Problem Chosen	2023 MCM/ICM Summary Sheet	Team Control Number
D		2304611

Summary

The United Nations has set 17 sustainable development goals, which are interconnected and interdependent, forming a complex network system. Achieving these goals will benefit the sustainable development of all humanity. However, these goals are often seen as vague and difficult to quantify. This paper proposes a set of network, evaluation, and prediction models to provide theoretical support for the United Nations and governments in achieving sustainable development goals in the present and the next decade and helping other companies or organizations with goal prioritization management.

In Model 1, we created a relationship network that covers all sustainable development goals and visualized the relationships between the 17 goals while considering the development situation in different regions. Firstly, based on the fact that different backgrounds have different goal priorities, we divided the world into three types of economies ——developed countries, upper-middle-income developing countries, and lower-middle-income developing countries. Three representative countries were selected from each economic entity. Then for a single country, we conducted a **canonical correlation analysis** on every two goals among the 17 SDGs. Finally, we averaged the data based on population proportions to obtain the matrix of world typical correlation coefficient. **We used each SDG as a node, calculated the correlations between each SDG as weights, and constructed three network graphs based on the three types of economies and a global average network graph.**

In Model 2, we determine the development priority of each goal from a global perspective. We used three dimensions, namely, **eigenvector centrality**, the rate of completion, and adjustment parameters based on different economies to obtain priorities and predicted the goals that could be reasonably achieved in the next decade using **LSTM**. Firstly, we calculated the eigenvector centrality of the node to measure its long-term influence. Then, based on the UN database, we obtained the global completion rate weighted by the population proportion. In addition, we set adjustment parameters for different economies based on the experience of expert groups. Finally, we obtained the global average goal priority and the goal priority of three economies. **In short, the higher the centrality and the lower the completion rate of a goal, the greater its impact and urgency, and the higher its priority.** Among them, SDG-3, SDG-9, and SDG-0 are the common vision for the long-term development of all humanity. After the priority items are launched, the accuracy of the prediction is continuously improved by iteration. To conclude, the completion rate of each goal shows **a steady upward trend from 2023 to 2032, and the probability of achieving goals that have a higher initial priority is greater after 10 years.**

In Model 3, when SDG-1 is implemented, first, we reconstruct a network with SDG-1 removed according to the principle of Model 2. On the basis of this network, the eigenvector centrality of all nodes and the completion degree of nodes strongly correlated with SDG-1 are recalculated. **In conclusion, when SDG-1 is implemented, the priority of SDG-9 rises significantly, while the priority of SDG-17 falls obviously.**

As for **Model 4** and **Extension of Model 2** ,the principles of Model 1, 2 are applied according to the principles of Model 1, 2. And, we test the stability of Model 2 and the test passes. It proves that our model is stable.

Keywords: SDGs ; Canonical correlation analysis ; Centrality of eigenvectors ; LSTM ; Newey West

Contents

1 Introduction

"Fifty years after the first UN Conference on the human environment in Stockholm in 1972, the bedrock SDG principles of social inclusion, clean energy, responsible consumption, and universal access to public services are needed more than ever to respond to the major challenges of our time."

——Prof. Jeffrey D. Sachs, President of the SDSN and first author of the report

Fig. 1 17 Sustainable Development Goals

1.1 Problem background

The adoption of the 2030 Agenda for Sustainable Development at the United Nations Development Summit in 2015 set out a new and ambitious blueprint for international development cooperation by establishing the Sustainable Development Goals (SDGs). However, **taken as a whole**, the SDGs are a comprehensive and closely linked whole, universal goals that are interconnected in a complex web of interactions. Their universality means that no single SDG is prioritized, and the multidimensional nature of their integration leads to complex feedback among the SDGs. For example, progress on one goal may drive or hinder the achievement of other goals. **In terms of individual goals**, the Sustainable Development Goals (SDGs) are difficult to measure and ambiguous. They are multidimensional structures of people planet prosperity peace partnership, and most studies discuss concepts and frameworks for the interconnectedness of the goals, but do not present potential theoretical models and analytical evidence.

One of the objectives of this paper is the appropriate method to construct a suitable network to study the interlinkages among goals to detect trade-offs (negative feedback) and synergistic effects (positive feedback) among SDGs with the aim of providing theoretical support for policy formulation for the implementation of SDGs in each country. In the temporal dimension, the dynamic network shows temporal variability to cope with the current multiple chain reactions and cross-crisis dominated by COVID-19 , climate change and regional conflicts. In the spatial dimension, the diagram conveys a sense of unequally woven networks, with some goals linked to many others and others less connected to the rest of the network, based on consideration of different national and regional development contexts. We believe that further work on SDG interpretation will lead to greater scientific resonance in the future regarding performance measurement, operationalization, and interconnectedness of the SDGs.

1.2 Restatement of the Problem

- **Problem 1**: Create a network covering 17 SDGs.
- **Problem 2**: Assess the priority of each goal from multiple perspectives and test the validity of the priorities by selecting data from a typical country or region.

- **Problem 3**: When one of the goals is achieved, how should the network structure and goal priorities change? Should we also include other objectives?
- **Problem 4**: Consider the impact of external international crises on the networks and goal priorities we have created, and how do external crises affect progress on the UN Sustainable Development Goals?
- **Problem 5**: Extend and transfer our network structure and prioritization model to other companies or organizations.

1.3 Our Work

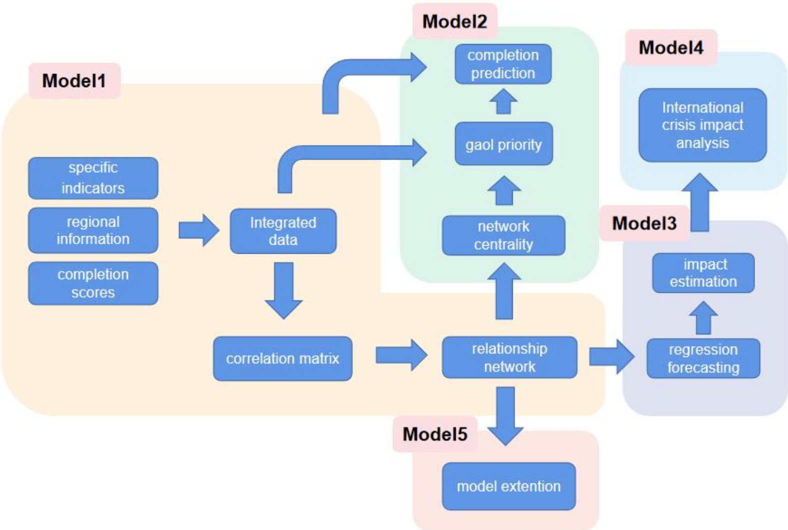


Fig. 2 Frame of our work

2 Assumptions

- The 17 SDGs and their sub-goals remain virtually unchanged over the decade.
- The literature reviewed is more comprehensive and scientific, and the analysis is consistent with facts.
- In problem 2, after the priorities have been set in motion and considering what is reasonably achievable in the next decade, we assume no shocks from the international crisis in the next decade.
- The serial numbers targets in this paper are consistent with those in the UN report.

3 Notations

Table 1: Notations used in this paper

Symbol	Description
V	Node
$D_{17\times 17}(d_{ij})$	Matrix of typical correlation coefficient
$G(V_i\ ,\ E\ )$	Diagram of network
x'	Eigenvector centrality
y'	Degree of target completion after normalization
c	Adjustment parameters
w	Scores of priority

4 Model 1 : Network of 17 SDGs

4.1 Analysis of problem

Problem 1 requires the construction of a network that illustrates the interrelationships between the 17 SDGs. In their paper, [1]Bail Swain and Ranjula (2021) argue that benchmarking the SDGs is only meaningful within a regional context. This is because different regions, such as OECD countries, South Asia, and Sub-Saharan Africa, have unique characteristics that make universal global goals or benchmarks unsuitable. As a result, this paper constructs separate networks for developed countries, upper middle-income developing countries, low-income developing countries, and the world as a whole.

Each of the constructed networks requires an analysis of the association between each of the 17 objectives, including the strength and direction of the relationship. Since each objective contains multiple sub-indicators, a simple correlation analysis between a single X and Y variable is insufficient as it fails to account for the correlation between variables within each group. Therefore, a multivariate statistical method, such as typical correlation analysis, is used to identify the relationship between two groups of variables for each pair of objectives. This method considers the linear combination of the two groups of variables and studies their typical correlation coefficients ***p***.

Each country-specific network diagram requires the calculation of typical correlation coefficients for each pair of objectives, resulting in a total of C_{17}^{2} =136 typical correlation analyses. As an example, we will detail the correlation between SDG-4 and SDG-11 for Australia.

4.2 Processing of data

Upon searching for data, we find that some years have missing data, which is insufficient for analysis. To address this issue, we process this missing data by PCHIP (piecewise cubic Hermite interpolating). The simulation produces some new but more reliable values to meet the needs, and this is where interpolation comes into play. However, direct use of Hermite interpolation yields a high number of polynomials and also suffers from the Runge phenomenon. Therefore, in practice, PCHIP is often used.

By using Matlab, we get the result as shown in Fig. 2. Value1 refers to sdg4_tertiary of Germany; Value2 represents sdg13_co2gcp of China; Value3 is sdg7_cleanfuel of Papua New Guinea.

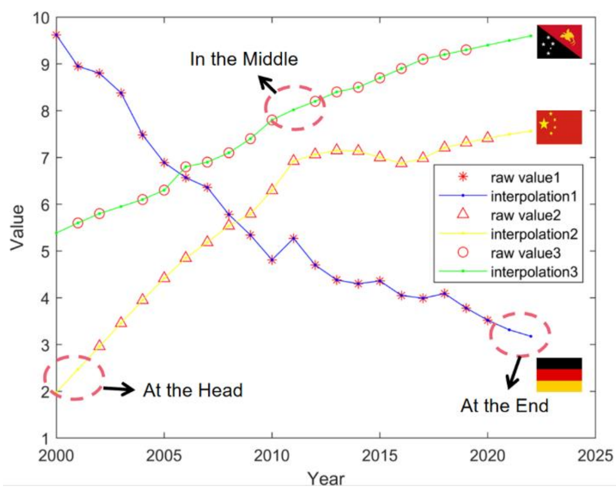


Fig. 2 Results of interpolation fitting

4.3 Establishment of Model 1

- **Step 1：build the original matrix；**

  Based on the data presented in section 4.2, we select three sub-goals in SDG-4 as vector X=(X1,X2,X3)', and two sub-goals in SDG-9 as vector Y=(Y1,Y2,Y3)'. Z is a 3-by-3 matrix that encompasses the overall 23 degrees of centralized observation data.

$$Z=\begin{bmatrix} X_{1,1} & X_{1,2} & X_{1,3} & Y_{1,1} & Y_{1,2} & Y_{1,3} \\ \vdots & \vdots & \vdots & \vdots & \vdots & \vdots \\ X_{23,1} & X_{23,2} & X_{23,3} & Y_{23,1} & Y_{23,2} & Y_{23,3} \end{bmatrix}=\left(\underset{23\times 3}{X},\underset{23\times 3}{Y}\right)' \tag{1}$$

- **Step 2: normalization transformation of the original data and correlation coefficient matrix.**

  We first normalize the data for the sub-goals of the two objectives, then calculate the correlation coefficient matrix R between the two samples, and divide R as follows:

$$R=\begin{bmatrix} R_{11} & R_{12} \\ R_{21} & R_{22} \end{bmatrix} \tag{2}$$

- Where R11 and R22 are the correlation coefficient matrices within SDG-4 and SDG-11, respectively, and R12 and R21 are the correlation coefficient matrices between SDG-4 and SDG-11.
- **Step 3: finding typical correlation coefficients and typical variables.**

  First, find the eigenroot λ_i^2 and the eigenvector $S_1\alpha_i$ of $A=R_{11}^{-1}R_{12}R_{22}^{-1}R_{21}$; the eigenroot λ_i^2 and the eigenvector $S_2\beta_i$ of $B=R_{22}^{-1}R_{21}R_{11}^{-1}R_{12}$. Then, we have:

$$\alpha_i=S_1^{-1}(S_1\alpha_i),\beta_i=S_2^{-1}(S_2\beta_i) \tag{3}$$

  Then the typical correlation coefficient between the random variable X (SDG-4) and the random variable Y (SDG-11) is λ , and the typical variables are:

$$\begin{bmatrix} V_1=\alpha_1'X \\ W_1=\beta_1'Y \end{bmatrix};\begin{bmatrix} V_2=\alpha_2'X \\ W_2=\beta_2'Y \end{bmatrix};\dots;\begin{bmatrix} V_t=\alpha_t'X \\ W_t=\beta_t'Y \end{bmatrix} \quad (t\le 23) \tag{4}$$

- **Step 4: test the significance of each typical correlation coefficient.**

  The typical correlation coefficients are tested for significance. Before conducting a typical correlation analysis of two sets of variables X and Y, it should first be tested whether the two sets of variables are correlated. If there is no correlation, then the typical correlation between the two sets of variables is meaningless.

4.4 Solutions of Model 1 and visualization

4.4.1 Take Australia as an example

A total of 6 sub-targets for 23 years were selected for 3 Goals 4 and 3 sub-goals for Goal 11 for Australia from 2000 to 2022.

Table. 2 : Selected 6 sub-goals

X_1	X_2	X_3	Y_1	Y_2	Y_3
earlyedu	primary	tertiary	pm25	pipedwat	transport

The typical correlation coefficients of the first and second pairs of variables are 0.787644, 0.398143, and 0.134586 respectively. The typical correlation coefficient of the first pair is greater than 0.5, indicating that this typical pair of variables has relatively strong explanatory power. However, to determine the significance of the correlation of typical variables, a significance test of the typical correlation coefficient is needed. Results show that the p-values are 0.0005, 0.2718, and 0.5646, respectively. The correlation between the first pair of typical variables is significant at the 0.05 level of significance. After normalizing the coefficients of the variables, we can develop a typical correlation model as follows.

$$\begin{aligned} V_1&=0.2128X_1-0.0595X_2+0.8613X_3 \\ W_1&=-0.2327Y_1+0.8537Y_2-0.1020Y_3 \end{aligned} \tag{5}$$

The coefficients in front of each component of the typical variable represent their degree of importance. As per formula (5), it is evident that for SDG-4, the most crucial influencing factor is sub-goal 3. The correlation coefficient between X3 and V1 is significantly higher than the correlation coefficients between X1, X2, and V1. This implies that V1 primarily represents the degree of access to higher education. On the other hand, for SDG-11, sub-goal 2 is the most important influencing factor. The correlation coefficient between Y2 and W1 is significantly greater than the correlation coefficients between Y1, Y3, and V1. The absolute value of the correlation coefficients indicates that W1 mainly represents access to an improved water source with a pipe.

The pair of typical variables explains 70.25% and 87.31% of the cumulative proportion of variance in the set of paired variables, respectively. We followed the same approach to compute the typical correlation coefficients for the nine representative countries' two pairs of 17 indicators. If more than one pair of typical correlation coefficients passed the significance test, we will calculate a weighted average based on their contribution margin.

4.4.2 Calculate the matrixes of canonical correlation coefficient

We follow the same procedure to select 3 developed countries (Group A: USA, Germany, Australia), 3 middle-income and high-income developing countries (Group B: China, Malaysia, Brazil), and three low-income developing countries (Group C: Mozambique, Papua New Guinea, Sierra Leone), based on the different national or regional development of the data.
First, we solve the typical correlation coefficients between each two of the 17 indicators for each country separately, and obtain a total of nine 17×17 adjacency matrices. Then, we arithmetically average each set of data separately to obtain 3 adjacency matrices. Finally, we obtain a global matrix based on these 3 matrices, which is weighted by the population share. The global matrix is shown below.

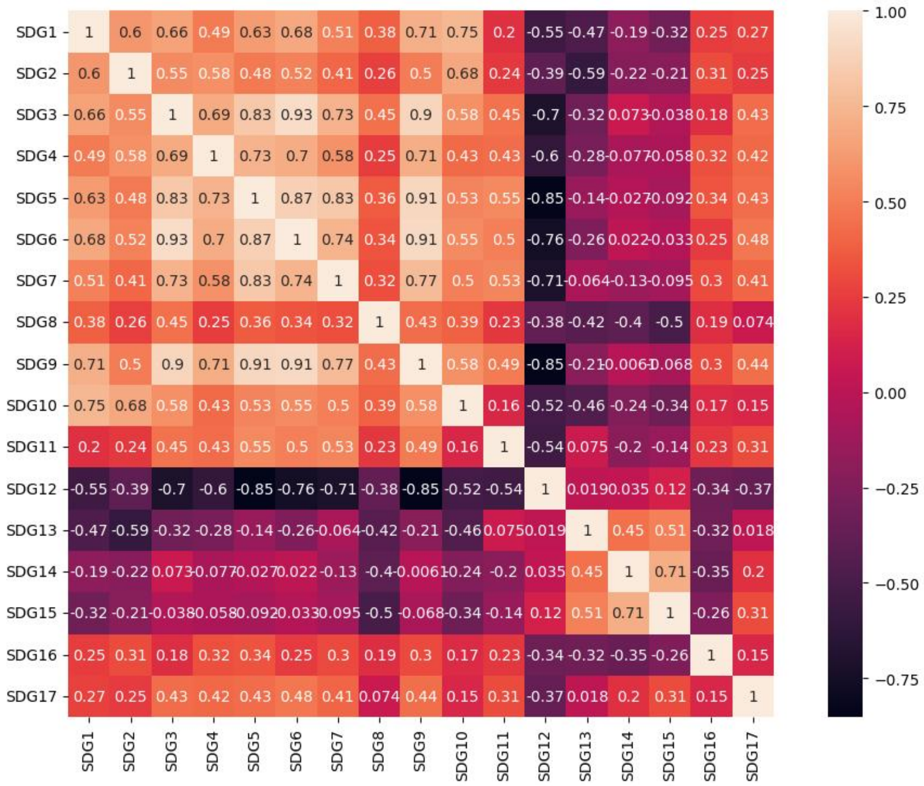


Fig. 3 Matrix of typical correlation coefficients for the global average

The figures below show the network of interlinkages among developed countries, upper-middle-income developing countries, low-income developing countries, and all the world.

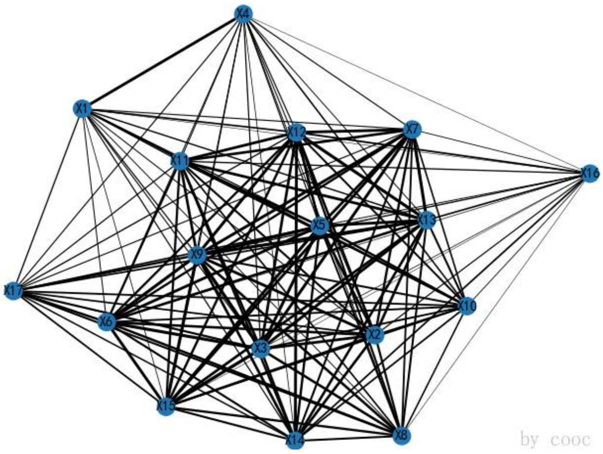


Fig. 4 Network G1 of developed countries

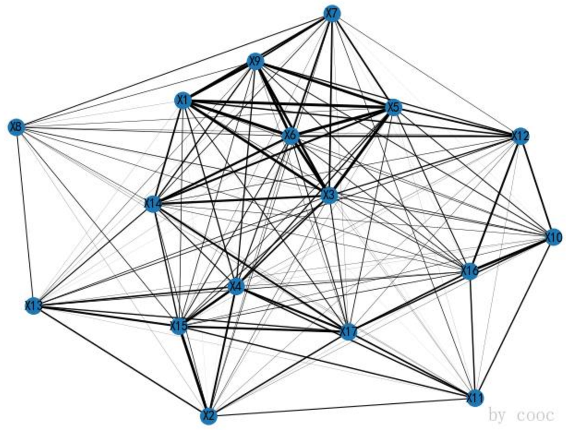


Fig. 5 Network G2 of upper middle-income developing countries

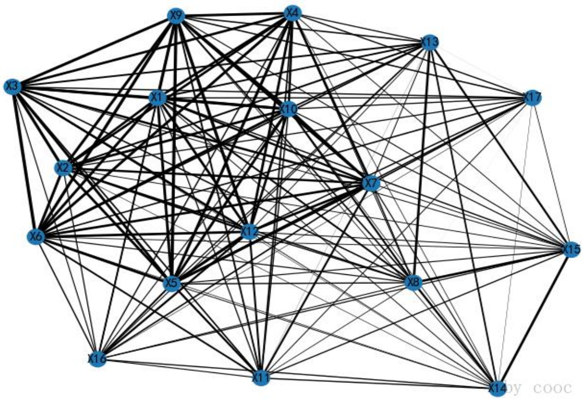


Fig. 6 Network G3 of upper middle-income developing countries

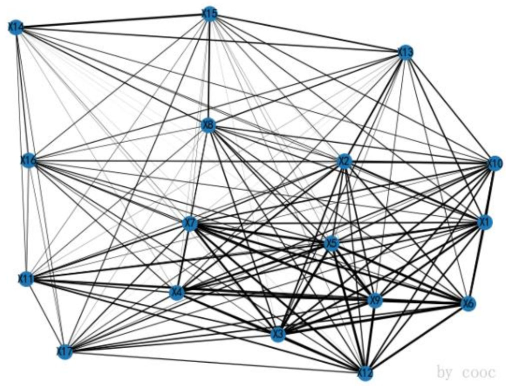


Fig. 7 Network G4 of all the world

The diagram depicts 17 nodes, each corresponding to one of the 17 SDGs, with lines connecting nodes that are related to each other. To ensure clarity, the goals and their corresponding numbers are consistent with those used in the UN report. The thickness of the lines indicates the strength of the link between nodes, which varies depending on the development status and implementation of the goals in different countries and regions. Moreover, each node in the three groups has distinct priorities, which we will explain in further detail below.

5 Model 2 : Evaluating priorities of goals and forecasting

5.1 Analysis of problem

The concept of priority is qualitative and needs to be quantified using indicators for each of the 17 goals. We calculate the eigenvector centrality of each node in the network and assesses the degree of completion of each goal using data from the global SDG database of the UN and World Bank building blocks. Next, adjustment parameters are set based on experience and common sense, considering the different development situations of countries and regions. The aim is to highlight the most urgent constraints for each of the three categories of countries mentioned above and reflect the urgency of completing different goals. In other words, countries in different development situations prioritize goals that offer greater scope for success by redeploying existing capacity (high centrality) and those that are highly urgent but poorly accomplished. These 3 indicators are given equal weight in the priority score, and the final 17 goals are ranked based on their score.

This paper first prioritizes the goals for each of the three categories of countries and regions. Then, [2] the base priorities of the 17 goals are assigned based on the global population weight of the three categories of countries and regions to obtain the average global centrality and rank them. This approach allows us to identify the core SDGs that are better connected to the network and are fundamental for redeploying and expanding the SDG delivery mechanism. This, in turn, can contribute to the successful realization of the entire SDG agenda.

5.2 Evaluate priorities

5.2.1 Feature vector centrality: measuring the impact of a single goal in a network

The evaluation of important nodes is applied to discover high-priority goals based on the constructed influence network structure. The centrality of an SDG, which is the sum of the pairwise proximity of all its SDGs, is used as a measure of relatedness. SDGs with high centrality have many SDGs in their vicinity, indicating that a country successful with such SDGs can use its current capacity to diversify the focus of its SDG goal implementation. On the other hand, if an SDG is located in the sparse part of the network, opportunities for success may be limited. [2] The aim is to prioritize goals that should be addressed by different types of countries, thus increasing the efficiency of SDG implementation. (Reference 2)

While traditional centrality methods only consider the density of a node's position in the network, we believe that assessing the impact of an individual goal should also take into account the node itself. Eigenvector centrality takes into consideration not only the number of neighbors of a node (i.e., its degree), but also the quality of its connections, which affects the importance of the node. From a propagation perspective, eigenvector centrality is suitable for describing the long-term influence of a node.

Taking G4 (Vi, E) as an example, we will use the canonical correlation coefficient of the adjacency matrix D as the edge weight for each node to provide a detailed description of the model.

• Let x_i be the importance measure of node Vi and $\mathbf{a_{ij}}$ be the edge weight of the edge of the node, then:

$$EC(i)=x_i=c\sum_{j=1}^{n}a_{ij}x_{j\,\mathrm{p}} \tag{6}$$

Where c is a proportionality constant.

•Let x= $[x_1\ ,\ x_2\ ,\ \cdots\ ,\ x_n]^T$ and after multiple iterations, when reaching a steady state, it can be written in matrix form as follows:

$$\mathbf{x}=c\,\mathbf{Dx} \tag{7}$$

The expression indicates that x is the eigenvector corresponding to the eigenvalue c^{-1} of the matrix $\boldsymbol{D}$.

• The basic method for computing the vector x is to give an initial value x(0) and use an iterative algorithm until the normalized $x'(t)=x'(t-1)$. Then we have $\mathbf{x}=\lambda^{-1}\mathbf{Dx}$ .So, $c=\lambda^{-1}$.

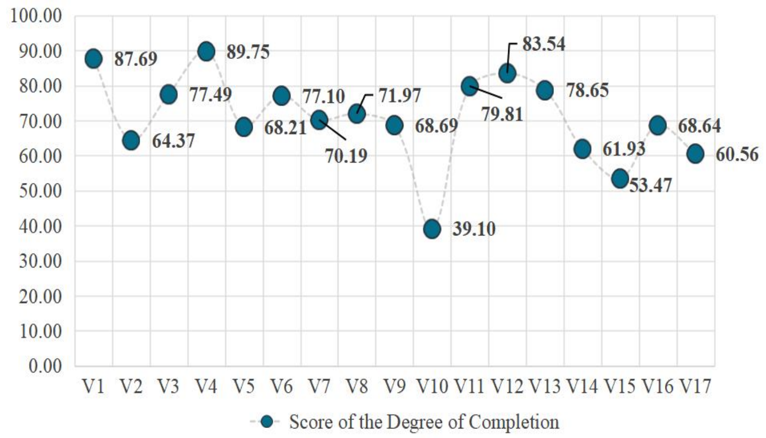
•The eigenvectors are normalized using the following formula, where mx , mn respectively represent the maximum and minimum values of the solved eigenvector centrality.

$$\mathbf{x}' = \frac{\mathbf{x} - mn}{(mx - mn) + 1} \tag{8}$$

• Finally, the normalized $\mathbf{x'}$ is mapped to intervals [1 , 2]. The purpose is to prevent the centrality from having too much influence in the final calculation of the priority. The results are shown in Appendix 1.

5.2.2 Collect the degree of SDGs completion y'

Based on official UN statistics on the implementation of the 17 goals in each country, data was collected for nine countries with target completion data. These countries were selected as representatives from developed, middle- and high-income developing, and low-income developing categories, as described in Section 4.4.2. The global average target achievement was obtained by weighting the population share of the three categories of countries. The figure below illustrates this weighting and shows the global average target achievement.



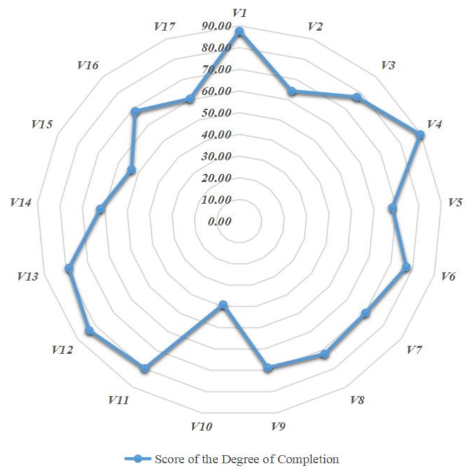


Fig. 8 Global SDG index scores in 2022 **Fig. 9 Global SDGs completion in 2022**

Notice: Author's calculations

The completion data were normalized using the following formula.

$$y' = \frac{2}{y/100 + 1} \tag{9}$$

Denote the normalized degree of completion as $\boldsymbol{y'}$, $\boldsymbol{y}$ represents the original degree of completion. The data is normalized to the interval [1, 2] to prevent it from having an excessive impact on the final priority results. The normalization formula ensures that lower completion degrees correspond to higher priority levels, reflecting the greater urgency for completing those goals.

5.2.3 Regional differences: setting adjustment parameters c

[i]Spaiser et al. highlighted the fact that different economies have unique national circumstances and development trajectories, which means that a one-size-fits-all approach to sustainable development may not be effective. Similarly, [2]Nicolai (2017) argued that in the context of the MDGs, universal goals that ignore imbalances and inequalities within countries can be counterproductive to overall human development in the long run. This suggests that unless regional contexts and development are taken into account, setting universally applicable targets or benchmarks may be counterproductive to the achievement of multidimensional SDGs.Therefore, we set adjustment parameters c for each of the 17 SDGs for the three types of economies through a combination of questionnaire surveys and expert group experience.

$$\begin{cases} c_1 = 1.3 & , \textit{more urgent} \\ c_2 = 1.0 & , \textit{less urgent} \end{cases} \tag{10}$$

The adjustment parameters are obtained through questionnaire surveys and expert group experience, and are specific to each of the 17 SDGs for the three types of economies. To obtain global SDG adjustment parameters that reflects the diversity of national circumstances, these parameters were weighted by[3] weight of population and a weighted average was calculated. The resulting global SDG adjustment parameter is shown below. The purpose of this adjustment parameter is to take into account the unique circumstances and development trajectories of different economies, and to ensure that the SDGs are adapted to each country's context and needs.

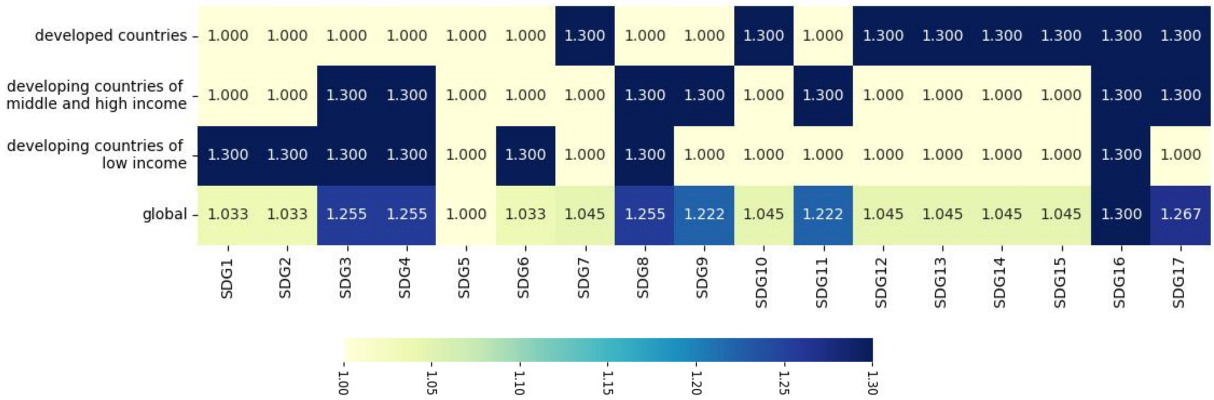


Fig. 10 Adjustment parameters for global SDGs

5.2.4 Calculate the final target priority value w

The final priority value ***w*** of each target is calculated by the following formula. And normalize them.

$$w = x' \times y' \times c \tag{11}$$

Table 3 Global SDGs prioritization

SDGs	*V9*	*V10*	*V3*	*V17*	*V16*	*V8*	*V4*	*V11*
Global Scores of Priority	0.0742	0.0721	0.0717	0.0699	0.0671	0.0657	0.0647	0.0620
V7	*V5*	*V2*	*V6*	*V1*	*V15*	*V14*	*V13*	*V12*
0.0607	0.0605	0.0597	0.0593	0.0539	0.0466	0.0451	0.0375	0.0292

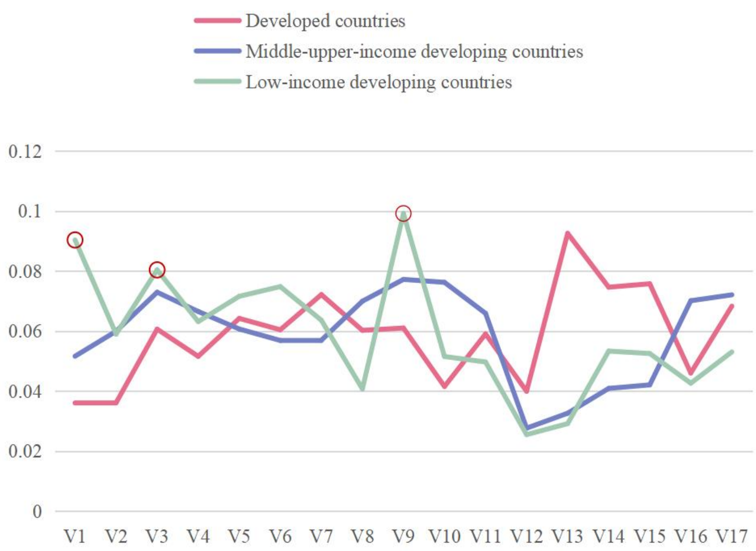


Table 4 Global SDGs prioritization

According to the table, SDG-9, SDG-10, and SDG-3 are identified as key priorities for long-term global development. Governments should provide strong support in terms of policies and funding for these goals. On the other hand, SDG-14, SDG-13, and SDG-12 are of lower priority and can be considered as secondary objectives when countries plan for sustainable development. However, specific implementation programs should be analyzed and designed according to the unique national conditions of each economy.

5.3 Projection of what can reasonably be achieved in the next decade

5.3.1 Analysis of problem

In this paper, we make the assumption that global sustainable development will remain on a desirable trajectory in the next decade, free from any major international crises. To predict the reasonable goals that can be achieved in the next decade, we adopt the LSTM model, a deep learning neural network that is well-suited for handling time series tasks. Compared to other deep learning models like CNN and RNN, LSTM can solve the long-term dependency and "gradient disappearance" problem more effectively. Additionally, the small data set used in this paper avoids the problems of time-consuming training and inapplicability of LSTM to longer series data. By predicting the target scores for each year in the next decade and incorporating them into the training set, the model dynamically adjusts to the expected target completion. As a result, we can predict reasonable goals for the next ten years that take into account the expected completion of current goals.

Establishment of model

•Step 0 : There are 23 sets of data from 2000 to 2022（noted as $\Omega{:}DA_{1\sim23}$）; Training set:$Tr{:}DA_{1\sim18}$; Testing set:$Te{:}DA_{19\sim23}$; Two hyper-parameters: α : number of layers of the neural network and β : time span of prediction

• Step 1 : Use the test set ***Te*** to verify the accuracy of ***LSTMa*** (the result of 2021 and 2022 projections are seen in Fig. 12) and adjust the 2 hyperparameters to improve the accuracy to obtain $LSTM_b$.

• Step 2 :Then, the $LSTM_b$ is used to combine 23 data sets to predict the achievement of the SDGs in 2023.

• Step 3 : After obtaining the data of 2023 , a comparison with the data of 2022 reveals the amount of change $\Delta W=\sum_{i=1}^{17}\Delta\omega_i$, $i=1\sim17$, for each goal. Take $\frac{\Delta W}{2}$ and allocate it to each goal in 2023 in proportion to priority.

• Step 4 : A total of 24 sets of data from 2000 to 2023 are used to project the situation in 2024.

• Step 5 : Repeat the above steps for 10 times to get the reasonable progress of completion of 17 SDGs in the next decade. The flow diagram is as follows.

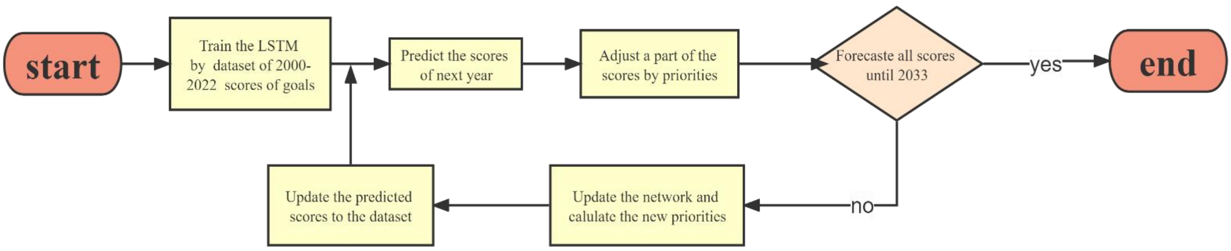


Fig. 11 Flowchart of predicating reasonably achievable objectives

5.3.2 Solutions and visualization

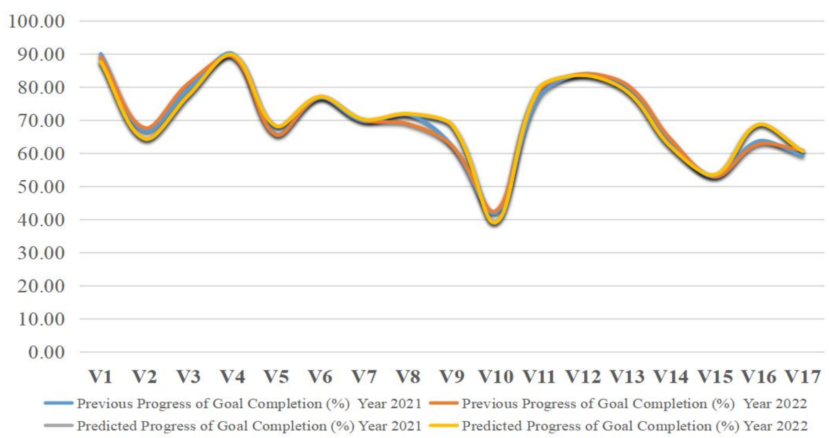


Fig. 12 Performance of LSTM in sets of testing in 2021 and 2022

As shown in Figure 12, the LSTM model constructed in this study performs well on the testing set. The predicted data align more closely with the actual data compared to the original data.

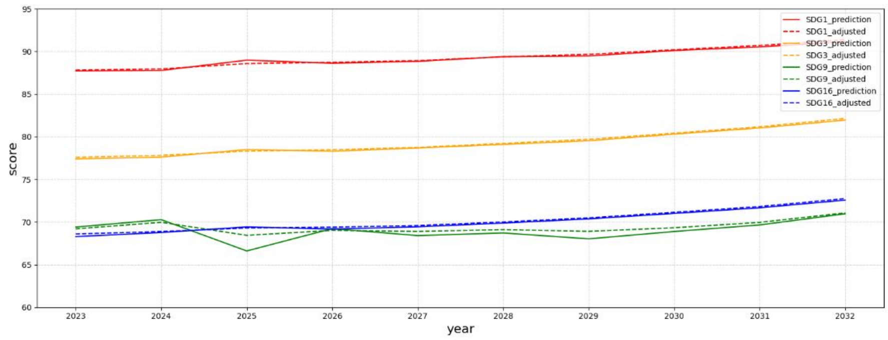


Fig. 13 Forecasted and adjusted values of the completion degree for some goals between 2023 and 2032

To conclude, within the period of 2023-2032, the completion degree of each goal shows a steady upward trend. Furthermore, the higher the initial priority level of a goal, the greater the probability of achieving it in ten years.

6 Model 4 : changes of network and priority when one goal is achieved

6.1 Analysis of problem

We randomly chose SDG-1 (no poverty) as a goal that has been achieved and removed it from the network to observe how it would affect the network structure. We repeated the steps in Model 1 to reconstruct the SDG interdependence network without SDG-1 and observed the resulted changes in the network structure, which went from G_4（V_i , E） to G_5（V_i , E）.

We then quantified the impact of removing SDG-1 on the network structure by calculating the change in eigenvector centrality of all nodes and the change in completion of strongly correlated nodes. Additionally, we calculated the change in priority of the remaining 16 SDGs using equation 11.

6.2 Establishment of model

6.2.1 Comparison of networks before and after removing node V_1

Based on figures 13 and 14, the left panel highlights the eight nodes (V2, V3, V5, V6, V7, V9, V10, V12) that are strongly correlated with V1 (as per the matrix, typical correlation coefficients are qualitatively strongly correlated) and their synergistic utility (positive correlation) and trade-offs (negative correlation) with V1 based on the network. Out of these nodes, only V1 and V7 show a negative correlation (represented by a blue edge), while the remaining seven nodes are positively correlated with V1 (represented by orange edges). The right panel shows the structure of the network after removing V1, and the changes are quantified in sections 6.2.2 and 6.2.3.

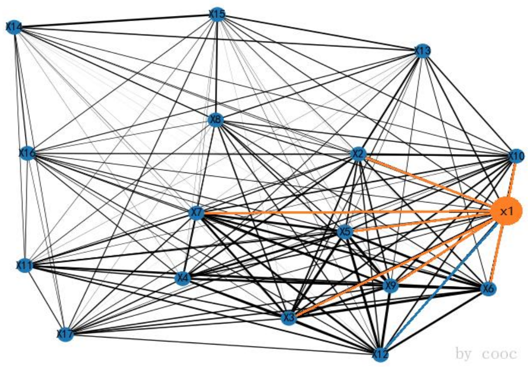


Fig. 14 Network before removing V_1

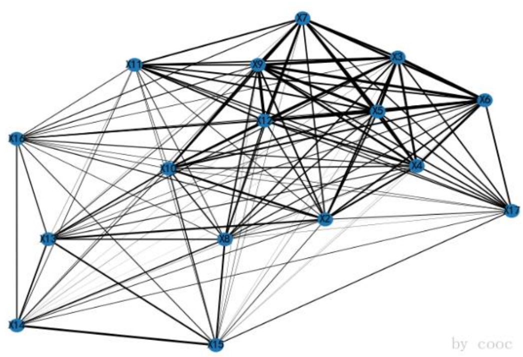


Fig. 15 Network after removing V_1

6.2.2 Variation in the centrality of eigenvectors 16 nodes

Repeating the steps in 5.2.1, we obtain the centrality of the eigenvector for the new 16 nodes x'' （seen in Appendix 1）.

6.2.3 Change in completion of strongly correlated nodes based on Newey West regression model

· OLS

To analyze relationships between the completion of different SDGs, a multiple linear regression model is used in this study. Specifically, the completion data from 2000 to 2022 for V_1 are used as explanatory variables (X_i, i=2,3,5,6,7,9,10,12), while the completion data from the other 8 strongly correlated target nodes are used as response variables (Y_j,j=2,3,5,6,7,9,10,12). Ordinary least squares linear regression is then conducted to obtain eight regression equations, as presented in Table 4.

· Hypothesis testing

As time series data are prone to autocorrelation issues, and cross-sectional data are prone to heteroskedasticity issues, we conduct hypothesis testing on panel data. To conduct the test, we construct the Durbin-Watson (DW) statistic, and the results of the test are presented below.

Table 5 The results of testing and regression

Hypothesis Testing		NO.	Regression Equations	
DW	F		Intercept Term	Coefficient of SDG1
0.7111	0.0265	1	44.6548	0.2250
0.2359	0.0089	2	22.1029	0.6296
0.3501	0.2594	3	12.0416	0.6282
0.2187	0.0420	4	45.2000	0.3591
0.4272	0.4881	5	40.4261	0.3326
0.4295	0.9587	6	-166.8114	2.6494
0.4842	0.1159	7	8.9270	0.3493
0.2109	0.0006	8	91.9962	-0.0937

Upon conducting a DW test on the panel data, it is observed that with n=23 and k=2, the DW values for the eight regression equations are lower than the value of dL=1.168. This indicates that the original hypothesis of no autocorrelation is not valid, and the equations are found to have autocorrelation issues. In addition, the P-values obtained from the White's test for the third, fifth, and sixth equations are above 0.05, suggesting that the original hypothesis of no heteroskedasticity in these equations is acceptable. However, the P-values for the other equations are less than 0.05, indicating that the original hypothesis is rejected, and these equations are found to have heteroskedasticity problems.

• Newey West

Due to the presence of autocorrelation and heteroskedasticity problems in the aforementioned eight regression equations, the OLS estimates of the parameters no longer exhibit the minimum variance property. Additionally, the standard errors of the OLS estimates are no longer unbiased estimates of the true standard errors. Consequently, the results of confidence intervals and hypothesis tests of the regression parameters and predicted values cannot be relied upon. To overcome this issue, we have decided to use the [4] Newey-West method to address the problems of heteroskedasticity and autocorrelation.

Compared to other methods, the Newey West method does not require specifying sequence-dependent function forms. Instead, it replaces the variance with the sum of squared OLS residuals and adds the product of the OLS residuals, where $\boldsymbol{p}$ represents the maximum order of serial correlation we wish to assume. The consistency of this method depends on $\boldsymbol{p}$, which is relatively small compared to the number of observations n since the effect of the perturbation terms in previous periods on the perturbation terms in the current period decays rapidly.

We conduct the eight regression equations using Newey West in Eviews. After completing goal 1, we provide the predicted values of several other strongly correlated goals and their prediction intervals at the 95% confidence level. The point estimates and interval estimates are shown in Appendix 2.

6.3 Solutions and visualization

After substituting the point estimation of completion degree, we used equation 11 again to calculate the new priority degree. The results shown in the following two figures indicate that the left graph displays the priority of the 17 SDGs before achieving goal 1, and the right graph displays the priority of the remaining 16 SDGs after achieving goal 1. The change in priority before and after is shown in Appendix 2.

As seen from figures 15 and 16, when SDG-1 (no poverty) is achieved, the priority of SDG-9 (industry, innovation and infrastructure) rises significantly, while the priority of SDG-17 (partnerships to achieve the goals) falls sharply. The priority of other goals does not change much, or their priority ranking remains the same.

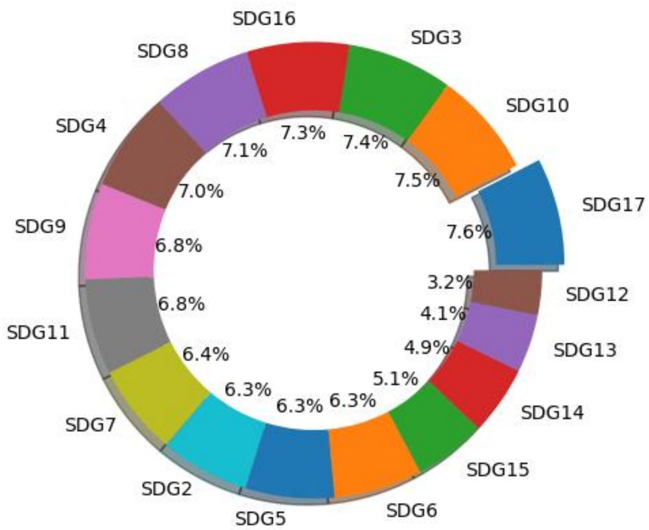


Fig. 15 Priority of not deleting V_1

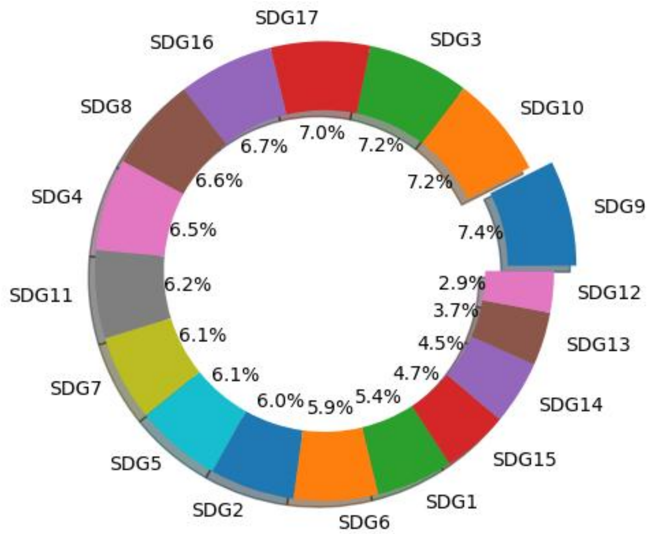


Fig. 16 Priority of deleting V_1

When SDG-1 is completed, reducing the gap between rich and poor can be considered for inclusion in the Sustainable Development Goals. This is because eliminating poverty does not mean that there is no gap in economic income. If the gap is too large, it can cause social harms such as higher crime rates and social stratification.

According to official data, it is known that in 2020, adults with \$1 million account for 1.1% of the world, and they hold a cumulative \$191.6 trillion in assets, accounting for 45.8% of global wealth, while the 55% of the population with less than \$10,000 owns only 1.3% of global wealth, or about \$5.5 trillion cumulatively. Such a large gap between the rich and the poor clearly runs counter to the concept of the Sustainable Development Goals (SDGs), and it is therefore recommended that the reduction of the gap between the rich and the poor be included in the SDGs.

7 Model extension

[ii]Bill Hewlett, one of the founders of Hewlett-Packard, said that Hewlett-Packard has never made profit maximization our business goal, but it has also never put profit outside of all considerations. **We have seven goals: cultivating and developing loyal customers; reasonable profits (exceeding the industry average); industry leadership (staying in first and second place and building scale), sustained growth (having momentum and potential), employee development, team leadership improvement, and social responsibility.** These goals bite together like cogs and move the whole picture with one hair. Our highest priority is to nurture and develop loyal customers. A successful network is a structured strategic plan. We have selected these 7 goals as network nodes to promote the application of our network and assessment model to different for-profit organizations.

Applying the principles of model 1 and model 2, the steps are as follows.

Step 1 :7 business objectives are defined as 7 multidimensional vectors $V_{1\sim7}$ in turn. And the respective sub-objectives are defined as v_i^j ; $i=1\sim7$, $j=1\sim n$. Matlab is used to analyze the typical correlation between the 7 business objectives to obtain the matrix of typical correlation coefficient adjacency $Q_{7\times7}(q_{ij})$; $i=1\sim7$, $j=1\sim7$. Use this matrix to construct the network of targets $G_{7\times7}^{6}$.

Step 2 : Here we adapt equation 11 to the following formula 12 in order to better fit the needs of a for-profit organizations in business scenario.

$$NW = NY \times \frac{PI}{CO} \times C_t \tag{12}$$

NW is the priority score; ***NY*** is the normalized eigenvector centrality; ***PI/CO*** represents the expected rate of return over a certain period of time, where ***PI*** is the expected return over a certain period of time and ***CO*** is the amortized cost over a certain period of time; ***Ct*** is the adjustment parameter, where Ct=1.2 when the target is considered to be more urgent. Otherwise, Ct=1.0.

Clearly, equation 12 indicates that the higher the priority score of this goal node when the higher the contribution of this goal node, the higher the expected profitability and the urgent need for completion. This is also in line with the reality. Profitable organizations need to consider both short-term profitability needs and the construction of long-term strategies when managing goals.

8 How to respond to changes internationally

In this section, we will discuss the impact of other international crises, such as technological advances, global pandemics, and climate change on the network we have built. For a specific type of international crisis, we determine several sustainable development goals that are most relevant to it by reviewing relevant information. We refer to these highly correlated goals as the direct impact goals of the crisis. For these direct impact goals, we roughly determine how the crisis will affect them by consulting experts and reviewing literature. Through the relationship network we established earlier, these direct impact goals may have a certain indirect impact on other goals. We will analyze the impact on other points in the network in a way similar to model three and provide rough estimates of the degree of impact.

As an example, we will discuss the impact of global pandemics on the SDGs. Through research, we have found that the most direct impacts of this crisis on the SDGs are as follows: For decades, the number of people in global poverty has increased for the first time, and a large number of people have fallen into extreme poverty due to the pandemic. The global food system is under threat, and small-scale food producers are severely affected. The improvements in global health over the past decade have been reversed, and the number of child deaths has increased significantly. Global economic growth has stagnated, per capita GDP has declined, and many people face the risk of unemployment. Global peace and security are threatened, with the number of people fleeing war, persecution, and conflict exceeding the highest recorded value. These situations correspond to SDG1, SDG2, SDG3, SDG8, and SDG16. We believe that these goals are the direct impact goals of global pandemics.

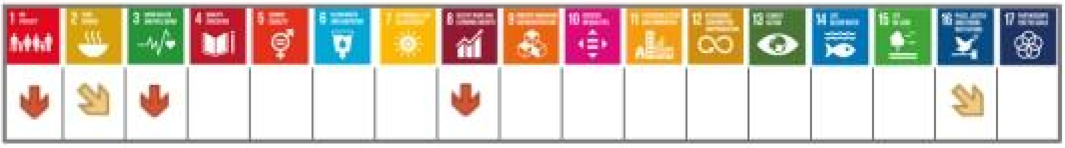


Fig. 17 directly impact of global pandemic

Through the relationship network, we observe the impact of these direct impact goals on other targets. If a goal has strong connections to many direct impact goals, then it will also be affected to a certain extent. We use the network and actual situations to determine the degree of impact on these indirectly affected goals.

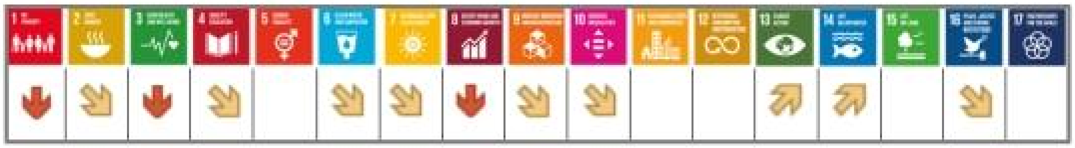


Fig. 18 impact of global pandemic

We will evaluate the direct impact targets and the corresponding approximate impact levels of technological progress, climate change, regional conflicts, and refugee activities using a similar method.

Fig. 19 directly impact of international crises

Estimated table of overall impacts derived from relationship network and actual situations：

Fig. 20 impact of international crises

The impact of international conflicts will be reflected in the completion status of the goals and will affect the priority calculated by the model. Negative impacts will decrease the growth of completion status and increase the proportion of priority, while positive impacts will increase the growth of completion status and decrease the proportion of priority. Therefore, the estimated impact table mentioned above can also provide an estimated table of relative changes in priority.

Fig. 19 impact of international crises for priority

The impact of international crises on the development of many SDGs cannot be ignored. Major events can seriously slow down the progress of SDGs. In response to these situations, adjustments should be made to the corresponding regional development strategies to minimize the negative impact of these events. The development strategy and priorities of SDGs should be dynamic, regional, and timely.

9 Stability test

In this paper, stability tests are performed for **Model 2**. In model 2, in order to obtain the priority of SDGs, we calculates the product of the centrality of eigenvectors , progress of completion, and adjustment parameters of SDGs, w. The priority of development is determined according to the magnitude of the product ***w***.

However, in practice, we cannot ignore the objective existence of errors. When we conduct a typical correlation analysis of the SDGs, the data of each sub-goal utilized will inevitably have errors with the true values. Thus, when calculating the typical correlation coefficient matrix of the sustainable development goals, there will be fluctuations in d_{ij} within a certain range. In order to investigate whether the error has an impact on the priority of the targets, we use formula 13 to make the elements on the non-diagonal line of the typical correlation coefficient matrix fluctuate within a range of 10% (the same fluctuation of elements symmetrical about the diagonal line).

$$\tilde{d}_{ij} = d_{ij} \times (1 + e),\ e \sim U(-0.1, 0.1) \tag{13}$$

We repeat the test for 1000 times and obtain the results of testing, which are shown in the following figures.

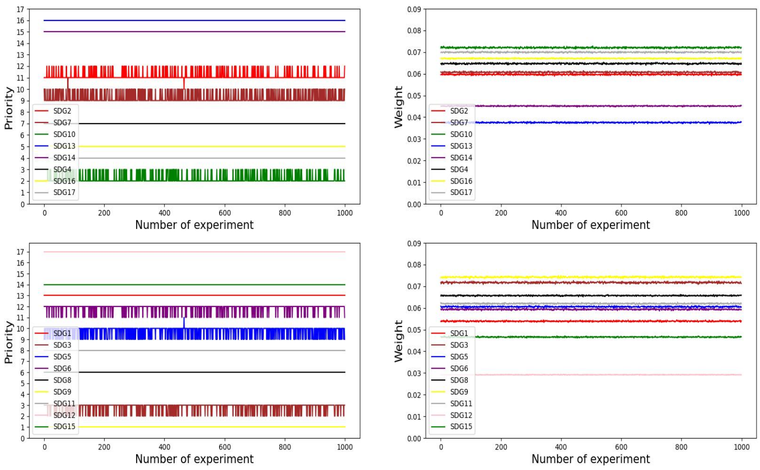


Fig. 17 The results of stability test

From the above figures, it can be seen that the priority ranking of goals (SDG-2, SDG-3, SDG-5, SDG-6, SDG-7 and SDG-10) fluctuates when the typical correlation coefficient fluctuates. However, the range of fluctuation is small and the development weights change slightly. It indicates that the random fluctuations of the elements in the typical correlation coefficient matrix of this paper within a certain range do not affect the priority ranking of the 17 SDGs. This precisely proves the stability of Model 2.

10 Model Evaluation and Further Discussion

10.1 Strengths

- Taking into account the different development situations of countries and regions, their goal priorities may differ.
- When calculating the global average goal completion rate, it is weighted by the population percentage of the three types of economies.
- The reference data and literature are widely available and authoritative, ensuring that our results are consistent with reality.

10.2 Weaknesses

- ·The discussion on the impact of external factors such as pandemics and wars on the model adopts a relatively subjective approach due to the large volume of data and the difficulty in establishing causal relationships.

11 Conclusion

We have found that a series of international crises have led countries to focus some of their funds that could be used for sustainable development on certain SDGs. The most pressing issue facing sustainable development currently is the lack of funding, which has slowed down or even halted progress in some SDGs. However, as we can see from the network we have constructed, the goals are interdependent and interconnected. "Just as crises connect them in their complexity, so must solutions." (2022 Sustainable Development Report) Only by taking action to strengthen social protection and promote social equality can we solve problems such as good health and well-being, peace and justice.

Reference

[1] Ranjula Bali Swain , Shyam Ranganathan，Modeling interlinkages between sustainable development goals using network analysis. World Development 138 (2021) 105136.

[2] LeBlanc, D. (2015), Towards integration at last? The sustainable development goals as a network of targets. Sustainable Development, 23, 176–187

[3] https://zh.wikipedia.org/zh-hans/developing countries

[4] Whitney K. Newey , Kenneth D. West, A Simple, Positive Semi-Definite, Heteroskedasticity and Autocorrelation Consistent Covariance Matrix. *Econometrica* , Vol. 55, No. 3 (1987), 703-708.

[5] Sachs, J., Lafortune, G., Kroll, C., Fuller, G., Woelm, F., From Crisis to Sustainable Development: the SDGs as Roadmap to 2030 and Beyond. Cambridge University Press, (2022).

[6] https://www.un.org/zh/ Sustainable Development Report 2022.

[7] https://www.un.org/zh/ Cambridge: Cambridge University Press.

Appendices

Appendix 1

Introduce: centrality of eigenvectors

Before deleting V_1

1	2	3	4	5	6	7	8	9
1.9106	1.8534	1.9798	1.9104	1.9875	1.9857	1.9302	1.7568	2.0000
10	11	12	13	14	15	16	17	
1.8710	1.7810	1.0000	1.2512	1.3648	1.3350	1.6986	1.7301	

After deleting V_1

2	3	4	5	6	7	8	9
1.8459	1.9811	1.9170	1.9920	1.9868	1.9376	1.7558	2.0000
10	11	12	13	14	15	16	17
1.8568	1.7974	1.0000	1.2713	1.3750	1.3512	1.7027	1.7384

Appendix 2

Introduce: The point estimate and the interval estimate of the new priority

No.Target Node	Point Estimate of Priority	Interval Estimates of Priority	
V2	2.2814	2.2565	2.3069
V3	2.687	2.6261	2.7508
V4	2.5372	-	
V5	2.2784	2.2434	2.3145
V6	2.2664	2.2413	2.2921
V7	2.3315	2.3022	2.3616
V8	2.56	-	
V9	2.4671	2.444	2.5926
V10	2.6976	2.6661	2.7299
V11	2.4451	-	
V12	1.1444	1.1411	1.1477
V13	1.4868	-	
V14	1.7747	-	
V15	1.8407	-	
V16	2.6329	-	
V17	2.7458	-	

Note : The gray shading shows the nodes that are strongly related to V1.

i