrenfeld, J. R. (2004). Searching for sustainability: No quick fix. *Reflections*, *5*, 1–13.

Ervural, B. C., Zaim, S., & Delen, D. (2018). A two-stage analytical approach to assess sustainable energy efficiency. *Energy, 164*, 822–836.

Ester, M., Kriegel, H.-P., Sander, J., Xu, X., & others. (1996). A density-based algorithm for discovering clusters in large spatial databases with noise. *kdd*, *96*, pp. 226–231.

Esty, D. C., Levy, M. A., Srebotnjak, T., de Sherbinin, A., Kim, C. H., & Anderson, B. (2006). *Pilot 2006 environmental performance index.* Tech. rep.

Fang, L. (2020). Opening the “black box” of environmental production technology in a nonparametric analysis. *European Journal of Operational Research*.

Färe, R., & Grosskopf, S. (1996). Productivity and intermediate products: A frontier approach. *Economics letters, 50*, 65–70.

Färe, R., & Grosskopf, S. (2009). A comment on weak disposability in nonparametric production analysis. *American Journal of Agricultural Economics, 91*, 535–538.

Färe, R., & Whittaker, G. (1995). An intermediate input model of dairy production using complex survey data. *Journal of Agricultural Economics, 46*, 201–213.

Farrell, M. J. (1957). The measurement of productive efficiency. *Journal of the Royal Statistical Society: Series A (General), 120*, 253–281.

Feng, C., Zhang, H., & Huang, J.-B. (2017). The approach to realizing the potential of emissions reduction in China: An implication from data envelopment analysis. *Renewable and Sustainable Energy Reviews, 71*, 859–872.

Førsund, F. R., & Sarafoglou, N. (2002). On the origins of data envelopment analysis. *Journal of Productivity Analysis, 17*, 23–40.

Fraedrich, R., Schneider, J., & Westermann, R. (2009). Exploring the millennium run-scalable rendering of large-scale cosmological datasets. *IEEE Transactions on Visualization and Computer Graphics, 15*, 1251–1258.

Freeman III, A. M., Haveman, R. H., & Kneese, A. V. (1973). *Economics of environmental policy*. New, York, NY: John Wiley and Sons, Inc.

Friedman, J. H. (2017). *The elements of statistical learning: Data mining, inference, and prediction.* Springer open.

Fuglestvedt, J., Berntsen, T., Myhre, G., Rypdal, K., & Skeie, R. B. (2008). Climate forcing from the transport sectors. *Proceedings of the National Academy of Sciences, 105*, 454–458.

Ganley, J. A., & Cubbin, J. S. (1992). *Public sector efficiency measurement: Applications of data envelopment analysis.* Elsevier Science Inc.

Gao, T., Erokhin, V., & Arskiy, A. (2019). Dynamic optimization of fuel and logistics costs as a tool in pursuing economic sustainability of a farm. *Sustainability, 11*, 5463.

García-Cascales, M. S., & Lamata, M. T. (2012). On rank reversal and TOPSIS method. *Mathematical and Computer Modelling, 56*, 123–132.

Georgiou, A. C., Thanassoulis, E., & Papadopoulou, A. (2021). Using Data Envelopment Analysis in Markovian Decision Making. *European Journal of Operational Research*.

Gharakhani, D., Eshlaghy, A. T., Hafshejani, K. F., Mavi, R. K., & Lotfi, F. H. (2018). Common weights in dynamic network DEA with goal programming approach for performance assessment of insurance companies in Iran. *Management Research Review, 41*, 920–938.

Ghasemi, M.-R., Ignatius, J., & Emrouznejad, A. (2014). A bi-objective weighted model for improving the discrimination power in MCDEA. *European Journal of Operational Research, 233*, 640–650.

Gibson, R. B. (2006). Sustainability assessment: basic components of a practical approach. *Impact assessment and project appraisal, 24*, 170–182.

Gonzalez-Garcia, S., Manteiga, R., Moreira, M. T., & Feijoo, G. (2018). Assessing the sustainability of Spanish cities considering environmental and socio-economic indicators. *Journal of Cleaner Production, 178*, 599–610.

Guo, X.-D., Zhu, L., Fan, Y., & Xie, B.-C. (2011). Evaluation of potential reductions in carbon emissions in Chinese provinces based on environmental DEA. *Energy Policy, 39*, 2352–2360.

Hajiagha, S. H., Hashemi, S. S., & Mahdiraji, H. A. (2016). Fuzzy C-means based data envelopment analysis for mitigating the impact of units' heterogeneity. *Kybernetes, 3*, 536-551.

Halkos, G. E., Tzeremes, N. G., & Kourtzidis, S. A. (2014). A unified classification of two-stage DEA models. *Surveys in operations research and management science, 19,* 1–16.

Halkos, G. E., Tzeremes, N. G., & Kourtzidis, S. A. (2015). Regional sustainability efficiency index in Europe: an additive two-stage DEA approach. *Operational Research, 15,* 1–23.

Han, C., Pervez, A., Wu, J., Shen, X., & Zhang, D. (2020). Home-Delivery-Oriented Agri-Food Supply Chain Alliance: Framework, Management Strategies, and Cooperation Stability Control. *Sustainability, 12*, 6547.

Hassanzadeh, A., Yousefi, S., Saen, R. F., & Hosseininia, S. S. (2018). How to assess sustainability of countries via inverse data envelopment analysis? *Clean Technologies and Environmental Policy, 20*, 29–40.

He, J., Wan, Y., Feng, L., Ai, J., & Wang, Y. (2016). An integrated data envelopment analysis and emergy-based ecological footprint methodology in evaluating sustainable development, a case study of Jiangsu Province, China. *Ecological indicators, 70*, 23–34.

Hoang, V.-N., & Alauddin, M. (2012). Input-orientated data envelopment analysis framework for measuring and decomposing economic, environmental and ecological efficiency: an application to OECD agriculture. *Environmental and Resource Economics, 51*, 431–452.

Hoang, V.-N., & Rao, D. P. (2010). Measuring and decomposing sustainable efficiency in agricultural production: A cumulative exergy balance approach. *Ecological economics, 69*, 1765–1776.

Hu, J.-L., Sheu, H.-J., & Lo, S.-F. (2005). Under the shadow of Asian Brown Clouds: Unbalanced regional productivities in China and environmental concerns. *The International Journal of Sustainable Development & World Ecology, 12*, 429–442.

Hu, M. Y., Guo, W., Zhao, N., & Wang, L. (2012). Empirical Analysis on IT Investments Efficiency and Influence Factors of Equipment Manufacturing Industry Based on DEA-CART Model. *Applied Mechanics and Materials, 197*, pp. 638–642.

Huang, J., Yang, X., Cheng, G., & Wang, S. (2014). A comprehensive eco-efficiency model and dynamics of regional eco-efficiency in China. *Journal of Cleaner Production, 67*, 228–238.

Huppes, G., & Ishikawa, M. (2005). Eco-efficiency and Its xsTerminology. *Journal of Industrial ecology, 9*, 43–46.

Hwang, C.-L., & Yoon, K. (1981). Methods for multiple attribute decision making. In *Multiple attribute decision making* (pp. 58–191). Springer.

Ishizaka, A., & Nemery, P. (2013). *Multi-criteria decision analysis: methods and software.* John Wiley & Sons.

Jahanshahloo, G. R., Lotfi, F. H., Khanmohammadi, M., Kazemimanesh, M., & Rezaie, V. (2010). Ranking of units by positive ideal DMU with common weights. *Expert Systems with Applications, 37*, 7483–7488.

James, G., Witten, D., Hastie, T., & Tibshirani, R. (2013). *An introduction to statistical learning.* New York: Springer.

Joint Research Centre-European Commission. (2008). *Handbook on constructing composite indicators: methodology and user guide.* Paris: OECD publishing.

Kam, H. T., & others. (1995). Random decision forest. *Proceedings of the 3rd international conference on document analysis and recognition*, *1416*, p. 278282.

Kamvysi, K., Gotzamani, K., Georgiou, A. C., & Andronikidis, A. (2010). Integrating DEAHP and DEANP into the quality function deployment. *The TQM Journal*.

Kao, C., & Hung, H.-T. (2005). Data envelopment analysis with common weights: the compromise solution approach. *Journal of the Operational Research Society, 56*, 1196–1203.

Kao, C., & Hwang, S.-N. (2008). Efficiency decomposition in two-stage data envelopment analysis: An application to non-life insurance companies in Taiwan. *European journal of operational research, 185*, 418–429.

Kates, R. W., Clark, W. C., Corell, R., Hall, J. M., Jaeger, C. C., Lowe, I., . . . others. (2001). Sustainability science. *Science, 292*, 641–642.

Khalili-Damghani, K., & Fadaei, M. (2018). A comprehensive common weights data envelopment analysis model: ideal and anti-ideal virtual decision making units approach. *Journal of Industrial and Systems Engineering, 11*, 281–306.

Kiaei, H., & Matin, R. K. (2020). Common set of weights and efficiency improvement on the basis of separation vector in two-stage network data envelopment analysis. *Mathematical Sciences, 14*, 53–65.

Kissinger, M. (2012). International trade related food miles–The case of Canada. *Food policy, 37*, 171–178.

Kočišová, K. (2015). Application of the DEA on the measurement of efficiency in the EU countries. *Agricultural Economics, 61*, 51–62.

Kong, F., Wei, W., & Gong, J.-H. (2016). Rank reversal and rank preservation in ANP method. *Journal of Discrete Mathematical Sciences and Cryptography, 19*, 821–836.

Krishna, K., & Murty, M. N. (1999). Genetic K-means algorithm. *IEEE Transactions on Systems, Man, and Cybernetics, Part B (Cybernetics), 29*, 433–439.

Kritikos, M. N. (2017). A full ranking methodology in data envelopment analysis based on a set of dummy decision making units. *Expert Systems with Applications, 77*, 211–225.

Kuosmanen, T. (2005). Weak disposability in nonparametric production analysis with undesirable outputs. *American Journal of Agricultural Economics, 87*, 1077–1082.

Kuosmanen, T., & Kortelainen, M. (2005). Measuring eco-efficiency of production with data envelopment analysis. *Journal of Industrial Ecology, 9*, 59–72.

Kwakkel, J. H., & Pruyt, E. (2013). Exploratory Modeling and Analysis, an approach for model-based foresight under deep uncertainty. *Technological Forecasting and Social Change, 80*, 419–431.

Lee, C.-Y., & Cai, J.-Y. (2020). LASSO variable selection in data envelopment analysis with small datasets. *Omega, 91*, 102019.

Lee, S. (2010). Using data envelopment analysis and decision trees for efficiency analysis and recommendation of B2C controls. *Decision Support Systems, 49*, 486–497.

Lertworasirikul, S., Charnsethikul, P., & Fang, S.-C. (2011). Inverse data envelopment analysis model to preserve relative efficiency values: The case of variable returns to scale. *Computers & Industrial Engineering, 61*, 1017–1023.

Lewis, M., Li, H., & Sycara, K. (2021). Deep learning, transparency, and trust in human robot teamwork. In *Trust in Human-Robot Interaction* (pp. 321–352). Elsevier.

Li, A., Peng, D., Wang, D., & Yao, X. (2017). Comparing regional effects of climate policies to promote non-fossil fuels in China. *Energy, 141*, 1998–2012.

Li, A., Zhang, A., Zhou, Y., & Yao, X. (2017). Decomposition analysis of factors affecting carbon dioxide emissions across provinces in China. *Journal of Cleaner Production, 141*, 1428–1444.

Li, F., Liu, X., Hu, D., Wang, R., Yang, W., Li, D., & Zhao, D. (2009). Measurement indicators and an evaluation approach for assessing urban sustainable development: A case study for China's Jining City. *Landscape and urban planning, 90*, 134–142.

Li, K., & Lin, B. (2015). Metafroniter energy efficiency with CO2 emissions and its convergence analysis for China. *Energy Economics, 48*, 230–241.

Li, L.-B., & Hu, J.-L. (2012). Ecological total-factor energy efficiency of regions in China. *Energy Policy, 46*, 216–224.

Li, N., Jiang, Y., Mu, H., & Yu, Z. (2018). Efficiency evaluation and improvement potential for the Chinese agricultural sector at the provincial level based on data envelopment analysis (DEA). *Energy, 164*, 1145–1160.

Li, X.-B., & Reeves, G. R. (1999). A multiple criteria approach to data envelopment analysis. *European Journal of Operational Research, 115*, 507–517.

Liang, L., Cook, W. D., & Zhu, J. (2008). DEA models for two-stage processes: Game approach and efficiency decomposition. *Naval Research Logistics (NRL), 55*, 643–653.

Liang, L., Yang, F., Cook, W. D., & Zhu, J. (2006). DEA models for supply chain efficiency evaluation. *Annals of operations research, 145*, 35–49.

Likas, A., Vlassis, N., & Verbeek, J. J. (2003). The global k-means clustering algorithm. *Pattern recognition, 36*, 451–461.

Lin, T.-Y., & Chiu, S.-H. (2018). Sustainable Performance of Low-Carbon Energy Infrastructure Investment on Regional Development: Evidence from China. *Sustainability, 10*, 4657.

Long, R., Wang, H., & Chen, H. (2016). Regional differences and pattern classifications in the efficiency of coal consumption in China. *Journal of Cleaner Production, 112*, 3684–3691.

Longo, J., & McNutt, K. (2018). From policy analysis to policy analytics. *Policy Analysis in Canada*, 367–389.

Lotfi, F. H., Noora, A. A., Jahanshahloo, G. R., & Reshadi, M. (2011). One DEA ranking method based on applying aggregate units. *Expert Systems with Applications, 38*, 13468–13471.

Lovell, C. K., & Pastor, J. T. (1995). Units invariant and translation invariant DEA models. *Operations research letters, 18*, 147–151.

Mahdiloo, M., Jafarzadeh, A. H., Saen, R. F., Tatham, P., & Fisher, R. (2016). A multiple criteria approach to two-stage data envelopment analysis. *Transportation Research Part D: Transport and Environment, 46*, 317–327.

Masternak-Janus, A., & Rybaczewska-Błażejowska, M. (2017). Comprehensive regional eco-efficiency analysis based on data envelopment analysis: The case of Polish regions. *Journal of Industrial Ecology, 21*, 180–190.

Mavi, R. K., Saen, R. F., & Goh, M. (2019). Joint analysis of eco-efficiency and eco-innovation with common weights in two-stage network DEA: A big data approach. *Technological Forecasting and Social Change, 144*, 553–562.

McCune, T. J. (2006). *The population problem: Moving beyond the Ehrlich-Commoner debate.* Ph.D. dissertation, Southern Illinois University at Carbondale.

McInnes, L., Healy, J., & Astels, S. (2017). hdbscan: Hierarchical density based clustering. *Journal of Open Source Software, 2*, 205.

Moallemi, E. A., Kwakkel, J., de Haan, F. J., & Bryan, B. A. (2020). Exploratory modeling for analyzing coupled human-natural systems under uncertainty. *Global Environmental Change, 65*, 102186.

Moraffah, R., Karami, M., Guo, R., Raglin, A., & Liu, H. (2020). Causal interpretability for machine learning-problems, methods and evaluation. *ACM SIGKDD Explorations Newsletter, 22*, 18–33.

Moutinho, V., Madaleno, M., & Robaina, M. (2017). The economic and environmental efficiency assessment in EU cross-country: Evidence from DEA and quantile regression approach. *Ecological Indicators, 78,* 85–97.

Moutinho, V., Madaleno, M., Robaina, M., & Villar, J. (2018). Advanced scoring method of eco-efficiency in European cities. *Environmental Science and Pollution Research, 25*, 1637–1654.

Munda, G., & Nardo, M. (2009). Noncompensatory/nonlinear composite indicators for ranking countries: a defensible setting. *Applied Economics, 41*, 1513–1523.

Munda, G., & Saisana, M. (2011). Methodological considerations on regional sustainability assessment based on multicriteria and sensitivity analysis. *Regional Studies, 45*, 261–276.

Munksgaard, J., Wier, M., Lenzen, M., & Dey, C. (2005). Using input-output analysis to measure the environmental pressure of consumption at different spatial levels. *Journal of Industrial Ecology, 9*, 169–185.

Murtagh, F., & Contreras, P. (2012). Algorithms for hierarchical clustering: an overview. *Wiley Interdisciplinary Reviews: Data Mining and Knowledge Discovery, 2,* 86–97.

Nahra, T. A., Mendez, D., & Alexander, J. A. (2009). Employing super-efficiency analysis as an alternative to DEA: An application in outpatient substance abuse treatment. *European Journal of Operational Research, 196*, 1097–1106.

Nandy, A., & Singh, P. K. (2020a). Application of fuzzy DEA and machine learning algorithms in efficiency estimation of paddy producers of rural Eastern India. *Benchmarking: An International Journal*.

Nandy, A., & Singh, P. K. (2020b). Farm efficiency estimation using a hybrid approach of machine-learning and data envelopment analysis: Evidence from rural eastern India. *Journal of Cleaner Production, 267*, 122106.

Nissi, E., & Rapposelli, A. (2013). Performance Measurement of Italian Provinces in the Presence of Environmental Goals. In *Statistical Models for Data Analysis* (pp. 251–258). Springer.

O'Leary-Kelly, S. W., & Flores, B. E. (2002). The integration of manufacturing and marketing/sales decisions: impact on organizational performance. *Journal of operations management, 20*, 221–240.

OECD. (1998). *Ecoefficiency*. Paris: OECD.

Opricovic, S., & Tzeng, G.-H. (2004). Compromise solution by MCDM methods: A comparative analysis of VIKOR and TOPSIS. *European journal of operational research, 156*, 445–455.

Örkcü, H. H., Özsoy, V. S., Örkcü, M., & Bal, H. (2019). A neutral cross efficiency approach for basic two stage production systems. *Expert Systems with Applications, 125*, 333–344.

Pang, J., Chen, X., Zhang, Z., & Li, H. (2016). Measuring eco-efficiency of agriculture in China. *Sustainability, 8*, 398.

Papathanasiou, J. (2021). An example on the use and limitations of MCDA: The case of fuzzy VIKOR. *Examples and Counterexamples, 1*, 100001.

Pedraja-Chaparro, F., Salinas-Jimenez, J., & Smith, P. (1997). On the role of weight restrictions in data envelopment analysis. *Journal of Productivity Analysis, 8*, 215–230.

Pedregosa, F., Varoquaux, G., Gramfort, A., Michel, V., Thirion, B., Grisel, O., . . . Duchesnay, E. (2011). Scikit-learn: Machine Learning in Python. *Journal of Machine Learning Research, 12*, 2825–2830.

Peng, Y., Zhang, Y., Tang, Y., & Li, S. (2011). An incident information management framework based on data integration, data mining, and multi-criteria decision making. *Decision Support Systems, 51*, 316–327.

Picazo-Tadeo, A. J., Gómez-Limón, J. A., & Reig-Martínez, E. (2011). Assessing farming eco-efficiency: a data envelopment analysis approach. *Journal of environmental management, 92*, 1154–1164.

Piña, A., William, H., & Pardo Martínez, C. I. (2016). Development and urban sustainability: An analysis of efficiency using data envelopment analysis. *Sustainability, 8*, 148.

Poore, J., & Nemecek, T. (2018). Reducing food's environmental impacts through producers and consumers. *Science, 360*, 987–992.

Pope, J., Annandale, D., & Morrison-Saunders, A. (2004). Conceptualising sustainability assessment. *Environmental impact assessment review, 24*, 595–616.

Pozo, C., Limleamthong, P., Guo, Y., Green, T., Shah, N., Acha, S., . . . Guillén-Gosálbez, G. (2019). Temporal sustainability efficiency analysis of urban areas via Data Envelopment Analysis and the hypervolume indicator: Application to London boroughs. *Journal of Cleaner Production, 239*, 117839.

Pradhan, P., Kriewald, S., Costa, L., Rybski, D., Benton, T. G., Fischer, G., & Kropp, J. P. (2020). Urban food systems: how regionalization can contribute to climate change mitigation. *Environmental Science & Technology, 54*, 10551–10560.

Rajaeifar, M. A., Akram, A., Ghobadian, B., Rafiee, S., & Heidari, M. D. (2014). Energy-economic life cycle assessment (LCA) and greenhouse gas emissions analysis of olive oil production in Iran. *Energy, 66*, 139–149.

Ramanathan, R. (2002). Combining indicators of energy consumption and CO2 emissions: a cross-country comparison. *International Journal of Global Energy Issues, 17*, 214–227.

Rebolledo-Leiva, R., Angulo-Meza, L., Iriarte, A., & González-Araya, M. C. (2017). Joint carbon footprint assessment and data envelopment analysis for the reduction of greenhouse gas emissions in agriculture production. *Science of the Total Environment, 593*, 36–46.

Rees, W. E. (2000). Eco-footprint analysis: merits and brickbats. *Ecological Economics, 32*, 371–374.

Rees, W. E., & Wackernagel, M. (1999). Monetary analysis: turning a blind eye on sustainability. *Ecological Economics, 29*, 47–52.

Rezaee, M. J., Izadbakhsh, H., & Yousefi, S. (2016). An improvement approach based on DEA-game theory for comparison of operational and spatial efficiencies in urban transportation systems. *KSCE Journal of Civil Engineering, 20*, 1526–1531.

Robaina-Alves, M., Moutinho, V., & Macedo, P. (2015). A new frontier approach to model the eco-efficiency in European countries. *Journal of Cleaner Production, 103*, 562–573.

Robinson, J. (2004). Squaring the circle? Some thoughts on the idea of sustainable development. *Ecological economics, 48*, 369–384.

Rogge, N., De Jaeger, S., & Lavigne, C. (2017). Waste performance of NUTS 2-regions in the EU: a conditional directional distance benefit-of-the-doubt model. *Ecological Economics, 139*, 19–32.

Roll, Y., & Golany, B. (1993). Alternate methods of treating factor weights in DEA. *Omega, 21*, 99–109.

Roll, Y., Cook, W. D., & Golany, B. (1991). Controlling factor weights in data envelopment analysis. *IIE transactions, 23*, 2–9.

Ross, A. D. (2000). Performance-based strategic resource allocation in supply networks. *International journal of production economics, 63,* 255–266.

Ross, A., & Droge, C. (2002). An integrated benchmarking approach to distribution center performance using DEA modeling. *Journal of Operations Management, 20,* 19–32.

Saaty, T. L., & Sagir, M. (2009). An essay on rank preservation and reversal. *Mathematical and Computer Modelling, 49,* 1230–1243.

Salehi, V., Veitch, B., & Musharraf, M. (2020). Measuring and improving adaptive capacity in resilient systems by means of an integrated DEA-Machine learning approach. *Applied ergonomics, 82*, 102975.

Samoilenko, S., & Osei-Bryson, K.-M. (2008). Increasing the discriminatory power of DEA in the presence of the sample heterogeneity with cluster analysis and decision trees. *Expert Systems with Applications, 34,* 1568–1581.

Santana, N. B., Mariano, E. B., Camioto, F. d., & Rebelatto, D. A. (2015). National innovative capacity as determinant in sustainable development: a comparison between the BRICS and G7 countries. *International Journal of Innovation and Sustainable Development, 9,* 384–405.

Sayed, H., Hamed, R., Hosny, S. H., & Abdelhamid, A. H. (2018). Avoiding ranking contradictions in human development index using goal programming. *Social Indicators Research, 138,* 405–442.

Sayed, H., Hamed, R., Ramadan, M. A.-G., & Hosny, S. (2015). Using meta-goal programming for a new human development indicator with distinguishable country ranks. *Social Indicators Research, 123*, 1–27.

Seiford, L. M., & Zhu, J. (1999). Profitability and marketability of the top 55 US commercial banks. *Management science, 45*, 1270–1288.

Sexton, T. R., Silkman, R. H., & Hogan, A. J. (1986). Data envelopment analysis: Critique and extensions. *New Directions for Program Evaluation, 1986*, 73–105.

Shi, G.-M., Bi, J., & Wang, J.-N. (2010). Chinese regional industrial energy efficiency evaluation based on a DEA model of fixing non-energy inputs. *Energy policy, 38,* 6172–6179.

Shin, Y. B., & Lee, S. (2013). Note on an approach to preventing rank reversals with addition or deletion of an alternative in analytic hierarchy process. *US-China Education Review A, 3,* 66–72.

Simar, L., & Wilson, P. W. (2007). Estimation and inference in two-stage, semi-parametric models of production processes. *Journal of econometrics, 136*, 31–64.

Sinuany-Stern, Z., & Friedman, L. (1998). DEA and the discriminant analysis of ratios for ranking units. *European Journal of Operational Research, 111*, 470–478.

Sneddon, C., Howarth, R. B., & Norgaard, R. B. (2006). Sustainable development in a post-Brundtland world. *Ecological economics, 57*, 253–268.

Sohn, S. Y., & Moon, T. H. (2004). Decision tree based on data envelopment analysis for effective technology commercialization. *Expert systems with Applications, 26*, 279–284.

Soltanifar, M., & Shahghobadi, S. (2014). Survey on rank preservation and rank reversal in data envelopment analysis. *Knowledge-Based Systems, 60*, 10–19.

Song, M., Peng, J., Wang, J., & Zhao, J. (2018). Environmental efficiency and economic growth of China: A Ray slack-based model analysis. *European Journal of Operational Research, 269*, 51–63.

Sueyoshi, T., & Yuan, Y. (2016). Returns to damage under undesirable congestion and damages to return under desirable congestion measured by DEA environmental assessment with multiplier restriction: Economic and energy planning for social sustainability in China. *Energy Economics, 56*, 288–309.

Sueyoshi, T., & Yuan, Y. (2017). Social sustainability measured by intermediate approach for DEA environmental assessment: Chinese regional planning for economic development and pollution prevention. *Energy Economics, 66*, 154–166.

Sueyoshi, T., & Yuan, Y. (2018). Measuring energy usage and sustainability development in Asian nations by DEA intermediate approach. *Journal of Economic Structures, 7*, 6.

Sueyoshi, T., Yuan, Y., & Goto, M. (2017). A literature study for DEA applied to energy and environment. *Energy Economics, 62*, 104–124.

Sueyoshi, T., Yuan, Y., Li, A., & Wang, D. (2017). Methodological comparison among radial, non-radial and intermediate approaches for DEA environmental assessment. *Energy Economics, 67*, 439–453.

Sun, J., Wu, J., & Guo, D. (2013). Performance ranking of units considering ideal and anti-ideal DMU with common weights. *Applied Mathematical Modelling, 37*, 6301–6310.

Tajbakhsh, A., & Shamsi, A. (2019). Sustainability performance of countries matters: A non-parametric index. *Journal of cleaner production, 224*, 506–522.

Taskin, F., & Zaim, O. (2000). Searching for a Kuznets curve in environmental efficiency using kernel estimation. *Economics letters*, *68*, 217–223.

Thaker, K., Charles, V., Pant, A., & Gherman, T. (2021). A DEA and random forest regression approach to studying bank efficiency and corporate governance. *Journal of the Operational Research Society*, 1–28.

Thanassoulis, E. (2001). *Introduction to the theory and application of data envelopment analysis.* Dordrecht: Kluwer Academic Publishers.

Tofallis, C. (2014). Add or multiply? A tutorial on ranking and choosing with multiple criteria. *INFORMS Transactions on education, 14,* 109–119.

Toma, E., Dobre, C., Dona, I., & Cofas, E. (2015). DEA applicability in assessment of agriculture efficiency on areas with similar geographically patterns. *Agriculture and Agricultural Science Procedia, 6*, 704–711.

Toma, P., Miglietta, P. P., Zurlini, G., Valente, D., & Petrosillo, I. (2017). A non-parametric bootstrap-data envelopment analysis approach for environmental policy planning and management of agricultural efficiency in EU countries. *Ecological Indicators, 83*, 132–143.

Tone, K. (2017). *Advances in DEA theory and applications: with extensions to forecasting models.* John Wiley & Sons.

Tone, K., & Tsutsui, M. (2009). Network DEA: A slacks-based measure approach. *European journal of operational research, 197*, 243–252.

Troutt, M. D. (1988). Rank reversal and the dependence of priorities on the underlying MAV function. *Omega, 16,* 365–367.

Tsai, C.-F., & Tsai, J.-H. (2010). Performance Evaluation of the Judicial System in Taiwan Using Data Envelopment Analysis and Decision Trees. *2010 Second International Conference on Computer Engineering and Applications*, *2,* pp. 290–294.

Tsaples, G. (2021). *A novel Multi-level Data Envelopment Analysis variation for the construction of composite indicators and the measurement of sustainability (PhD Thesis currently in progress).* Ph.D. dissertation, University of Macedonia.

Tsaples, G., & Papathanasiou, J. (2020a). Using multi-level DEA to go beyond the three dimensions of sustainability. *International Conference on Decision Support System Technology*, (pp. 15–29).

Tsaples, G., & Papathanasiou, J. (2020b). Multi-level DEA for the construction of multi-dimensional indices. *MethodsX, 7*, 101169.

Tsaples, G., & Papathanasiou, J. (2021). Data envelopment analysis and the concept of sustainability: A review and analysis of the literature. *Renewable and Sustainable Energy Reviews, 138*, 110664.

Tsaples, G., Papathanasiou, J., Georgiou, A. C., & Samaras, N. (2019). Assessing multidimensional sustainability of European countries with a novel, two-stage DEA. *International Conference on Decision Support System Technology*, (pp. 111–122).

Tsoukias, A., Montibeller, G., Lucertini, G., & Belton, V. (2013). Policy analytics: an agenda for research and practice. *EURO Journal on Decision Processes, 1*, 115–134.

Tyteca, D. (1998). Sustainability indicators at the firm level: pollution and resource efficiency as a necessary condition toward sustainability. *Journal of Industrial Ecology, 2*, 61–77.

Tzeng, G.-H., & Huang, J.-J. (2011). *Multiple attribute decision making: methods and applications.* CRC press.

UNDP. (1990). *Concept and measurement of Human Development, Human Development Report.* Tech. rep., Oxford: Oxford University Press.

Vayssières, M. P., Plant, R. E., & Allen-Diaz, B. H. (2000). Classification trees: An alternative non-parametric approach for predicting species distributions. *Journal of vegetation science, 11*, 679–694.

Wang, C. H., Gopal, R. D., & Zionts, S. (1997). Use of data envelopment analysis in assessing information technology impact on firm performance. *Annals of operations Research, 73*, 191–213.

Wang, H., Zhou, P., & Zhou, D. Q. (2013). Scenario-based energy efficiency and productivity in China: a non-radial directional distance function analysis. *Energy Economics, 40*, 795–803.

Wang, K., & Wei, Y.-M. (2014). China’s regional industrial energy efficiency and carbon emissions abatement costs. *Applied Energy, 130*, 617–631.

Wang, K., Wei, Y.-M., & Zhang, X. (2012). A comparative analysis of China’s regional energy and emission performance: Which is the better way to deal with undesirable outputs? *Energy Policy, 46*, 574–584.

Wang, Q., Zhao, Z., Zhou, P., & Zhou, D. (2013). Energy efficiency and production technology heterogeneity in China: a meta-frontier DEA approach. *Economic Modelling, 35*, 283–289.

Wang, X., & Triantaphyllou, E. (2008). Ranking irregularities when evaluating alternatives by using some ELECTRE methods. *Omega, 36*, 45–63.

Wang, Y.-M., & Luo, Y. (2009). On rank reversal in decision analysis. *Mathematical and Computer Modelling, 49*, 1221–1229.

Wang, Y.-M., Luo, Y., & Lan, Y.-X. (2011). Common weights for fully ranking decision making units by regression analysis. *Expert Systems with Applications, 38*, 9122–9128.

Wang, Z., Feng, C., & Zhang, B. (2014). An empirical analysis of China's energy efficiency from both static and dynamic perspectives. *Energy, 74*, 322–330.

WBCSD. (2000). Measuring eco-efficiency: A guide to reporting company performance. *Measuring eco-efficiency: A guide to reporting company performance*. World Business Council for Sustainable Development Geneva.

Weber, C. L., & Matthews, H. S. (2008). Food-miles and the relative climate impacts of food choices in the United States. *Environmental Science & Technology, 42*(10), 3508–3513.

Wei, C., Ni, J., & Du, L. (2012). Regional allocation of carbon dioxide abatement in China. *China Economic Review, 23*, 552–565.

Wei, Q., Zhang, J., & Zhang, X. (2000). An inverse DEA model for inputs/outputs estimate. *European Journal of Operational Research, 121*, 151–163.

White, L., & Bourne, H. (2007). Voices and values: Linking values with participation in OR/MS in public policy making. *Omega, 35*, 588–603.

Winfield, M., Gibson, R. B., Markvart, T., Gaudreau, K., & Taylor, J. (2010). Implications of sustainability assessment for electricity system design: The case of the Ontario Power Authority's integrated power system plan. *Energy Policy, 38*, 4115–4126.

Wu, D. (2009a). Supplier selection: A hybrid model using DEA, decision tree and neural network. *Expert systems with Applications, 36*, 9105–9112.

Wu, D. (2009b). Performance evaluation: an integrated method using data envelopment analysis and fuzzy preference relations. *European Journal of Operational Research, 194*, 227–235.

Wu, F., Fan, L. W., Zhou, P., & Zhou, D. Q. (2012). Industrial energy efficiency with CO2 emissions in China: A nonparametric analysis. *Energy Policy, 49*, 164–172.

Wu, J., An, Q., Yao, X., & Wang, B. (2014). Environmental efficiency evaluation of industry in China based on a new fixed sum undesirable output data envelopment analysis. *Journal of cleaner production, 74*, 96–104.

Wu, J., Yin, P., Sun, J., Chu, J., & Liang, L. (2016). Evaluating the environmental efficiency of a two-stage system with undesired outputs by a DEA approach: An interest preference perspective. *European Journal of Operational Research, 254*, 1047–1062.

Xing, Z., Wang, J., & Zhang, J. (2018). Expansion of environmental impact assessment for eco-efficiency evaluation of China's economic sectors: An economic input-output based frontier approach. *Science of the Total Environment, 635*, 284–293.

Xu, J., Li, B., & Wu, D. (2009). Rough data envelopment analysis and its application to supply chain performance evaluation. *International Journal of Production Economics, 122*, 628–638.

Yang, F.-C., Kao, R.-H., Chen, Y.-T., Ho, Y.-F., Cho, C.-C., & Huang, S.-W. (2018). A common weight approach to construct composite indicators: The evaluation of fourteen emerging markets. *Social Indicators Research, 137*, 463–479.

Yang, L., & Zhang, X. (2018). Assessing regional eco-efficiency from the perspective of resource, environmental and economic performance in China: A bootstrapping approach in global data envelopment analysis. *Journal of Cleaner Production, 173*, 100–111.

Yang, L., Ouyang, H., Fang, K., Ye, L., & Zhang, J. (2015). Evaluation of regional environmental efficiencies in China based on super-efficiency-DEA. *Ecological Indicators, 51*, 13–19.

Yao, X., Zhou, H., Zhang, A., & Li, A. (2015). Regional energy efficiency, carbon emission performance and technology gaps in China: A meta-frontier non-radial directional distance function analysis. *Energy Policy, 84*, 142–154.

Yin, K., Wang, R., An, Q., Yao, L., & Liang, J. (2014). Using eco-efficiency as an indicator for sustainable urban development: A case study of Chinese provincial capital cities. *Ecological Indicators, 36*, 665–671.

Zaim, O., & Taskin, F. (2000). Environmental efficiency in carbon dioxide emissions in the OECD: A non-parametric approach. *Journal of Environmental Management, 58*, 95–107.

Zhang, A., Li, A., & Gao, Y. (2018). Social sustainability assessment across provinces in China: An analysis of combining intermediate approach with data envelopment analysis (DEA) window analysis. *Sustainability, 10*, 732.

Zhang, N., Kong, F., & Choi, Y. (2014). Measuring sustainability performance for China: A sequential generalized directional distance function approach. *Economic Modelling, 41*, 392–397.

Zhang, N., Kong, F., & Yu, Y. (2015). Measuring ecological total-factor energy efficiency incorporating regional heterogeneities in China. *Ecological indicators, 51*, 165–172.

Zhang, Y.-J., Hao, J.-F., & Song, J. (2016). The CO2 emission efficiency, reduction potential and spatial clustering in China's industry: Evidence from the regional level. *Applied Energy, 174*, 213–223.

Zhao, L., Zha, Y., Zhuang, Y., & Liang, L. (2019). Data envelopment analysis for sustainability evaluation in China: Tackling the economic, environmental, and social dimensions. *European Journal of Operational Research, 275*, 1083–1095.

Zhou, H., Yang, Y., Chen, Y., & Zhu, J. (2018). Data envelopment analysis application in sustainability: The origins, development and future directions. *European Journal of Operational Research, 264*, 1–16.

Zhou, P., Ang, B. W., & Han, J. Y. (2010). Total factor carbon emission performance: a Malmquist index analysis. *Energy Economics, 32*, 194–201.

Zhou, P., Ang, B. W., & Poh, K. L. (2007). A mathematical programming approach to constructing composite indicators. *Ecological economics, 62*, 291–297.

Zhou, P., Ang, B. W., & Poh, K. L. (2008). Measuring environmental performance under different environmental DEA technologies. *Energy Economics, 30*, 1–14.

Zhou, P., Ang, B. W., & Zhou, D. Q. (2010). Weighting and aggregation in composite indicator construction: A multiplicative optimization approach. *Social indicators research, 96*, 169–181.

Zhou, P., Poh, K. L., & Ang, B. W. (2007). A non-radial DEA approach to measuring environmental performance. *European journal of operational research, 178*, 1–9.

Zhou, P., Sun, Z. R., & Zhou, D. Q. (2014). Optimal path for controlling CO2 emissions in China: a perspective of efficiency analysis. *Energy Economics*, *45*, 99–110.

Zhu, J. (2014). *Quantitative models for performance evaluation and benchmarking: data envelopment analysis with spreadsheets* (Vol. 213). Springer.

Zhu, Q., Wu, J., Li, X., & Xiong, B. (2017). China's regional natural resource allocation and utilization: a DEA-based approach in a big data environment. *Journal of Cleaner Production*, *142*, 809–818.

Zofío, J. L., & Prieto, A. M. (2001). Environmental efficiency and regulatory standards: the case of CO2 emissions from OECD industries. *Resource and Energy Economics, 23*, 63–83.

Appendix A

All the DEA variations were programmed in the Julia Programming Language[16]. Julia is free and open-source for scientific computing. Furthermore, the language is compiled and not interpreted that offers a faster performance time. It has a simple and understandable syntax and it combines the advantages of static and dynamic typing. Finally, the community around the language is expanding daily and the code developed for the current dissertation could contribute to the community.

Two-stage, independent, DEA CCR code

```
# TWO-STAGE, INDEPENDENT CCR DEA MODEL
using JuMP
using SCS
using CSV
using DataFrames
using XLSX



#Empty dataframes to be used
df = DataFrame(Country = [], E0 = Float16[])
df1 = DataFrame(Country = [], E1 = Float16[])
df2 = DataFrame(Country = [], E2 = Float16[])

# Load data from CSV file
data = CSV.read("C:/Users/…/data_overall2.csv


countries = data[:,1]
inputs =  data[:, [7,8
inp=size(inputs)[2]

intermediate= data[:,[10]]
md = size(intermediate)[2]
outputs = data[:, [30,13,15]]
rs=size(outputs)[2]

# scale is the number of DMU, dimension is the total number of inputs
and outputs
scale, dimension = size(data)
generalEff = []
generalfirst = []
generalsecond = []

for t = 1 : scale

    ### Modeling section
    # Here is the CRS input-oriented model (CCR model) to evaluate
DMU t
```

[16] https://julialang.org/

```
    model = Model(SCS.Optimizer) # SCS is used as the LP solver here. 
Users can choose their favorite solver.

    @variable(model, gamma[1:rs]>=0)
    @variable(model, omega[1:inp]>=0)

    # Objective function
    @objective(model, Max, sum(gamma[k]*outputs[t,k] for k = 1: rs))

    # Constraints
    @constraint(model,  sum(omega[i]*inputs[t,i] for i=1:inp) == 1 )
    @constraint(model, [j=1:scale], sum(gamma[r]*outputs[j,r] for 
r=1:rs)- sum(omega[i]*inputs[j,i] for i=1:inp)<=0)


    ### Problem solving
    optimize!(model)
    E1_0 = JuMP.objective_value.(model)
    push!(generalEff, JuMP.objective_value.(model))
    push!(df, [countries[t], JuMP.objective_value.(model)])

    #Calculating the efficiencies of the first stage
    model2 = Model(SCS.Optimizer)

    @variable(model2, mu[1:md]>=0)
    @variable(model2, omega[1:inp]>=0)

    # Objective function
    @objective(model2, Max, sum(mu[d]*intermediate[t,d] for d =1:md))

    #Constraints
    @constraint(model2,    sum(omega[i]*inputs[t,i] for i=1:inp ) == 
1 )
    @constraint(model2, [j=1:scale], sum(mu[d]*intermediate[j,d] for 
d =1:md) - sum(omega[i]*inputs[j,i] for i=1:inp) <= 0 )

    optimize!(model2)
    E1_1 = JuMP.objective_value.(model2)
    push!(generalfirst, JuMP.objective_value.(model2))
    push!(df1, [countries[t], JuMP.objective_value.(model2)])


    #Calculating the efficiencies of the second stage
    model3 = Model(SCS.Optimizer)

    @variable(model3, mu[1:md]>=0)
    @variable(model3, gamma[1:rs]>=0)

    #Objective function
    @objective(model3, Max, sum(gamma[k]*outputs[t,k] for k = 1: rs))

    #Constraints
    @constraint(model3,    sum(mu[d]*intermediate[t,d] for d =1:md) 
== 1 )
    @constraint(model3, [j=1:scale], sum(gamma[k]*outputs[t,k] for k 
= 1: rs)-sum(mu[d]*intermediate[j,d] for d =1:md)<= 0 )
```

```
    optimize!(model3)
    E1_2 = JuMP.objective_value.(model3)
    push!(generalsecond, JuMP.objective_value.(model3))
    push!(df2, [countries[t], JuMP.objective_value.(model3)])
end
sub1= innerjoin(df,df1,df2, on = :Country, makeunique=true )
CSV.write("C:/Users/…/outputfile.csv",sub1)
```

Two-stage, independent, DEA VRS code

```
# TWO-STAGE, INDEPENDENT VRS DEA MODEL
using JuMP
using SCS
using CSV
using DataFrames
using XLSX


#Empty dataframes to be used
df = DataFrame(Country = [], E0 = Float16[])
df1 = DataFrame(Country = [], E1 = Float16[])
df2 = DataFrame(Country = [], E2 = Float16[])

# Load data from CSV file
data = CSV.read("C:/Users/…/data_overall2.csv


countries = data[:,1]
inputs =  data[:, [7,8
inp=size(inputs)[2]

intermediate= data[:,[10]]
md = size(intermediate)[2]
outputs = data[:, [30,13,15]]
rs=size(outputs)[2]

# scale is the number of DMU, dimension is the total number of inputs
and outputs
scale, dimension = size(data)
generalEff = []
generalfirst = []
generalsecond = []
for t = 1:scale

    ### Modeling section
    # Here is the CRS input-oriented model (CCR model) to evaluate
DMU t
    model = Model(SCS.Optimizer) # SCS is used as the LP solver here.
Users can choose their favorite solver.

    @variable(model, gamma[1:rs] >= 0)
    @variable(model, omega[1:inp] >= 0)
    @variable(model, u1)

    # Objective function
    @objective(model, Max, sum(gamma[k] * outputs[t, k] for k = 1:rs)
+ u1)

    # Constraints
    @constraint(model, sum(omega[i] * inputs[t, i] for i = 1:inp) ==
1)
    @constraint(model,[j = 1:scale], sum(gamma[r] * outputs[j, r] for
r = 1:rs) - sum(omega[i] * inputs[j, i] for i = 1:inp) + u1 <= 0)
```

```
    ### Problem solving
    optimize!(model)
    E1_0 = JuMP.objective_value.(model)
    #JuMP.objective_value.(model)
    push!(generalEff, JuMP.objective_value.(model))
    push!(df, [countries[t], JuMP.objective_value.(model)])

    #Calculating the efficiencies of each stage
    model2 = Model(SCS.Optimizer)

    @variable(model2, mu[1:md] >= 0)
    @variable(model2, omega[1:inp] >= 0)
    @variable(model2, u3)

    #Objective function
    @objective(model2, Max, sum(mu[d] * intermediate[t, d] for d =
1:md) + u3)

    #Constraints
    @constraint(model2, sum(omega[i] * inputs[t, i] for i = 1:inp) ==
1)
    @constraint(model2,[j = 1:scale], sum(mu[d] * intermediate[j, d]
for d = 1:md) -sum(omega[i] * inputs[j, i] for i = 1:inp) + u3 <= 0)

    optimize!(model2)
    E1_1 = JuMP.objective_value.(model2)
    push!(generalfirst, JuMP.objective_value.(model2))
    push!(df1, [countries[t], JuMP.objective_value.(model2)])



    model3 = Model(SCS.Optimizer)
    @variable(model3, mu[1:md] >= 0)
    @variable(model3, gamma[1:rs] >= 0)
    @variable(model3, u24)

    #Objective function
    @objective(model3, Max, sum(gamma[k] * outputs[t, k] for k =
1:rs) + u24)

    #Constraints
    @constraint(model3, sum(mu[d] * intermediate[t, d] for d = 1:md)
== 1)
    @constraint(model3,[j = 1:scale],sum(gamma[k] * outputs[t, k] for
k = 1:rs) -sum(mu[d] * intermediate[j, d] for d = 1:md) + u24 <= 0)

    optimize!(model3)
    E1_2 = JuMP.objective_value.(model3)
    push!(generalsecond, JuMP.objective_value.(model3))
    push!(df2, [countries[t], JuMP.objective_value.(model3)])
end
sub1 = innerjoin(df, df1, df2, on = :Country, makeunique = true)
CSV.write("C:/Users/…/outputfile.csv",sub1,)
```

Chen et al. (2012) code

```
# TWO-STAGE MODEL BASED ON CHEN ET AL. (2012)
using JuMP
using SCS
using CSV
using DataFrames
using XLSX

# Empty dataframes to be used lated
df = DataFrame(Country = [], E0 = Float16[])
df1 = DataFrame(Country = [], E1 = Float16[])
df2 = DataFrame(Country = [], E2 = Float16[])
df3 = DataFrame(Country = [], obj = Float16[])

# Load data
data = CSV.read("C:/Users/…/agri_data2.csv")

countries = data[:,1]
inputs =  data[:, [2,3,4]]
inp=size(inputs)[2]

intermediate= data[:,[5,6,7]]
md = size(intermediate)[2]

outputs = data[:, [8,9]]
rs=size(outputs)[2]

# scale is the number of DMU, dimension is the total number of inputs
and outputs
scale, dimension = size(data)
generalEff = []
generalfirst = []
generalsecond = []

# The two-stage DEA otpimization
for t = 1 : scale
    ### Modeling section
    ### SCS is used as the LP solver here. Users can choose their
favorite solver.
    model = Model(SCS.Optimizer)
    @variable(model, mu[1:md]>=0.0000000000000000000000000000001)
    @variable(model, gamma[1:rs]>=0.00000000000000000000000000001)
    @variable(model, omega[1:inp]>=0.0000000000000000000000000001)

    ### Objective function
    @objective(model, Max, sum(gamma[k]*outputs[t,k] for k = 1: rs))

    ### Constraints
    @constraint(model,  sum(omega[i]*inputs[t,i] for i=1:inp) == 1 )
    @constraint(model, [j=1:scale], sum(gamma[r]*outputs[j,r] for
r=1:rs) - sum(omega[i]*inputs[j,i] for i=1:inp) <=0)
    @constraint(model, [j=1:scale], sum(mu[d]*intermediate[j,d] for d
= 1:md) - sum(omega[i]*inputs[j,i] for i=1:inp) <= 0 )
    @constraint(model, [j=1:scale], sum(gamma[r]*outputs[j,r] for
r=1:rs) - sum(mu[d]*intermediate[j,d] for d = 1:md) <= 0)
```

```
    ### Problem solving
    optimize!(model)
    E1_0 = JuMP.objective_value.(model)
    Eff=sum(getvalue.(gamma)[r]*outputs[t,r] for r =
1:rs)/sum(getvalue.(omega)[i]*inputs[t,i] for i = 1:inp)
    E1_ast = sum(getvalue.(mu)[d]*intermediate[t,d] for d =
1:md)/sum(getvalue.(omega)[i]*inputs[t,i] for i = 1:inp)
    E2_ast= sum(getvalue.(gamma)[r]*outputs[t,r] for r =
1:rs)/sum(getvalue.(mu)[d]*intermediate[t,d] for d = 1:md)


# Write the efficiency of each country into the appropriate dataframe

    push!(df3, [countries[t], JuMP.objective_value.(model)])
    push!(df, [countries[t], sum(getvalue.(gamma)[r]*outputs[t,r] for
r = 1:rs)/sum(getvalue.(omega)[i]*inputs[t,i] for i = 1:inp)])
    push!(df1, [countries[t], sum(getvalue.(mu)[d]*intermediate[t,d]
for d = 1:md)/sum(getvalue.(omega)[i]*inputs[t,i] for i = 1:inp)])
    push!(df2, [countries[t], sum(getvalue.(gamma)[r]*outputs[t,r]
for r = 1:rs)/sum(getvalue.(mu)[d]*intermediate[t,d] for d = 1:md)])

end

# Connect the dataframes into one
sub= innerjoin(df,df1, df2, on = :Country, makeunique=true )

# Write the results into a CSV file
CSV.write("C:/Users/…/outputfile.csv",sub)
```

Proposed, two-stage DEA variation code

```
######PROPOSED, TWO-STAGE DEA VARIATION

using JuMP
using SCS
using CSV
using DataFrames
using XLSX

# Empty dataframes to be used lated
df = DataFrame(Country = [], E0 = Float16[])
df1 = DataFrame(Country = [], E1 = Float16[])
df2 = DataFrame(Country = [], E2 = Float16[])
df3 = DataFrame(Country = [], obj = Float16[])

# Load data
data = CSV.read("C:/Users/…/agri_data2.csv")

countries = data[:,1]
inputs =  data[:, [2,3,4]]
inp=size(inputs)[2]

intermediate= data[:,[5,6,7]]
md = size(intermediate)[2]

outputs = data[:, [8,9]]
rs=size(outputs)[2]

# scale is the number of DMU, dimension is the total number of inputs
and outputs
scale, dimension = size(data)
generalEff = []
generalfirst = []
generalsecond = []

# The two-stage DEA otpimization
for t =1:scale
    model = Model(SCS.Optimizer) # SCS is used as the LP solver here.
Users can choose their favorite solver.
    @variable(model, mu[1:md]>=0.0000000000000000001)
    @variable(model, gamma[1:rs]>=0.0000000000000000001)
    @variable(model, omega[1:inp]>=0.0000000000000000001)
    @variable(model, n[1:scale]>=0)
    @variable(model, nminus[1:scale]>=0)
    @variable(model, d[1:scale]>=0)
    @variable(model, dminus[1:scale]>=0)



    #OBjective function
    @objective(model, Min, n[t] + nminus[t]+d[t] + dminus[t])

    #Constraints
    @constraint(model, sum(mu[d]*intermediate[t,d] for d = 1:md) ==1)
```

```
    @constraint(model, [j=1:scale], -sum(mu[d]*intermediate[j,d] for
d = 1:md) + sum(omega[i]*inputs[j,i] for i=1:inp) >= 0 )
    @constraint(model, [j=1:scale], -sum(mu[d]*intermediate[t,d] for
d = 1:md) + sum(omega[i]*inputs[t,i] for i=1:inp) - delta[t] +n[t] ==
0)
    @constraint(model, [j=1:scale], -sum(gamma[r]*outputs[j,r] for
r=1:rs) + sum(mu[d]*intermediate[j,d] for d = 1:md)  >= 0)

    @constraint(model, [j=1:scale], -sum(gamma[r]*outputs[t,r] for
r=1:rs) + sum(mu[d]*intermediate[t,d] for d = 1:md) - deltaminus[t]+
nminus[t] == 0)

    optimize!(model)
    E1_ast = sum(getvalue.(mu)[d]*intermediate[t,d] for d =
1:md)/sum(getvalue.(omega)[i]*inputs[t,i] for i = 1:inp)
    E2_ast= sum(getvalue.(gamma)[r]*outputs[t,r] for r =
1:rs)/sum(getvalue.(mu)[d]*intermediate[t,d] for d = 1:md)

    push!(df1, [countries[t], sum(getvalue.(mu)[d]*intermediate[t,d]
for d = 1:md)/sum(getvalue.(omega)[i]*inputs[t,i] for i = 1:inp)])
    push!(df2, [countries[t], sum(getvalue.(gamma)[r]*outputs[t,r]
for r = 1:rs)/sum(getvalue.(mu)[d]*intermediate[t,d] for d = 1:md)])
    E_ast = (E1_ast + E2_ast)/2.0
    push!(df, [countries[t],E_ast])
end
sub2= innerjoin(df,df1, df2, on = :Country, makeunique=true )
CSV.write("C:/Users/…/outputfile1.csv",sub2)
```

Appendix B

Table 14 below summarizes the parameters along with the major descriptive statistics of the data that were used in the proposed computational framework for the calculation of sustainability.

Table 14 Parameters that are used in the computational framework and summary statistics

	Gross fixed capital at current prices (PPS)	***Total Labour force (x1000 persons)***	***GDP per capita in PPS-Index (EU28 = 100)***	***Median equivalised net income [Purchasing power standard (PPS)]-2018***	**Final consumption expenditure of households [Current prices, million euro]**	**Population- 2018**
Mean	93.7035714	8954.214286	97.60714286	15896.46429	309080.5929	17970379.21
Standard Error	25.2418411	2219.976821	7.815724769	1144.509393	89744.88872	4373182.198
Median	35.85	4360.45	84.5	16372.5	122017.75	8846162.5
Standard Deviation	133.567268	11747.01317	41.35692811	6056.174452	474885.314	23140705.07
Sample Variance	17840.2152	137992318.5	1710.395503	36677249	2.25516E+11	5.35492E+14
Kurtosis	4.02051893	2.281925743	8.29166899	0.182552646	2.792124589	1.428904095
Skewness	2.11001276	1.770197969	2.325043795	0.448817209	1.981594431	1.622982
Range	526.2	44245.7	215	25880	1664272.9	81388230
Minimum	1.7	193.3	46	6278	6506.1	414027
Maximum	527.9	44439	261	32158	1670779	81802257

	Gross electricity production [Thousand tonnes of oil equivalent (TOE)]-2018	Domestic material consumption [Thousand tonnes]-2018	Final energy consumption [Terajoule] -2018	Terrestrial protected area (km2)-2018	Share of renewable energy in gross final energy consumption	Greenhouse gas emissions (in CO2 equivalent)
Mean	10048.5454	246185.8875	424169.6536	28009	18.53214286	9.228571429
Standard Error	2697.45258	54633.58413	110868.6098	5782.166667	2.211867065	0.624962584
Median	4853.78	148400.71	214985.875	16821.5	15.75	8.4
Standard Deviation	14273.5774	289093.7537	586661.5397	30596.35008	11.70410038	3.306991152
Sample Variance	203735012	83575198424	3.44172E+11	936136638.3	136.9859656	10.93619048
Kurtosis	4.39672619	4.383577372	3.565649211	4.921733029	0.957448781	3.456735643
Skewness	2.17413267	2.014708432	2.041196627	1.92448185	0.958061175	1.586879483
Range	54998.36	1239884.92	2309837.03	137974	48.5	14.9
Minimum	168.71	6499.1	3892.59	42	3.5	5.4
Maximum	55167.07	1246384.02	2313729.62	138016	52	20.3

	Total expenditure [Euro per inhabitant]	Mean consumption expenditure of private households on cultural goods and services by COICOP consumption purpose [Purchasing power standard (PPS)]	Patent applications to the European patent office (EPO) by priority year	Overall life satisfaction	Satisfaction with living environment	Percentage of females in total labor population-2018
Mean	7182.62071	25608.67857	1951.743214	6.971428571	7.221428571	68.53214286
Standard Error	993.301805	1686.520727	746.3961868	0.135581946	0.148480589	1.349613442
Median	4696.11	26815	277.075	7.05	7.55	68.6
Standard Deviation	5256.0591	8924.228852	3949.55738	0.717432223	0.785685425	7.141483069
Sample Variance	27626157.3	79641860.6	15599003.5	0.514708995	0.617301587	51.00078042
Kurtosis	-0.1173526	0.381582258	12.94505502	1.597491057	0.06125311	1.020390146
Skewness	0.82586947	0.374300212	3.389370605	-0.877951855	-0.844073462	-0.959481566
Range	19595.89	38416	18875.07	3.2	3.2	31.1
Minimum	1248.76	11422	6.63	4.8	5.2	49.1

Maximum	20844.65	49838	18881.7	8	8.4	80.2

	Satisfaction with financial situation	**Intramural R&D expenditure (GERD) by sectors of performance [Euro per inhabitant]**	**Pupils and students enrolled All ISCED 2011 levels excluding early childhood educational development**	**Participation rate in education and training (last 4 weeks) by sex and age From 25 to 64 years Percentage**	**Life expectancy at birth**	**Urban population exposure to air pollution by particulate matter [Particulates < 2.5µm]**
Mean	5.90714286	488.85	3872286.429	11.56785714	80.225	13.275
Standard Error	0.19134228	86.49583843	942453.4172	1.463814449	0.523864986	1.018423216
Median	5.8	281.75	1792249	9.4	81.5	12.95
Standard Deviation	1.01248816	457.6929559	4986994.728	7.745777993	2.772032948	5.388989117
Sample Variance	1.02513228	209482.8419	2.48701E+13	59.99707672	7.684166667	29.0412037
Kurtosis	-0.3674502	-0.514620128	1.455757741	0.550769838	-0.864099668	0.149557265
Skewness	-0.0743634	0.888998451	1.644603126	1.007501236	-0.778315177	-0.212994905
Range	3.9	1479.7	16110645	30.5	8.5	24.3
Minimum	3.7	27.9	82343	0.9	75	0

Maximum	7.6	1507.6	16192988	31.4	83.5	24.3

	Carbon dioxide [thousand tonnes]	**People at risk of poverty or social exclusion [thousand persons]**	**Final energy consumption [Million tonnes of oil equivalent (TOE)]**
Mean	112130.516	3924	40.14642857
Standard Error	30703.5494	967.6611291	10.04674016
Median	43570.865	1667	17.725
Standard Deviation	162467.912	5120.381402	53.16235191
Sample Variance	2.6396E+10	26218305.7	2826.235661
Kurtosis	6.71719926	1.04478003	3.634446153
Skewness	2.39557289	1.556049423	1.987717623
Range	728021.85	16352	214.71
Minimum	-1974.29	89	0.66
Maximum	726047.56	16441	215.37