World Development 140 (2021) 105269

Contents lists available at ScienceDirect

World Development

journal homepage: www.elsevier.com/locate/worlddev

Systematic prioritisation of SDGs: Machine learning approach

Atie Asadikia *, Abbas Rajabifard, Mohsen Kalantari

SDGs Research Group, Centre for Spatial Data Infrastructures and Land Administration, Department of Infrastructure Engineering, The University of Melbourne, Parkville, VIC 3010, Australia

ARTICLE INFO

Article history:
Accepted 20 October 2020
Available online 8 November 2020

Keywords:
Synergetic SDGs
SDGs Interlinkage
Systematic prioritisation
Machine learning
Data mining
SDGs Global prioritisation

ABSTRACT

The Sustainable Development Goals (SDGs) framework is recognised throughout the world as a significant global agreement that has been adopted by all UN members. These goals represent a solution to sustainability problems dealing with nations' economies, the natural environment and societies. However, making progress towards achieving these goals has not been as effective as originally intended. One of the major concerns is whether the SDGs will be achieved globally by 2030. Given the current damage wrought internationally by the COVID-19 pandemic, what is required is a coordinated global effort to achieve the SDGs. In this uncertain time, an era of corona virus outbreaks, when countries' resources are finite and the deadline is fast approaching, prioritisation is necessary to allocate resources effectively. Several attempts have been made to prioritise SDGs by quantifying synergies. However, systematic methods to identify the magnitude of how to enhance the SDG index by improving individual SDGs is lacking. The objective of this paper is to identify synergetic SDGs using Boosted Regression Trees model which is a machine learning and data mining technique. In this study, contributions of all SDGs to form the SDG index are identified and a "what-if" analysis is conducted to understand the significance of goal scores. Findings show that SDG3, "Good health and well-being", SDG4, "Quality education", and SDG7, "Affordable and clean energy", are the most synergetic goals, when their scores are >60%. The findings of this research will help decision-makers implement effective strategies and allocate resources by prioritising synergetic goals.

1. Introduction

In September 2015, it was acknowledged by world leaders at the United Nations (UN) Summit meeting that it is necessary to transform the world into a more sustainable one. On this issue, the 2030 Agenda by the UN proposed universally applicable goals and formally titled the Sustainable Development Goals (SDGs) (UN, 2015). There are in total 17 goals, 169 targets, and an expanding list of 231 unique indicators (UN Statistical Commission, 2020). This universal framework covers three dimensions of economic development, social inclusion, and environmental sustainability. It came into effect in 2016 and is known as one of the most significant global agreements. Although SDGs have received their share of criticism, they are an effective roadmap to measure and improve sustainability (Weitz, Persson, Nilsson, & Tenggren, 2015). Despite various attempts and projects being undertaken around the world to achieve SDGs, the progress rate is not as rapid as expected. Consequently, one of the main concerns is whether the SDGs will be globally achieved by 2030.

The integrated nature of the SDGs is important and needs to be considered in order to attain them (UN, 2018). Moving closer towards achieving a goal can have a negative (trade-offs) or positive (synergies) impact on other goals. The existence of these relationships is the primary reason why the SDGs are such a complex system. Although these complexities create some difficulties, they can in fact also be advantageous. Recent literature indicates that maximising synergies between goals can accelerate goal fulfilment (Breuer, Janetschek, & Malerba, 2019). Given that all SDGs are synergetic but not necessarily at an equal level (Nilsson, Griggs, & Visbeck, 2016), identification of those levels and the improvement in goals that impact positively on other goals provides more detailed insights for resource allocations to maximise goal attainment. In an era of pandemic diseases that have plagued modern societies since the late 1990s, the most lethal outbreak now universally known as COVID-19, has forced governments to extensively allocate huge budgets and resources to control the virus. There is as a consequence a high possibility of overlooking or now discarding SDGs as a priority even when sustainability will benefit all societies. In circumstances when there are high levels

* Corresponding author.
E-mail addresses: asadikiaa@student.unimelb.edu.au (A. Asadikia), abbas.r@unimelb.edu.au (A. Rajabifard), mohsen.kalantari@unimelb.edu.au (M. Kalantari).

https://doi.org/10.1016/j.worlddev.2020.105269
0305-750X/

of uncertainty, and governments may become more conservative or reactive in their policies, having prior knowledge that improving some goals can automatically improve the others and result in positive outcomes, is essential. To ensure smart resource allocation, it is vital to prioritise synergetic SDGs if they are to be achieved.

Similar to the Human Development Index (HDI), which is used to rank countries in a single composite measure (Dervis & Klugman, 2011), the SDG Index measures the progress of each country's SDGs (Sachs, Schmidt-Traub, Kroll, Lafortune, & Fuller, 2016). This index is lauded as one of the most comprehensive attempts at measuring sustainable development (Biggeri, Clark, Ferrannini, & Mauro, 2019). This single composite measurement not only enables an international comparison, but also significantly contributes to "identifying priority areas for action" (Biggeri et al., 2019). Although SDGs are not a legally binding agreement, nations which adopted this framework are willing to improve their sustainability and consequently increase their SDG index. Given the wide variety of SDG-related policies and programs, uncertainties around pandemics, and the time remaining to achieve the goals, the question arises as to which goals require greater attention to boost the SDG index.

A growing body of literature investigates relationships between SDGs by various analytical approaches. Qualitative methods of network analysis used to map the SDGs and based on target wordings, (e.g. Le Blanc, 2015; Mainali, 2018; Singh et al., 2018) served to find relationships between SDGs. Subsequently, this was followed by reviewing documents (Vladimirova & Le Blanc, 2016). To quantify the synergies and trade-offs among SDGs, a Nexus approach, (e.g. Hazarika & Jandl, 2019; Howe, 2019; ICSU, 2017; Nilsson et al., 2016; Weitz, Nilsson, & Davis, 2014) which is evidently one of the most popular approaches, has been used in many studies. A quantitative network analysis to estimate the system of SDG interactions was employed in the study conducted by Lusseau and Mancini (2019). Multi-Criteria Decision Analysis (MCDA) (Jayaraman, Colapinto, La Torre, & Malik, 2015), or identification based on the "willingness to pay" (Barbier & Burgess, 2019) were also methodologies to investigate the relationships between goals.

There are a number of different starting points to explore the relationships, which include starting with: a goal, (e.g. Vladimirova & Le Blanc, 2016); a group of goals/targets (Allen et al., 2017); or across all goals, (e.g. Pradhan, Costa, Rybski, Lucht, & Kropp, 2017). In addition to identifying the relationships among SDGs, Allen Metternicht & Wiedmann (2018) applied MCDA to prioritise SDGs based on their "level of urgency", "systematic impact", and "policy gaps". This research was conducted for 43 targets with 57 indicators. The combined approaches, such as implementing a mix of the cross-impact matrix and network theory (Weitz, Carlsen, Nilsson, & Skånberg, 2017), or a survey-based and network analysis (Jaramillo et al., 2019), are also applied for the same purpose. Although previous studies focused on building the fundamentals of prioritising SDGs, the majority are qualitative in nature or include subjective perspectives. Among the relevant literature there are statistical attempts, which explore the synergies and trade-offs among SDGs by calculating pairwise correlations (Mainali, 2018; Pradhan et al., 2017). Both papers offer valuable insights, although neither paper explicitly addresses the interaction of other goals on the pairs, nor do they seek to calculate how much improvement is required to trigger a positive impact on the other goals.

To summarise, an extensive literature review on prioritising SDGs revealed a number of deficiencies in our knowledge: (a) Identification of goal contributions and patterns that form a higher SDG index; (b) Prioritisation of synergetic goals without subjective or judgmental selection; and (c) Investigation of the goal's score where the synergetic effect begins to exert itself. The findings of this research address and answer identified gaps and the recent question raised in the literature, namely "Which SDGs can be enablers?" (Pincet, Okabe, & Pawelczyk, 2019). An enabler can be interpreted as a synergetic goal which has a high positive and low negative impact on other goals' attainment when improved.

This study offers a combination of machine learning and data mining techniques, the purpose of which is to identify those synergetic goals that are enablers. This is determined by assessing their contribution and importance to the SDG index, using SDG data. Machine learning (ML) techniques are capable of learning and analysing patterns, making predictions, and producing data-driven insights. ML algorithms start the model by learning from the relationship between the response (dependent variable) and its predictors (independent variable) by finding the patterns (Breiman, 2001). As SDGs are complex, ML can facilitate analysis that requires "what-if" roles to be identified. To gain a better understanding of the relationships among SDGs we used Boosted Regression Trees (BRTs). The outstanding characteristic of BRTs are their ability to handle correlated independent variables (Elith, Leathwick, & Hastie, 2008), making it perfect for the analysis of complex behavioural data (Buston & Elith, 2011). There are also other reasons that fit the BRTs to identify synergetic goals which are described in the second section of this study.

In this study, the scores of 17 goals (predictors or independent variables) are used to predict the behaviour of the SDG index (response or dependent variable). In doing so we investigated: (i) the relative influence of each goal on the SDG index when interactions between SDGs are involved; (ii) data-driven identification of synergetic goals; and (iii) the goal's score, which generates the synergetic effect.

Section 2 describes the materials and methods of analysis used in this study. This is followed by Section 3 which elaborates the results, Section 4 which discusses the findings of this research, and Section 5 which presents the conclusions.

2. Materials and methods

In this section, we discuss data resources and methods of analysis used for this study.

2.1. Source of data collection

There are a number of resources that provide SDG data such as the official Global SDG indicators by the United Nations Statistics Division (United Nations, 2020), the World Bank Atlas of Sustainable Development Goals (World Bank, 2020), and Bertelsmann Stiftung and Sustainable Development Solutions Network (SDSN) (Jeffrey Sachs, Guido Schmidt-Traub, Christian Kroll, David Durand-Delacre, & Katerina Teksoz., 2017; Jeffrey Sachs, Guido Schmidt-Traub, Christian Kroll, Guillaume Lafortune, & Grayson Fuller, 2018, 2019) which will be referred to as the BS-SDSN dataset in this study. In our analysis, the BS-SDSN dataset was selected based on four reasons. The first reason referred to the consistency of indicators that are used for calculating a goal's scores. Electing different and inconsistent indicators can change the score of the goals and the SDG index (Pradhan et al., 2017). Although there are more time-series indicators in other data sources, not all of them are applicable to all countries. In the BS-SDSN dataset, there are 86 global indicators and 35 indicators of data for OECD member states which are internationally comparable and applicable to all countries.

The second and third reasons are related to the availability of SDG scores and the consistency of the method used to calculate them. In this dataset, the same method is applied to compute all goal scores and the SDG index. There are other methods of calculation such as the one suggested by Biggeri et al. (2019), one that captures the interaction between goals by rewarding synergies and penalising trade-offs. That calculation is suitable to compute the index of countries when interactions are involved. However, the "simple arithmetic mean" is more suitable for the purpose of this paper, as goals are weighted equally and the equality of weights is applied to all countries. This consistency of data is essential to identify the impact of SDGs when by score calculation they are treated as equivalent. The transparency of the methodology which is used along with the expert-based calculation of scores in the BS-SDSN dataset constitutes the fourth reason that supports our decision to select this dataset.

It is important to note that in this dataset, the SDG data for a minimum of 156 countries out of 193 UN members from 2017 to 2019 is available. About two-thirds of this data is from "official data providers (World Bank, OECD, WHO, FAO, ILO, UNICEF, other)" and the remainder originates from non-official data sources such as household surveys, research institutions, universities, and the civil societies who have more comprehensive datasets (Sachs et al., 2017, 2018, Sachs et al., 2019). The detailed explanation of the data sources and score calculations are explained in greater detail in their methodology report (Sachs et al., 2019).

In this study, SDG scores and the SDG index of 157 countries in 2017, 156 in 2018, and 161 in 2019 are used. The range of all scores is between 0% and 100%. To investigate as whether a temporal analysis should be conducted for the values of SDG index, ANOVA analysis has been undertaken for 2017, 2018 and 2019. Since the means of SDG index for various years are not significantly different ($p = 0.279$), values of the SDG index over three years have been aggregated into a single variable. In other words, the parameter of year/time is excluded from this analysis. The records of countries with missing SDG index are eliminated and the score of goals which are missed are replaced by *"NA"*. In total, 474 observations are used in our analysis.

2.2. Method of analysis

Regression models are often used to measure the relationship between dependent and independent variables. In order to predict the SDG index and identify the contribution of each goal, an efficient regression-based machine learning approach which minimises the prediction errors is required. In doing so, an ensemble machine learning and data mining technique called Boosted Regression Trees (BRT) is applied.

2.2.1. Why BRT?

There are four main reasons why BRT is selected for this study:

i) The existence of correlations between goals is proven statistically (Pradhan et al., 2017). Therefore, typical regression models are not suitable for modelling and exploring SDGs as they often assume that independent variables do not have a correlation with each other (Breiman, Friedman, Olshen, & Stone, 2017). Ignoring this can potentially cause problems such as multicollinearity in which reducing the prediction power impacts directly on the result interpretation (Kim, 2019).

ii) In our dataset, there are missing goal scores for some countries and there may also be goals which do not make a significant contribution to the SDG index. Tree-based methods can accommodate any type of data (i.e. quantitative and qualitative), missing values in the predictors as well as handling uninformative predictors (Friedman & Meulman, 2003).

iii) An individual regression tree has limited prediction power (Breiman et al., 2017). On the other hand, the ensemble of trees is more effective, flexible, and less data-sensitive for prediction. Boosting is an ensemble technique that fits thousands of trees in a sequential and stage-wise fashion (Elith et al., 2008). This explains why machine learning techniques which utilise a single tree to make the prediction are not used in this study.

iv) In BRT, the level of interactions and model complexity can be defined (tree complexity) which can result in a larger number of trees that generally tend to have superior predictive performance (Elith et al., 2008). This explicit control over the model parametrisation does not exist in similar tree-based models like Random Forest (RF). For RF it is specifically suggested to fit deep trees (Breiman, 2001) and this means there is no control over the level of interactions. BRT explicitly controls the level of interactions and unlike RF which uses bagging, it employs boosting. Boosting function makes BRT models progressive and stage-wise learners besides handling very shallow trees by fitting more trees. Since the aim of this study is interpreting relationships between SDGs and the SDG index, BRT is a more suitable model than RF because we have more control over the model fitting.

This powerful boosted regression technique provides robust predictive performance and automatically accounts for the interaction between independent variables. It also reliably selects influential predictors without being affected by multicollinearity (Buston & Elith, 2011). These comprehensive features of the BRT make it a perfect option to investigate influential factors on the dependent variable, especially when dealing with complex non-linear relationships.

2.2.2. How does BRT operate?

There are two algorithms that BRT uses. Boosting combines and builds collections of models and regression trees which are categorised in the classification and regression tree group models. Regression trees, also known as decision trees, are non-linear models that recursively partition (split) the predictor space, in order to model the pattern in the response data (Breiman et al., 2017; Friedman & Meulman, 2003). This data splitting consists of a series of "if-then" statements in the predictors. BRT tries to minimise the prediction errors in each stage by adding another tree. To elaborate, on this, in each stage a regression tree is fitted on top of the previously fitted trees, focusing on poorly predicted observations by those trees (Elith et al., 2008). The algorithm of building a regression tree is summarised in Algorithm 1 which is adopted from James, Witten, Hastie, & Tibshirani (2013, p. 309).

Algorithm 1: Building a Regression Tree

1. Use recursive binary splitting to grow a large tree on the training data, stopping only when each terminal node has fewer than some minimum number of observations.
2. Apply cost complexity pruning to the large tree in order to obtain a sequence of best subtrees, as a function of α.
3. Use K-fold cross-validation to choose α. That is, divide the training observations into K folds. For each $k = 1,\ldots,K$:

a.

Repeat Steps 1 and 2 on all but the kth fold of the training data.

b.

Evaluate the mean squared prediction error on the data in the left-out kth fold, as a function of α.

Average the results for each value of α, and pick α to minimise the average error.

4. Return the subtree from Step 2 that corresponds to the chosen value of α.

Boosting models use bootstrap to create multiple copies of the dataset, combines a large number of decision trees which are grown sequentially using information from last fitted trees. Growing trees by learning from previous trees means that the boosting model learns slowly. This prevents overfitting which might occur in a single large decision tree fitting. The algorithm of boosting for regression trees is explained in Algorithm 2 which is adopted from James, Witten, Hastie, & Tibshirani (2013, p. 323).

Algorithm 2: Boosting for Regression Tree

1. Set $\hat{f}(x) = 0$ and $r_i = y_i$ for all i in the training set.
2. For $b = 1, 2, \cdots, B$, repeat:
   a. Fit a tree $\hat{f}^b$ with d splits (d + 1 terminal nodes) to the training data (X, r).
   b. Update $\hat{f}$ by adding in a shrunken version of the new tree:
   $$\hat{f}(x) = 0 \leftarrow \hat{f}(x) + \lambda \hat{f}^b(x).$$
   c. Update the residuals,
   $r_i \leftarrow r_i - \lambda \hat{f}^b(x_i)$.
   d. Output the boosted model,
   $\hat{f}(x) = \sum_{b=1}^{B} \lambda \hat{f}^b(x)$.

where B is the number of trees, λ is the shrinkage parameter and d is the interaction depth.

Fitting and visualisation of all models are conducted in R programming language (R Core Team, 2019). The BRT model is fitted using the gbm.step function in the dismo package (Hijmans, Phillips, Leathwick, & Elith, 2017). To fit a BRT model, two main parameters must be specified in order to estimate the optimal number of trees required for prediction. Those parameters are tree complexity and learning rate. The first parameter controls the number of nodes for each individual tree (also known as tree depth). This value is directly related to the level of interaction allowed in the model. When the tree complexity is set to 1, it indicates that each tree has only a single node or decision rule. This is called a 'stump' that fits an additive model (Elith et al., 2008). A value of 2 means each tree has two nodes and makes possible a two-way interaction and so on. We used a tree complexity of 1, 2 and 5 to examine the contribution of SDGs with no interaction, two levels of interactions, and high interaction among goals. The other parameter is the rate at which the model complexity increases (Elith et al., 2008). We used 0.001 for the learning rate since the lower number of this parameter results in a larger number of trees that generally tend to have a superior predictive performance (Elith et al., 2008).

Using BRT, the relative influence of each predictor (SDG) on the response variable (the SDG index) is estimated when high levels of interactions are set (tree complexity 5). The result is then scaled so that the sum adds up to 100%, and then displayed in a horizontal bar chart. To compare the difference between the relative influence of each predictor on the response variable, we apply the BRT with a tree complexity of 1 (no interaction between predictors) and 5 (high level of interaction between predictors). This helps us to compare the result when interactions between goals are ignored whilst they are included.

2.3. The impact of each SDG on the SDG index

The partial dependence plots (PDP, also known as the response curve) are used to visualise the behaviour of each goal score on forming the SDG index when no interactions are included (tree complexity of 1). The PDP illustrates whether the relationship between the predictor and the response variable is linear, monotonic or more complex. It also provides us with an insight into how the response variable is impacted by the increase in the value of an interest predictor variable (Friedman & Meulman, 2003). PDP can provide a detailed picture of how changes in each goal score impact on the SDG index, when all other goals are hold at a constant value (i.e. their mean).

2.4. Interaction among synergetic SDGs

In order to illustrate the interactions between goals and their impact on the SDG index, three-dimensional partial dependence plots (3DPDPs) are used. 3DPDPs help to interpret interactions between two predictors with the response variable and is useful for models that allow fitting interaction terms (Elith et al., 2008). For this visualisation we applied the tree complexity of 2 (two levels of interactions) in the BRT model to all observations. This feature of 3DPDPs allows us to investigate the possible interactions between those goals where changing their values might result in a higher or lower index. For this article, the interaction of those exerting more influence on the SDG index will be visualised by 3DPDPs.

2.5. Evaluating the model's performance

One of the most commonly used techniques to statistically evaluate prediction performance is known as hold-out cross-validation (Granholm, Noble, & Kall, 2012). In this technique data will be divided into two parts. One part is used for training purposes and another part serves to test the prediction accuracy. However, the significant limitation of this method is how appropriately the data is divided (Zhao & Wang, 2014). Unsystematic data split and improper division leads to inconsistency in fitting the model. A growing body of literature recommends the K-fold cross-validation (CV) technique to combat the limitation of hold-out validation (Valavi, Elith, Lahoz-Monfort, & Guillera-Arroita, 2019). The simplicity and universality of the CV makes it one of the most popular approaches of model validity. Using this technique, data is divided into K parts where K-1 is used for fitting the model, with the result then tested on the remaining part. CV provides a method for progressively developing and testing models on the hold-out portion of the data. This ensures that the final model is general enough to have an accurate prediction on the withheld data (Valavi et al., 2019).

To validate the prediction model with k-fold CV, any number of folds can be used. However, 5 and 10 folds are the most popular values recommended in research studies. Based on our dataset which has less than 500 records, a 5-fold CV is selected to develop our BRT model. This selection prevents an over-fitting of our model. In a 5-fold CV our data is divided into five parts that are all used simultaneously in modelling. In each fold, one part of the data is used for the testing and the rest are used to train the machine. To examine whether the model is over-fitting or not, we evaluated the performance of our model using CV over the years. This helped us to assess the out-of-sample accuracy of our model. In this CV data for two years is used to train the machine and the remaining year used to test forecasting performance. This process is repeated three times to cover all sample data related to 2017, 2018 and 2019.

3. Results

The BRT technique is applied to 474 observations to find the contributions of goals which are on the SDG index. The results are presented in more details below.

3.1. Model fitting by BRT

Using the BRT technique with the tree depth of 5 and the learning rate value of 0.001, the total number of 7500 trees is selected and used by CV to model the SDG index. As can be seen in Fig. 1, the true and predicted values of the SDG index are perfectly falling on the straight line. The coefficient of determination (or R^2) is an evaluation measure of the agreement between the predicted values and actual data. When the model is fitted, 98% of the variance of the dependent variable (the SDG index) is explained by the variability of the independent variable (SDGs). This indicates that the fitted model produced a very accurate prediction on the hold-out data with only 0.02 error rate. The error rate between observed and predicted data for Root Mean Squared Log Error (RMSE), Mean Squared Error (MSE) and Mean Absolute Error (MAE) are 0.76, 0.57 and 0.59, respectively.

The CV to assess the out-of-sample accuracy over years resulted in the very similar R^2 with only 0.025 error rate. Fig. 2 displays the plot resulting from the observed and predicted values for the SDG index over years. This outcome is a valid indication that the model is not overfit by BRT and can be perfectly generalised.

3.2. Contributions of SDGs to the SDG index

In predicting the response variable (the SDG index) not all predictors (goals) have relative or equal importance. According to the relative influence plot with a tree complexity of five in Fig. 3, there are three SDGs which had a particularly significant contribution to predicting the SDG index. These include: SDG3, "Good health and well-being"; SDG4, "Quality education"; and SDG7, "Affordable and clean energy". These three goals account for 75% of predicting the SDG index.

Among those goals, "Good health and well-being" is the salient factor comprising 42% of the relative influence on the SDG index. On the other hand, Goal16, "Peace, justice and strong institutions", wields the minimum influence with a 0.12% contribution to predict the SDG index.

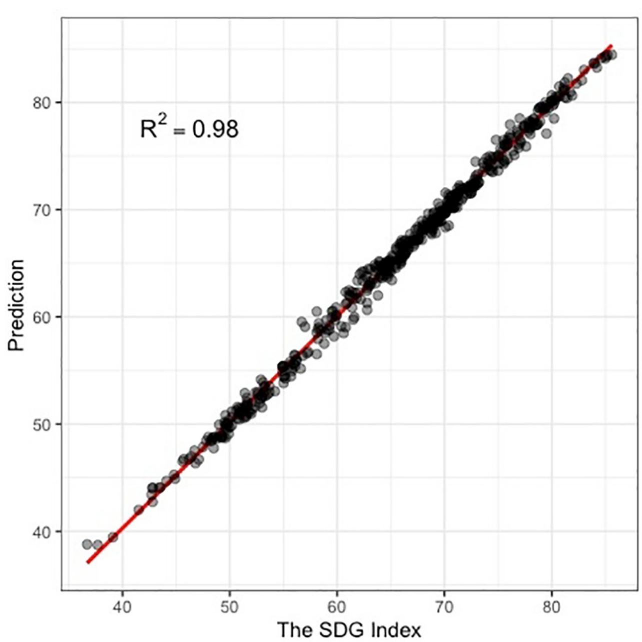

Fig. 1. Scatter plot showing the actual versus the predicted values based on 5-fold cross-validation using all observations.

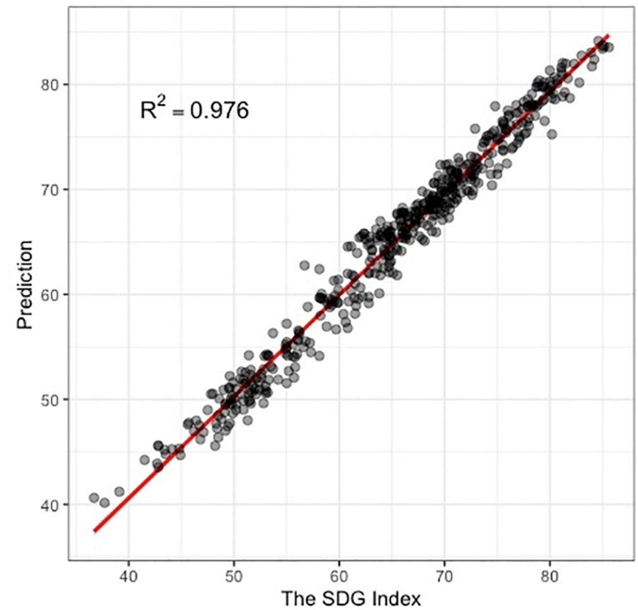

Fig. 2. Scatter plot showing the actual versus the predicted values over years.

The SDG12, "Responsible consumption and production", SDG13, "Climate action", SDG14, "Life below water", and SDG17, "Partnerships for the goals" are contributing less than one percent to form the SDG index. The total contribution of these five SDGs is less than 3%, and the rest of the SDGs are in between.

3.3. Relevant influence of SDGs based on the level of SDG index

To further our understanding of the driving forces of the relevant influence of SDGs on the SDG index, we divided the main dataset into two sub-datasets based on the SDG index score of countries: Group A with a SDG index less than or equal to the mean ($\leq$65.3); and Group B with a SDG index more than mean (>65.3). We then applied the same settings (tree complexity 5, learning rate 0.001, and 5-fold CV) to each group to evaluate the change of relative influence based on the SDG index. Fig. 4 shows the different contributions made by SDGs in each group.

The result indicates SDG3 is one of the highly influential goals in both groups. SDG4 still is a highly ranked goal in Group A along with SDG1. It is interesting that SDG1 which is ranked third among influencers in Group A, is among the least important influencers in Group B. SDG9 and SDG10 are among the top three goals with high relative influence in Group B and SDG16 is still one of the smallest contributors in both groups. To improve the transparency of comparison, Table 1 lists the top three SDGs from highest to lowest contributors in each category (All Observations, Group A and Group B). In all categories, SDG3 and SDG9 are among the top five contributors to the SDG index.

3.4. Contributions by different levels of interactions

The relative influence of SDGs on the SDG index displays different results when goals are studied in isolation when compared to including all goals' interactions. Fig. 5 compares the contribution of SDGs to the SDG index in isolation and when high levels of interaction are included. SDGs which have less than 1% relevant influence are excluded from this chart.

As can been seen, SDG1, "No poverty", SDG2, "Zero hunger", SDG4, "Quality education", SDG6, "Clean water and sanitation", SDG8, "Decent work and economic growth", SDG9, "Industry, inno-

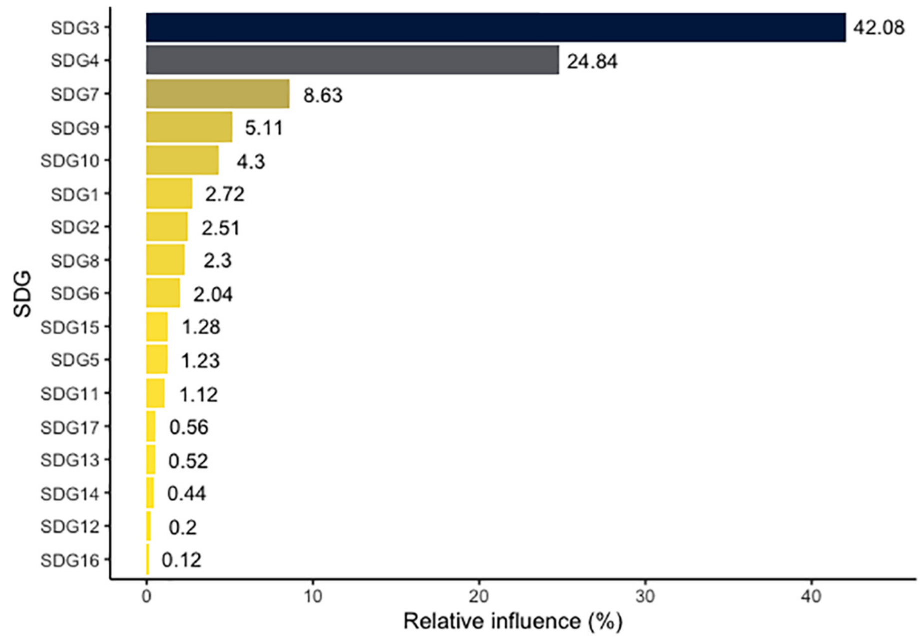

Fig. 3. The relative influence (%) of SDGs on the SDG index based on all observations. The values show the percentage for the importance of each goal applying BRT with the high level of interactions (tree complexity of 5).

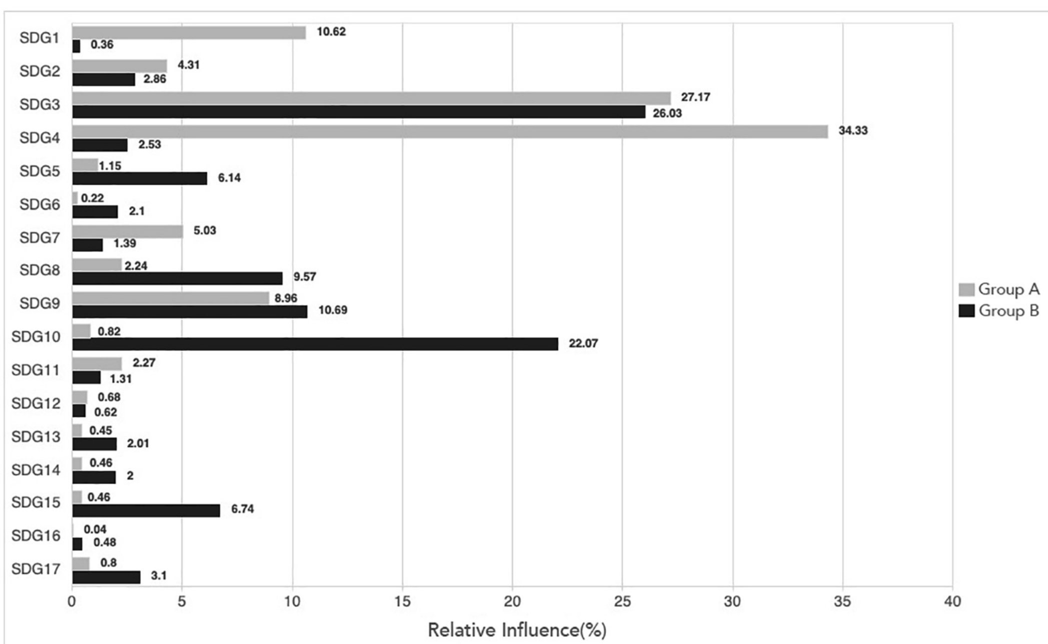

Fig. 4. Relative influence (%) of the SDGs on the SDG index result based on BRT model with tree-complexity 5, learning rate 0.001 and 5-fold CV for Group A (Countries with SDG index ≤ 65.3), and Group B (Countries with SDG index > 65.3).

Table 1
SDGs with high relative influence on the SDG index for all observations, Group A (SDG index ≤ 65.3) and Group B (SDG index > 65.3). Top three SDGs sorted from highest contributors to the lowest.

SDGs Relative Influence (%) on the SDG index					
All Observations		Group A		Group B	
SDG3	42.08	SDG4	34.33	SDG3	26.03
SDG4	24.84	SDG3	27.17	SDG10	22.07
SDG7	8.63	SDG1	10.62	SDG9	10.69

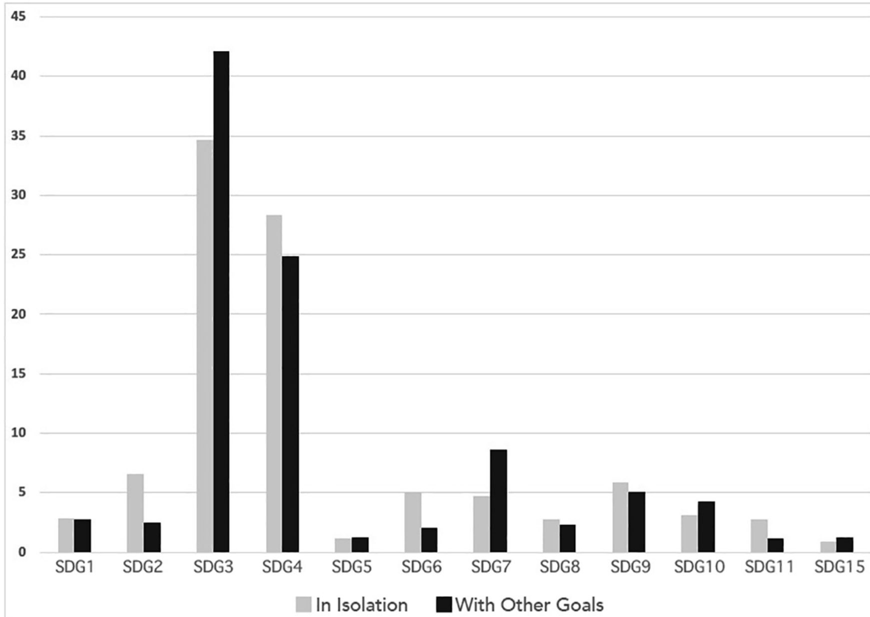

Fig. 5. Comparison of SDGs contribution on the SDG index with and without interactions among goals.

vation and infrastructure", and SDG11, "Sustainable cities and communities", are waning in terms of their contribution when the level of interaction increases. The opposite is the case, however, in SDG3, "Good Health and Well-being", SDG5, "Gender equality", SDG7, "Affordable and clean energy", SDG10, "Reduced inequalities", and SDG15, "Life on land". These have more influence on the SDG index when interacting with other goals. SDG9, "Industry and innovation and infrastructure", with no interaction makes a higher contribution to the SDG index than SDG7, "Affordable and clean energy". Yet when all goals are involved, SDG9 contributes less than SDG7.

3.5. SDG scores and their impact on the SDG index

Fig. 6 illustrates the PDP when BRTs with no interactions (tree complexity 1) are fitted. All SDGs wield a positive upward effect on the SDG index but with a different magnitude. The SDGs with a higher contribution or relative importance have steeper curves and vice versa. The PDP of the SDG3, "Good health and well-being", reveals a notable pattern.

It starts and continues with a sharp rise and involves two abrupt jumps at around 62% and 90% with a substantial increase in the predicted values compared to other goals. This pattern means that increasing the score of health and well-being to >60% can improve the SDG index considerably. SDG4, "Quality education", starts with a gentle upward slope up to 62%, and then it continues with a rapid rise up to the end. This goal is the second contributor to the SDG index. The SDG7, "Affordable and clean energy", however, begins with a slight increase and continues steadily up to the same point, then experiences another slight rise. Most of the other goals have no significant effect on the SDG index up to a certain score. The SDG16, "Peace, justice and strong institutions" and SDG12, "Responsible consumption and production", illustrate no noticeable effect on improving the SDG index. These two makes a very small contribution to the prediction of the SDG index.

3.6. Interactions between SDGs and their impact on the SDG index

Interaction happens when the SDG index is affected simultaneously by more than one goal. The most robust interactions are between the SDG4, "Quality education", with SDG3, "Good health and well-being", and SDG10, "Reduced inequalities". The last but not least of the interactions occur between SDG9, "Industry, innovation and infrastructure", and SDG4. Here we visualised the first two pairs interactions in Fig. 7.

The right-hand side plot displays the interaction of SDG3, "Good health and well-being", SDG4, "Quality education", and the SDG index. The SDG index is at its intermediate level when SDG3 is at its maximum value (corner B) and the SDG4 is at its minimum (corner A). The SDG index stands at its highest level whilst both goals are at their maximum value. This indicates the high synergy that is generated when SDG3 and SDG4 are interacting. This means that increasing the SDG3 score does not lead to a high SDG index if SDG4 is not simultaneously high.

The left-hand side plot in Fig. 7 displays the interaction between SDG10, "Reduced inequalities", and SDG4, "Quality education". In this example, however, higher values of the SDG index are associated with the higher values of SDG4 (corner C) compared to the higher values of SDG10 (corner B). This indicates the fact that SDG4 contributes more than SDG10 to form the higher SDG index. The SDG index is relatively high when SDG4 is high and is not impacted by the values of SDG10. On the contrary, with high values of SDG10 (corner B), the SDG index is at an intermediate level when SDG4 is at its minimum (corner A). This changes dramatically as the value of SDG4 increases (corner C).

It is important to note that the SDG index values have a higher value when interaction between goals is included in comparison to when no interaction is involved. In PDP the SDG index is in a 60% to 65% range, while in 3DPDPs it is between 60% and 70%.

4. Discussion

The contribution of SDGs and their interactions to form a higher SDG index plays a crucial role in prioritising the goals. Applying the Boosted Regression Trees model in this study resulted in a high level of accuracy between the predicted SDG index and observed values, which confirms how fit the BRT model is for our dataset. The result based on the global data indicates that when the high level of interaction is considered the SDG index is influenced

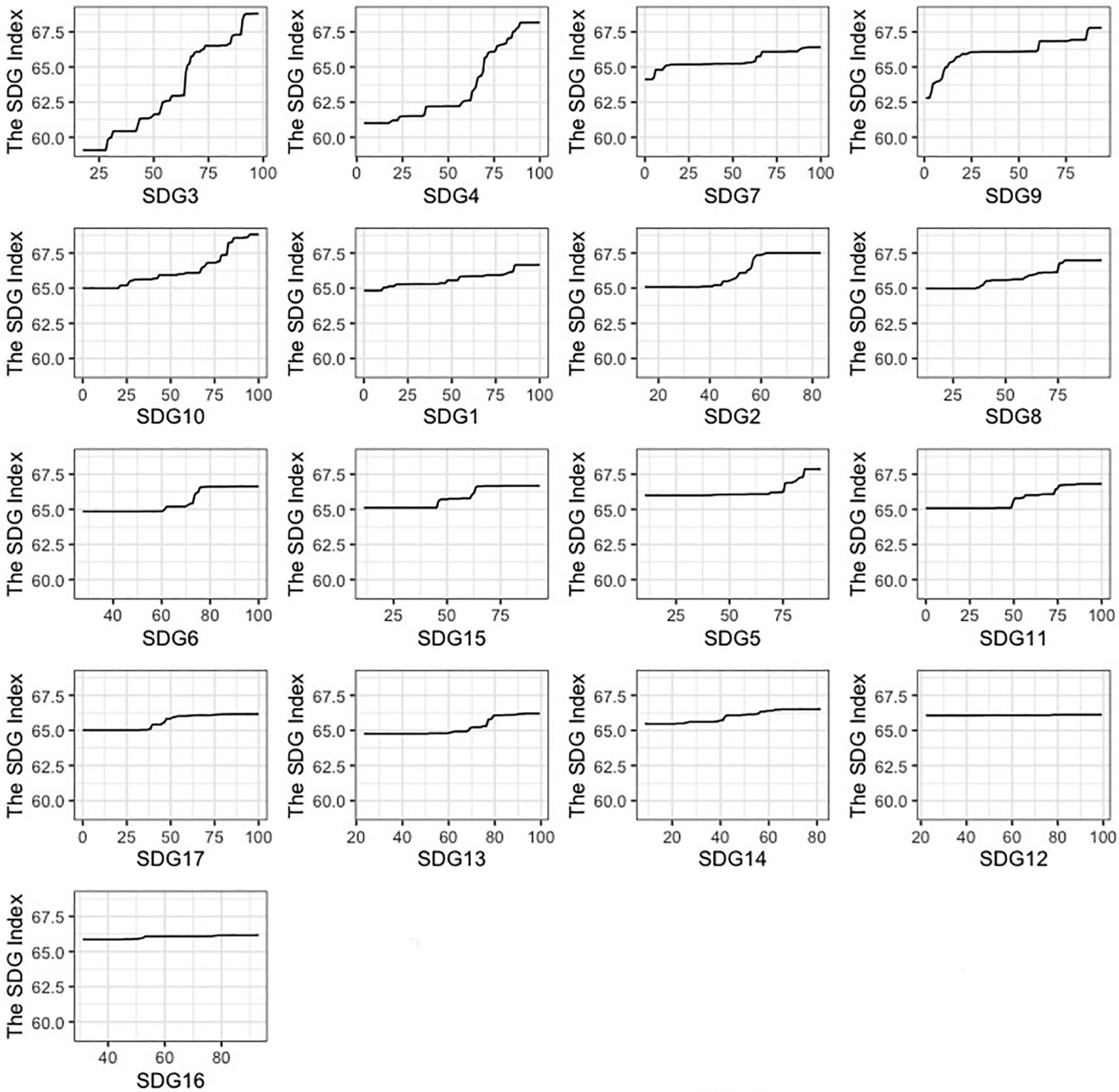

Fig. 6. Partial dependence plots for SDGs showing the changes in SDG index resulting from variations of each goal score within its range when all other goals are hold at a constant value (i.e. their mean) The plots are sorted based on goal contributions (the percentage value).

almost 75% by three goals, SDG3, "Good health and well-being", SDG4, "Quality education", and SDG7, "Affordable and clean energy". In a recent study which suggests six transformations keys to achieve the SDGs, education, health and energy are identified among the first three keys which can enable SDGs attainment (Sachs et al., 2019). These suggested keys are driven from an extensive literature review using a completely different approach. This study has clearly pointed out the contributions of education, health and energy to meet the SDGs.

According to the scientific perspective provided by the Council of Science (2015), SDG3 is the only goal whose targets are linked to all other goals, and "is related to every other aspect of development either as an input or as a consequence of activity in other goals". This can also be supported by the fact that international goals regarding public health have a significant role to play in the goals' complex development (McArthur & Rasmussen, 2019). In a quantitative time-series analysis which investigated the correlation among SDGs indicators, SDG3 has ranked as the top synergetic goal (Pradhan et al., 2017). The result of prior analysis along with this study has significantly and clearly agreed about the impact of SDG3, "Good health and well-being", on the other SDGs and the SDG index. This sheds some light on how catastrophic the impact of pandemic disease such as COVID-19 on SDGs might be. The adverse effects of such an outbreak on SDG3 can directly or indirectly influence other goals and consequently on sustainable development.

After good health and well-being, quality education earns the second rank of importance to form the SDG index. Literature has justified the broad positive impacts of education on all dimensions of SDGs (society, economy, and environment) (Bengtsson, Barakat,

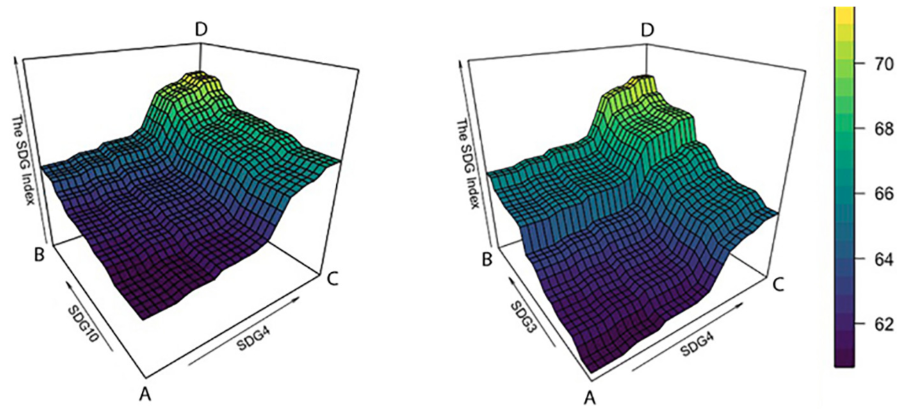

Fig. 7. Three-dimensional PDP illustrates the interaction between SDG 3 and SDG 4 (right) along with SDG 10 and SDG 4 (left), fitted by BRT model of depth 2. Note that the magnitude of the colours in the plot show different ranges of the SDG index.

& Muttarak, 2018). Education can promote economic growth (Jeong, Lee, & Kang, 2020), and economic growth contributes to eliminating extreme poverty (Dollar & Kraay, 2002), provides decent work or opportunities to find employment (Singh, 2005), helps to overcome gender and other inequalities (Karlsson, 2009), and many more positive outcomes. It is also identified as a critical factor in "addressing environmental and sustainability issues and ensuring human well-being" (Council of Science, 2015; WHO et al., 2018). Therefore, it can be concluded that positive direct and indirect impacts of SDG4 convert it into a high contributor to the SDG index. The positive synergy that SDG4 generates when interacting with SDG3 increases the SDG index substantially.

The role of energy in sustainability is a well-established concept. In our study, affordable and clean energy (SDG7) is identified as the third significant contributor of the SDG index based on the global data and among top five contributors for countries with SDG index $\leq$ 65.3. This goal's contribution to the SDG index increases when interactions between goals are included. There are potential chain-effects between SDG3, SDG4, and SDG7 (Collste, Pedercini, & Cornell, 2017) that might cause the multi-fold positive impact on other goals and as a result on the SDG index.

It is notable that SDG12, "Responsible consumption and production", is the least predictor when no interaction is involved. In a systematic quantitative analysis, which assessed the correlation between time-series indicators, this goal is mentioned in 7 of the top 10 goals which have a negative correlation (trade-offs) with other goals (Pradhan et al., 2017). In the same time-series study, SDG12 has been identified as a goal having negative correlations with 10 other goals (SDGs 1–7, 9, 10, 17). This might be the reason that changing the SDG12 score does not have a significant impact on predicting the SDG index in our analysis. The same theory cannot be applied to SDG16 which is the least contributor to the SDG index when the level of interaction is high. There may be other reasons, such as the relationship between peace and justice with sustainability that requires further investigation. After SDGs12 and 16, SDG14, "Life below water" is one of the least contributors in our global analysis. This might be related to this goal's applicability to all countries. SDG14 targets mostly cover ocean health and marine life. Almost one-fifth of the world's countries have no direct access to an ocean and many countries have a limited fisheries industry. This may also explain why SDG14 has many missing scores in the BS-SDSN dataset.

Based on the level of the SDG index, improvement of SDG1, SDG3, and SDG4 for countries with SDG index $\leq$ 65.3, and SDG3, SDG9, and SDG10 for countries with SDG index > 65.3 can significantly boost their SDG index. According to the sustainable development report of 2019, more than two-thirds of countries with SDG index $\leq$ 65.3, score below 60% for their most contributing goals (SDG1, SDG3, and SDG4) (Sachs et al., 2019). The same report indicates similar rate (two-thirds) for SDG9 and one-third for SDG10 in countries with SDG index > 65.3 which score less than 60% in their synergetic goals. It is notable that SDG10, "Reduced Inequalities" is the most influential goal after SDG3 for countries with a more than average SDG index. There are many synergetic targets in SDG10 which wield a positive impact on other targets and goals. In a recent study which investigated the impact of SDG10 in high-income countries, this goal has been identified as having the most positive effect on all other goals (Lusseau & Mancini, 2019). In our study almost two-thirds of countries with the SDG index > 65.3 can be categorised as being in the high-income group. It is interesting that in the same study, SDG1 has been identified as one of the most influential goals for low-income countries, which is aligned with the result of our further analysis based on the level of SDG index. This can help global and national decision-makers to prioritise policies and allocate resources which make it possible to achieve of those goals and boost their SDG index significantly.

Our study based on the global data suggests that the most interactions are among SDG3-SDG4, and SDG4-SDG10. Apart from SDG3 and SDG4 which were covered above, when delving into the literature, many studies discuss the positive impact of SDG10 on other goals. The majority of targets of SDG10 are aiming for social protection systems and labour standards. Social protection eases access of society's vulnerable groups to health and education services. Good health improves education (McArthur & Rasmussen, 2018; Shonkoff, Radner, & Foote, 2017) and education improves health (WHO etl, 2018). Both SDG3 and SDG4 reinforce SDG10 (Bengtsson et al., 2018). The circle of cause and effect among these goals can support the high level of interactions characterising the goals.

It is interesting to note that the importance of goals to predict the SDG index changes when the relationships among goals are considered. The contribution of some goals when the interaction is none is higher than when 5 levels of interaction are included. This can be explained by the negative impacts of SDGs on each other. Some goals are indicating a higher contribution when the tree depth is increased. This relates to the positive impacts of goals on each other. As well, the maximum value of the SDG index changed when goals were studied in isolation or along with other goals. This can be a good indication of how synergies and trade-offs among goals can increase or decrease SDG contributions to the SDG index.

According to patterns found in the PDP and the 3DPDPs, traditional linear regressions, pair analysis, or correlations among goals may not be the best approach for the purpose of achieving SDGs. Even improving goals does not have a linear impact on the SDG index. Given the non-linearity relationship between goals and the index, forecasting how countries are progressing towards achieving SDGs might not provide correct and precise output without a systematic prediction which considers how changing the score of each goal alters the SDG index.

5. Conclusion

Given the heterogeneous policies and programs of each country toward achieving the SDGs with their limited resources, and considering that the deadline of 2030 is fast approaching, it is necessary to prioritise synergetic goals. This study reveals that prioritising goals without including the interaction among them may not be suitable for SDGs. In other words, attempts to achieve goals in isolation may need more resources in comparison to when the synergies of goals are activated. This study identified both the contributions of SDGs and the impact of different goal scores on the SDG index without defining any judgmental criteria. Identifying how much improvement is required for each goal to maximise the synergy among them plays a pivotal role in developing an effective framework for policy-makers. This helps to make informed decisions on how to allocate limited resources to achieve the maximum SDG index.

The findings of this investigation indicate that "Good health and well-being" is one of the most important synergetic goals. There have been disease outbreaks throughout history that have been so deadly, widespread and costly to manage that their impact has been devastating throughout society and the economy. As SDGs try to balance social, economic and environmental aspects, outbreaks like COVID-19 can have a massive impact on the prospects of achieving the goals. An increase in mortality rate and negative impacts on health services can reduce the SDG3 score. Adding the impact of such pandemics to the findings of this research, which identified SDG3 as the most influential goal on the SDG index, brings to light how significant the impact of COVID-19 can be on sustainable development. This can change the sustainability index of countries that have been highly impacted by pandemic disease. Even solutions that are considered to control the pandemic, such as social distancing and self-isolation, which negatively impact on the entertainment industry, travel bans, which have a tremendous impact on the tourism industry, and a reduction in each nation's spending due to the fear of a recession, can all adversely impact on the economy. This is in addition to the allocation of a considerable amount of financial aid and resources to control the economic impact on society, by providing help to people who are encouraged to stay at home in order to curtail the spread of deadly diseases. When a pandemic strikes, achieving targets of goals becomes more challenging than in non-pandemic times. It is evident that during and soon after an outbreak, government resources may not be allocated as before. This is because priorities will change after an outbreak or, due to the limited budget, resource allocation will be affected.

If the consequences of the pandemic are more detectable on the least influential goals such as SDG16, "Peace, justice and strong institutions", SDG12, "Responsible consumption and production", SDG13, "Climate action", SDG14, "Life below water", or SDG17, "Partnerships for the goals", it might not significantly impact on other goals and the SDG index. However, if SDG3, "Good health and well-being", SDG4, "Quality education", and SDG7, "Affordable and clean energy", SDG9, "Industry, innovation and infrastructure", SDG8, "Decent work and economic growth", and SDG10, "Reduced Inequalities" are being impacted then it might affect the other goal scores and consequently the SDG index. Further research is required to evaluate the magnitude of pandemic disease on the synergetic goals.

It is worth mentioning that SDG targets are expected to be met by 2030, a mere decade away. The findings of this research can extensively help decision-makers who aim to achieve SDGs and have smart resource allocation. By prioritising synergetic goals instead of random, politically salient selections, or the ease with which the goal can be attained, more positive results can be achieved.

Utilising the data-driven and machine learning analysis in this research identified patterns which provide better information for policy-makers to prioritise goals which can maximise SDG achievement by utilising synergies among them. This analysis is conducted based on global data, however, for more context-based and relevant insights, future studies can investigate goals which are enablers for different levels of income, and at the local and national contexts. To provide comprehensive, relative, and more in-depth insights for decision-makers, the contribution of targets and indicators to the SDGs can also be examined. Considering the non-linear relationships between goals and the SDG index, further research is required to assess how to include interactions to forecast the SDG index when the goal scores are changing to having a more accurate estimation of where countries are standing in 2030.

The major limitation in this research is related to data. In SDGs data that is used in this study, some tier 2 and tier 3 indicators are not included when calculating goal scores. There might be more or less synergies by involving those indicators in the score computation. It is also important to note that data is for the last three years and findings in this analysis might not be applicable in the future if there are substantial changes in the data. Finally, evolving data quality, integrity, and availability might alter results in the future.

CRediT authorship contribution statement

Atie Asadikia: Conceptualization, Methodology, Software, Investigation, Formal analysis, Visualization, Validation, Writing - original draft. **Abbas Rajabifard:** Writing - review & editing, Project administration, Supervision. **Mohsen Kalantari:** Writing - review & editing, Supervision.

Declaration of Competing Interest

The authors declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper.

References

Allen, C., Metternicht, G., & Wiedmann, T. (2019). Prioritising SDG targets: Assessing baselines, gaps and interlinkages. *Sustainability Science, 14*(2), 421–438. https://doi.org/10.1007/s11625-018-0596-8.

Allen, C., Nejdawi, R., El-Baba, J., Hamati, K., Metternicht, G., & Wiedmann, T. (2017). Indicator-based assessments of progress towards the sustainable development goals (SDGs): A case study from the Arab region. *Sustainability Science, 12*(6), 975–989. https://doi.org/10.1007/s11625-017-0437-1.

Barbier, E. B., & Burgess, J. C. (2019). Sustainable development goal indicators: Analyzing trade-offs and complementarities. *World Development, 122*, 295–305. https://doi.org/10.1016/j.worlddev.2019.05.026.

Bengtsson, S. E. L., Barakat, B., & Muttarak, R. (2018). The role of education in enabling the sustainable development agenda. Routledge.

Biggeri, M., Clark, D. A., Ferrannini, A., & Mauro, V. (2019). Tracking the SDGs in an 'integrated' manner: A proposal for a new index to capture synergies and trade-offs between and within goals. *World Development, 122*, 628–647. https://doi.org/10.1016/j.worlddev.2019.05.022.

Breiman, L. (2001). Statistical modeling: The two cultures. *Statistical Science, 16*(3), 199–215. https://doi.org/10.1214/ss/1009213726.

Breiman, L., Friedman, J. H., Olshen, R. A., & Stone, C. J. (2017). Classification and regression trees. *Classification and Regression Trees*, 1–358. https://doi.org/10.1201/9781315139470.

Breuer, A., Janetschek, H., & Malerba, D. (2019). Translating Sustainable Development Goal (SDG) interdependencies into policy advice. *Sustainability (Switzerland), 11*(7). https://doi.org/10.3390/su1102092.

Buston, P. M., & Elith, J. (2011). Determinants of reproductive success in dominant pairs of clownfish: A boosted regression tree analysis. Journal of Animal Ecology, 80(3), 528–538. https://doi.org/10.1111/j.1365-2656.2011.01803.x

Collste, D., Pedercini, M., & Cornell, S. E. (2017). Policy coherence to achieve the SDGs: Using integrated simulation models to assess effective policies. *Sustainability Science, 12*(6), 921–931. https://doi.org/10.1007/s11625-017-0457-x.

Council of Science. (2015). Review of the SDGs :The Science Perspective.

Dervis, K., & Klugman, J. (2011). Measuring human progress: the contribution of the Human Development Index and related indices. Revue d'économie Politique, 121(1), 73.

Dollar, D., & Kraay, A. (2002). Growth is good for the poor. *Journal of Economic Growth, 7*(3), 195–225. https://doi.org/10.1023/A:1020139631000.

Elith, J., Leathwick, J. R., & Hastie, T. (2008). A working guide to boosted regression trees. *Journal of Animal Ecology, 77*(4), 802–813. https://doi.org/10.1111/j.1365-2656.2008.01390.x.

Friedman, J. H., & Meulman, J. J. (2003). Multiple additive regression trees with application in epidemiology. *Statistics in Medicine, 22*(9), 1365–1381. https://doi.org/10.1002/sim.1501.

Granholm, V., Noble, W., & Käll, L. (2012). A cross-validation scheme for machine learning algorithms in shotgun proteomics. *BMC Bioinformatics, 13*(16), S3. https://doi.org/10.1186/1471-2105-13-S16-S3.

Hazarika, R., & Jandl, R. (2019). The Nexus between the Austrian forestry sector and the sustainable development Goals: A review of the interlinkages. *Forests, 10*(3), 205. https://doi.org/10.3390/f10030205.

Hijmans, R. J., Phillips, S., Leathwick, J., & Elith, J. (2017). Dismo: Species distribution modeling.

Howe, P. (2019). The triple nexus: A potential approach to supporting the achievement of the Sustainable Development Goals?. *World Development, 124*, 104629. https://doi.org/10.1016/j.worlddev.2019.104629.

ICSU (2017). A Guide To SDG Interactions : From Science to Implementation, Internatio, 1–239. https://doi.org/DOI: 10.24948/2017.01 ICSU

James, G., Witten, D., Hastie, T., & Tibshirani, R. (2013). An introduction to statistical learning. https://doi.org/10.1007/978-1-4614-7138-7_8

Jaramillo, F., Desormeaux, A., Hedlund, J., Jawitz, J. W., Clerici, N., Piemontese, L., ... Åhlén, I. (2019). Priorities and interactions of Sustainable Development Goals (SDGs) with focus on wetlands. Water (Switzerland), 11(3). https://doi.org/10.3390/w11030619

Jayaraman, R., Colapinto, C., Torre, D. L., & Malik, T. (2015). Multi-criteria model for sustainable development using goal programming applied to the United Arab Emirates. *Energy Policy, 87*, 447–454. https://doi.org/10.1016/j.enpol.2015.09.027.

Jeffrey Sachs, Guido Schmidt-Traub, Christian Kroll, David Durand-Delacre, & Katerina Teksoz (2017). SDG Index and Dashboards Report 2017. New York. Retrieved from https://www.sdgindex.org/reports/sdg-index-and-dashboards-2017/

Jeffrey Sachs, Guido Schmidt-Traub, Christian Kroll, Guillaume Lafortune, & Grayson Fuller. (2018). SDG Index and Dashboards Report 2018. New York. Retrieved from https://www.sdgindex.org/reports/sdg-index-and-dashboards-2018/

Jeffrey Sachs, Guido Schmidt-Traub, Christian Kroll, Guillaume Lafortune, & Grayson Fuller. (2019). Sustainable Development Report 2019. Bertelsmann Stiftung and Sustainable Development Solutions Network (SDSN). New York. Retrieved from https://sdgindex.org/reports/sustainable-development-report-2019/

Jeong, S., Lee, Y., & Kang, S. H. (2020). Government spending and sustainable economic growth: Based on first- and second-level COFOG data. *Public Money & Management, 40*(2), 140. https://doi.org/10.1080/09540962.2019.1651035.

Karlsson, J. (2009). Understandings about the interrelationship of gender inequality, poverty and education, and gender-based strategies to reduce poverty: Some findings from two case studies in the South African education sector. *Agenda: Empowering Women for Gender Equity, 81*, 71.

Kim, J. H. (2019). Multicollinearity and misleading statistical results Retrieved from. *Korean Journal of Anesthesiology, 72*(6), 558–569 http://10.0.16.1/kja.19087.

Le Blanc, D. (2015). Towards integration at last? The sustainable development goals as a network of targets. Department of Economic and Social Affairs, 1(141), 1–17.

Lusseau, D., & Mancini, F. (2019). Income-based variation in Sustainable Development Goal interaction networks. *Nature Sustainability, 2*(March), 242–247. https://doi.org/10.1038/s41893-019-0231-4.

Mainali, B. (2018). Evaluating Synergies and Trade-Offs among Sustainable Development Goals (SDGs): Explorative Analyses of Development Paths in South Asia and Sub-Saharan Africa. Sustainability, 10(3). https://doi.org/10.3390/su10030815

McArthur, J. W., & Rasmussen, K. (2018). Change of pace: Accelerations and advances during the Millennium Development Goal era. *World Development, 105*, 132–143. https://doi.org/10.1016/j.worlddev.2017.12.030.

McArthur, J. W., & Rasmussen, K. (2019). Classifying Sustainable Development Goal trajectories: A country-level methodology for identifying which issues and people are getting left behind. *World Development, 123*, 104608. https://doi.org/10.1016/j.worlddev.2019.06.031.

Nilsson, M., Griggs, D., & Visbeck, M. (2016). Policy: Map the interactions between Sustainable Development Goals. *Nature, 534*(7607), 320–322.

Pincet, A., Okabe, S., & Pawelczyk, M. (2019). Linking Aid to Sustainable Development Goals: A machine learning approach. OECD Development Co-operation Working Papers (OECD Publishing). Paris.

Pradhan, P., Costa, L., Rybski, D., Lucht, W., & Kropp, J. P. (2017). A Systematic Study of Sustainable Development Goal (SDG) Interactions. *Earth's Future, 5*(11), 1169–1179. https://doi.org/10.1002/2017EF000632.

R Core Team (2019). R: The R Project for Statistical Computing. Retrieved December 22, 2019, from https://www.r-project.org/index.html.

Sachs, J., Schmidt-Traub, G., Kroll, C., Lafortune, G., & Fuller, G. (2016). SDG Index and Dashboards - Global Report. New York: Bertelsmann Stiftung and Sustainable Development Solutions Network (SDSN). Retrieved from https://www.sdgindex.org/reports/sdg-index-and-dashboards-2016/

Sachs, J. D., Schmidt-Traub, G., Mazzucato, M., Messner, D., Nakicenovic, N., & Rockström, J. (2019). Six transformations to achieve the sustainable development goals. *Nature Sustainability, 2*(9), 805–814. https://doi.org/10.1038/s41893-019-0352-9.

Shonkoff, J. P., Radner, J. M., & Foote, N. (2017). Expanding the evidence base to drive more productive early childhood investment. *The Lancet, 389*(10064), 14–16.

Singh, G. G., Cisneros-Montemayor, A. M., Swartz, W., Cheung, W., Guy, J. A., Kenny, T.-A., ... Ota, Y. (2018). A rapid assessment of co-benefits and trade-offs among Sustainable Development Goals. *Marine Policy, 93*, 223–231.

Singh, M. (2005). *Meeting basic learning needs in the informal sector Integrating education and training for decent work, empowerment and citizenship*. Springer.

UN (2015). Transforming our world: the 2030 Agenda for Sustainable Development. General Assembley 70 Session, 16301(October), 1–35. https://doi.org/10.1007/s13398-014-0173-7.2

UN (2018). The sustainable development goals report 2018. Retrieved from https://unstats.un.org/sdgs/files/report/ 2018/TheSustainableDevelopmentGoalsReport2018.pdf

UN Statistical Commission (2020). Global Indicator Framework for the Sustainable Development Goals and Targets of the 2030 Agenda for Sustainable Development. In *Work of the Statistical Commission Pertaining to the 2030 Agenda for Sustainable Development* (pp. 1–21).

United Nations (2020). SDG indicators, United Nations Global SDG Database Retrieved from https://unstats.un.org/sdgs/indicators/database/.

Valavi, R., Elith, J., Lahoz-Monfort, J. J., & Guillera-Arroita, G. (2019). blockCV: An R package for generating spatially or environmentally separated folds for k-fold cross-validation of species distribution models. Methods in Ecology and Evolution https://doi.org/10.1111/2041-210X.13107

Vladimirova, K., & Le Blanc, D. (2016). Exploring links between education and sustainable development goals through the lens of UN flagship reports. *Sustainable Development, 24*(4), 254–271. https://doi.org/10.1002/sd.1626.

Weitz, N., Carlsen, H., Nilsson, M., & Skånberg, K. (2017). Towards systemic and contextual priority setting for implementing the 2030 agenda. *Sustainability Science, 13*(2), 531–548. https://doi.org/10.1007/s11625-017-0470-0.

Weitz, N., Nilsson, M., & Davis, M. (2014). A Nexus approach to the post-2015 agenda: Formulating integrated water, energy, and food SDGs. *SAIS Review of International Affairs, 34*(2), 37–50. https://doi.org/10.1353/sais.2014.0022.

Weitz, N., Persson, Å., Nilsson, M., & Tenggren, S. (2015). Sustainable Development Goals for Sweden : Insights on Setting a National Agenda, 1–57. Retrieved from https://www.sei-international.org/mediamanager/documents/Publications/SEI-WP-2015-10-SDG-Sweden.pdf

WHO et al. (2018). Towards a global action plan for healthy lives and well-being for all: uniting to accelerate progress towards the health-related SDGs. Geneva PP - Geneva: World Health Organization. Retrieved from https://apps.who.int/iris/handle/10665/311667

World Bank (2020). SDG Atlas 2020 Retrieved from https://databank.worldbank.org/source/sustainable-development-goals-(sdgs.

Zhao, Y., & Wang, K. (2014). Fast cross validation for regularized extreme learning machine. *Journal of Systems Engineering And Electronics, 25*(5), 895–900. https://doi.org/10.1109/JSEE.2014.000103.