Science of Remote Sensing 12 (2025) 100298

Contents lists available at ScienceDirect

Science of Remote Sensing

journal homepage: www.sciencedirect.com/journal/science-of-remote-sensing

A perspective on the interpretability of poverty maps derived from Earth Observation

Gary R. Watmough [a,b,*], Dan Brockington [c,d,e], Charlotte L.J. Marcinko [f], Ola Hall [g], Rose Pritchard [h], Tristan Berchoux [i], Lesley Gibson [j], Enrique Delamonica [k], Doreen Boyd [l,m], Reason Mlambo [a], Seán Ó Héir [n], Sohan Seth [n]

[a] School of Geosciences, University of Edinburgh, UK
[b] Global Academy of Agriculture and Food Systems, University of Edinburgh, UK
[c] Institut de Ciència i Tecnologia Ambientals de la Universitat Autònoma de Barcelona (ICTA-UAB), Spain
[d] Department of Private Law, Universitat Autònoma de Barcelona, Spain
[e] 3 ICREA, Pg. Lluís Companys 23, Barcelona, Spain
[f] Government Actuary's Department, UK
[g] The Department of Human Geography, Lund University, Sweden
[h] Global Development Institute, The University of Manchester, UK
[i] Mediterranean Agronomic Institute of Montpellier (CIHEAM-IAMM), University of Montpellier, France
[j] Department of Civil Engineering, Stellenbosch University, South Africa
[k] Division of Data, Research and Policy, UNICEF, USA
[l] School of Geography, The University of Nottingham, UK
[m] Rights Lab, University of Nottingham, UK
[n] School of Informatics, University of Edinburgh, UK

A B S T R A C T

The use of Earth Observation Data and Machine Learning models to generate gridded micro-level poverty maps has increased in recent years, with several high-profile publications. producing some compelling results. Poverty alleviation remains one of the most critical global challenges. Earth Observation (EO) technologies represent a promising avenue to enhance our ability to address poverty through improved data availability. However, global poverty maps generated by these technologies tend to oversimplify the complex and nuanced nature of poverty preventing progression from proof-of-concept studies to technology that can be deployed in decision making. We provide a perspective on the EO4Poverty field with a focus on areas that need attention. To increase the awareness of what is possible with this technology and reduce the discomfort with model-based estimates, we argue that the EO4Poverty models could and should focus on explainability and operationalizability alongside accuracy and robustness. The use of raw imagery in black-box models results in predictions that appear highly accurate but that are often flawed when investigated in specific local contexts. These models will benefit for incorporating interpretable geospatial features that are directly linked to local context. The use of domain expertise from local end users could make model predictions accessible and more transferable to hard-to-reach areas with little training data.

1. Introduction

The Earth Observation for Poverty ('EO4Poverty') field has emerged in recent years with studies examining if and how Earth Observation data can be used to estimate and predict poverty. EO4Poverty, in part, is a consequence of the '*Data Revolution*' that the UN says is needed to report progress towards the Sustainable Development Goals (IEAG, 2014; Watmough and Marcinko, 2024). But the deficiencies and challenges facing EO4Poverty techniques, and its conceptualisation of poverty, are not properly appreciated. Three recent reviews have examined how EO data are used in poverty mapping (Burke et al., 2021; Newhouse, 2024; Lamichhane et al., 2025) and we do not seek to re-review the literature and refer interested readers to these reviews. Instead, we focus on trends in the literature that need attention if EO poverty maps are to progress from proof-of-concept studies to being recognized as a genuine tool in policy implementation.

EO4Poverty models could and should focus on explainability, operationalizability alongside accuracy and robustness - but recent publications have focused on the latter aspects. Furthermore, models should not assume that the relation between Earth Observation and poverty is

* Corresponding author. School of Geosciences, University of Edinburgh, UK
E-mail address: Gary.watmough@ed.ac.uk (G.R. Watmough).

https://doi.org/10.1016/j.srs.2025.100298
Received 23 May 2025; Received in revised form 29 September 2025; Accepted 30 September 2025
Available online 30 September 2025

static over space and time. These models will benefit from incorporating geospatial features alongside raw imagery as well as domain expertise from local end users to ensure accessibility. For these challenges to be addressed, a more systematic approach to building these maps is required, which can be readily adapted from existing technology assessment procedures.

Technology readiness levels (TRLs) are used to assess a product's maturity and suitability for use in the market. These levels help to ensure that the technology is checked for appropriateness and robustness and assesses performance before the eventual deployment of the technology more widely (Lavin, 2022; NASA, 2020). There are ten TRLs, broadly covering three main domains of product creation: research, development and deployment. These domains include *first principles* (embedded with research) where new ideas are explored, *proof-of-concept* where the technology or algorithm is demonstrated in real scenarios (development domain) and *deployment* where the technology is operationalized in the real world and its performance monitored (Lavin, 2022). Most of the work on EO4Poverty to date suggests that estimating poverty from Earth Observation data is at, or around, the proof-of-concept level - level 4 in Lavin (2022).

1.1. Why do we need EO derived poverty maps?

Rapid, frequent and precise satellite data have much to offer, given the deficiencies of current survey and census data (Jerven, 2017), which are rarely sufficient for monitoring the SDG1 indicators (Fig. 1). Census data can be used to estimate multidimensional poverty, but the available indicators tend to be too few. Furthermore, these data are collected infrequently (typically once a decade), which cannot capture the changes required in a 15-year SDG cycle (Fig. 2). They often cannot be used to estimate monetary poverty because they rarely include income and consumption data. In addition, the gaps between census surveys mean that, it can be difficult to use these data to understand the drivers of change (Watmough and Marcinko, 2024). Household surveys that target representative samples of a population provide in-depth information on specific focal areas (e.g., demographics, health, expenditures). They are more frequently collected than census (typically conducted once every 3–5 years) but, although cheaper than a census, these surveys can still be expensive and increasing the number will be prohibitively expensive (Jerven, 2017). Therefore, it is unlikely traditional surveys will be increased in frequency and will mean that gaps in data will remain in most countries for the foreseeable future.

The theory behind using Earth Observation data to map poverty is that landscape characteristics are often indicative of socioeconomic conditions (Okwi et al., 2007), and these data can be used to '*see*' some of these characteristics (Watmough and Marcinko, 2024). Therefore, EO data can be used to create metrics that represent these characteristics which can be statistically linked to aspects of local poverty and well-being and subsequently used to predict poverty and/or wellbeing (Watmough et al., 2016).

Earth Observation data have several features that make them suitable for contributing to the SDG1 data revolution and poverty mapping in general: (1) They are, relatively speaking, available in standard formats for most of the globe and can be joined/integrated with traditional survey data relatively easily (Watmough et al., 2016); (2) They are collected frequently, e.g., daily, weekly or monthly, and could potentially fill gaps between traditional surveys in a cost-effective manner [8]; (3) They can provide data in locations where none are available (Chi et al., 2022; Smythe and Blumenstock, 2022; Hall et al., 2023; Lee and Braithwaite, 2022; McCallum et al., 2022; Watmough et al., 2019). As such, there have been several studies examining how Earth Observation data can be used to estimate poverty and wealth in small regions of Kenya and India (Watmough et al., 2016; Watmough et al., 2019; Marcinko et al., 2022), as well as country-level models in Sri Lanka (Engstrom et al., 2022) and Bangladesh (Steele et al., 2017). Each was able to link a measure of poverty with proxies derived from Earth Observation data or by using raw imagery sources.

2. How is EO used for poverty mapping?

There are many models that have been used to map poverty using EO data, including linear models (Watmough et al., 2013), structural

Target	Indicator	Progress	Reporting Frequency	Future EO role
1.1 Eradicate extreme poverty	1.1.1 % of population living below extreme poverty line	Increase of 70 million people in extreme poverty since 2019. Projected to be 575 million in 2030	Annual, but can be projections as require NSO to release national HH surveys which are not often conducted annually.	Extreme poverty line is defined by $ per day. Most EO measures look at asset based wealth. Examine if EO can accurately estimate income/flow based poverty.
1.2 Halve all forms of poverty	1.2.1 % of population living below national poverty line	Current trends mean 67% of countries will not achieve this.	See above.	National poverty line is define by $ per day. Most EO measures look at asset based wealth. Examine if EO can accurately estimate income/flow based poverty.
	1.2.2 % of population living in any form of poverty	Only a few countries have adopted multi-dimensional poverty indicators	Annual – EU and some LA, most Low income countries = once in last 10 years.	Alkire-Foster (2009) method commonly used. EO can proxy aspects of living standards and access/travel time to health and education.
1.3 Social protection for all	1.3.1 % of population with access to social protection systems	Increased from 45% in 2015 to 47% in 2020. Slow progress, unlikely to be achieved.	Annual, but can be projections as require NSO to release national HH surveys which are not often conducted annually.	Governance and systems difficult to see from space. Distribution and access to institutions (banks, govt, coops) may be proxied using EO based metrics showing, roads, rail, building density etc.
1.4 Equal rights & access to basic services	1.4.1 % of population living in HH with access to basic services	Increased basic service investment from 47% of total govt expenditure to 53% in 2021. But its 62% in high incomes countries and 20% lower in lower income countries.	Every 2-5 years data are collected.	Physical access to services is partly determined by transport infrastructure. Roads can be detected and classified in satellite imagery. But new roads and improved roads are not commonly measured yet.

Fig. 1. Sustainable Development Goal 1 Targets, Indicators and current reporting frequency of data collection/publication. All data taken from the SDG meta data portal from December 2023 https://unstats.un.org/sdgs/metadata/EU = European Union; LA = Latin America; NSO = National Statistics Office; HH = Household; EO = Earth Observation; govt = government, coops = cooperatives (Alkire and Foster, 2009).

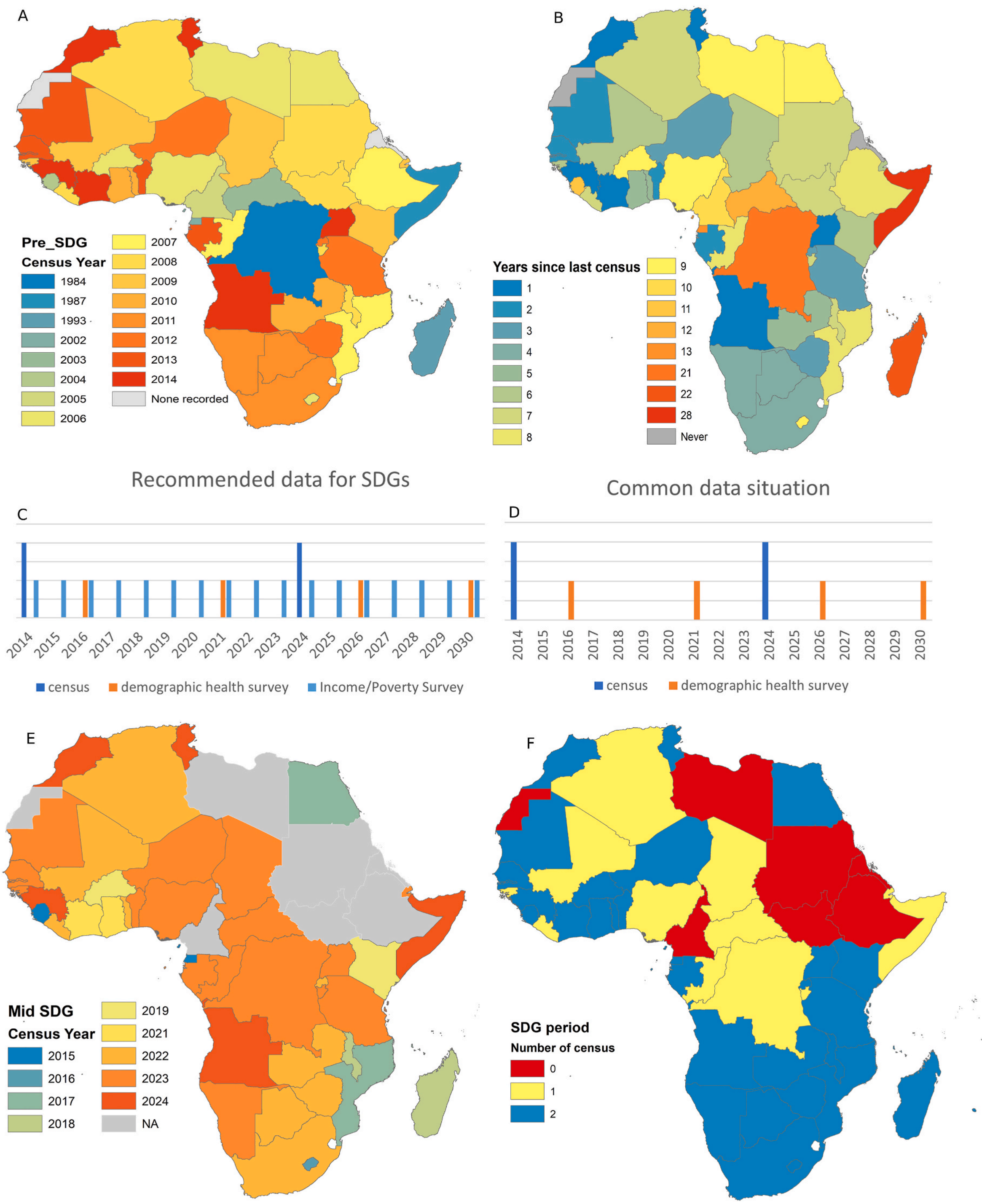

Fig. 2. (A) Showing the most recent year of national census data available prior to the SDGs start date in 2015; Eritrea and Western Sahara have never conducted a national census, South Sudan has not conducted a census but was included in the Sudan Census in 2008 before independence; (B) years since last census calculated from 2015 baseline – so for example 1 indicates the last census was in 2014; (C) Morten Jerven (2013) and others (United Nations, 2015) recommend that countries collect a census every 10 years, health surveys every 3–5 years and income/poverty surveys annually in order to monitor the SDGs (D) the majority of countries however collect a decadal census and then have Demographic and Health Surveys once every 5 years. (E) the expected year of a census during the SDG period, those countries in green may have a second census during the SDG time window. Several countries have no planned census during the SDG time period including Sudan, Cameroon, Libya and Ethiopia has had to repeatedly delay the planned 2017 census due to security concerns; (E) the number of census surveys that a country could complete and use for policy during the SDG time window – we have extended this to 2010 to allow for a baseline, the ideal situation would be for each country to have 2 census during this time. 7 countries (12 %) have zero census surveys available, 12 countries (21 %) will have only a single census whilst 39 (75 %) will have 2 – note here we assume that countries with a planned census in 2015–2019 will perform another 10 years later; Source: the authors with data from UNSTATs.

equation models (Steele et al., 2017; Van der Weide et al., 2024), tree-based methods (Watmough et al., 2019; Lee and Braithwaite, 2022), random forest models (Krennmair and Schmid, 2022), augmented small-area estimations (Masaki et al., 2022) and deep learning (Chi et al., 2022; Jean, 2016). We acknowledge that each of these models has its respective strengths and limitations in the context of accuracy, explainability, robustness and usability. In this perspective, we focus on the way in which the EO data is used in these models. We recognise that there are two distinct ways: the first is the use of interpretable features where EO data is processed and analysed to create a series of proxy metrics for specific ground features related to poverty and wealth, and the second is the use of raw satellite imagery entered directly into the model.

When using interpretable EO features to estimate poverty, satellite imagery is converted to land cover classes or other variables of interest such as a time series of vegetation greenness over agricultural fields (Steele et al., 2017; Watmough 2016; Watmough et al., 2019). These data are linked to human settlements using boundaries of villages, towns, counties, districts. Using a GIS, metrics can be created such as the proportion of different land cover types within a settlement, the length of the agricultural growing period, counts of buildings, types and density of road infrastructure. Although few studies go to this level of detail to describe the inclusion of each predictor (Newhouse, 2023). These metrics are then statistically related to household, village or district level poverty measures (Engstrom et al., 2022; Watmough 2016; Steele et al., 2017). The advantage to this approach is that hypotheses can be tested for specific EO variables and then described and explained, building a level of confidence in the models. The disadvantage to this approach is that providing hypotheses for individual EO variables might require local expert opinion and means model outputs apply to smaller areas.

More recently, several studies have approached mapping poverty using Earth Observation data using deep learning and raw imagery (Chi et al., 2022; Lee and Braithwaite, 2022; McCallum et al., 2022; Jean, 2016; Zheng et al., 2025), alongside potentially using traditional machine learning models on embeddings from pre-trained deep models. These have resulted in gridded micro-level poverty data outputs spanning multiple countries which are easy to download and analyse. The accuracy of the poverty estimates from these models can be compelling as they are often in the range of 70–90 % (Chi et al., 2022; Jean, 2016; Yeh et al., 2020; Lee and Braithwaite, 2022). However, multi-country models such as the relative wealth index (Chi et al., 2022) have been found to give inaccurate estimates in specific locations (Gualavisi and Newhouse, 2024). Part of the reason for this is the large, generalised model smoothing local patterns and context. We argue that if the models are to be used in downstream decision-making and adopted into SDG reporting dashboards, they also need to be explainable, i.e., the user must be able to account for the patterns they show because explanation of the patterns is required for informing appropriate policy responses. These models usually use a convolutional neural network or a transformer architecture that processes raw satellite imagery tiles to predict the poverty index. The advantage is that raw satellite imagery can be processed directly without extracting hand-crafted features. The disadvantage is that it is difficult to understand what aspects of the imagery led to a decision on a specific poverty value. This lack of clarity in why the poverty map looks a particular way will limit the uptake in policy and decision-making and prevent the technology from progressing towards deployment.

3. EO4Poverty ambitions

Two recent developments in the field of EO4Poverty raise questions about the readiness of the technology for providing the explanations that would support policy-decision making and SDG monitoring. The first is the use of Earth Observation and Machine Learning to support geographic targeting of poverty reduction measures, including pro-poor cash transfer schemes (Smythe and Blumenstock, 2022). The second is the interest in using Earth Observation and Machine Learning in locations where little or no data are published, such as the recent example in North Korea (Ahn et al., 2023). Considering the aims of these recent studies alongside the ideas set out by the Sustainable Development Solutions Network (O'Connor et al. 2020), three key ambitions for *EO4Poverty* appear to be emerging:

1. To support geographic targeting of poverty reduction measures;
2. To improve the frequency of updates on wealth and poverty in support of the SDG1 reporting, so that they can yield annual updates (Hall et al., 2022).;
3. To provide estimates of poverty and wealth in data-poor countries or regions.

We provide suggestions for how the field of EO4Poverty could move forward to operationalise such laudable ambitions. We are particularly focused on the models that combine black-box machine learning approaches with raw EO imagery to provide micro-level (gridded data outputs at 1–3 km resolution) estimates of poverty rather than meso-level (administrative boundaries) estimates, as these seem to have the most focus in the literature and are more widely applicable to larger areas, meaning they can reach TRL9 deployment.

4. What needs to change?

The use of Earth Observation data in poverty monitoring and decision making is rare (Hall et al., 2022). To arrive at technology deployment (level 9) requires a period of evidence gathering from multiple locations and using a range of data sources to identify:

1. What are EO derived poverty maps predicting?
2. Are the maps providing generalised patterns, or do they consider local context?
3. Which aspects of EO4Poverty maps are not being explored?

4.1. What are EO derived poverty maps predicting?

Poverty is multi-facetted. It can cover, *inter alia*, aspects of health, education, well-being more generally, wealth, assets, expenditure or income. If we are to know about data poor regions, or geographically target poverty alleviation schemes, then it is essential to be specific about what form of poverty we are talking about. For example, cash-transfers are anti-poverty policies that should be targeted using reliable poverty statistics that accurately identify specific forms of poverty.

It is often unclear what Earth Observation indices are measuring because terms such as wealth, welfare, wellbeing and poverty are used in a variety of ways and often interchangeably (Chi et al., 2022; McCallum et al., 2022; Watmough et al., 2019; Yeh et al., 2020). Different forms of wealth, well-being and poverty are correlated but different (Johnston, 2016) and how they are defined and measured can lead to differences in who is classed as being poor (Pu et al., 2024). *Wealth* can measure the historical acquisition of assets (from durable consumption goods like houses and cars, financial instruments and liquid savings to ownership of capital goods) over a relatively longer period. *Poverty* is measured in terms of income, diet or consumption (to afford a minimum standard of living) (Johnston, 2016; Pu et al., 2024). It is possible for people to be wealthy in terms of their assets, but relatively poor in terms of diet or consumption, for example when families are restricting their expenditure to invest in land, houses education or their businesses.

Household income and expenditure surveys are not as widely available as those measuring durable household items, and as such most Earth Observation approaches use wealth indices based on assets rather than poverty statistics to train models. Even then, asset indices can rank households differently than consumption-based poverty measures (Ngo and Christiaensen, 2019). Cash-transfers focus on cash-income and thus, using an asset-based wealth index to geographically target a cash-poverty policy is not optimal and may go some way to explain why almost a third of eligible recipients of a theoretical cash transfer scheme in Indonesia were missed when the policy was targeted using the Earth Observation-derived relative wealth index (Sartirano et al., 2023). If a policy is targeting cash poverty, then the data should measure income or expenditure poverty and not asset-based wealth, as the two are related but different and it would not be possible to defend allocations of cash resources using a wealth or asset-based dataset.

4.2. Are the maps providing generalised patterns, or do they consider local context?

There might be insufficient training data to generate micro-level poverty maps using machine learning for individual countries. To overcome this, models combine multiple countries into the same training set. When applied across dozens of countries, models effectively assume that the relationships between Earth Observation-derived metrics and wealth are static, i.e., the Earth Observation features or proxies have fixed relationships with wealth/poverty across space and time. However, assets do not have universal value to people, and the same asset can be worth different things to different people (Jonhston and Abreu, 2016; Steinert, 2016) and therefore simply identifying the presence of an asset is not always enough to estimate wealth (Gallemore et al., 2022). Ultimately, this means that when multi-country models are produced local context is often being ignored. Many of the deep learning models take a relatively standard approach to the construction of a wealth index from Earth Observation data (Filmer and Pritchett, 2001) whereby, multiple features or latent representations derived from Earth Observation data are statistically combined into a single index value (Chi et al., 2022; jean et al., 2016; Yeh et al., 2020; Hall et al., 2022). For example, the RWI (Chi et al., 2022) uses 2048-dimensional latent representation reduced using PCA to around 200 components. It is very difficult to identify (1) what is driving the model output when 200 PCs are used as predictors and (2) how local context is being considered.

Understanding spatial diversity is particularly important when using EO4Poverty to explore data-poor regions. The recent attempt to estimate economic development in North Korea (Ahn et al., 2023) using geospatial data demonstrated a key strength of the EO4Poverty field. Some countries cannot or will not release data publicly. Earth Observation data are uniquely placed to provide information on places without requiring a physical presence (albeit a limited one if no ground data is available to be combined with the Earth Observation data). Ordinarily, transferring a model developed on one location should require the identification of comparable countries. But which metrics should be used to show countries are comparable? Arguably all countries will differ when considering a combination of economic, social, historical and climatic factors. But can they be compared in this way using thousands of raw-imagery variables that have been reduced in a statistical manner.

Temporal dynamics are vital for understanding the relationship between different forms of poverty and their causes. If the dynamics of an area change, for example, the main economic activity changes from agriculture to industry, this is highly likely to be missed using large models with raw-imagery variables derived from multiple years. Moreover, the performance of these models can deteriorate in short periods of time (months), and they may perform very unevenly across sub-national units (Sartirano et al., 2023). To improve results, the focus should be on developing approaches that account for local drivers of poverty and how changes in these can be identified using Earth Observation.

As SDG reporting requirements specify, Earth Observation data can provide data annually in support of SDG 1, but to date, there have been no studies examining how estimates derived from Earth Observation capture annual changes in poverty or wealth. Kondmann and Zhu (2020) found that the transfer learning methods commonly used in predicting relative poverty using Earth Observation were able to predict relative wealth in Rwanda at two fixed time points using separate models. However, they were unable to predict change in wealth using changes in Earth Observation metrics over the same period for the same areas. In fact, the transfer learning model predicted a reduction in wealth in several areas despite the household survey data used to train the model showing the opposite outcome. The authors suggest that the features extracted from Earth Observation changed slowly, meaning that they do not provide a strong enough signal for models (Kondmann and Zhu, 2020). However, the model used Landsat 7 data with a 30m spatial resolution, which cannot be used to identify small-scale changes such as building roof material types (Table 1) that might be indicative of poverty (Brockington and NOE, 2021; Östberg, 2018).

A further challenge of EO4Poverty for annual updates to SDG1.1 and 1.2 is that most of the approaches so far have prioritised accuracy by maximising the amount of data used for prediction (Chi et al., 2022; McCallum et al., 2022; Jean, 2016). This has served to create some compelling overall results ranging from 70 to 90 % and products such as the RWI (Chi et al., 2022; Lee and Braithwaite, 2022). However, this approach assumes that, first, the relationship between Earth Observation and poverty does not change over time. As we have already discussed, this is unreasonable. Second, it requires that both data (Earth observation and household survey) have been acquired at consistent time. Given the datasets are often collected at different time intervals, e. g., household surveys can span over months while high-resolution satellite products can pool data over years, it makes it difficult to examine if the approaches can detect subtle changes in wealth/poverty reliably on

Table 1
Examples of domain-knowledge being used in the selection of geospatial datasets and how the reason for their inclusion can help to make models mode explainable.

Geospatial data	Reason for inclusion (explainable)
Nighttime lights	The nighttime lights have been correlated with GDP and wealth in the past. It can be used in combination the settlement extents to estimate the % of lit area which has been used for predicting a wealth index (McCallum et al., 2022).
Length of growing period	Shorter growing periods can mean lower amounts of food available in markets with subsequent impacts on market activity and local economies (Watmough et al., 2019).
Building roof material (improved/metal)	Improved material is indicative of better-quality building and higher incomes (Watmough et al., 2019).
Proportion of agricultural land	Having some agricultural land is expected to be positive as it should mean more food available, easier local access. But it can vary, for example large commercial plantations have been found to lead to increased local population poverty for some rural areas with lower human capital to diversify livelihoods (Berchoux et al., 2019)
Proportion of shrubs, closed and open forest and herbaceous vegetation	Typically expect to see these positively associated with wealth and livelihoods as a mixture of different land cover types indicates opportunities for different ecosystem services etc.
Proportion of bare ground	Typically this is a negative in resource dependent areas as it indicates unproductive land that either isn't being used or cannot be used which limits local agriculture and food etc.
Road density	Often expect more roads to be a positive influence on wealth and deprivation as they allow for easier access to assets such as school, health and markets
Walking travel time to health centre	Longer time to travel is a negative for health as it can mean treatments are delayed or not sought at all for children (Watmough et al., 2022).

an annual basis while using a range of different datasets included from a variety of different periods (i.e., Earth Observation, household survey, census etc.).

5. Which aspects of EO4Poverty maps are not being explored?

It is often not possible to identify which features are driving the prediction of the wealth indices because of the black-box nature of the models combined with the use of raw –imagery in many models. Therefore, these models in their current formulations do not provide the transparency required to understand the mechanisms that are driving the changes in wealth or poverty (Sekara et al., 2024). Until this is achieved, the field might not progress from TRL4 – proof-of-concept.

Explaining model outputs is hard because 'features' extracted by an algorithm are neither tangible nor legible to potential users. Ultimately, poverty is not directly related to the amount of radiance detected in the red portion of the electro-magnetic spectrum (to give an example of one such raw feature used in some models). Instead, it is associated with physical access to agricultural land (as well as land tenure rights), road surfaces/access to services, agricultural productivity, and the interactions between each of these. Trust in these Earth Observation-driven methods requires transparent and explainable results, which have been lacking in several recent papers.

Alongside transparency, another crucial aspect to consider is uncertainty quantification, particularly when the model is being deployed to areas it has not been trained on. Rather than returning a single value of poverty for an area, uncertainty estimates provide information on how confident the methods are in predicting these values. Considering the model's uncertainty helps the end user judge whether the resulting poverty values should be trusted and acted upon, particularly when transferring these models to areas they have not been seen before.

Our focus has been on the quantitative aspects of EO4Poverty, while various other qualitative facets can also influence their operationalizability and adoption in practice, including the lack of awareness of these tools and, as a result, a possible lack of trust from policymakers. A detailed discussion on this has been beyond the scope of this perspective, and we acknowledge that further studies might explore the awareness, accessibility, expertise, opportunity and accountability required to facilitate the safe integration of these tools in decision-making.

5.1. Conceptualising improvements to EO4Poverty using TRL

Earth Observation-driven representations of poverty need to be both accurate and explainable. The complexity and opacity of current EO4Poverty modelling approaches leaves policymakers and NGOs struggling to explain the basis on which decisions were being made, potentially compromising the trust and support of electorates and donors (von Eschenbach, 2021). We have first-hand experience of UNICEF questioning how the poverty maps were created and without providing interpretable model features they do not feel they can be fully incorporated into operational processes. There is also a risk of disenfranchising people who do not have the technical knowledge to challenge the ways that decisions are being made, particularly if Earth Observation-based monitoring approaches are coupled with any kind of automated decision-making techniques (Adams, 2018). We can use the TRL approach to specify the pathways through which the required improvements required may be possible.

Reliance on obscure black boxes and offering incomprehensible variables as explanations put some elements of EO4Poverty in TRL Level 0 *'first principles'* (Lavin, 2022). This is because the technology is unlikely to be accepted in downstream decision-making without establishing how EO4Poverty works by ensuring that models can be explained. More reliable outputs and trust in them could be addressed by improving the way models are trained, rethinking the level at which poverty can be predicted using these models (e.g. meso-level models estimating population at the administrative units as opposed to micro-level models estimating population at a gridded level) and better inclusion of domain knowledge and context-specific expertise. Meso-level (administrative boundaries) poverty predictions typically use Bayesian models and are conditioned on the surveys but the existing micro-level (gridded poverty predictions we focus on here) are affected by the fact that there is very little training data because most pixels in a country have no training label available. Table 1 gives a theoretical example of how domain-knowledge can be used to identify geospatial data for use in a model and provide reasons for the use of data. This is how the authors have approached this with stakeholders. Working out how this domain-knowledge can be generalised or not is a key question for the future of the EO4Poverty field and should not be based solely on model performance. To date, there have been only one or two attempts to use domain-knowledge (Sartirano et al., 2023; Watmough et al., 2013) to define the Earth Observation proxies, and both had limited spatial extents. Providing descriptions of which geospatial variables were used to generate poverty predictions and what the relationships look like (e.g., areas with longer growing periods are predicted to be wealthier) helps link local expert knowledge with the modelling approaches and will help to build confidence in downstream users. These relationships can be explored with various linear and nonlinear (e.g., tree-based models) alongside model-agnostic explainability tools such as the Shapley decomposition or SHAP values (Lundberg et al., 2020). Domain expertise can be used to identify geospatial indicators relevant for poverty mapping in a particular location. It is likely that a sub-sample designed to create proxies for specific aspects **of** poverty would be selected from a list of globally available indicators. The interpretation of how these proxies are linked to poverty would also benefit from local domain-expertise (Alatas et al., 2012).

Building on the explainable models and inclusion of domain knowledge could allow the technology to be integrated into downstream decision-making for regions or countries that have been unable to publish statistics on their populations. This would allow the technology to reach TRL 6 *'Application Development'* where the approach can be applied to specific use-cases (Lavin, 2022) prior to it being '*Deployed*' (TRL Level 9).

Ahn (2023) provide a good case study in North Korea using a model that combined human expertise with models that were trained on other countries that publish data. In this example, specific domain-knowledge was acquired from human experts on how poverty or wealth is determined. These narratives of poverty/wealth were used to identify geospatial data sets to act as proxies for the local determinants. We suggest that this could represent a way for EO4Poverty to move from First Principles (level 1) to Proof-of-concept (level 4) and beyond because a model using these predictors is more explainable to the end-user and therefore more likely to be adopted. This is because direct connections can be drawn between the determinant of poverty and the geospatial variable used in the model. Should the model not produce an expected output, this can be explored further with experts, perhaps identifying how to weight variables differently, calculate them in different ways or finding new variables to use instead.

An advantage of combining interpretable features with expert interpretation is that it places power and control back into the hands of the local experts that can be characteristic of the highest TRL levels. The 'panoptic gaze' of satellites (Davis et al., 2021) creates an abundance of opportunities, but with these opportunities comes the risk of further disempowering people who are already disadvantaged or marginalised. Making sure that EO4Poverty approaches are contextualised and open to challenge, by represented people themselves where possible and by area experts where not, is an important step to mitigating the risk that represented people lose even more agency over the ways their lives are represented and governed. There are still important conversations to be had, however, about whether it is equitable or fair for the power to represent poverty to be concentrated among people external to the contexts in which poverty is directly experienced.

Limited understanding of temporal and spatial dynamics puts many

EO4Poverty initiatives on TRL Level 2 *'Proof of Principle'*. To improve the field needs to develop models for specific years using household survey and Earth Observation from the same period and repeat this for multiple years to see how consistent the relationships are and establish the effectiveness of development monitoring through Earth Observation. The field also needs to identify if Earth Observation data can capture specific metrics, that are informative of different aspects of wealth and poverty, at certain resolutions and if it can detect changes in particular indicators, and crucially, if these indicators are related to changes in poverty or wealth in the local populations being studied. These models should also be preferably dynamic in nature to accommodate inconsistencies in data collection, and to allow the relationship between Earth Observation and poverty to change over time.

6. Conclusion

There has been great progress made in the field of EO4Poverty in estimating poverty and wealth from Earth Observation data. Several data-driven models have been published demonstrating the proof-of-concept that Earth Observation can indeed be used to estimate some aspects of poverty and wealth. Recent studies have started to consider how the approaches can be integrated into downstream decision-making by contributing to geographic targeting of anti-poverty policies. Further ambitions are suggested in the literature of using the technology to update annual poverty and wealth estimates and providing estimates of wealth in countries that have no published statistics. The ambition is laudable but there are some limitations in the current set of approaches that need to be overcome before the Earth Observation data can be considered as part of the solution for mapping and monitoring of SDG1 indicators. The models need to be more explainable to allow decision makers to understand the drivers of poverty, make decisions on local contextual information and justify their choices for allocating scarce resources. Increasing the ability of purely data-driven models to account for domain knowledge will likely help cement Earth Observation's role in the Sustainable Development Goal 1 data revolution.

CRediT authorship contribution statement

Gary R. Watmough: Writing – review & editing, Writing – original draft, Visualization, Supervision, Investigation, Formal analysis, Conceptualization. **Dan Brockington:** Writing – review & editing, Writing – original draft, Methodology, Conceptualization. **Charlotte L. J. Marcinko:** Writing – original draft, Investigation, Formal analysis, Conceptualization. **Ola Hall:** Writing – original draft, Methodology, Conceptualization. **Rose Pritchard:** Writing – original draft, Conceptualization. **Tristan Berchoux:** Writing – original draft, Conceptualization. **Lesley Gibson:** Writing – original draft, Conceptualization. **Enrique Delamonica:** Writing – original draft, Investigation, Conceptualization. **Doreen Boyd:** Writing – review & editing, Writing – original draft, Investigation, Conceptualization. **Reason Mlambo:** Writing – original draft, Investigation. **Seán Ó Héir:** Writing – review & editing, Validation. **Sohan Seth:** Writing – review & editing, Writing – original draft, Formal analysis, Conceptualization.

Funding

This work was supported by NERC through funding for SENSE CDT scholarships two co-authors (Reason Mlambo and Seán Ó Héir) and their supervisors (Gary Watmough and Sohan Seth).

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.

Data availability

No data was used for the research described in the article.

References

Adams, et al., 2018. Conservation by algorithm. Oryx 52 (1), 1–2. https://doi.org/10.1017/S0030605317001764.
Ahn, et al., 2023. A human-machine collaborative approach measures economic development using satellite imagery. Nat. Comms 14 (6811).
Alatas, V., Banerjee, A., Hanna, R., Olken, B.A., Tobias, J., 2012. Targeting the poor: evidence from a field experiment in Indonesia. Am. Econ. Rev. 102 (4), 1206–1240.
Alkire, S., Foster, J., 2009. Counting and Multidimensional Poverty Measurement, OPHI. *OPHI Working Paper No. 32* Oxford Poverty & Human Development Initiative, Oxford. ISBN 978-1-907194 16-0.
Berchoux, T., et al., 2019. Agricultural shocks and drivers of livelihood precariousness across Indian rural communities. Landsc. Urban Plann. https://doi.org/10.1016/j.landurbplan.2019.04.014.
Brockington, D., Noe, C. (Eds.), 2021. Prosperity in Rural Africa? Insights from Longitudinal Research in Tanzania. Oxford University Press, Oxford.
Burke, M., Driscoll, A., Lobell, D.B., Ermon, S., 2021. Using satellite imagery to understand and promote sustainable development. Science 371 (1219).
Chi, et al., 2022. Microestimates of wealth for all low- and middle-income countries. Proc. Natl. Acad. Sci. USA 119 (3). https://doi.org/10.1073/pnas.2113658119.
Davis, et al., 2021. The aerial panopticon and the ethics of archaeological remote sensing in sacred cultural spaces. Archaeol. Prospect. 28 (3), 305–320.
Engstrom, R., et al., 2022. Poverty from space: using high resolution satellite imagery for estimating economic well-being. World Bank Econ. Rev. 36 (2), 382–412.
Filmer, D., Pritchett, L.H., 2001. Estimating wealth effects without expenditure data or tears: an application to educational enrollments in states of India. Demography 38 (1), 115–132. https://doi.org/10.2307/3088292.
Gallemore, et al., 2022. The livelihood impacts of sustainability partnerships in south-east Tanzania, chapter 10. In: Ponte, et al. (Eds.), Contested Sustainability. James Currey, Suffolk, UK. ISBN 978-1-80010-562-1.
Gualavisi, M., Newhouse, D., 2024. Integrating survey and geospatial data for geographical targeting of the poor and vulnerable: evidence from Malawi. World Bank Econ. Rev. lhae025.
Hall, O., et al., 2022. A review of explainable AI in the satellite data, deep machine learning, and human poverty domain. Patterns 3 (10), 1–15. https://doi.org/10.1016/j.patter.2022.100600.
Hall, et al., 2023. A review of machine learning and satellite imagery for poverty prediction: implications for development research and applications. J. Int. Dev. 1–16. https://doi.org/10.1002/jid.3751.
IEAG, 2014. A world that counts: mobilising the data revolution for sustainable development. Independent Expert Advisory Group on a Data Revolution for Sustainable Development. United Nations, New York.
Jean, et al., 2016. Combining satellite imagery and machine learning to predict poverty. Science 353 (6301), 790–794. https://doi.org/10.1126/science.aaf7894.
Jerven, M., 2013. Poor Numbers: How we are Misled by African Development Statistics and what to Do About it. Cornell University Press. ISBN: 978-0-8014-7860-4 p.187.
Jerven, M., 2017. How much will a data revolution in development cost? Forum Dev. Stud. 44, 31–50. https://doi.org/10.1080/08039410.2016.1260050.
Johnston, Abreu, 2016. The asset debates: how (Not) to use asset indices to measure well-being and the middle class in Africa. Afr. Aff. 115 (460), 399–419.
Krennmair, P., Schmid, T., 2022. Flexible domain prediction using mixed effects random forests. J R Stat Soc Ser C 71, 1865–1894.
Kondmann, Zhu, 2020. Measuring changes in poverty with deep learning and satellite images. In: International Conference on Learning Representations (ICLR) 2020, Practical ML for Developing Countries Workshop, pp. 1–6. https://elib.dlr.de/137108/2/camera_ready.pdf.
Lamichhane, B.R., Isnan, M., Horanont, T., 2025. Exploring machine learning trends in poverty mapping: a review and meta-analysis. Sci. Remote Sens. 11, 100200.
Lavin, et al., 2022. Technology readiness levels for machine learning systems. Nat. Commun. 13 (6039). https://www.nature.com/articles/s41467-022-33128-9.
Lee, Braithwaite, 2022. High-resolution poverty maps in Sub-Saharan Africa. World Dev. 159, 106028. https://doi.org/10.1016/j.worlddev.2022.106028.
Lundberg, S.M., et al., 2020. From local explanations to global understanding with explainable AI for trees. Nat. Mach. Intell. 2, 56–67.
McCallum, et al., 2022. Estimating global economic well-being with unlit settlements. Nat. Commun. 13 (2459).
Marcinko, et al., 2022. Earth observation and geospatial data can predict the relative distribution of village level poverty in the Sundarban biosphere reserve, India. J. Environ. Manag. 313, 114950. https://doi.org/10.1016/j.jenvman.2022.114950.
Masaki, T., et al., 2022. Small area estimation of non-monetary poverty with geospatial data. Stat. J. IAOS 38, 1035–1051.
NASA, 2020. Technology Readiness Assessment: Best Practices Guide. National Aeronautics and Space Administration Office of the Chief Technologist [SP-20205003605].
Newhouse, D., 2023. Small area estimation of poverty and wealth using geospatial data: what have we learned so far? World Bank Policy Res. Work. Paper 10512.
Newhouse, D., 2024. Small area estimation of poverty and wealth using geospatial data: what have we learned so far? Calcutta Statist. Assoc. Bullet. 76 (1), 7–32.

Ngo, D.K., Christiaensen, L., 2019. The performance of a consumption augmented asset index in ranking households and identifying the poor. Rev. Income Wealth 65 (4), 804–833.

O'Connor, B., et al., 2020. Earth observation for SDG; compendium of Earth observation contributions to the SDG targets and indicators. Eur. Space Agency. https://eo4sdg.org/release-of-a-compendium-of-eo-contribution-to-the-sdgs/.

Okwi, et al., 2007. Spatial determinants of poverty in rural Kenya. Proc. Natl. Acad. Sci. USA 104 (43), 16769–16774.

Östberg, et al., 2018. Tracing improving livelihoods in rural Africa using local measures of wealth: a case study from central Tanzania, 1991–2016. Land 7 (2), 44.

Pu, C.J., et al., 2024. How poverty is measured impacts who gets classified as impoverished. Proc. Natl. Acad. Sci. USA 121 (7), e2316730121.

Sekara, et al., 2024. Opportunities, limitations, and challenges in using machine learning technologies for humanitarian work and development. Advances in Complex Systems 2440002. https://doi.org/10.1142/S0219525924400022.

Smythe, Blumenstock, 2022. Geographic microtargeting of social assistance with high-resolution poverty maps. Proc. Natl. Acad. Sci. USA 119 (32), e2120025119. https://doi.org/10.1073/pnas.2120025119.

Steele, J.E., et al., 2017. Mapping poverty using Mobile phone and satellite data. J. R. Soc. Interface 14 (127). https://doi.org/10.1098/rsif.2016.0690.

Steinert, et al., 2016. One size fits all? The validity of a composite poverty index across urban and rural households in South Africa. Soc. Indic. Res. 136, 51–72.

United Nations, 2015. Data for development: a needs assessment for SDG monitoring and statistical capacity development. https://sdgs.un.org/publications/data-development-needs-assessment-sdg-monitoring-and-statistical-capacity-development.

von, Eschenbach, 2021. Transparency and the black box problem: why we do not trust AI. Philoso. Tech. 34, 1607–1622.

Watmough, et al., 2013. Predicting socioecomomic conditions from satellite sensor data in rural developing countries: a case study using female literacy in Assam, India. Appl. Geogr. 44, 192–2000.

Watmough, et al., 2016. Understanding the evidence base for poverty–environment relationships using remotely sensed satellite data: an example from Assam, India. World Dev. 78, 188–203.

Watmough, et al., 2019. Socioecologically informed use of remote sensing data to predict rural household poverty. Proc. Natl. Acad. Sci. USA 116 (4). https://doi.org/10.1073/pnas.1812969116.

Watmough, G.R., et al., 2022. Using open-source data to construct 20 metre resolution maps of children's travel time to the nearest health facility. Sci. Data 9 (1).

Watmough, G.R., Marcinko, C.L.J., 2024. EO for poverty: developing metrics to support decision making using Earth observation. In: Comprehensive Remote Sensing: Volume 9 Remote Sensing Applications. ISBN: 9780443239496.

Yeh, C., et al., 2020. Using publicly available satellite imagery and deep learning to understand economic well-being in Africa. Nat. Comms 11 (1), 1–11.

Zheng, Z., et al., 2025. Dynamic, high-resolution Wealth Measurement in data-scarce Environments. World Bank Policy Research Working Paper, 11058.