

February
2021UNCTAD Research Paper No. 82
UNCTAD/SER.RP/2022/2**Daniel
Hopp**Associate Statistician
Division on
Globalisation and
Development
Strategies, UNCTAD
daniel.hopp@unctad.org**Emily
Fu**Intern
Division on
Globalisation and
Development
Strategies, UNCTAD
EmilyFu@mail.carleton.ca**Anu
Peltola**Acting Head
Development
Statistics and
Information, UNCTAD
anu.peltola@unctad.org

Feasibility of nowcasting SDG indicators: a comprehensive survey

Abstract

The 2030 Agenda and accompanying Sustainable Development Goals (SDGs) are vital in guiding national and global policy. However, many of the SDG indicators used to measure progress toward those goals suffer from long publication lags. Nowcasting has the potential to address this problem and generate more timely estimates of those indicators. This paper provides resources for achieving that potential by 1) carrying out a comprehensive nowcasting feasibility survey of all SDG indicators to assess their potential to be nowcast, and 2) performing a case study of indicator 9.4.1 to illustrate and shed light on the process of performing a nowcasting exercise. There exist 231 SDG indicators, but due to only examining Tier 1 indicators and the fact that many indicators have multiple sub-indicators, 362 indicators and sub-indicators were eventually surveyed. Of those 362, 150 were found highly likely to be suitable candidates for nowcasting, 87 were found to be likely, and 125 were found to be unsuitable.

Key words: Nowcasting, Sustainable Development Goals, SDG indicator, survey



The findings, interpretations, and conclusions expressed herein are those of the author(s) and do not necessarily reflect the views of the United Nations or its official Member States. The designations employed and the presentation of material on any map in this work do not imply the expression of any opinion whatsoever on the part of the United Nations concerning the legal status of any country, territory, city, or area or of its authorities, or concerning the delimitation of its frontiers and boundaries.

This paper represents the personal views of the author(s) only, not the views of the UNCTAD secretariat or member States. The author(s) accept sole responsibility for any errors. Any citation should refer to the author(s) and not the publisher. This paper has not been formally edited.

Contents

1. Introduction	3
2. Background	5
2.1 SDG Indicators.....	5
2.2 Nowcasting in the SDG context.....	5
3. Nowcasting feasibility survey	7
3.1 Description of methodology	7
3.2 Survey results	9
Goal 1: No Poverty	9
Goal 2: Zero Hunger.....	10
Goal 3: Good Health and Well-Being:	10
Goal 4: Quality Education	11
Goal 5: Gender Equality	11
Goal 6: Clean Water and Sanitation	11
Goal 7: Affordable and Clean Energy.....	12
Goal 8: Decent Work and Economic Growth.....	12
Goal 9: Industry, Innovation and Infrastructure	13
Goal 10: Reduced Inequalities	13
Goal 11: Sustainable Cities and Communities	14
Goal 12: Responsible Consumption and Production.....	14
Goal 13: Climate Action	15
Goal 14: Life Below Water	15
Goal 15: Life on Land.....	16
Goal 16: Peace, Justice and Strong Institutions	16
Goal 17: Partnerships for the Goals.....	17
4. Empirical Pilot.....	18
4.1 Indicator and data.....	18
4.2 Methodology.....	19
4.3 Results.....	21
5. Conclusion.....	24
References.....	25
Appendix	32

Acknowledgements

We would like to thank Bojan Nastav, Katalin Bokor, and Nour Barnat for their valuable comments and feedback.

1. Introduction

The Sustainable Development Goals (SDGs) of the 2030 Agenda were adopted by the United Nations General Assembly in 2015 in recognition of the need for an organized international framework to help address the myriad challenges facing the world in the 21st century (UN, 2015). The goals transformed considerably when compared with the 2000 Millennium Development Goals (MDGs), reflecting an increasing pace of technological, economic, and social change and applying to all countries globally instead of developing economies only. Some issues have remained timeless and appear in both the SDGs and MDGs, such as poverty, hunger, and education, while others were substantially expanded upon, such as those concerning the environment and climate change. Still many others, such as clean energy, were newly added.

Despite the ambitious aims of the 2030 Agenda, the UN recognized that its impact would be limited without proper means of measuring and quantifying progress on its goals. Consequently, the General Assembly asked the UN Statistical Commission to coordinate the substantive and technical work to develop the SDG indicator framework to measure targets selected for each goal, currently with a total of 231 indicators spread across the 17 goals (UNDESA, 2021a). While some indicators are similar or identical to existing statistics compiled and published by national statistical offices or other national authorities, such as unemployment rate, others needed to be newly defined and collected specifically for the 2030 Agenda. Furthermore, countries' data gaps vary greatly, and they have had to put in place special efforts to enable more comprehensive reporting on the indicator framework.

In addition to filling data gaps, national statistical authorities have been challenged by increasing pressure to provide more up-to-date information as evidence for policy makers so that they have enough time to influence progress towards achieving the goals of the Agenda by 2030. Poor timeliness is a common issue for many SDG indicators (UNSD, 2020). Indicators are of limited use to policy makers in terms of both planning and programming assessment if they are published with significant lags. As noted in the 'a World that Counts' report (United Nations, 2014), data delayed is data denied.

Recently, the rise of new technological possibilities and emerging digital data sources have enabled the compilation of timelier statistics. Numerous statistical offices have quickly responded to the demand for timely data during the COVID-19 pandemic, including with the use of non-traditional data sources and new statistical techniques. One such tool that could help address issues of timeliness in SDG indicators is nowcasting. Nowcasting is the estimation of the current, or near-current, value of a target series using information from more timely series. In a world awash in data from both a plethora of new sources and from new ways of storing old data (Einav and Levin, 2014), nowcasting can help leverage that information to obtain advance estimates of lower velocity indicators. As noted by MacFeely (2021), while nowcasting has generally been well received, many questions regarding the robustness of the methodologies employed need to be solved, as well as concerns over the validity of using a wide variety of data sources, including both hard and soft indicators. Concerns have also been flagged about the impact of revisions in the underlying data, dissemination strategies, potential confusion for users, division of work between international and national agencies, and relevance to some areas of sustainable development. To date, timeliness of SDG indicators has been the responsibility of the many separate SDG indicator custodian agencies, some

managing better than others to improve timeliness in collaboration with national statistical authorities. Furthermore, as the indicators are mainly compiled by national statistical authorities, common approaches, methods and rules are needed. To increase collaboration on nowcasting among official statisticians, the United Nations Conference on Trade and Development (UNCTAD) held a nowcasting workshop with United Nations Industrial Development Organization (UNIDO) in February 2020. The meeting discussed case studies on nowcasting exercises carried out in official statistics. In 2021, UNCTAD shared its experience on new nowcasting methodologies based on neural networks with the global statistical community at a UN Brown Bag seminar. This paper aims to help turn the potential to nowcast SDG indicators into reality by firstly providing a comprehensive survey of the nowcasting feasibility of all SDG indicators and secondly by fully documenting the process of nowcasting an SDG indicator via a case study. This work was carried out in the context of an informal ‘nowcasting network’, chaired by UNIDO and set up by UN Chief Statisticians.

For this paper, 362 SDG indicators and sub-indicators were surveyed for nowcasting feasibility. This number differs from the 231 mentioned above due firstly to only examining the 130 Tier 1 SDG indicators, and secondly to the fact that some indicators have several sub-indicators. See section 2.1 for more information on Tier 1 indicators. Of those 362, 150 received a classification of “Highly likely” able to be nowcast, 87 received a classification of “Likely”, and 125 received a classification of “Unlikely”. See sections 3.1 and 3.2 for more information on survey methodology and factors determining classification. Most indicators and sub-indicators were found to be recorded at an annual frequency, with publication lags ranging between one to three years, though these lags may differ by data source, which often varies by country or region of coverage. The existence of potential explanatory variables for use in modelling is unlikely to pose a problem for most indicators. Some indicators contain sets of sub-indicators surrounding a core subject matter. These sub-indicators typically share the same data availability and publication lags, with similar sets of likely explanatory variables.

The rest of the paper proceeds as follows: section 2 will provide further background on the SDG indicators and nowcasting; section 3 will describe the approach taken to complete the survey and report on general findings for each SDG; section 4 will present the empirical case study illustrating one approach to nowcasting an SDG indicator; section 5 will conclude, summarizing main results and recommendations going forward.

2. Background

2.1 SDG Indicators

In 2012, as the target year for the MDGs approached, work began on the development of a post-2015 development agenda (UNEP, 2012). The result was the 2030 Development Agenda, adopted by the UN General Assembly on the 25th of September, 2015 (UN, 2015). In contrast to the eight goals of the MDGs, the 2030 Agenda outlined 17 goals, called the Sustainable Development Goals (SDGs), which are accompanied by a varying number of targets per goal, for a total of 169 targets (UN, 2015). Each target in turn has one or more indicators to aid in monitoring progress towards accomplishment of each target and goal, for a total of 231 indicators (UNDESA, 2021a). Each of these indicators in turn has a varying number of sub-indicators. The 2030 Agenda serves as a policy framework to help tackle such issues as the eradication of poverty, reducing inequality, and addressing climate change, among many others. For more information on specific targets and indicators, see UNSD (2021a) and Ritchie et al. (2018).

Defining priorities and agreeing on a framework in 2015 was only part of the story, as a further two years were required to make the stated goals actionable by developing the targets and indicators for each goal (UN, 2017). Measurement is a vital aspect of the 2030 Agenda, both in terms of guiding and informing policy decisions at a national and international level, as well as in quantifying progress towards the goals. Developing indicators for such an expansive agenda with a diverse array of interconnections is no small task and highly dependent on the target and type of data available. As such, there are three tiers of SDG indicator:

- ***Tier 1:** Indicator is conceptually clear, has an internationally established methodology and standards are available, and data are regularly produced by countries for at least 50 per cent of countries and of the population in every region where the indicator is relevant.*
- ***Tier 2:** Indicator is conceptually clear, has an internationally established methodology and standards are available, but data are not regularly produced by countries.*
- ***Tier 3:** No internationally established methodology or standards are yet available for the indicator, but methodology/standards are being (or will be) developed or tested (UNSD, 2021b).*

In simpler terms, Tier 1 indicators are well defined and already being produced, Tier 2 indicators are defined, but not yet being produced, while Tier 3 indicators are still being defined. It should also be noted that even indicators classified as Tier 1 may still be plagued by issues of data coverage, availability, and/or timeliness. Assessment of progress towards the achievement of the 2030 Agenda relies heavily upon reliable, accurate, disaggregated, and timely indicators. Nowcasting, examined in the next section, has the potential to address the last of those characteristics.

2.2 Nowcasting in the SDG context

The term “nowcast” itself is a portmanteau of “now” and “forecast”. Nowcasting as a term and discipline originated in meteorology in the 1980s (WMO, 2017) but began to appear in economic literature in the 2000s. Nowcasting in the economic sense, and in the sense relevant for SDG indicators, refers to the estimation of the current

value of a target variable based on timelier data and information. The distinction with forecasting comes from the fact that estimations are produced for time periods that either have already concluded or that are currently running, as opposed to periods in the future.

The intuition or justification for nowcasting is best explained by way of example using gross domestic product (GDP). GDP is a frequently nowcast target variable due to three of its characteristics. First, GDP is often published with a significant lag due to the many data sources needed and the complex accounting and aggregation procedures necessary for its calculation. GDP figures for a given quarter or year are often published many months after the conclusion of the period, even though all economic activity measured in the eventual figure has already occurred. Second, GDP usually has a long publication history, meaning there exist sufficient observations to estimate a model on historical information. Finally, there exist numerous potential explanatory variables which are published on a much timelier basis which can be used as inputs for a nowcast. Series such as consumer price indices, industrial production indices, consumer and business confidence indices, and retail trade figures, among many others, are typically published with a significantly shorter time lag than GDP, so can be used to obtain an estimate of GDP well before final figures are made available. These characteristics, together with the salience and relevance of GDP as an indicator, have made GDP the target variable for many nowcasting applications and papers. A GDP nowcasting model could then be fit on historical data and fed the latest information of timelier indicators to obtain both an estimate of GDP months before final figures are published, as well as monitor GDP outlook during the period by rerunning the model on the latest data continuously. For examples of GDP nowcasting applications see Morgado et al. (2007), Rossiter (2010), or Bok et al. (2018).

Nowcasting is relevant for SDG indicators because many face issues with timeliness. In order to successfully implement the 2030 Agenda for Sustainable Development, it is essential that policy makers have access to timely information as it relates to SDG indicators, a primary means of monitoring and evaluating progress and guiding policy interventions. The United Nations Statistics Division (UNSD) and UNCTAD have identified nowcasting as a key means of meeting this timeliness challenge (UNSD, 2020). The existing literature on nowcasting specifically as it relates to SDGs is sparse, but two notable works include Bierbaumer-Polly et al. (2019), where a comprehensive nowcasting exercise of SDG indicators using dynamic factor models is performed for Austria, and Hughes et al. (2021), where the International Futures forecasting system is used to nowcast many SDG-related indicators for more than 180 countries.

Nowcasting is however no panacea. It is only applicable for obtaining timelier estimates of an already produced SDG indicator. That restricts its application to Tier 1 indicators, where data are produced. For a given indicator to be suitable for nowcasting, a further two conditions need to be satisfied: the indicator needs a sufficiently long time series to be able to train a nowcasting model, and there need to exist sufficient related and timely explanatory variables. In order to assess the nowcasting feasibility of SDG indicators, these conditions were applied in carrying out the survey explained in greater detail in the next section. Complete survey results are available in Appendix 2, with a visualization of results available in Appendix 1.

Care should be taken if nowcasts are eventually adopted as advanced estimates of SDG indicators. Their status as data-based, quantitative estimates, liable to revision

as the data outlook changes, should be clear to users, as well as when a figure has changed from a nowcast to its actual recorded value.

3. Nowcasting feasibility survey

3.1 Description of methodology

The first step in nowcasting an SDG indicator is determining whether it is even applicable to the case. That is, do the characteristics of the indicator fulfill the data requirements of nowcasting outlined in the previous section. This was the goal of this feasibility survey: to provide a comprehensive overview of every Tier 1 SDG indicator and their sub-indicators and their potential to be nowcast. The results of the survey could help statisticians and custodian agencies know at a glance whether their indicators have the potential to be nowcast and provide a springboard from which to launch their own investigations. There are no hard and fast rules for applying the earlier mentioned three conditions of nowcasting, which depend rather on the indicator. A one-month lag for an economic series may be considered a short lag, while for an epidemiological series it could be considered a long one, etc. Rather, each indicator needed to be examined individually and evaluated for nowcasting suitability on a holistic basis. It is also worth mentioning that some indicators may be composed of a combination of two or more series, for instance indicator 8.4.2 (*Domestic material consumption, domestic material consumption per capita, and domestic material consumption per GDP*). In these cases, the indicator's publication delay may be due to one of its constituent series, and better results may be obtained from nowcasting just this series rather than the entire indicator itself.

As Tier 2 and 3 indicators lack the publication of any historical data, they could immediately be classified as not suitable for nowcasting due to the second condition. As such, they were excluded from the survey and do not appear in the table in Appendix 2. Restricted to Tier 1 indicators, the survey was conducted in the following manner: the main sources for information on Tier 1 indicators were the SDG Indicators Metadata Repository and the Global SDG Indicator Database (UNSD, 2021a; UNDESA, 2021b). The SDG Indicators Metadata Repository ideally includes information on data characteristics relevant to nowcasting feasibility. However, the type of information included in each SDG indicator metadata file tends to vary, despite ongoing work to standardize the contents. For instance, some files do not include any information on when the data are collected or released. Some metadata files are unfinished, have missing parts, or require updating or reviewing. The SDG Indicator Database displays data for each indicator but does not always reflect the data availability described for each indicator in the respective metadata file.

Custodian agency databases also usually provide access to data for their SDG indicators. Information from these databases was used as a direct source of indicator characteristics or used to validate metadata. Data availability in the SDG Indicator Database is usually up to date with the custodian agency databases but may also be vastly different in content. The survey combines data available in the two sources if the years do not completely overlap. Otherwise, the source with the longest time series provides the information in Appendix 2. The SDG Indicator Database is also limited to displaying annual data. Indicators with monthly or quarterly data, for instance, are only displayed as annual. Other database sources must be used to get

information about these indicators. Data availability is not easily describable for some indicators and sub-indicators. In many cases the length of time series varies greatly by country or region. This is indicated in Appendix 2.

For an indicator to have Tier 1 classification, data must be available for over 50% of relevant countries. There are a few indicators that can variably be classified as Tier 1 or Tier 2, depending on sub-indicator, and some sub-indicators included appear to not meet the Tier 1 requirement stated above. The number of countries or territories covered by each indicator across all possible sources is difficult to confirm due to inaccurate metadata information and inconsistencies between data sources. Many SDG indicators are also not classified at the sub-indicator level, or existing classifications have changed over time without metadata updates. Some are broadly considered Tier 1 indicators despite having sub-indicators that may not meet the Tier 1 requirements, while others are classified as either Tier 1 or Tier 2 by sub-indicator, but without specifying which sub-indicators belong to which classification.

The amount of data available for a particular indicator can vary by country, location, or aggregate grouping, so the data availability described in Appendix 2 generally focused on the data availability of a world or global aggregate, if available. When the global aggregate did not have enough data for nowcasting purposes or a global aggregate was not available for a particular indicator, for the purposes of this survey, a general summary of the relevant countries and other aggregations was used. Information on publication lags for each indicator often had to be inferred from the data that appeared to be available and the existing metadata information.

A further consideration for nowcasting feasibility is the existence of explanatory variables for a particular indicator. Each indicator or sub-indicator was given a score of “Highly likely,” “Likely,” or “Unlikely” for this area. An indicator got a label of “Highly likely” if explanatory variables would likely be easy to find. Many SDG indicators are macroeconomic variables, such as GDP, or fall under poverty, health, education, environment or ecological topics. Variables like these are frequently modelled and existence and availability of explanatory variables is well-documented in literature. An indicator gets a label of “Likely” if potential explanatory variables may not be closely related to the behavior of the indicator or had limited data available. For instance, variables related to Official Development Assistance (ODA) or certain government spending decisions were listed in this category, since values for these indicators are generally pre-determined by government decisions, although there are a variety of socioeconomic factors that may still contribute to initial spending decisions themselves. An indicator got a label of “Unlikely” if it may prove

预览已结束，完整报告链接和二维码如下：

https://www.yunbaogao.cn/report/index/report?reportId=5_8483

