Optimizing Sustainable Agriculture with EO-Derived Data for Soil Health Assessment

By

Leizel G. De la Cruz

Submitted to

Central European University

Department of Environmental Sciences and Policy

In partial fulfillment of the requirements for the degree of

MSc in Environmental Sciences and Policy

Supervisor: Prof. Viktor Lagutov

Vienna, Austria

AUTHOR'S DECLARATION

I, the undersigned, Leizel De la Cruz, candidate for the MSc degree in Environmental Sciences and Policy declare herewith that the present thesis is exclusively my own work, based on my research and only such external information as properly credited in notes and bibliography. I declare that no unidentified and illegitimate use was made of the work of others, and no part of the thesis infringes on any person's or institution's copyright. I also declare that no part of the thesis has been submitted in this form to any other institution of higher education for an academic degree.

Vienna, 28 June 2024

Leizel De la Cruz

Abstract

Healthy soil is critical for food production, but traditional methods to assess it are expensive and time-consuming. This study is conducted to evaluate the potential of open, free, and off-the-shelf Earth Observation (EO) derived data for assessing various soil health in agricultural lands. By assessing data quality and its effectiveness, this study can help in developing cost-effective, data-driven methods for sustainable agriculture.

A systematic literature review was done to identify the appropriate indicators for assessing soil health and promoting sustainable agriculture, as guided by the "Management Goal" approach. This led to the selection of nine (9) EO platforms offering data on soil health indicators. To understand the strengths and weaknesses of this EO data from a user's perspective, the study examined data quality factors: intrinsic characteristics, context, representation, and accessibility of the platforms. This analysis was conducted at both global and in Central Asia regional scales to determine the applicability of EO data across different geographic areas. The findings were then compiled into a table inventory, highlighting the capabilities and value of EO data for assessing soil health indicators.

Results of the research found a limited number of EO platforms offering free, high-quality soil data in both global scale and Central Asian regional case study. The free and readily available EO datasets for soil properties can be limited with the spatial resolution and temporal coverage that often presented a single year, even if historical data exists. Despite these limitations, the study presented a variety of EO data available that can be used depending on the user's ability in remote sensing and highlights the potential of EO data for soil health assessment in different geographic ranges. By understanding these limitations, users can choose the right platform, interpret data accurately, and ultimately make data-driven decisions for sustainable agriculture.

Overall, this research demonstrates the potential of freely available EO data in soil health assessment, ultimately promoting data-driven decision making for sustainable agricultural land management.

Keywords: Soil, Soil Health, Soil Health Indicators, Agricultural land, Earth Observation Platform, Earth Observation, EO Derived Data, Off-the-shelf data, readily available data, Data Quality assessment, user-centric perspective

Acknowledgements

The journey of this research, though swift, has been incredibly productive. This achievement wouldn't have been possible without the unwavering encouragement and understanding of those around me.

Central European University (CEU) and its professors, I am grateful for the opportunity to have honed my knowledge and skills in my area of interest at such University.

Dr. Viktor Lagutov, my deepest gratitude to my advisor. Your guidance, support, motivation, and patience were instrumental in shaping my research and especially valuable during the writing process of this thesis.

MESP and MESPOM 2023 classmates and friends, though our journey together was brief, it was filled with camaraderie and happiness. Thank you for your support and sense of community.

My parents and siblings, you are my unwavering source of inspiration, and your love has sustained me throughout this endeavor.

Finally, to God, all the glory.

Table of Contents

Contents

List of Tables

List of Figures

1. INTRODUCTION

1.1 Background

Soil, the foundation of life on earth, is a complicated system that intricately linked to the earth's systems and plays a crucial role in food security and sustainability (Lehmann et al. 2020). It has seen in recent years that there is an increase of interest in evaluating the quality and health of soil resources along the increasing awareness that soil is an important component of the biosphere for food production and maintenance of environmental quality (John W. Doran and Zeiss 2000). Alarmingly, assessments of agricultural land productivity reveal that nearly 40% globally suffers from human-caused degradation due to intensive agriculture, unsustainable land-use practices and worsening climate change (Tziolas et al. 2021). Cardoso et al. (2013) highlight that there is an intensifying concern from recent years regarding sustainable agricultural systems because the world is approaching the limits of agricultural expansion, making it more difficult to balance food production for a growing population with environmental responsibility. Driven by the growing pressure to conserve the environment, sustainable agricultural practices are gaining traction, as it relies on healthy soil for both economic and environmental benefits. Consequently, maintaining or restoring soil health is crucial for achieving sustainable yields.

Soil health, as defined by US Department of Agriculture, "is the continued capacity of soil to function as a vital living ecosystem that sustains plants, animals, and humans"(US Department of Agriculture 2024). While Kibblewhite et al. (2008) described it as a holistic measure reflecting how well soil responds to agricultural practices while ensuring continuous support for food production and other ecosystem services. The concept and terminology of soil health are still under development, but a unifying thought is its consideration of how soil health impact water quality, plant health, animal health within ecosystems (Lehmann et al. 2020) and has bigger effect on agricultural productivity. The concept of soil health or soil quality as used in the study of Kibblewhite, Ritz, and Swift (2008) has two categories: First is the reductionist approach, using independent measures of physical, chemical, and biological properties and integrated approach considering the overall impact on the ecosystem. This approach to soil health assessment aligns well with the adopted concept of soil quality or soil health (Cardoso et al. 2013). Hence, this study is focused on the selection and use of different biological, chemical, and physical indicators combined to create a holistic assessment of soil health in agricultural lands.

Systematic collection of soil properties is crucial for monitoring soil health, but traditional field surveys and lab analysis can be time-consuming and expensive. Remote sensing (RS) techniques, utilizing sensors on satellites, aircraft, or the ground, offer a promising solution by providing extensive and efficient capabilities to assess spatial and temporal variations in soil properties and conditions (Abdulraheem et al. 2023). However, utilizing Earth observation (EO) data can be complex and requires specific skills. Over the past decade, there has been a surge in development of ready-made or "off-the-shelf" EO-based datasets that are processed and analyzed by experts or scientists providing new monitoring opportunities for end users (Viloria 2023). These EO platforms offer comprehensive and consistent data on various physical, biological, and socioeconomic variables at global or regional scales. However, as Viloria (2023) highlights, understanding a product's characteristics and generation process is crucial before using any off-the-shelf geospatial product. Like any data source, existing EO-based data may have limitations related to applicability, local representativeness, or data consistency. Given the potential of EO derived data, this study intends to bridge the gap in soil health measurement by developing a comprehensive assessment of available remotely sensed data. This also focuses on recent, free, open, and user-friendly EO platforms for soil health monitoring, while evaluating the limitations and strengths of each EO platform to guide users for an informed selection.

To date, there are a small number of studies that have been conducted on evaluating EO platforms specifically for assessing agricultural soil health. This gap highlights the need for a comprehensive review and evaluation of freely available, off-the-shelf geospatial data in assessing soil health in agricultural lands. This study addresses this need by developing a comparative table that analyzes the characteristics of various EO platforms relevant to key soil health parameters.

1.2. Research Questions and Objectives

Generally, this research aims to evaluate the potential of open, free, and off-the-shelf Earth Observation (EO) derived data for assessing various soil health parameters in agricultural lands. By examining data quality and its effectiveness in specific soil indicators, this research can contribute to developing cost-effective and data-driven methods for sustainable agriculture land use.

Specifically, the following research questions and specific objectives are addressed:

1. Which specific soil health parameters can be potentially evaluated or estimated using freely available EO data??
    a. Identify the agricultural soil health parameters that can be effectively assessed using EO data.
    b. Determine the freely available EO data that are currently available for evaluation of soil health in agricultural lands.
2. How do the characteristics of freely available Earth Observation (EO) data products vary across different platforms in both global and Central Asia regional context?

a. Identify the data quality characteristics of EO Platforms based on the user's perspective.
b. Characterize and compare these freely available EO data products considering the user's needs.
c. Create inventory of freely available EO data products suitable for assessing parameters of soil health in agriculture.
d. Identify potential limitations and biases in the characteristics of freely available EO data products.
e. Develop a resource guide or comparison table for researchers and practitioners interested in utilizing freely available EO data for soil health condition.

1.3. Significance of the Research

Agricultural soil is globally under threat, and we urgently need a way to keep it healthy. As Parr et al. (1992) pointed out, there's a lack of central system to continuously monitor and assess the quality of soil health is tracked to be linked alongside air and water quality, as data for basis of international programs to combat global warming and protect biodiversity. Thus, this research is timely to determine the global progress with the available soil health data.

This research has significant implications for sustainable agriculture and food security. This approach democratizes soil health assessment, making it accessible to the public, particularly farmers and agricultural institutions. By enabling cost-effective assessment, even small-scale organizations in developing countries can be informed on the status of agricultural soil without relying on expensive traditional methods. This data empowers farmers to adjust management practices for increased crop yields, reduced reliance on synthetic fertilizers and pesticides, and improved carbon sequestration in healthy soils, mitigating climate change. Additionally, EO derived data enables broader spatial coverage (regional or national) for soil health assessment as compared to the time-consuming and labor-intensive localized field measurements and traditional methods.

Furthermore, this research contributes to the standardization of geospatial data through the development of a resource guide or comparison table for EO platforms. This facilitates data sharing among researchers and practitioners, fostering further advancements in soil monitoring methods.

The evaluation of existing EO platforms provides valuable insights on the strengths and limitations of using EO derived data for soil health monitoring. This information can guide future development efforts to better serve the agricultural sector. Finally, this research offers recommendations for policymakers, researchers, and practitioners on utilizing EO data to promote sustainable agricultural practices and improve crop productivity. This allows for the implementation of science-based policies and projects for a more sustainable future.

1.4. Scope and Limitations

This research focuses on the evaluation of thirty (30) web platforms containing Earth Observation (EO) data. Nine (9) of these platforms have global extent and “off the shelf data” specifically targetting soil properties relevant to soil health assessment. Also, the criteria used in selection of indicators for soil health was based on the "management goal approach" as explained in the methodology. This ensures the chosen indicators directly relate to practical goals for improving soil health. Likewise, the evaluation of the EO datasets employed a "user's perspective" approach, prioritizing factors that would be most relevant to potential users, such as data accessibility, usability, and ease of use.

It is relevant to clarify that this research does not endorse or promote specific EO web platforms. The primary objective is to demonstrate the potential of EO data derived from various platforms for assessing soil health. This highlights the broader applicability of EO technology in agriculture.

1.5. Thesis Outline

This thesis is structured into seven chapters. Chapter 1 introduces the research, outlining its objectives, research questions, and the significance of the study. Limitations and scope are also defined in this chapter. Chapter 2 explores the related literature, discussing the concept of soil health, approaches for selecting soil health indicators, and various soil properties. It then explores the use of remote sensing for acquiring Earth Observation (EO) data, the benefits of analysis-ready data and "off-the-shelf" datasets, and the different categories of EO platforms, highlighting their unique functionalities. Chapter 3 presents the research framework and methodology. It details the process of selecting soil health indicators and the data quality assessment approach, based on the user's perspective adopted by Wand and Wang (1996). This chapter also explains how available datasets for soil health indicators will be evaluated based on different data quality components. Chapter 4 presents the results and discussions, followed by Chapter 5 which details a case study applying the method in a regional scale, which is the Central Asia and the chosen EO platforms. Chapter 6 discusses the research outcomes in relation to the objectives and drawn conclusions, and Chapter 7 gives recommendations for future research and improvements.

II. LITERATURE REVIEW

2.1. Overview of Agricultural Soil Health

Agriculture is a crucial component in ensuring global food security, but its productivity is now at risked due to various environmental problems driven by human-caused climate change and human activities, including soil degradation (Wuepper et al. 2021). The rapid decline in soil health threatens our ability to grow food – not only losing soil fertility and productivity puts the world's food supply at threat but also harms ecosystems and hinders sustainable development (Kogut 2023). Currently, around one-third of the world's soil is already classified with moderate to severe degradation, to which 40% of degraded soil is found to be in Africa and areas experiencing poverty and food insecurity (FAO 2015). The critical link between food security and soil health needs urgent strategic action (Kogut 2023) especially that there is a rapid increase of human growth. Thus, increased demand for food along the increasing per capita calorific consumption and changed dietary patterns puts significant pressure on soils globally (Kopittke et al. 2019). Initially, the growth in food production primarily relied on expanding agricultural land which has accelerated over the past three centuries. However, starting from the mid-20th century, the shift in strategy coined "agricultural intensification" was implemented to meet the needs of the rapidly expanding global population (Kopittke et al. 2019). This refers to enhancing crop yield production per unit of land area, rather than further expanding agricultural areas. Therefore, soil plays a vital role in food production by providing essential nutrients, supporting crop growth, and influencing crop yields. If we understand soil as a system and how its parts interact, we can better analyze how well it functions as a whole (John W. Doran and Zeiss 2000).

2.1.1. Soil Health Concept

A better understanding of the soil health concept, its system and components is needed to have an integrated approach in addressing soil degradation and provide sustainable agriculture. In this research, the concept of soil health and soil quality is used equally and interchangeably, adopting the definition of USDA-NRCS "Soil health is defined as the continued capacity of soil to function as a vital living ecosystem that sustains plants, animals, and humans."

The soil health concept started even during ancient times. The Roman philosophers are aware of the importance of healthy soil to agricultural yield. Their writings focused on farm management, stressing the value of soil fertility. Unlike modern methods that can artificially boost yields, Roman agriculture relied heavily on working with the natural fertility of the land. They understand that there are different soil types with limitations on their ability to grow certain crops, recommending practices that maximized efficiency within those boundaries. They also had methods for evaluating soil health even used some of the same indicators we use at the present (J.W. Doran, Sarrantonio, and Liebig 1996).

In the 1800s, the rapid population growth sparked fears of food shortage so chemists studied the connection between soil and plants. In the study of Doran (1996), he mentioned that Wallerius, Thaer, and von Wullfen believes that healthy soil needed high amount of humus, the organic matter in soil, for better production. Their research showed that farming causes depletion of humus and soil exhaustion negatively affecting the crop yields in Europe (J.W. Doran, Sarrantonio, and Liebig 1996).

However, this humus concept for soil fertility was argued by Liebig saying that plants needed a balance of various elements, not just humus. Liebig believed farmers could achieve this balance by adding artificial fertilizers, mimicking a natural nutrient cycle and creating a sustainable system (J.W. Doran, Sarrantonio, and Liebig 1996). This marks the new paradigm in agriculture and by the 1900s, farming had transformed into a large-scale production industry. As mentioned in the study of Doran (2011), the writing of Hafner (2003) and Smil (2011) refers to this as the Green Revolution when the use of synthetic fertilizers and pesticides increased the yields on wheat and rice harvests by four times.

During the late 20th and early 21st centuries, concerns arises on the negative impacts of chemicals on soil health, which led to the development of the concepts of "soil quality" and "soil health" (J.W. Doran, Sarrantonio, and Liebig 1996) as a way to assess the overall soil status.

At the 1991 Conference on Assessment and Monitoring of Soil Quality, Dr. David White argued that a good definition of soil health should consider the following four key points: 1. Recognize that soil is a living ecosystem; 2. Should encompass all the ecosystem services that soil plays; 3. Compare soil's condition to its natural potential within its specific climate, landscape, and vegetation; and 4. Should allow us to track changes in soil health over time (J.W. Doran, Sarrantonio, and Liebig 1996). Given these factors, soil health can be defined as: “the continued capacity of soil to function as a vital living system, within ecosystem and land-use boundaries, to sustain biological productivity, maintain the quality of air and water environments, and promote plant, animal, and human health” (J.W. Doran, Sarrantonio, and Liebig 1996). Likewise, soil health is now described by most agencies, such as the US Department of Agriculture to be more of the functionality of soil to provide for the sustenance of life of plants, animals and humans (US Department of Agriculture 2024) .

There are several other related concepts exist when referring to soil health, it includes soil fertility, soil quality, and soil security which also focuses on the functions of soil in the agriculture and environment (Toor et al. 2021). Soil quality is defined as a broader concept that considers soil's ability to function for agriculture as well as its impact on the surrounding environment, including water quality and the health of plants and animals (Toor et al. 2021).

Kibblewhite et al. (2008) defines soil health in consideration of agricultural sustainability. It means the soil must be able to produce enough high-quality food and fiber to meet our needs, while also continuing to provide other ecosystems services essential for life. This definition is in parallel with John W. Doran and Zeiss (2000) with emphasis on treating the soil as a living system and a part of ecosystems hence, soil health is “the capacity of soil to function as a vital living system, within ecosystem and land-use boundaries, to sustain plant and animal productivity, maintain or enhance water and air quality, and promote plant and animal health”.

Lehmann et al. (2020) argued that soil health goes beyond its function to serve humans and encompasses the bigger picture of overall sustainability, including the health of the entire planet. Soil quality, on the other hand, typically focuses on the services a healthy soil provides for human needs within an ecosystem.

Some definitions are based on the characteristics of a healthy soil, generally, it should have the following: (i) good soil tilth, (ii) sufficient depth of roots to access water and nutrients, (iii) adequate supply (but not excess) of nutrients, (iv) optimal pH, (v) low population of pathogens and insect pests, (vi) high and diverse population of beneficial organisms for organic matter decomposition and nutrient cycling and soil structure maintenance, (vii) low weed pressure, (viii) free of harmful chemicals and toxins, and (ix) resistant to degradation or resilient soils (Toor et al. 2021).This characteristics are dependent to the physical, chemical, and biological properties of the soils and should have common ground as shown in Figure 1. As mentioned in Toor et al.’s (2021) writings, Hurni et al. (2015) and Larson and Pierce (1991) - soil quality is a unique property of soils intrinsically linked reflecting how productive the soil is and how human activities impact these soil characteristics (physical, chemical, and biological).

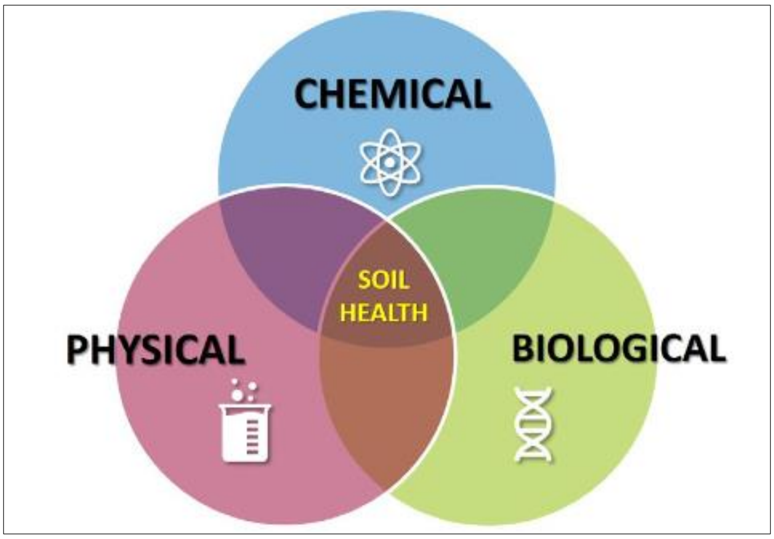

Figure 1. Soil Health diagram of interaction among the chemical, biological and physical properties (Toor et al. 2021).

The definition, understanding, and practical application of soil health are still under development. The terms “soil health” and “soil quality” are used interchangeably in scientific literatures and publications like Haberern (1992), Doran and Parkin (1994), Larson and Pierce, 1991) as mentioned by (Toor et al. 2021). Some prefer to use the term "soil health" because it emphasizes soil as a living, dynamic organism that functions holistically and not just sand, silt, and clay while others prefer "soil quality" with its focus on measurable properties, physical, chemical, and biological characteristics (J.W. Doran, Sarrantonio, and Liebig 1996).

Understanding soil health is tricky because it is a complex system. The physical, chemical and biological characteristics and processes plays a role and influence each other that’s why the concepts "soil quality" and "soil health" is adopted to capture the big-picture view (Rinot et al. 2019) using different soil health parameters.

2.1.2. Soil Health Indicators for Sustainable Agriculture

Adapting the definition of soil health as the continued capacity of the soil to function as a living ecosystem important to sustain life on earth underscores that soil is much more than just a growing medium, rather a complex and dynamic system full of life. In this perspective, it highlights the need for assessment of soil health using appropriate and fitting indicators, especially for agricultural purposes. The goal of devising soil health indicators is to track changes in soil health by comparing them to a natural baseline, drawing insights from various disciplines to represent overall soil conditions(Toor et al. 2021).

In a broader context, soil health can be monitored using the evaluation of the physical, chemical, or biological properties of the soil (Lehmann et al. 2020). These categories sometimes overlap because many soil properties are the result of a combination of these processes. Aside from that, a good soil health indicator needs to have several qualities according to Lehmann et al. (2020). It should be relevant to the health of the soil ecosystem and the ecosystem services it provides and needs to be sensitive with changes noticeably when soil health improves or declines (John W. Doran and Zeiss 2000). Ideally, the indicator should also be cost-effective and measurable (Lehmann et al. 2020;Rinot et al. 2019). Soil health indicators should also be informative for a wide range of users, including farmers, agricultural managers, and policymakers for sound decision making (John W. Doran and Zeiss 2000). In the study of Toor et al. (2021) and Rinot et al. (2019), they added that indicators should also consider the accuracy and repeatability of data, and that it can be interpreted on its own or combined with other indicators for a comprehensive assessment.

During the early soil assessment, the major indicators used to assess soil health are categorized into three groups: physical, chemical, and biological (Toor et al. 2021 cited by Moebius-Clune et al., 2016). Figure 2 illustrates the simplified overview of selected soil health indicators, where the physical indicator focuses on water availability, the chemical indicator measures the nutrient availability and the biological indicators that reflects the complex life within the soil and the processes that are crucial in achieving healthy soil.

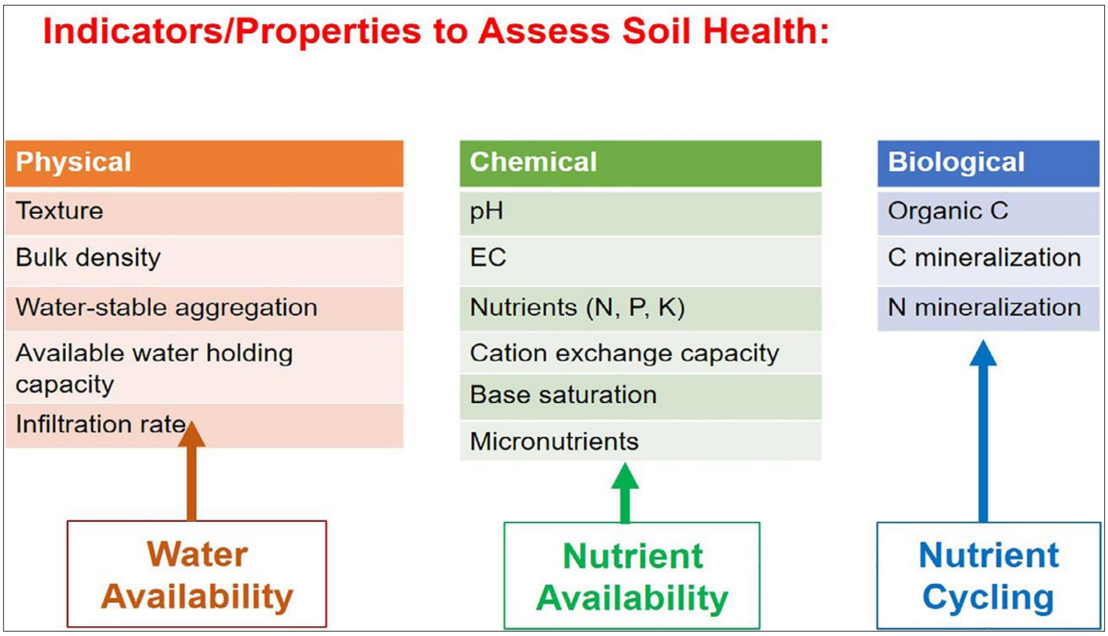

Figure 2.The overview of soil health indicators using 3 major soil attributes. (Source:Toor et al. 2021).

A. *Physical Attributes*

Physical indicators of soil health are often simple, quick, and affordable to measure. The physical measures are related with hydrological processes impacting water and air balance and stability (Cardoso et al. 2013). These indicators include the texture and bulk density (Cardoso et al. 2013; Douglas L. Karlen, Ditzler, and Andrews 2003; Toor et al. 2021), porosity and aggregate stability (Cardoso et al. 2013; D.L. Karlen, Eash, and Unger 1992), soil moisture and water-stable aggregates (Toor et al. 2021). According to Cardoso et al. (2013), citing Dexter (2004), several signs that soil has poor health from a physical perspective is when it exhibits characteristics such as slow water absorption, increased surface water flow, weak structural integrity, inadequate aeration, limited root growth, and challenges for agricultural machinery operation.

Another physical indicator of soil quality as suggested by is the soil tilth, defined "as the physical condition of a soil described by bulk density, porosity, structure, roughness, and aggregate characteristics, as related to several phenomena: transport of water, nutrients, heat and air, the stimulation of microbial and micro- fauna populations and processes; and the impedance to seedling emergence and root penetration". They proposed that soil tilth is a strong indicator of overall soil health since it reflects how management practices impact physical, chemical, and biological aspects (D.L. Karlen, Eash, and Unger 1992).

B. Chemical Attributes

The chemical attributes of soil are directly related to its ability to supply essential nutrients that plants need to grow while also influencing its role in environmental protection by sequestering potentially toxic or damaging chemicals (Cardoso et al. 2013). Through the conduct of analyzing key chemical aspects of soil indicators that directly impacts plant yields, it allows for quick improvements through treatments like liming and fertilization (Cardoso et al. 2013), ultimately benefiting soil health and agricultural sustainability. Hence, understanding the chemical properties of soil explains is relevant to know the interaction among soil water, nutrients, physical attributes, and levels of soil contaminants and that is tolerable for plants (APEX Publishers, n.d.).

According to Parikh (2012), there are seventeen (17) essential elements needed by plants to survive and any shortages of each one can limit the crop yield. The primary macronutrients are Nitrogen, Phosphorus, and Potassium that are used the most and are missing in farmland. Secondary macronutrients like Calcium, Magnesium, and Sulfur are also needed but in lesser amounts. There are also micronutrients like Iron, Manganese, Zinc, Copper, Boron, Molybdenum, Chlorine, and Nickel that are needed in exceedingly small amounts and if in excess can actually be toxic to plants. In addition to the chemical nutrients as soil health indicators, it is also essential to monitor soil pH, cation exchange capacity (CEC), and organic matter content as it affects the soil capacity for production of crop yields (Cardoso et al. 2013).

Authors in the study of Cardoso et al. (2013) concluded that the most important chemical parameters to be assessed were soil pH, level of organic matter, Phosphorus, Potassium, Copper, Manganese, and Zinc. Chemical soil health indicator’s main concern is the potential for nutrient availability, like water availability for the physical soil attributes, in agricultural system.

C. Biological Attributes

For a long time, soil health has mainly focused on the physical and chemical properties but with the current focus on soil as a system- the measurement of soil’s biological aspect is possible (Cardoso et al. 2013). Numerous biological processes play crucial roles in key soil functions, including organic matter decomposition, mineralization and recycling nutrients, fixing nitrogen, cleansing pollutants, preserving soil structure, and naturally controlling plant pests and parasites (Parikh 2012b). Some of the more frequent biological indicators that have been proposed are microbial biomass, Nitrogen mineralization and Earthworm density (Bünemann et al. 2018). He also included the root development and potential rooting depth, though not assessed frequently. In addition, Cardoso et al. (2013) proposed measurement of soil respiration as it is extensively used as bioindicator in forestry and agricultural soil health.

Relying solely on specific biochemical properties to assess soil health can be inaccurate leading to conflicting results across studies because these properties can fluctuate greatly depending on factors like climate, season, location, and even the soil's history (Cardoso et al. 2013).

D. *Faunal Indicators*

Recently, studies emphasized the value of faunal components as indicators of soil health. This refers to the species living in the soil medium and plays a crucial role in soil structure, decomposition of organic matter and interrelationship with other microorganisms (Cardoso et al. 2013). As cited by Decaëns et al. (2004) and Eggleton et al. (2005) in Cardoso et al. (2013) study, the diversity, abundance, biomass and density of soil fauna are useful indicators for assessing the impacts of changes on terrestrial ecosystems because soil fauna are closely linked to the physical, chemical, and microbiological properties of soil. They stated that the evaluation of the taxonomic diversity of soil fauna at the order, class or key species level is a straightforward, simple, and cost-effective method for assessing soil health. Monitoring the changes in the diversity and frequency of soil organisms provides valuable inferences into the overall health and functioning of the soil ecosystem.

In addition to the above-mentioned characteristics, it is also important to consider other indicators of soil health, Parr et al. (1992) proposed to include crop productivity (grain yield or biomass), plant vitality, and the quality of surface and groundwater. Particularly in some situations, water quality can have a significant impact on the chemical, physical, and biological aspects of soil and there are certain indicators of soil quality that may show sensitivity to changes, while others could exhibit more subtle changes. Researchers also argue to consider the climate land use changes and farming practices as they affect the physical, chemical, and biological variables of soils.

However, Toor et al. (2019) argued that even though this detailed information about soils helps in understanding soil behavior, farmers may not require such complexity to successfully cultivate crops. From farmer’s perspective, water and nutrient availability are the most critical needs since these essential resources are influenced by various factors, measuring these variables can serve as a proxy for their likelihood of being available to support healthy crop growth. To effectively monitor agricultural soil health, this research is focused on choosing practical and informative chemical, physical, and biological soil health indicators that can be readily used to monitor agricultural soil health. Taking into perspective that soil quality or health is one of the key factors to attain agricultural sustainability, Andrewsi and Carroll (2024) also highlight that soil quality assessment is a vital input for quantifying the sustainability of agricultural ecosystems (as shown in Figure 3). In their view, research that explores soil processes and mechanisms provides crucial information for soil quality assessment and advances the field of soil science altogether.

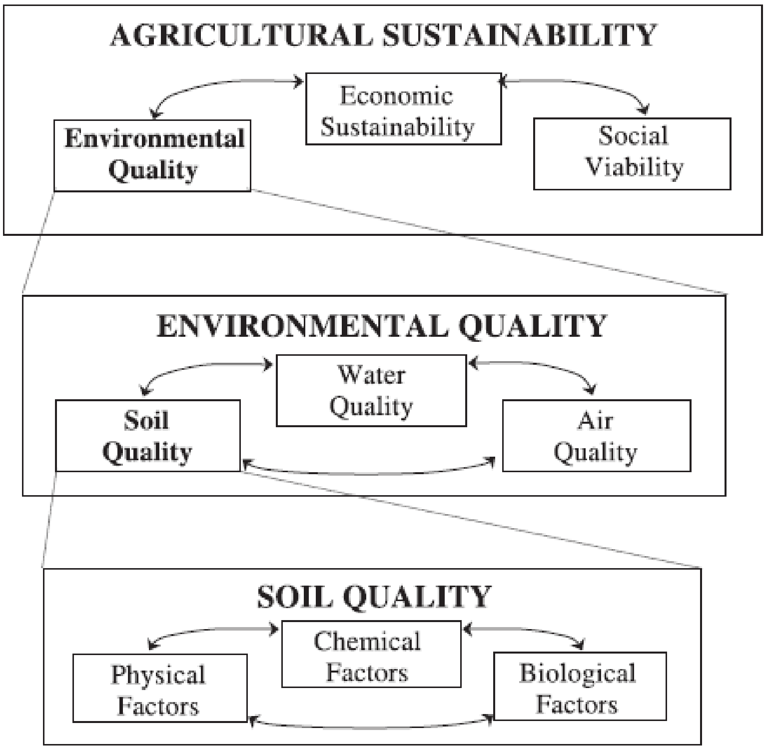

Figure 3.The relationship of soil health or quality to agricultural sustainability. (Source: Karlen et al. 2003).

2.1.3. Approaches on Assessing Soil Health

A plethora of soil quality assessment and monitoring tools have become available since the 1990s. Here, we give an overview of the main developments in different countries, before addressing aspects of soil quality indicators in more depth in section

Given the complex nature of soil systems, there is a lot of discussion on the best and appropriate methods to use for soil health assessment (Kibblewhite, Ritz, and Swift 2008). In the onset of 1990, there has been a surge in the development of tools and approaches for assessing and monitoring the soil health. Using Figure 4, the general overview of the development used by different countries on the assessment of soil quality indicators is illustrated.

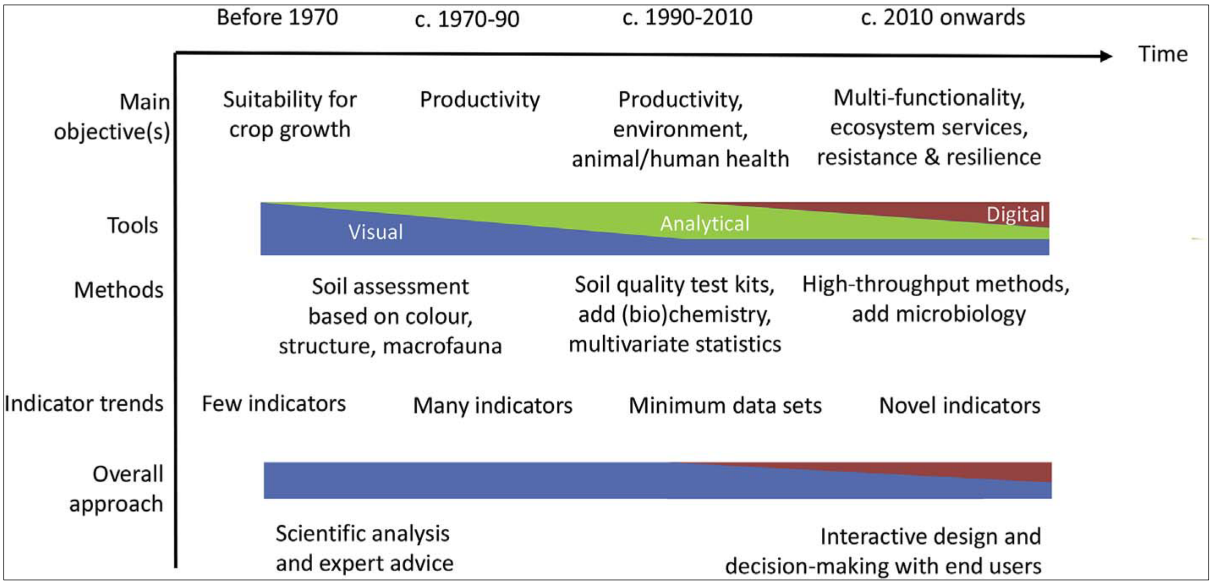

Figure 4. Soil Quality assessment development in terms of objectives, tools, methods, and overall approach. (Source: Bünemann et al. 2018).

Like to any ecosystem management approach, the initial stage in assessing soil health is the setting of objectives to inform management decisions, serve as an educational tool, or contribute to a monitoring program that is applicable on site (Andrewsi and Carroll 2024). Figure 4 also reflects the evolution of how the soil is perceived as a medium into a part of an ecosystem that is living and as a system itself. Likewise, successful adoption of any soil quality assessment approach hinges on identifying the target users from the very beginning and involving them in the process. This project exemplifies how stakeholder workshops were instrumental in defining the assessment methods, including the selection of soil functions and indicators (Douglas L Karlen et al. 2008).

During the 1990s, one of the first approaches to assess soil health is the use of scorecards (Douglas L Karlen et al. 2008, cited in Harris et al., 1996; Romig et al., 1996; Shepherd, 2000; Shepherd et al., 2000). The soil quality scorecards are developed by the NRCS-Soil Quality Institute (US Department of Agriculture 2024) having the primary purpose to raise awareness about soil health and provide a way for people without a scientific background to track their efforts towards improving soil quality. These scorecards, along with guidelines for creating them, were some of the first tools available.

Other approaches mentioned by Karlen et al. (2008) in their study include more practical understanding of soil health, done by digging soil pits for direct observation and using soil quality test kits developed by Doran et al. (1996) and Sarantonnio et al. (1996). These kits allowed users to measure various properties like water infiltration, density, and respiration, providing insights into how soil physical, chemical, and biological properties change spatially and temporally (Douglas L Karlen et al. 2008 cited in Doran et al., 1996; Liebig et al., 1996; USDA-NRCS, 1999).

There are two (2) prominent soil quality assessment methods designed for field plots developed in the USA, namely: the Soil Management Assessment Framework (SMAF) and the Cornell Soil Health Test.

A. The Soil Management Assessment Framework (SMAF) stands out for its flexibility in choosing indicators and based on the specific ecosystem service or management goal (Douglas L Karlen et al. 2008). With SMAF users can select a set of indicators from a pool of eighty-one options using defined selection rules. However, this can limit comparability between different sites if users deviate significantly from the recommended minimum dataset. The interpretation of individual indicator values relies on scoring curves, and an overall soil quality index can be calculated by adding the scores (Bünemann et al. 2018).

B. The Cornell Soil Health Test was launched in 2007 and has the following objectives: to educate about soil health, give guidance to farmers and land managers in better management practices, provide monitoring data for the NRCS and indirectly increase land values (Douglas L Karlen et al. 2008; Bünemann et al. 2018). Like the SMAF, this also measures the biological, chemical, and physical indicators and interpreted using score curves. According to Karlen (2008), the Cornell Soil Health Test is sensitive to management practices and aligns well with the concept of critical soil functions outlined by Doran and Parkin (1994). It is also consistent, reproducible, easy to use for sample collection, and cost-effective for soil testing laboratory to implement (Idowu et al. 2008).

   Scientists at Cornell University initially considered a wider range of 39 factors for a comprehensive soil health test but to make it more practical and affordable, they streamlined the test to include 12-13 key measurements on “physical (e.g. aggregate stability, penetration resistance, available water capacity) chemical (pH, P, K, micronutrients, organic matter content) and biological (soil proteins, soil respiration, soil pathogens) parameters” (Rinot et al. 2019).

Both the SMAF and the Cornell Soil Health Test approaches share a key principle: they assess "dynamic soil quality" that focuses on the current state and condition of the soil, reflecting the impact of management practices, rather than the soil's inherent qualities that are less influenced by human activity (Douglas L Karlen et al. 2008). Since these two methods often require laboratory analysis, it made them less accessible for some farmers. For these users, simpler educational methods are more suitable and allow for immediate results and easier communication between farmers and soils experts. Hence, the birth of a new method that relies on readily observable, qualitative indicators that farmers can assess in the field without needing a laboratory.

The Visual Soil Assessment (VSA) is a popular example of such farmer-friendly methods, with variations implemented around the world. As per Bünemann et al. (2018), most VSA methods focus primarily on soil structure, sometimes even linking it to crop productivity (Abdollahi et al., 2015; Mueller et al., 2013). For instance, the Spade-based methods are quicker than those requiring soil profiles, making them more practical for farmers (Bünemann et al. 2018 cited in Boizard et al., 2005). While a clear interpretation is a major benefit of VSA, it is important to remember that visual assessment alone cannot capture the full picture of the biological and chemical processes that affect the soil's ability to provide essential services.

As proposed in the study of Bünemann et al. (2018), combining the VSA and the dynamic soil quality methods can provide a more comprehensive picture of soil health promoting the wider use of visual assessment to be valuable for analyzing yield potential and developing effective land management program (McKenzie et al., 2015).

The next decade of soil health monitoring is poised for exciting technological advancements, particularly in remote sensing techniques (Lehmann et al. 2020). This method offers a fast and cost-effective way to measure various soil properties, including chemical composition, physical characteristics, and even biological activity (Bünemann et al. 2018). Remote sensing can be used either in the field or in a laboratory setting and companies are starting to offer commercial spectroscopy-based analyses. Combining this technique with laboratory-based methods and direct measurements is perceived to be beneficial globally. Another good thing with this approach is that remote sensing does not only map spatial variations in soil properties but also use mathematical models to assess management practices based on the impact on soil functions (Lehmann et al. 2020). While remote sensing offers a rapid and large-scale approach, it should be complemented by smaller scale for targeted assessments. With all the technological advancements nowadays, the development of sensors specifically for soil health monitoring is possible. These rapid screening and in-situ/remote monitoring technologies have the potential to significantly improve our understanding of soil health.

2.2. EO Data for assessing Agricultural Soil Health

As a strategy to cope with the increasing food demands, agricultural intensification is practiced that significantly changes the soil properties and causes modifications on the physical, chemical, and biological characteristics of soil, harming its ability to provide its functions. This soil degradation is one of the major problems we are facing, making protection and restoration of healthy soils a key part of the United Nation's Sustainable Development Goals (Safanelli et al. 2021). To determine the status, frameworks built on data collected by Earth observation satellites are used to monitor agricultural soil and its health.

Remote sensing is a powerful tool used to produce Earth Observation Data (EOD) in a short time interval and generate useful information for users (Diaz-Gonzalez et al. 2022). Generally, it works by gathering information from a distance - through satellite or airborne sensors or radars. The data captured is processed and analyzed to extract significant understandings about the earth (Abdulraheem et al. 2023). These techniques have revolutionized soil measurements by employing satellite, airborne, and ground-based methods, assessing soil erosion, determining areas with high moisture content, and mapping of soil nutrients is made possible (Abdulraheem et al. 2023). Remote sensing also offers a valuable toolbox for assessing various soil quality indicators in agro-industrial systems that encompasses both chemical aspects, like the levels of nitrogen, potassium, and phosphorus which is directly linked to the effect of fertilizer use, and biological aspects such as organic matter content (Diaz-Gonzalez et al. 2022). Furthermore, remote sensing aids in detecting soil contamination and evaluating overall fertility. The applications extend across various aspects of soil science, with near-infrared reflectance estimating vegetation cover and biomass production, and thermal infrared measurements gauging biomass production (Abdulraheem et al. 2023). As proven by Abdulraheem et al. (2023) in their study, EO data can provide an effective tool to explore the physical, chemical, and biological information of soils in a global context. Safanelli et al. (2021) agreed that EO data can monitor agricultural soils across large geographical areas, may it be in farmlot, national or continental coverage. RS has proven its usefulness particularly in areas where traditional soil sampling methods are limited or impractical and its ability capture data across various scales which depends on the resolution of imageries (Abdulraheem et al. 2023).

EO derived data is becoming a valuable source of information for overcoming the challenges in data collection (Andries et al. 2022) especially in the soil health which is an indicator used in monitoring Sustainable Development Goals.

2.2.1 RS methods in acquiring EO data for Soil Health

In agriculture, Remote Sensing (RS) is increasingly utilized to gather data on soil characteristics, soil moisture, erosion, and crop health across large areas to which the method uses various sensors and platforms to collect data from a distance, enabling large-scale, accurate, fast, and non-destructive analysis of soil characteristics (Abdulraheem et al. 2023). This section discusses different methods of analyzing soil using remote sensing techniques like spectral reflectance analysis, thermal infrared imaging, and radar remote sensing.

A. *Spectral Reflectance Analysis*

According to Abdulraheem et al. (2023), the use of spectral reflectance analysis is one of the most widely used RE techniques that involves measurement of the light reflectance of electromagnetic radiation in different wavelengths. Spectral Reflectance as described is the ratio of the amount of sunlight bounce back (upwelling) from a surface compared to the amount of sunlight that hits the surface (down-welling). This way of measuring light reflection is like human eyes adjusting to brightness. It makes it easier to understand and compare the reflectance of different things, rather than just focusing on the raw amount of light they reflect (Huete, n.d.). And since spectral patterns and values are different from each material, soil properties have unique spectral signatures too that allows the identification and quantification using remote sensing techniques. However, to fully benefit from remote sensing, understanding both how soil reflects light (spectral principles), and the limitation of this technique is needed. Most of the useful data for soil analysis are from the visible-near infrared (VNIR) and short-wave infrared (SWIR) parts of the light spectrum because the analysis how soil reflects light not how soil emits light, which is measured in the thermal infrared region (Ben-Dor, Irons, and Epema 1999). Even though thermal infrared readings can also reveal information about soil composition, VNIR (0.4-1.1μm) and SWIR (1.1-2.5μm) spectral regions are used due to its focus on reflectance (Ben-Dor, Irons, and Epema 1999). Taking the example given by Abdulraheem et al. (2023), the visible range (400–700 nm) can be used to detect organic matter and iron oxide minerals in the soil. Likewise, the Near infrared reflectance (700–1300 nm) can show sensitivity to soil moisture content and clay minerology while Shortwave infrared reflectance (1300–2500 nm) can be used to estimate soil organic carbon content and calcite or gypsum identification in soil.

B. *Thermal Infrared Imaging*

Although the spectral reflectance withing the VNIR-SWIR region shows its capability for assessing soil properties, the Thermal Infrared (TIR) wavelength withing the range of (8–14 μm) has the potential to extensive capabilities on soil (Eisele et al. 2012). This method particularly measures the heat radiation from an object (Abdulraheem et al. 2023) and transforms it into temperature without physical contact with the object (Khanal, Fulton, and Shearer 2017). Khanal et al. (2017) generalizes that objects with above absolute zero temperature emit radiation, "and the amount of radiation is a function of the emissivity of the surface and the surface temperature". Thus, the higher the temperature of an object, the higher intensity of the radiation it emits.

Thermal remote sensing has become a powerful tool for sustainable agriculture, providing information for crops and soil monitoring. With the use of this method, it is possible to estimate soil moisture levels and identify areas with different water availability, which is crucial for irrigation management, ensuring crops receive the right amount of water. It also proves useful in detecting variations in soil compaction. Thus, compacted soils have less space for air and water (porosity), hindering water infiltration and causing the surface to have higher temperature. By analyzing thermal patterns, experts can determine areas of compaction, which can negatively impact plant growth (Abdulraheem et al. 2023). Early detection of these soil health status allows for corrective measures and agricultural interventions to be taken, promoting crop productivity and sustainable agriculture.

C. *Radar Remote Sensing*

In comparison with the other two methods that use optical sensors, Radar Remote Sensing uses electromagnetic energy that bounces back (backscattered) from an object to extract physical properties of the ground (Prakash and Kumar 2022). The major advantage of this method is that it works anytime of the day and in any weather conditions.

Described in the study of Prakash et al. (2022), radar imaging works through transmission of microwave pulses towards an object at a certain frequency. The intensity and phase of the returning signal depend on the object's characteristics (e.g. texture and wetness), then the radar antenna picks up this reflected energy.

With the capacity of the microwaves to penetrate the soil surface, it can measure subsurface properties of the soil like moisture content and texture (Abdulraheem et al. 2023). It shows that wet soil has higher dielectric constant making the signal weak, making it possible to get information for soil moisture and with the differences of signal interaction to soil, the percentage of sand, silt and clay can be determined (Abdulraheem et al. 2023).

With the advancement of technologies, the global approach used to analyze the data obtained from remote sensing involves the utilization of machine learning and currently emerging technologies like Artificial Intelligence (AI). Some of the machine learning algorithms that are commonly used at present are the Random Forests (RF), Support Vector Machine (SVM) and Neural Networks (NN) for the development of predictive models based on available pedological soil data and other environmental auxiliary variables (Awais et al. 2023 cited by Naimi et al. 2022). However, it is still a challenge to produce global harmonized soil data given the differences in methods and standards used, mostly there are areas who do not have any data on soils at all. To augment this challenge, some geostatistical methods are being used like kriging and co-kriging to interpolate and extrapolate different soil properties in the unsampled locations. This also improves the spatial representation of data (Awais et al. 2023).

With the emergence of Artificial Intelligence models combined with remote sensing, it can significantly improve the analysis of soil data that will lead to a more efficient soil analysis, smarter decision-making and sustainable agricultural practices. But still, the AI models need to be refined to improve its importance in soil analysis application (Awais et al. 2023). Hence, further research is needed to advance RS methods for measuring soil properties that can be done through improvement of calibration and validation, sensor combinations and applications of ML and AI (Abdulraheem et al. 2023).

Theis rapid advancement of remote sensing technologies for soil measurements has led to the development of complex methods that may be challenging for non-technical individuals to utilize. To address this issue, there is a growing need for user-friendly platforms that provide readily available datasets that can be easily accessed and used by users which this research addresses by finding available EO data platforms and determining its potentials, strengths, and limitations in soil health applications.

2.2.2 Use of Open EO Data for Monitoring Soil Health

The availability of Earth Observation Data (EOD) is becoming extremely important for monitoring vast areas of cropland and soil health (Safanelli et al. 2021cited by Picoli et al.). Unlike the traditional methods that involve testing soil samples at individual locations, EOD can analyze soil in different geographical scales at once that is also open access and free of charge of the data to end-users. This was made possible due to the increasing EO data gathered for over forty years supported by open-access policies for data distribution as well as advancements in computing capabilities (Safanelli et al. 2021).

Free and open access to Earth Observation (EO) data is fundamental for monitoring soil health. Pioneering this movement, NASA and the U.S. Geological Survey (USGS) made Landsat imagery freely available since 2008; followed by ESA in 2010 by releasing data from missions like ERS, Envisat, and Meteosat (Andries et al. 2022). The Moderate Resolution Imaging Spectroradiometer (MODIS) with 20 years image collection is also made available for public use (Safanelli et al. 2021). Today, this trend continues with all data from the operational Sentinel missions being freely accessible through the Copernicus Programme (Andries et al. 2022). Safanelli et al. (2021) added that with the recent availability of cloud-based processing interfaces with the use of machine learning, it is easier to analyze massive amounts of EO data providing detailed information on soil health. These new EOD capabilities can significantly reduce the need for expensive and time-consuming traditional soil sampling in gathering data.

Open data is freely available for anyone to use, re-use, modify and share that allows widespread use, collaboration and innovation in fields like agriculture (Andries et al. 2022). Meanwhile, using commercial satellites typically requires users to pay and adhere to a license agreement but even it is expensive, it can provide substantial information on the users required resolution of the area and observation dates that an open EO data cannot directly address.

Despite making the EO data free and accessible to the public, translating this data into usable information remains a challenging task that requires substantial skills or expertise on handling the data and resources. Satellite data comes in various formats and requires specialized software and knowledge to process. Understanding the intricacies of different wavelengths, image corrections, and calibration techniques can be intimidating for non-experts. As a solution, the concept of having a pre-processed standardized stacks of satellite imagery coupled with adequate analytical tools and algorithms is a development in EO technology ("Master Report - ARD for Africa - May 2022.Pdf," n.d.). With this approach, even the novice users can use the data. Another one is the production of usable derived dataset or what I will refer to this study as the “off-the-shelf EO Derived data”. This data is products of processing made by experts for a certain type of users and theme. Generally, the EO derived data products are the “Ready-made solutions data” that don’t need any processing by end-users and can easily be utilized for interpretation.

A. *Analysis-Ready-Data*

Analysis Ready Data (ARD) refers to time-series stacks of overhead imagery that undergone pre-processing and already prepared for users to analyze directly, eliminating the need for users to spend hours on tedious pre-processing (Holmes 2018). For non-users of satellite imagery, a simple image requires significant effort to prepare for meaningful analysis. With the ARD, this complexity is already cut making satellite data accessible and user-friendly. The pre-processed format of ARD makes it ideal for training machine learning algorithms to identify patterns and relationships in the data (Dwyer et al. 2018). Some identified platforms where to download ARD are Google Earth Engine (GEE), NASA through AppEEARS, GLAD and CEOS-ARD for Sentinel.

In 2017, the USGS EROS Center made a significant contribution to Earth observation when the Landsat ARD was launched for a vast area covering the contiguous United States, Alaska, and Hawaii providing access to Landsat collection in format that can be easily used for monitoring and assessment of landscape changes (Dwyer et al. 2018). While there are established standards for producing Analysis-Ready Data (ARD) as outlined by Karagiannis (2023), these standards ensure several key characteristics: i. In tiled format, ii. Data is geo-registered, iii. Data undergone Top of atmosphere (TOA) and atmospherically corrected, iv. Has a defined and consistent projection, v. Includes spatially explicit quality assessment information (QA bands), vi. Non-target features (clouds) and poor-quality observations are flagged, vii. Data is accompanied by comprehensive metadata, providing vital information for further processing while ensuring traceability of the data's origin, viii. Geometrically and radiometrically consistent, and ix. Processed in a community endorsed method promoting widespread acceptance and use.

B. Off-the-shelf EO Derived Data

Existing or off-the-shelf EO-based products are datasets that transform the raw satellite data into a user-friendly data with valuable insights already extracted by the experts or scientists (Viloria 2023). As Aravind (2021) described it, as a “usable, useful product” which users can use the data without delving into the complexities of the underlying EO methods and technology. This type of product delivers clear, actionable information that can be used for decision-making requiring minimal configuration. The production of this off-the-shelf EO derived data product exemplifies the kind of user-centric approach that's essential to make EO data accessible to a wider audience. Planet is one of the pioneer organizations who produced this kind of data with their innovative Planetary Variables product (Aravind 2021).

According to Viloria (2023), the emergence of this satellite-based data products has given rise to new monitoring user-friendly tools that provide comprehensive and consistent data on physical, biological, and socioeconomic aspects at global or regional scales. Examples include soil moisture, forest health, biomass, rainfall patterns, and land use. While off-the-shelf geospatial products offer exciting possibilities, end users should understand the product’s characteristics, its metadata and method of generation to critically assess the underlying assumptions and limitations of the product (Viloria 2023). In doing so, users can determine if the off-the-shelf products are applicable for the study, fit for the purpose and compatible with the analysis.

2.2.3 Categories of EO Platforms

Along with the development in Earth Observation data, is the development of different web portals and services that cater to the need of different users, allowing for the discovery, access, and utilization of EOD products. These web portals are categorized based on the study of Andries et al. (2022) that is categorized in three main categories:

A. EO Data Portals

The eoPortal is a user-friendly gateway to access vast amounts of information and resources of different satellite mission ("Satellite Missions Catalogue - eoPortal," n.d.). It allows the user to discover historical and operational missions in the database and offers an exhaustive list of products of different missions.

As per Andries et al. (2022), around 25 data portals provide free access to medium-resolution satellite imagery like Sentinel and Landsat data. These portals also offer historical data from Very High-Resolution (VHR) missions. Some of these EO data portals are the USGS, GloVIS, NASA, ESA, ASTER, Copernicus, Sentinel Hub, and a lot more. Most of these are free and open access EO data that is produced by government or funded by the taxpayers and international space agencies and are provided for the public good. For users with specific needs, numerous commercial suppliers exist like Planet Labs, Airbus, Maxar Technologies, Pixera, etc., which offers EO satellite data and derived products tailored to client requirements, often via subscription plans or pay-per-use models.

B. *EO Processing, Visualization, and Cloud computing platforms*

This type of platform allows the users to do the EO processing, visualize, and do the analysis using the cloud computing tools (Andries et al. 2022) which give more flexibility for EO data users. However, this platform requires that users have knowledge of the back end' functionality of EO satellite images and most of the time requires expertise on programming languages. Some examples of this type of platform are Google Earth Engine that is run by Google, the Amazon Web Services Cloud (AWS) and Datacube.

EO processing and visualization platforms transform raw satellite data into valuable information products that undergone multi-stage processing like application of data correction, product generation and data integration where EO data is often combined with other georeferenced data sets like socio-demographic, economic, and environmental information (Andries et al. 2022).

C. *EO Derived Thematic products and Services.*

This type of EO platform takes a user-friendly approach to navigate around the enormous EO services categorized by theme. This taxonomy structure acts like a well-organized filing system grouped by subject (e.g., soil, agriculture, environment, disaster management). Similarly, the platform organizes EO services based on their thematic focus, making it easier for users to find what they need. Usually, these platforms provide ARD (Andries et al. 2022) and Off-the-shelf EO derived data that can be accessed with application programming interfaces (API) and web services. Some of the examples are ISRIC that provides information on the soils of the world and WAPOR by the Food and Agriculture Organization (FAO) to monitor Water Productivity through Open access of Remotely sensed derived data.

III. RESEARCH METHODOLOGY

3.1. Research Framework

Healthy soil is essential for sustainable agriculture, and EO data plays a vital role in monitoring and achieving sustainable land use practices. The concept of soil health in agricultural areas and its relationship with Earth Observation derived data is illustrated in Figure 5.

Following a "Management Goal" approach, a top-down design is employed to identify key soil indicators and environmental factors influencing soil health. Remote sensing allows for monitoring these identified indicators, serving as a valuable indicator for agricultural productivity. With the advanced analysis techniques, combined with ancillary or field data, it enables prediction and spatial mapping of soil health indicators. These maps and derived spatial data are often publicly free and available through the development of various EO web platforms.

However, these platforms may utilize different methodologies and datasets. This research addresses this gap by analyzing the differences and similarities among EO platform datasets using ten established metrics data quality based on user's perspective. This framework also identifies existing EO web platforms designed specifically for soil data.

By understanding the processes involved in generating EO-derived datasets, as outlined in this framework, readers can gain valuable insights into how this data can be used for assessing soil health indicators and promoting sustainable agricultural practices.

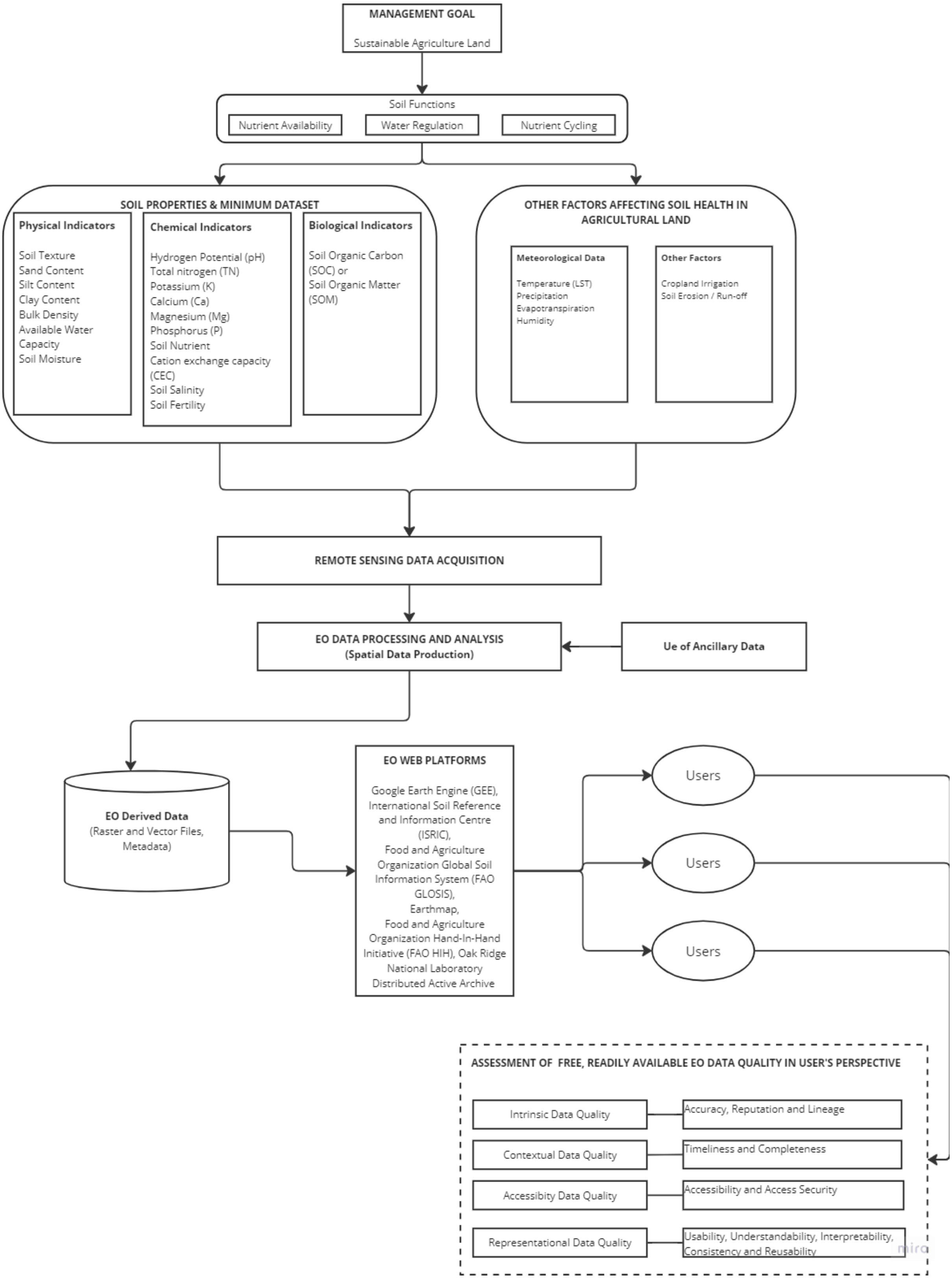

Figure 5. The research framework on assessment of EO derived data for soil health monitoring.

3.2. Methodology

The research methodology is based on systematic literature review of the appropriate soil health indicators for sustainable agricultural land using different earth observation derived data from different web platforms. This approach involves three key steps: Determination of Soil Health indicators, Identification of EO Platforms and suitable Data Quality elements based on user's perspective and lastly, the assessment of each soil health indicator dataset and platform. This method is used to the case study of Central Asia and evaluate its effectiveness in regional scale.

3.2.1. Determination of Soil Health indicators

First is the determination of the soil health indicators for agricultural lands. This is done through identifying the management goals or objectives and context of the area of interest first. In this research the management goal is achieving Sustainable Agricultural Land Use. Adapting the framework for selection of indicators for the minimum data set of (Douglas L. Karlen, Ditzler, and Andrews 2003) as shown in Figure 6. This framework involves identifying the key soil functions needed to achieve the management goal. Finally, based on those soil functions, the soil indicators are identified, which encompasses the essential physical, chemical, and biological characteristics of the soil. The soil functions are identified to achieve the management goal and lastly is to identify the indicators which are the different soil physical, chemical, and biological characteristics.

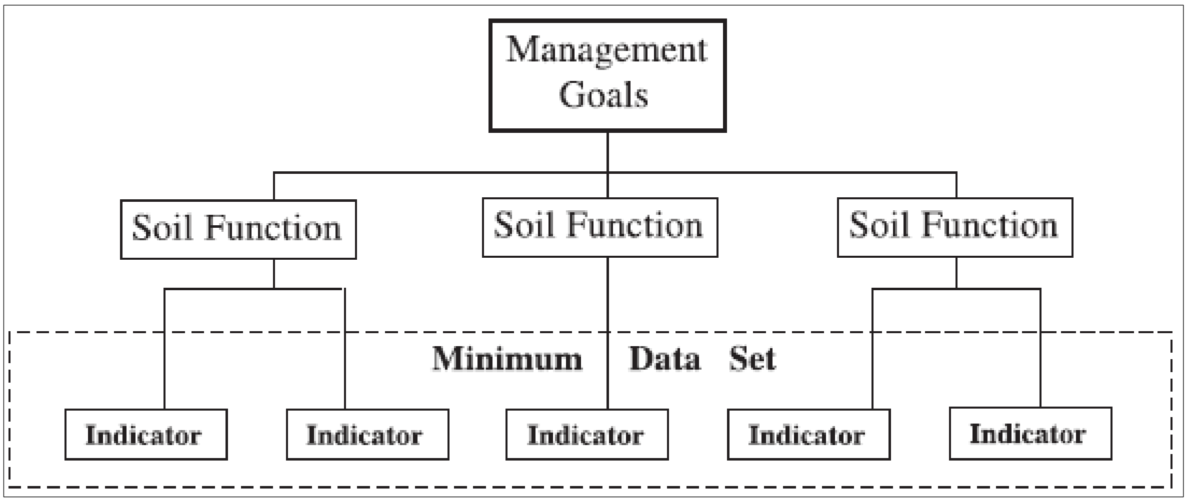

Figure 6. Framework for selecting indicators for a soil health indicator. (Source: Karlen et al. 2003)

3.2.2. Identification of EO Platforms and suitable Data Quality elements

The second step is the conduct of a systematic review of the known Earth Observation Platforms that the researcher did during her practicum at Central European University Environmental Systems Laboratory. There are twenty-seven (27) EO platforms (see Annex 1) that are identified and assessed. The review of each EO web platform is based on understanding the description of the Earth Observation metadata profile standard (Gasperi et al. 2016) and the basic set of data quality dimensions (Cappiello, Francalanci, and Pernici 2004) considering data consumers perspective. From a user's perspective, good data quality is all about whether the information is helpful for their specific needs. Borrowing from existing research on data quality, "good quality data" is defined as information that is usable by the people who need it (Wand and Wang 1996). This data quality is grouped into various aspects: Intrinsic Quality, Accessibility, Representability and Contextual Quality. These aspects are the building blocks that when combined, make the data high quality overall. These components encompass how information about the real world is presented in a product, including what data is shown, when it's shown, where it appears, and how it's formatted (Sweta and Bijker 2013). Table 1 shows the structure used in assessing Earth Observation web platforms.

Table 1. Template for Quality Assessment of Available EO web platforms.

No.	Data Quality Component	Field Name	Field Description
1	Intrinsic Data Quality	Platform Type	Type of web platform.
2		Platform Name	Name of the Web Application.
3		Platform Website	Refers to the web address or Uniform Resource Locator (URL)- the internet address.
4		Platform Description	Summary of the platform highlighting the usage, theme and key points.
5		Platform Owner	The individual or organization responsible for production and managing the platform.
6		Platform Partners	The individual or organization who give their services or collaborates to develop or enhance the platform.
7		EO Data Available	Refers to the available dataset in the platform (e.g. satellite images, soil map, etc.).
8		Use of Ancillary Data for validation	Refers whetherthe data generation process utilized additional data sources for validation purposes, such as ground-truth measurements and historical surveys.
9	Contextual Data Quality	Used Sensors/Radars with Return Period	Instrument (Sensor/Radar) name and the time it takes for a sensor to revisit the same location.
10		Temporal Coverage	Time period the data is collected and the frequency of data updates.
11		Spatial Resolution	Scale of the area covered (e.g. city level, national, regional, global).
12		Spatial Coverage	The geographical area covered by the data.
13	Accessibility Data Quality	Data Access	Indicates if the EO data available is free to downoad or requires purchase, or needs specific permissions
14		Cloud processing	Specifies whether data processing is performed in the cloud or on the user's device.
15	Representational Data Quality	Ease of Use (low, Moderate, High)	Refers to the platform's ease of use and navigation for data production.
16		Output Type	Available data format of the platform's product.
17		Language/s Used	Available language/s user can use in the platform
18		Platform Availability	Indicates if the platform is free to use, requires purchase, or needs specific permissions
19		Availability of Process Documentation	Refers whether there are written instruction manuals or process documentation explaining how the EO data are produced.

No.	Data Quality Component	Field Name	Field Description
20		Availability of User Guidelines	Refers whether there are user manuals, Tutorials, help documentation and communities that is readily accessible for users.

This assessment considers data quality elements and data quality overview elements as outlined by Wand and Wang (1996); Wang and Strong (1996); Joksic and Bajat (2004) which is a more comprehensive approach on providing general, non-quantitative information of data. Figure 7 shows the framework used by Wand and Wang (1996) on evaluating data quality.

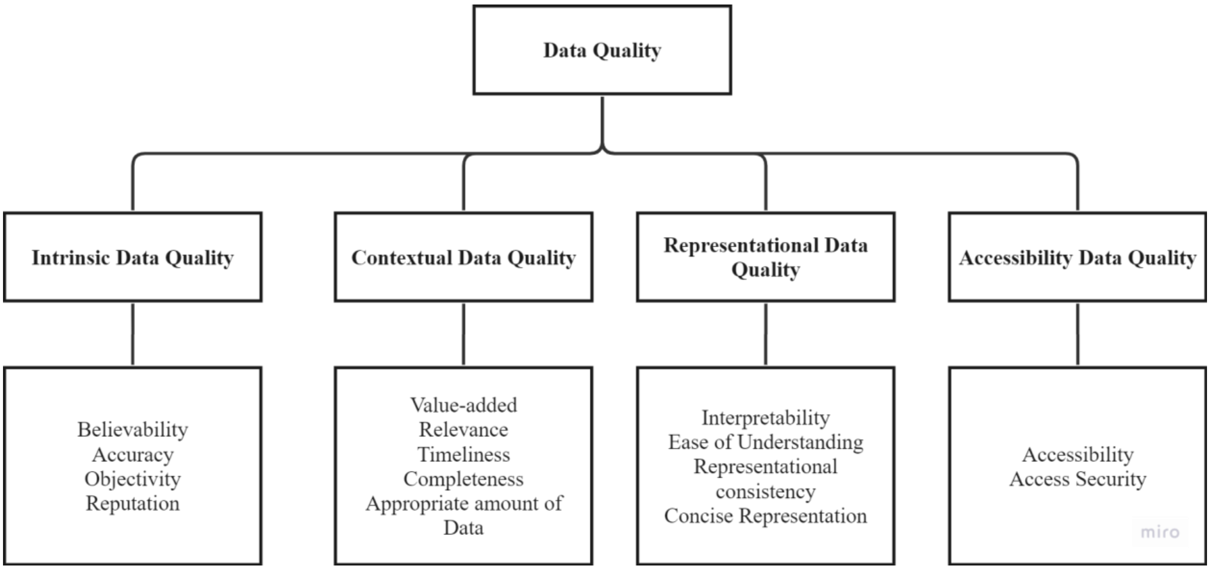

Figure 7. A Conceptual Framework of Data Quality. (Source: Wand and Wang 1996)

The Intrinsic Data quality goes beyond just accuracy and objectivity for users, it also incorporates the believability and reputation of the producer of the information (Wand and Wang 1996). For data users, one of the criteria to determine if data is of high quality is about who are the producers and the methods used in broader aspects. Intrinsic DQ focuses on the inherent quality of the data itself, independent of how it is used. On the other hand, the idea of Contextual Data Quality is beyond the completeness or timeliness, data quality depends on how it is utilized by users (Wand and Wang 1996). High quality data should consider the specific needs of the user and the context. While users have diverse needs, the information system should allow users to tailor data in the way they needed it to provide quality and better results. Representational Data Quality (DQ) is concerned with two aspects: the format of the data and the meaning of the data. This means it should be easy to understand and use, avoiding unnecessary complexity or ambiguity. Wand and Wang (1996) research show that users consider data accessibility (how easily users can get to the information) to be a key part of data quality and not a separate from information quality. In today's online world, users need data to be readily available through computer systems, making accessibility a crucial aspect of data quality. Both the Representational and Accessibility Data Quality is important in determining data quality as it provides the element how users can utilize the platform or system. The data should be presented clearly and understandable, and above all can easily be accessed by users (accessible systems) for good data quality. In other words, the system itself plays a crucial role.

After this evaluation, the web platforms that offer off-the-shelf data on global soil properties are selected and subjected to further assessment of spatial data products relating to soil health.

3.2.3. Assessment of Soil health indicator dataset and platform

The categories of data quality by Wang and Strong (1996), mentioned that the data quality dimensions that most researchers used are: accuracy, completeness, consistency, accessibility, interpretability, and timeliness which often includes related aspects like currency and volatility. These dimensions are the core aspects that make data high quality. Analyzing the Data Quality Components and the elements or dimensions that are based from the study of Sweta and Bijker (2013); Wand and Wang (1996); Wang and Strong (1996), there are ten (10) parameters used for this research. Table 2 shows the data elements and their corresponding components and the descriptors of each element. This served as a guide on evaluating the quality of EO data and the platforms.

Table 2. The Data Quality elements and the interpretation in this research.

Data Quality Elements	Description
Lineage	Refers to the description of data sources, methods used for dataset production and all data used in the process. This component contains all data important for both data sources and update of process.
Reputation	Refers to the producers of the data. Users judge the quality of data based on who produced the data itself.
Completeness	Refers to the ability of the platform to represent all possible real-world scenarios the data reflects, including optional data and how different data points interact with each other.
Timeliness	Refers to the delay between real-world changes and their reflection in the data platform and as to when the data was last updated.
Accuracy	Refers to the way the dataset reflects the real world-situation. In the case of EO data, this may refer to the accuracy on locations and object representation.
Accessibility	Refers to the straightforward way to get the data that users need. It means the dataset is easy to find in the platform, easy to retrieve and download and access it without technical difficulties.
Usability	Refers to the consideration that it can be used and processed easily. For example, the data format should be readily usable with commonly available software tools or platforms
Consistency	Refers to the adherence with the international standards of data quality as set by EO consortiums or ISO.
Interpretability	Refers to the ability or easiness of the data to be interpreted by its users. For example, the data are not represented in a foreign language.
Understandability	Refers to the way data is understood by users. Ensures that the data meaning is understood correctly even when exchanged between different databases and users.
Reusability	Refers to the clarity of documentation that allows users to understand and utilize the data effectively for different purposes.

Table 3 outlines the evaluation framework used in assessment of the data quality of soil indicators found in six (6) web platforms that meet the criteria: free, ready-to-use, global EO data. These platforms include Google Earth Engine (GEE), International Soil Reference and Information Centre (ISRIC), Food and Agriculture Organization Global Soil Information System (FAO GLOSIS), Earthmap, Food and Agriculture Organization Hand-In-Hand Initiative (FAO HIH), and the Oak Ridge National Laboratory Distributed Active Archive Center (ORNL DAAC).

Table 3. Template for Assessing Data Quality of Earth Observation (EO) Derived Products.

No	Data Quality Component	Data Quality Elements	Field Name	Field Description
1	Intrinsic Data Quality	Lineage	Name	Name of Dataset
2		Reputation	Dataset Provider	The individual or organization that provides dataset to user.
3		Reputation	Point of contact	The individual or organization that develops and maintain the dataset.
4		Accuracy	Methods of Processing EO Data	Type of method used in production of dataset.
5		Accuracy	Use of Ancillary Data	Refers whether the dataset produced additional data sources for validation purposes, such as ground-truth measurements and historical surveys.
6	Contextual Data Quality	Completeness	Sensor Used	Instrument (Sensor/Radar) name used in production of dataset.
7		Completeness	Spatial Resolution	Scale of the area covered (e.g. city level, national, regional, global).
8		Completeness	Spatial Coverage	The geographical area covered by the data.
9		Completeness and Timeliness	Temporal Resolution	Time period the data is collected.
10		Completeness and Timeliness	Temporal Coverage	The historical data used and the frequency of data updates.
11		Timeliness	Year of EO launch	The year of EO platform become available and functional
12	Accessibilty Data Quality	Accessibility	Data Access	Refer if the dataset available is free to downoad or requires purchase, or needs specific permissions

No	Data Quality Component	Data Quality Elements	Field Name	Field Description
13		Accessibility	Product Format	Available data format of the dataset for download.
14		Access Security	Security	Refers whether the platform support secure storage and use of personal data.
15	Representational Data Quality	Usability	Processing Tools	Refers whether the platform offer built-in tools for data analysis like image overlay, time series analysis, graphs.
16		Usability	User Technical Knowledge Requirement	Defines the minimum technical knowledge of user required to access and download data varies
17		Usability	Programming Languages	Refers whether it supports custom analysis using the user's preferred programming language (e.g., Python, R).
18		Usability	Cloud Processing	Specifies whether data processing is performed in the cloud or on the user's device.
19		Usability	Adaptability	Refers whether the platform is designed to handle high volume of data, including multiple downloads from global to community scale.
20		Interpretability	User Interface (UI)	Refers whether the interface is user-friendly and easy to navigate
21		Interpretability	Visualization Tools	Refers whether the platform offer built-in tools for visualization like graphs, charts.
22		Consistency	Open Standards	Refers whether it adhere to open data standards for interoperability with other platforms.

No	Data Quality Component	Data Quality Elements	Field Name	Field Description
23		Understandability and Reusability	Product Manual	Refers whether there are written instruction manuals or process documentation explaining how the dataset is produced.
24		Understandability and Reusability	User Support	Refers whether there are user manuals, Tutorials, help documentation and communities that is readily accessible for users.
25		Understandability and Reusability	Community	Refers whether there is an active user community for knowledge sharing and troubleshooting.

3.2.4. Assessment of EO Derived Data in Central Asia

After the review of relevant EO derived dataset for each identified soil health indicator on a global scale, the same assessment method is used to determine the status of Central Asia Region in terms of free and readily available soil health data. This analysis aimed to assess two key aspects: determine the EO derived data gaps in regional scale and determine the richness and quality of datasets in Central Asia Region. This regional scale study will help users to understand and consider the characteristics of satellite derived data on assessing soil health in a regional scale for sustainable agriculture land use.

IV. RESULTS AND DISCUSSION

4.1. Soil Health Indicators for Sustainable Agriculture Land Use

4.1.1. Basis for Selection of Soil Health Indicators

The definition of soil health as the continued capacity of the soil to function as a vital living ecosystem directly connects to the challenges of assessing it across different landscapes due to the diversity of agricultural systems, soil types and differences in climatic zones. This makes it difficult to develop a concrete set of soil indicators in assessing soil health conditions. As agreed by (Kibblewhite, Ritz, and Swift 2008), soil provides variety of benefits and services, and a good assessment method of approach should consider these various functions.

Adapted from Douglas L. Karlen, Ditzler, and Andrews (2003) and Andrewsi and Carroll (2024), the soil indicators to be used for the assessment should reflect the chosen management goals or objectives, which in this research is mainly on attainment of sustainable agriculture land use. Soil health is fundamentally linked to agriculture land use since it provides critical soil functions like nutrient cycling; water management; supporting plant growth and development; biodiversity and carbon sequestration (Lehmann et al. 2020). This research only focuses on the three (3) soil functions that are also selected by Food and Agriculture Organization (FAO) that directly affects sustainable agriculture land use, namely: a) Support of plant growth and development or nutrient availability, b) Nutrient cycling, and c) Water regulation.

The major indicators used to assess these three soil functions are classified in three (3) groups based on the soil properties: Physical, Chemical and Biological characteristics (Toor et al. 2021). Aside from that, an indicator should also be sensitive to changes in climate and management, easy to understand for users, practical enough to be measured and most likely a part of the existing soil database (J.W. Doran, Sarrantonio, and Liebig 1996). For this research, the Harmonized World Soil Database and FAO Soil Database are used as basis for selection of soil health indicator.

In addition to the soil properties, extrinsic factors such as climatic factors and management of site are also included as indicators because it affects the soil functions and management goals (Bünemann et al. 2018). Climatic factors like precipitation, temperature, humidity, and evapotranspiration data are available for longer range of time and can be used as a predictive input to understand the condition of soils through time. Having all these criteria, Figure 8 illustrates the selected soil health indicators for sustainable agricultural land use.

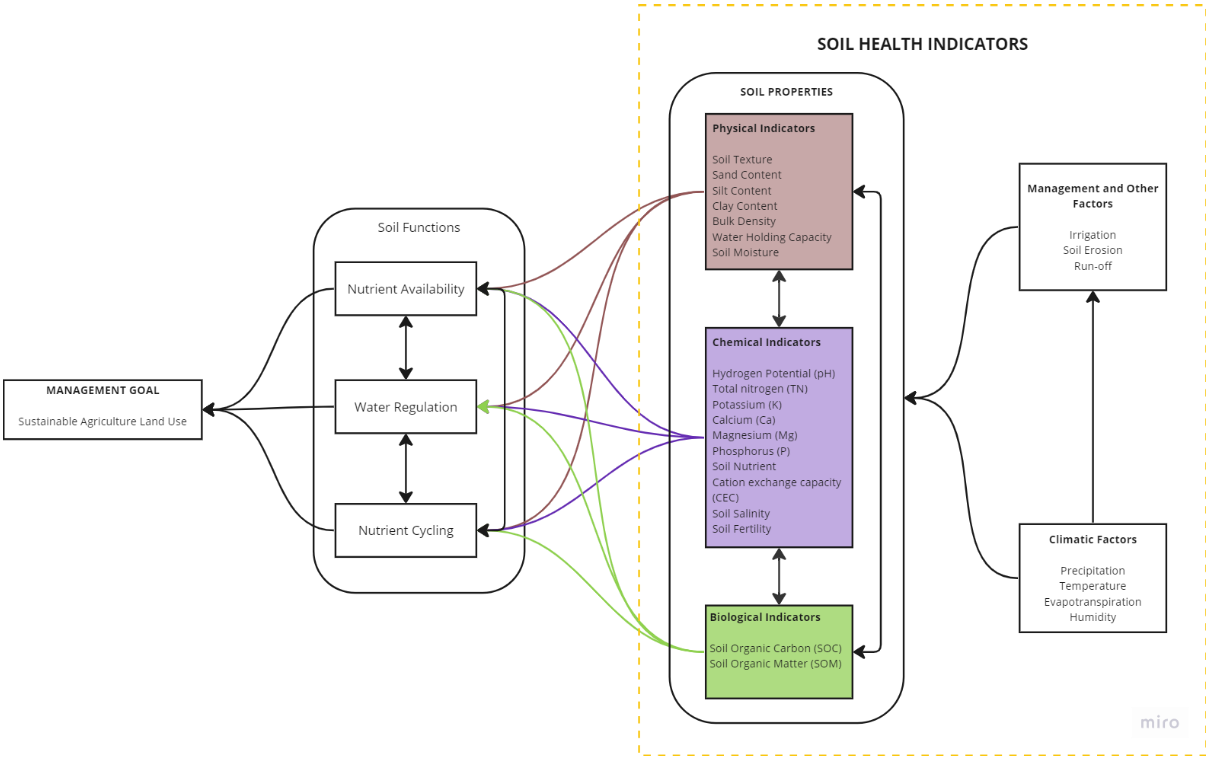

Figure 8.. The Framework of Selected Indicators using the Management Goal Approach.

4.1.2. Selected Soil Health Indicators for Agricultural Soil Health

Table 4 shows the functionality of each selected soil health indicator in terms of the categories. The importance of each indicator affects the efficiency and functionality of other indicators.

Physical properties like texture, bulk density, and structure influence the water and aeration availability while the chemical properties like pH, nutrient levels, and exchange capacity are the determinants of nutrient availability for plants and productivity. Chemical and physical properties were the focus of agricultural studies back then since it is directly related to crop yield and production. However, looking at the soil as a system, biological aspects like organic matter content and microbial activity are considered as indicators and highly impacting nutrient cycling. As Abdulraheem et al. (2023) mentioned, there are variety of biological processes that are responsible for important soil functions, like decomposition and detoxification of pollutants. While chemical and physical indicators have been common, the complexity of biological factors has made them less explored in assessing overall soil health (Toor et al. 2021). The advancement of microbiology field has made this all possible to measure and gather information.

Table 4. Soil Health Indicators and their importance as featured in the selected journals used.

SH Categories	SH Indicators	Importance	Article
Physical Properties	Soil Texture	Affects soil behavior retention capacity for nutrients and water.	Bünemann, Abdulraheem, Toor, Doran, Idowu, Sere, FAO
	Sand content	Affects aeration, water movement, conduction of heat, plant root growth, soil strength, hydraulic conductivity, permeability coefficients and resistance to erosion.	Diaz-Gonzales, Tziolas, Chen, Hengl, FAO
	Silt content	Indicator of soil structure. Affects nutrient retention, aeration, water movement, conduction of heat, plant root growth and resistance to erosion. Silt content affects cohesion and liquefaction potential.	Diaz-Gonzales, Tziolas, Chen, Hengl, FAO
	Clay content	Affects aeration,water movement, conduction of heat, plant root growth and resistance to erosion. Clay content stores	Diaz-Gonzales, Tziolas, Chen, Hengl, FAO

SH Categories	SH Indicators	Importance	Article
		organic matter particulates limiting microbial access.	
	Bulk Density (BD)	Potential for leaching, productivity, and erosivity. Limiting microbial activity, biochemical processes, and nutrient availability. It also impedes root penetration, leading to shallow plant roots, poor growth, and reduced crop yields	Diaz-Gonzales, Tziolas, Chen, Hengl, Lehmann, Bünemann, Toor, Andrews, Karlen, Sere, FAO
	Soil Moisture	Affects the rate of decomposition, structure and stability of the soil. Regulates evapotranspiration, which affects the temperature and humidity. Essential on microbial activities.	Ancy Stephen, Abdulraheem, Doran, Andrews, Sere, FAO
	Water Holding Capacity	Related to water retention, transport, and erosivity. Affects available water. Affected by soil bulk density, texture, and organic matter.	Chen, Bünemann, Toor, Andrews, Idowu, FAO
Chemical Properties	Hydrogen Potential (pH)	Affects availability of nutrients by influencing solubility. Influences microbial populations. Defines thresholds for these activities in the soil. High pH increases decomposition, releasing phosphorus, manganese, and calcium.	Diaz-Gonzales,Tziolas, Chen, Hengl, Lehmann, Bünemann, Toor, Doran, Andrews, Idowu, Karlen, Sere
	Electrical Conductivity	Affected by soil degradation. Define plant and microbial activity thresholds. Can be used for soil moisture and salinity indicator.	Doran, Karlen, Goldshleger, Toor, Chen
	Carbon	Influence on aggregate stability which affects water infiltration, retention, and runoff. Affects fertility, organic material that supports biological functions.	Karlen, Idowu, Toor, Abdulraheem, Lehmann, Hengl, FAO, Diaz-Gonzales

SH Categories	SH Indicators	Importance	Article
	Nitrogen (N)	Affects plant growth, soil structure and water penetration. increase soil biological. Influence plant available nutrients, activity, controls erosion and prevents surface sealing. Indicator of productivity and environmental quality indicators;	Diaz-Gonzales, Xu, Lehmann, Bünemann, Abdulraheem, Toor, Andrews, Sere
	Potassium (K)	Correlated with provision of nutrients for plants, productivity, and different soil indicators. Affects retention of other chemical elements or compounds and leaching.	Diaz-Gonzales, Bünemann, Abdulraheem, Toor, Idowu, Karlen
	Calcium (Ca)	Contribute to soil fertility by helping clay structure, aeration, and soil structure.	Diaz-Gonzales,Tziolas,Bünemann, Toor, Karlen
	Magnesium (Mg)	Correlated with the decomposition of organic matter. Affects soil productivity.	Diaz-Gonzales,Bünemann, Toor, Karlen
	Phosphorus (P)	Correlated with provision of nutrients for plants, productivity, and different soil indicators.	Diaz-Gonzales, Xu, Bünemann, Toor, Andrews, Idowu
	Soil Nutrient	Affects physical properties influencing rooting depth and volume. Plant growth and productivity determinants.	Diaz-Gonzales
	Cation exchange capacity (CEC)	Used as a measure of fertility, nutrient retention capacity, and the capacity to protect groundwater from cation contamination.	Diaz-Gonzales,Tziolas, Chen, Hengl, Lehmann, Bünemann, Toor, Andrews, Karlen
	Soil Salinity	Affects productivity as it influences uptake nutrients	Diaz-Gonzales, Ancy Stephen, Lehmann, Bünemann, Toor, Kopittke, Karlen
Biological Properties	Soil organic carbon (SOC)	Defines soil fertility, stability, and erosion.	Diaz-Gonzales, Tziolas, Chen, Hengl, Lehmann, Bünemann, Abdulraheem, Toor, Doran, Kopittke, Sere
	Soil microbial respiration (SMR)	Affects nutrient cycling, soil structure, and the degradation of organic matter.	Diaz-Gonzales, Xu, Toor

SH Categories	SH Indicators	Importance	Article
Climatic Factors	Temperature (LST)	Influence the speed of chemical reaction, soil aggregation and crop growth.	Ancy Stephen, Doran
	Precipitation		Ancy Stephen
	Evapotranspirati on		Ancy Stephen
	Humidity		Ancy Stephen
Management and Other Factors	Soil Erosion	Removes topsoil that lowers soil productivity and can impact water and nutrients avaiability.	Bünemann, Abdulraheem, Kopittke, , Karlen
	Irrigation	Affects the soil available water capacity (AWC), soil organic matter and leaching.	Ancy Stephen

There is no single perfect measure of soil health nor even feasible to measure everything since it wouldn't be practical or even essential (Kibblewhite, Ritz, and Swift 2008). Hence, this research used the readily available EO data to assess these indicators and determine its feasibility to monitoring soil health that leads to sustainable agriculture land use.

4.2. Earth Observation Derived Data of Soil Health Indicators

4.2.1. EO Web Platforms for Soil Data

To effectively evaluate the capabilities of the six chosen EO web platforms, this research investigates the features, functionalities, and underlying technical aspects. Particularly, in this section it addressed the questions: (i) Who developed (ii)What is the purpose; (iii) Is the data publicly available; (iv) How is the data stored and (v) What are the other functionalities? The platforms are described as follows:

A. Google Earth Engine (GEE)

Google Earth Engine (GEE) is a cloud-based platform with geospatial processing abilities that empowers users to implement large-scale geospatial analysis. Google developed it to have an interactive platform for development of geospatial algorithms efficiently, conduct data-driven studies with significant impact and address global challenges that require analyzing vast geospatial datasets. GEE offers a petabyte-scale catalog of public and free geospatial datasets, including over forty years of historical imagery and scientific data that is continuously updated and expanded daily (“Google Earth Engine,” n.d.).

As described by Gomes, Queiroz, and Ferreira (2020), the platform is developed from four technologies. These include Borg (large-scale computer cluster management), Bigtable and Spanner (distributed databases), Colossus (distributed file system), and FlumeJava (parallel pipeline execution framework), as illustrated in Figure 8.

One of the GEE’s core functionalities is that it provides the platform for machine learning, specialized algorithms (e.g., Landsat), and code editor features like charting, and data management. It also provides API for JavaScript and Python for data management and analysis. The web-based IDE (Integrated Development Environment) for JavaScript coding allows users to easily access extensive geospatial data, pre-processed imageries, visualization of data in real-time, and large repository of geospatial data. Similarly, a Python module provides the same functionality with a familiar structure (Gomes, Queiroz, and Ferreira 2020). It has extensive data catalog includes collections of satellite imagery, environmental data, weather forecasts, land cover, and socio-economic datasets, eliminating data management hurdles.

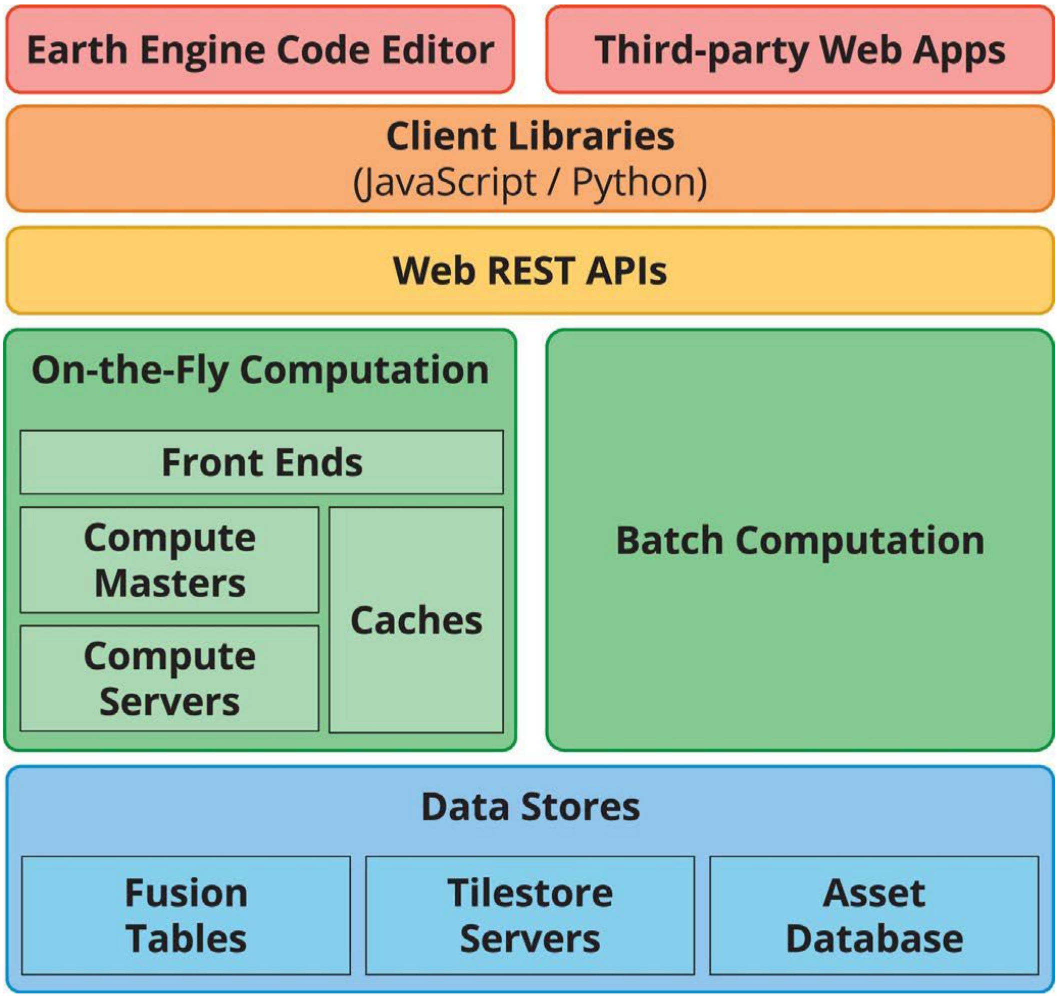

Figure 9. A simplified Google Earth Engine (GEE) system architecture diagram. (Source: Gomes et al. 2020)

Google Earth Engine (GEE) uses four key data objects (Gomes, Queiroz, and Ferreira 2020):

1. Images: These represent raster data, potentially with multiple bands (each with a name, data type, scale, and projection).
2. Image Collections: These represent stacks or time series of individual images, allowing analysis across different points in time.
3. Features: Used for vector data, this type combines geometry (point, line, or polygon) with associated attributes.
4. Feature Collections: These group related Features together, enabling functionalities like sorting, filtering, and visualization for comprehensive analysis.

GEE offers an operator library to manipulate data (public or user-uploaded) represented by its core object types (Image, Image Collection, Feature, Feature Collection). These operators work in a parallel processing, automatically splitting computations for efficient distributed execution on Google's infrastructure. Importantly, GEE objects in JavaScript or Python function as proxies, holding operation descriptions and input object references rather than the actual data itself (Gomes, Queiroz, and Ferreira 2020).

B. International Soil Reference and Information Centre (ISRIC) Data Hub

The International Soil Reference and Information Centre (ISRIC) acts as the global custodian of soil information with the mission of raising awareness and understanding of soil on global challenges. This is achieved by providing soil data and knowledge at various levels (global, national, sub-national) to promote sustainable soil and land management practices ("ISRIC — World Soil Information," n.d.).

There are three different components of ISRIC's cyber-infrastructure, as illustrated in Figure 8, namely:

1. WoSIS (World Soil Information Service): A database to find high-quality soil data that has been formatted in a consistent manner with international standards.

2. SoilGrids: This is a system that uses machine learning techniques to create predicted maps showing different soil properties in 250m resolution globally.

3. R-packages: These are special tools that can be used to analyze soil data and create maps automatically.

ISRIC created the World Soil Information Service (WOSIS) which is the database of the consolidated soil data in the world. There was a lot of overlap in this data, so WOSIS carefully sorted through it to remove duplicates and make it easier to use for everyone. This quality-assured data empowers digital soil mapping and various global assessments, making the WoSIS an important component of ISRIC's ever-evolving, searchable data infrastructure ("ISRIC — World Soil Information," n.d.).

The soil data in WoSIS comes from many diverse sources. To help users understand how reliable the data is for their specific needs, WoSIS provides three key pieces of information for each soil profile which includes the positional accuracy of the location where the soil sample was taken, the estimate of uncertainty the methods used and the date the soil sample was collected. Furthermore, data providers also include details about the specific lab methods used for analysis (like how they classified soil texture). This extra information paves the way for WoSIS's goal, to completely standardize all the data in the system.

Another system product of ISRIC is the SoilGrids (Global Gridded Soil Information) that is intended for digital soil mapping. It uses the existing soil data (WoSIS) and information about the environment (like temperature and rainfall) to create detailed maps that show soil types globally. SoilGrids is a system that uses machine learning to create detailed maps showing various soil properties globally. These maps are based on hundreds of thousands of real-world soil samples and consider factors like climate, vegetation, and landscape. SoilGrids provides these maps at high resolution (250 meters) for multiple depths underground. The maps also show how certain the predictions are in different areas. Anyone can access these maps for free.

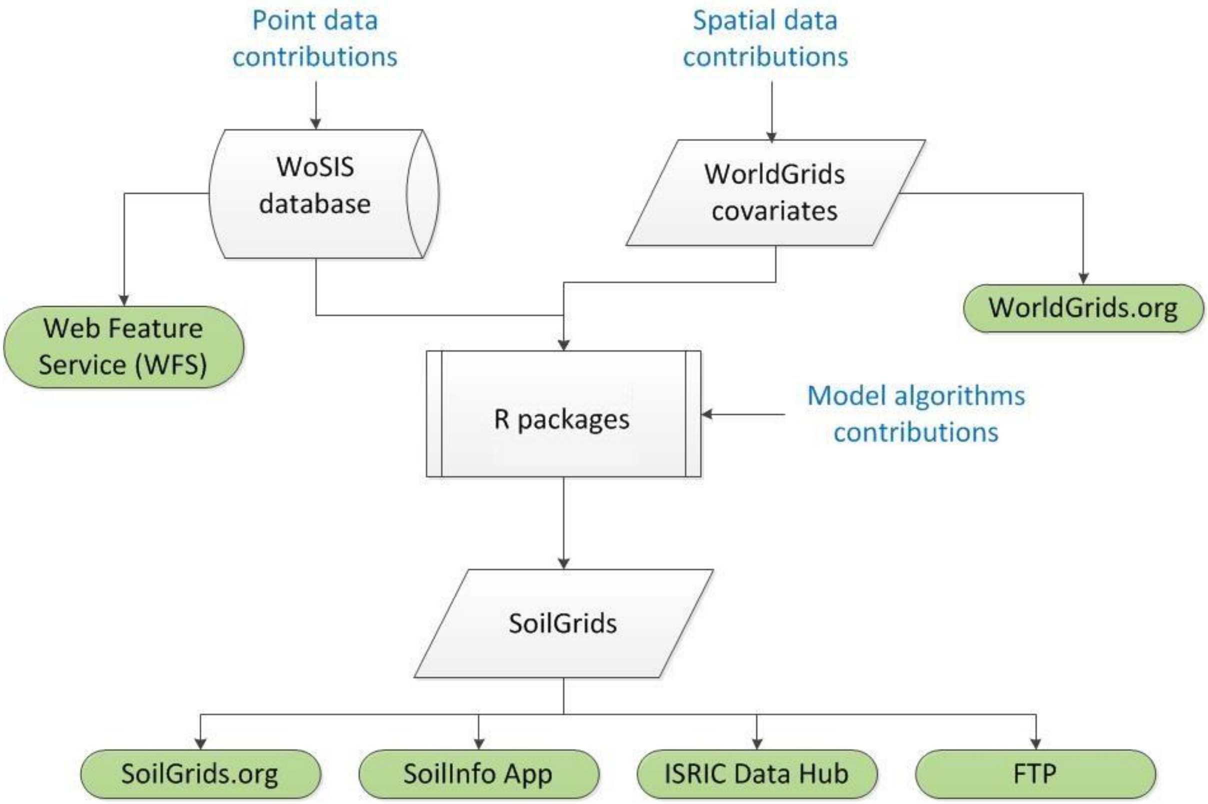

Figure 10. Simplified representation of ISRIC's cyber-infrastructure for soil data standardization, mapping and spatial modelling. Source: ISRIC

SoilGrids is known for providing soil data freely available to everyone based on collected existing soil information, old and new, and making sure it is consistent with the set standards, creating detailed maps at different scales. To create a collaborative and open approach of global soil data, there are six (6) key principles workstream used:

1. Crowdsourcing - Anyone can contribute data, and ownership still is with them.
2. Data license - Data is made freely available whenever possible, with clear licensing to ensure proper use.
3. Open-Source Software – The products are based on free and Open software like R, Google Earth, PostGIS and the likes.

4. International Standards – The cyber-infrastructure developed is tailored to facilitate global soil mapping projects and follows the FAIR principles (Findable, Accessible, Interoperable, and Reusable) for data sharing. Practices are also aligned with widely accepted international standards, such as: International System of Units (SI) for measurements; Internationally recognized soil classification systems; FAO guidelines; SoilML standards for data exchange; Open Geospatial Consortium (OGC) standards for geospatial data; World Geodetic System 1984 (WGS84) for geographic coordinates; and Other relevant global standards and conventions.

5. Reproducible research – Ensures data and method are transparent and up to date.

6. User Network: Foster a global network of soil information sharing.

SoilGrids data can be accessed in different data formats depending on the user's needs. Specifically, data can be visualized and overview using the Web Map Service (WMS), download a portion of a specific area using Web Coverage Service (WCS), whole map downloading in VRT format through WebDAV, and Google Earth Engine to access SoilGrids predictions directly within Google Earth Engine as a community contributed ("ISRIC — World Soil Information," n.d.)..

C. Food and Agriculture Organization Global Soil Information System (GLOSIS)

GloSIS, the Global Soil Information System of the FAO Global Soil Partnership is a central hub for global soil information and data created to provide easy access to dynamic soil resource information that is globally harmonized (FAO, n.d.-b). The International Network of Soil Information Institutions (INSII) is leading the development, with member countries contributing their data. GloSIS aims to be a game-changer in assessing soil resources, empowering policymakers with the knowledge needed to combat land degradation. It provides a globally harmonized platform from national data, fostering international collaboration on soil-related challenges (FAO, n.d.-b)As shown in Figure 9, GloSIS is designed as a decentralized system, that is made up of a federation of country or regional nodes. Each node represents a single Soil Information System (SIS) like a central location where a country can store and share its soil information according to data exchange standards. GloSIS also uses a bottom-up process that lets countries in charge of their own soil data. The process of creating or harmonizing a node will be carried out based on varying levels of participation and will be supported by the Global Soil Partnership (GSP) secretariat and the Soil Data Facility, which is managed by ISRIC (FAO, n.d.-b).

To "building blocks" of GloSIS are designed to have the Domain Model, Data Exchange, GloSIS Node, Support Node and Discovery Hub (FAO, n.d.-b). The Domain Model defines the structure and relationships between data, ensuring consistency and coherence across the system. Data exchange component enables national users to send and receive data in a standardized and harmonized manner. GloSIS Node represents a specific instance of the GloSIS system, where data is stored, processed, and shared. Support Node component provides additional functionality and services to support the operation and maintenance of GloSIS nodes. And the Discovery Hub facilitates the discovery and access of GloSIS nodes, allowing users to locate and interact with the system's various components.

Building a global system for sharing soil data is complex, done step-by-step, but mainly focuses on the standardization of data formats and ensuring countries have the technical know-how. To address these hurdles, we'll use a multi-phased approach: First is setting up the system and making some initial data available. Then work on making all the data from different countries completely compatible and in accordance with international standards (FAO, n.d.-b).

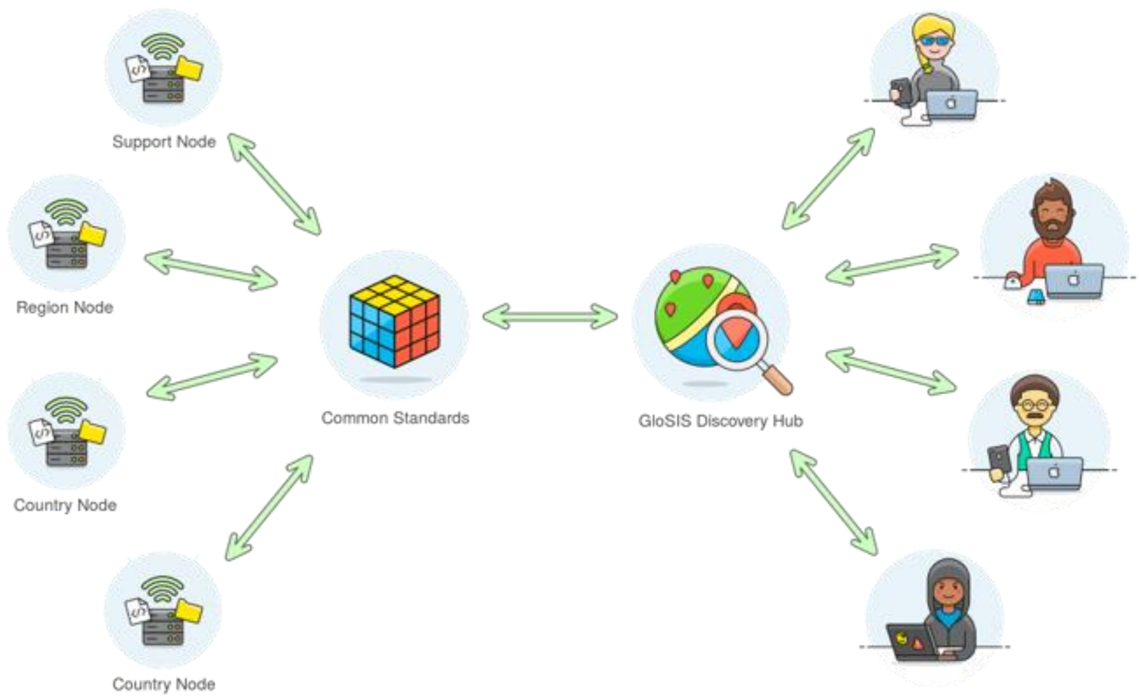

Figure 11. The GloSIS Concept. (Source: FAO-GloSIS)

D. Food and Agriculture Organization Hand-In-Hand Geospatial Platform (HIH)

The Hand-in-Hand Initiative is a powerful tool developed by the FAO (Food and Agriculture Organization) to address hunger and poverty in the world through empowerment of countries in transforming their own agricultural systems. With the aim of improving the overall food situation in regions facing limited resources and facing humanitarian crises, these regions will increase the quantity, quality, diversity, and accessibility of nutritious foods; and through empowering Rural Communities by identifying opportunities that can raise incomes for rural populations, making them more resilient and self-sufficient. This could involve supporting local farmers with improved agricultural practices, better market access, or creating new income streams through sustainable food production methods ("Hand-in-Hand Geospatial Platform | Food and Agriculture Organization of the United Nations," n.d.).

However, the HiH faces challenges in getting timely, detailed information, especially in areas with limited information. To overcome these challenges, the Data Lab has the role on developing text mining tools that gathers agricultural production data in HiH priority countries, specifically at a sub-national level, even when traditional data collection methods are lacking. For areas with no or outdated poverty maps, the Data Lab used the big data approach to produce vulnerability maps. The Hand-in-Hand Initiative, launched in 2019, has become a cornerstone of FAO's efforts, spearheading its fight against global hunger and poverty(“Hand-in-Hand Geospatial Platform | Food and Agriculture Organization of the United Nations,” n.d.).

In this initiative, FAO developed the open-access Hand-In-Hand Geospatial Platform, a supporting tool that provides advanced insights including food security indicators and agricultural statistics. This initiative utilizes cutting-edge geo-spatial modeling and analytics to determine the opportunities to increase income and reduce the inequalities and vulnerabilities of rural populations, who comprise the majority of the world's poor. By mapping these, it helps in determining suitable interventions that address specific needs. The platform acts as a central hub, unlocking millions of datasets that can be useful for various sectors like digital agriculture experts, economists, government and non-government agencies (“Hand-in-Hand Geospatial Platform | Food and Agriculture Organization of the United Nations,” n.d.).

The HiH Geospatial Platform has compiled datasets from over 20 different FAO departments including Animal Health to Trade and Markets, Soil, Land, Water, Climate, Fisheries, Livestock, Crops, Forestry, Trade, Social and Economics, and others. With this vast dataset, it allows users to analyze a wide range of factors, including land resources, water availability, and economic conditions. There are over a million spatial layers and thousands of statistics with metadata records. Data is sourced from FAO databases on food and agriculture with over 245 countries, UN agencies, NGO’s, universities, private sector and space agencies. This allows for a truly global perspective. Since its launch in 2020, the platform's impact has been growing. Over 65 countries and institutions have participated in workshops, learning how to leverage data and technology to improve agriculture and rural life (Montgomery, n.d.).

Aside from interactive data viewing, the users can produce data maps and create impactful stories, turning data into narratives and analyze all types of data from remote-sensed to statistical time series at a different scale (global, national, and even locals).

E. Earthmap

Earth Map is a state-of-the-art, free application created by the Food and Agriculture Organization of the United Nations (FAO), in collaboration with Google, to reform the way users interact with land and climate data (EarthMap, n.d.). This pioneering tool empowers users to visualize, process, and analyze vast amounts of data, fostering a deeper understanding of the complex relationships between land and climate systems. As per Morales et al. (2023), Earth Map has petabytes of multitemporal, multiscale, multiparametric, and quasi-real-time satellite imagery and geospatial datasets available to any user that is powered by Google Earth Engine.

The application has enhanced analytical capabilities that enable the detection, quantification, and visualization of global and local changes and trends on Earth's surface at a planetary scale. This allows for a comprehensive understanding of the dynamic interactions between the Earth's surface and the impact of global and local phenomena on the planet (Morales et al. 2023). It only has a point-and-click user interface with geospatial layers and statistics are generated on-the-fly. Although Earth Map uses the GEE API, it is not a GEE app and doesn't require any programming skills, therefore offers a better capacity for non-experts the ease to use it.

Earth Map's data is currently sub-divided into more than 15 thematic groups that cover agriculture, biodiversity, climate, greenhouse gas emissions, fire, forestry, geophysical, geosocial, hydrology, land use/land cover, land degradation neutrality, satellite images, soil, vegetation, and water. The continent- or region-specific thematic sections or layers only appear once the user has selected a specific area of interest (AOI).

Earth Map works in two parts: the Firebase-based backend and a web-based client (Morales et al. 2023) as illustrated in Figure 9. The Earth Map backend uses Google's Firebase platform to manage user accounts, store data, and run functions when needed. This acts as engine room, keeping things running smoothly. Meanwhile, the web interface is built using React and Material User Interface. It uses Google Maps to display the data (maps and statistics) in an easy-to-understand way. It refers to the dashboard to see the results. This is the part that uses the Google Maps API.

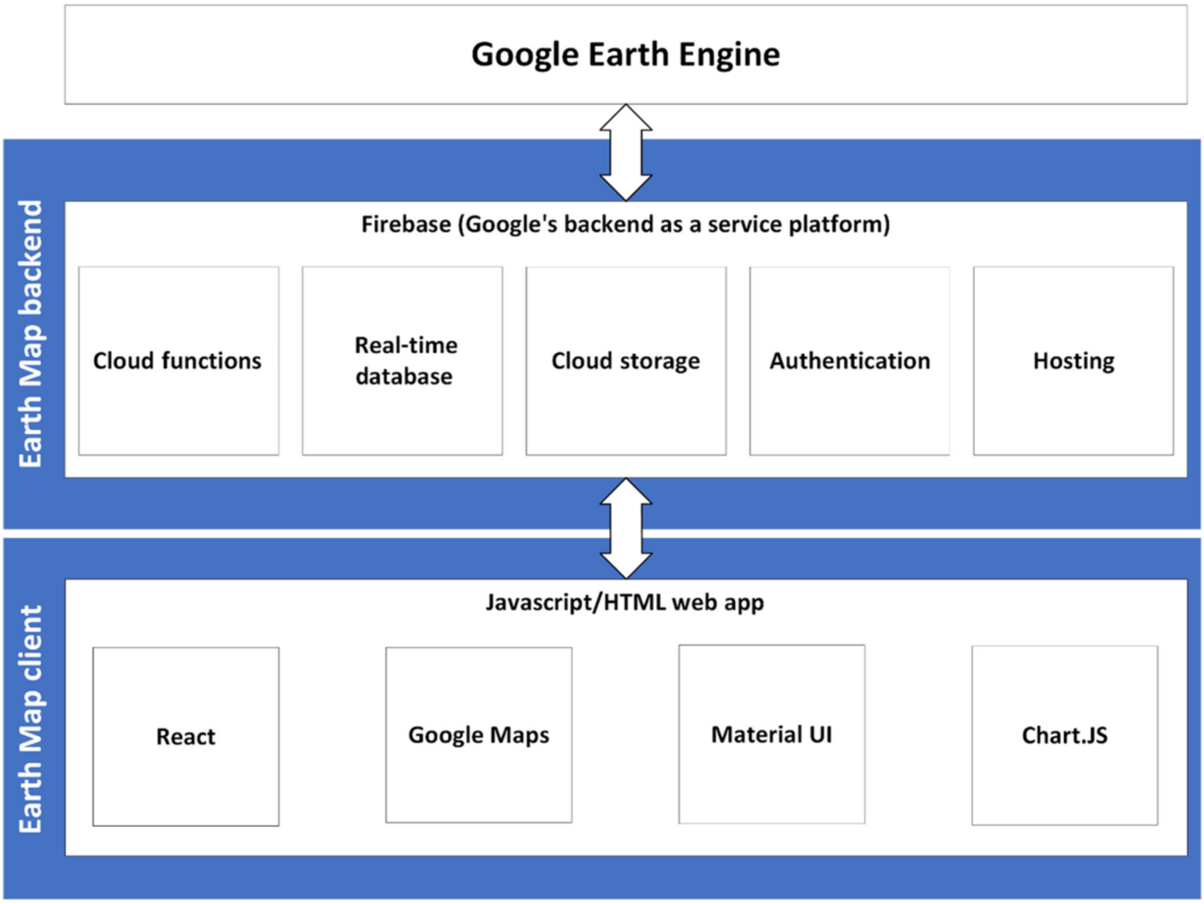

Figure 12. Earth Map's software architecture (Source: Morales et al. 2020).

F. Oak Ridge National Laboratory Distributed Active Archive Center (ORNL DAAC)

The Oak Ridge National Laboratory Distributed Active Archive Center (ORNL DAAC) is a NASA data center specializing in biogeochemical dynamics ("ORNL DAAC for Biogeochemical Dynamics," n.d.). Part of the Earth Observing System Data and Information System (EOSDIS), ORNL DAAC is managed by Oak Ridge National Laboratory in Tennessee and a member of the Remote Sensing and Environmental Informatics Group of the Environmental Sciences Division (ESD).

The ORNL DAAC focuses on archiving and distributing data on the chemical and physical properties of soils, including local studies to extensive global datasets. The ORNL DAAC's mission is to assemble and provide these resources, along with data services, to support research, education, and informed decision-making related to Earth's biogeochemistry and ecological health.

The ORNL DAAC serves as the primary archive for NASA's Earth observation related to Terrestrial Ecology and Carbon Cycle Science data and provides ground and airborne data to verify the accuracy of NASA's Earth Science missions. It also collaborates with NASA to develop best practices and tools for researchers working with this Earth science data. And by integrating diverse datasets, ORNL DAAC facilitates research by making it easier for scientists to find and analyze the information they need to address critical questions("ORNL DAAC for Biogeochemical Dynamics," n.d.).

The data formats possible to download in ORNL DAAC are the Data Product Development Guide (DPDG), GeoTIFF, ASCII, OGC KML, netCDF-4/HDF5, Cloud Optimized GeoTIFF (COG), NetCDF Classic that provides a data model, software libraries, machine-independent data format for geoscience data; and Zarr, which is a specification for storage of and access to multi-dimensional array data.

4.2.2. EO Derived Data for Soil Health Assessment

Out of the twenty-seven (27) EO web platforms evaluated (see in Annex 1), there are only six (6) platforms that are focused on soil datasets. With the twenty-five (25) selected soil health indicators to attain sustainable agriculture land use, there are twenty-two (22) datasets available in a single platform, which is Google Earth Engine, which are free and readily available for users to download and use for analysis. Although the EarthMap is powered by Google Earth Engine, it is the platform with the least available dataset that can be used in assessment of soil health. However, EarthMap offers a variety of readily available geospatial data that can be used in various environmental analysis. Table 5 indicates the available soil health datasets for each web platform. See Annex 2 for the dataset name that is used in the evaluation of agricultural soil health.

Table 5. The presence of available free data for Soil Health Assessment of the 6 web platforms.

CATEGORY	INDICATOR	EO WEB PLATFORMS AND NAME OF DATASET					
		GEE	ISRIC	FAO - GLOSIS	EARTH MAP	ORNL - DAAC	FAO - HIH
Physical Properties	Soil Texture	x	x	-	x	x	x
	Sand content (Sand-C)	x	x	x	-	x	x
	Silt content (Silt-C)	x	x	x	-	x	x
	Clay content (Clay-C)	x	x	x	-	x	x
	Bulk Density (BD)	x	x	x	-	x	x
	Soil Moisture	x	x	-	-	x	x
	Water Holding Capacity	x	x	x	x	x	x
Chemical Properties	Hydrogen Potential (pH)	x	x	x	-	-	x
	Electrical Conductivity	-	x	-	-	x	-

CATEGORY	INDICATOR	EO WEB PLATFORMS AND NAME OF DATASET					
		GEE	ISRIC	FAO - GLOSIS	EARTH MAP	ORNL - DAAC	FAO - HIH
	Nitrogen (N)	x	x	-	-	x	x
	Potassium (K)	x	x	-	-	-	-
	Calcium (Ca)	x	x	-	-	-	-
	Magnesium (Mg)	x	x	-	-	-	-
	Phosphorus (P)	x	x	-	-	x	-
	Soil Nutrient	x	x	x	-	-	x
	Cation Exchange Capacity (CEC)	x	x	x	-	-	x
	Soil Salinity	-	x	x	-	-	x
Biological Properties	Soil organic Carbon (SOC)	x	x	x	x	x	x
	Soil microbial respiration (SMR)	-	-	-	-	x	-
Meteorologic data	Temperature	x	-	-	-	x	x
	Precipitation	x	-	-	x	x	x
	Evapo - transpiration	x	-	-	x	x	x
	Humidity	x	-	-	-	x	x
Management and Other Factors	Soil Erosion	x	-	-	x	x	-
	Cropland Irrigation	x		--	x	x	-

Consequently, Figure 10 illustrates that among the soil health categories, Physical Properties has the highest number of datasets available across the 6 platforms. The GEE, ISRIC, ORNL-DAAC and FAO-HIH have the complete soil health indicator datasets (7 out of 7), with GloSIS having five out of seven due to its unavailability in Soil Texture and Soil Moisture. However, this does not completely discredit that it is non-existent because FAO treated the soil moisture as a separate entity and has its own special web application where users can look more closely at the details of the data. Even though GloSIS does not have the readily available data for soil texture, users who have knowledge on soils can decide it using the percentage data of sand, silt and clay which are available in the platform.

For the chemical attributes, ISRIC has the greatest number of indicators available, which is 9 out of 10 followed by GEE with 8 out of 10. ISRIC has the greatest number of data available for chemical categories since in the former years, people were more concerned on the production, thus, the monitoring of soil nutrient and chemical properties is very vital. Chemical properties serve as the basis for the management and agricultural interventions that can be implemented in an area. Another factor affecting the availability of datasets is the availability of pedological data since 1950 that can be used as auxiliary data for the model prediction of soil health indicators.

As mentioned in the Review of Literature, the Biological Indicators are only considered as a soil health indicator just recently where there are advancements made in the field of biochemistry. However, the indicator soil organic carbon (SOC) is available across the 6 platforms. The dataset for soil microbiological respiration is not yet available for all platforms.

Though meteorological factors like temperature, precipitation and humidity are not the basics of soil properties, they are directly affecting soil processes especially in the water availability and nutrient leaching. It is understandable that ISRIC and FAO-GloSIS do not have data on meteorological platforms since both platforms are developed for the soils alone. However, those 4 datasets are available in GEE, ORNL DAAC and FAO-HiH. While EarthMap can only provide 2 out of the 4 meteorological factors. But, these 4 platforms are designed to cater to different datasets on different environmental data.

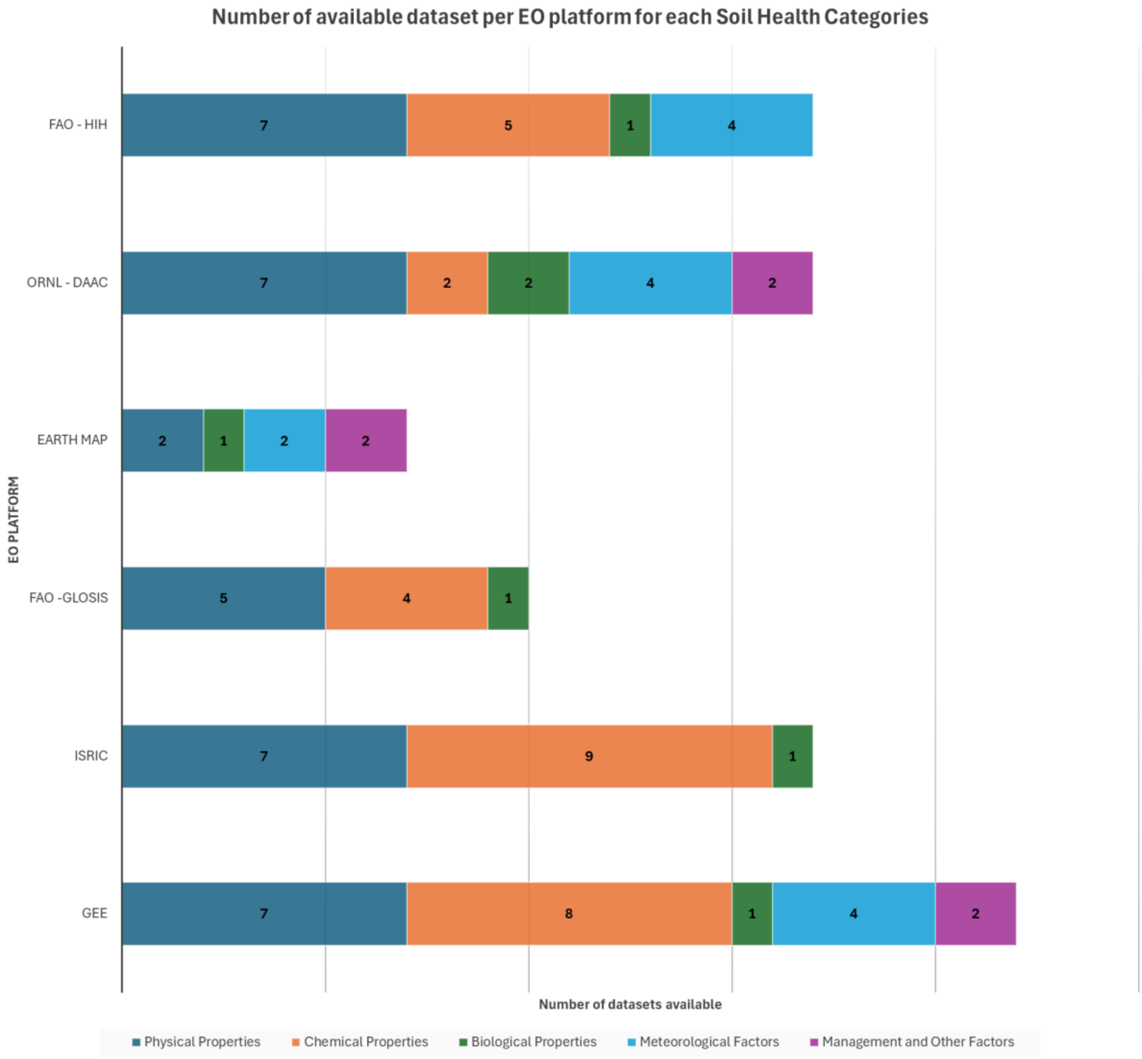

Figure 13. The number of available datasets for each EO Platform per Soil Health Categories.

Lastly, the soil erosion and cropland irrigation dataset are found vital in agricultural soil health assessment. As per Kibblewhite, Ritz, and Swift (2008), soil erosion is affecting the soil structure due to movement and changes on the percentages of sand, clay, loam and organic matter. Irrigation datasets are also considered nowadays as a factor in smart agriculture, hence inclusion of this as soil health indicators also helps in the proper management of farmlands. In this component, the 2 datasets are available in ORNL-DAAC, EarthMap and GEE.

The availability of these free and off-the-shelf EO datasets enables the assessment of the twenty-five (25) identified soil health indicators, suggesting their potential for operational soil health assessment. However, it is not enough that spatial data is just existing, it is important to determine the quality of each dataset in user's perspective so the data can be used extensively and appropriately for a certain application.

4.3. Assessment of EO Derived Data for Agricultural Soil Health

Traditionally, data quality has been defined and measured from a technical perspective, focusing on aspects like accuracy, completeness, and consistency. However, this approach often neglects the needs of the people who actually use the data – the data users. This research highlights the importance of understanding and valuing data quality from the perspective of data users. In this section, the developed framework on assessing data quality is presented and highlights the strengths and capabilities of the EO platform.

4.3.1. Intrinsic Data Quality

The intrinsic data quality dimension refers to the description of the data itself and the entity who produced the data. As mentioned by Wand and Wang (1996), in user's perspective – the inherent trustworthiness of the data is based on its source and production process, independent of how the data itself is used or analyzed.

As shown in Table 4, the datasets and platforms are produced by reputable organizations who are recognized globally and are known experts on the field of soil science, earth observation and agriculture, namely: Google, National Aeronautics and Space Administration (NASA), International Soil Reference and Information Centre (ISRIC) and United Nations – Food and Agriculture Organization (UN-FAO). Although they are the owners of these EO platforms, datasets are provided by different organizations, individuals and private entities like EnvirometriX Ltd., University of California, Global Soil Partnership and more that is indicated in Table 6. Although datasets were produced by different organizations, it is assured that it follows the data standards set by OGC, ISO and ISC World Data System.

It also indicates that all the datasets used by the platforms used auxiliary data in production, specifically these are the field measurement of soil properties of countries, and some datasets used the derivatives of other EO products.

Another data quality considered in the Intrinsic component is the methods of processing. It was observed that almost all datasets for all platforms are derived using different machine learning algorithms and model predictions. Since most of the datasets offer global coverage and there are areas that really don't have field soil data and information, machine learning and other model predictions are used to provide data for the areas lacking with field data. Some of the methods have available scripts online that can be referred to by users for further understanding of datasets or can be adapted by users if they want to implement and do their own data processing.

Consequently, you can refer to Annexes 3 to 8 for the complete documentation of each platform's characteristics and methods.

Table 6. The intrinsic component of EO web platforms.

Data Quality	DQ-Element	GEE	ISRIC	EARTHMAP	DAAC	FAO HIH	GLOSIS
Reputation	**Dataset Provider**	1. EnvirometriX Ltd 2. Innovative Solutions for Decision Agriculture Ltd. (iSDA) 3. Google and NSIDC 4. Commonwealth Scientific and Industrial Research Organisation 5. NASA LP DAAC at the USGS EROS Center 6. Climate Hazards Center 7. UC Santa Barbara 8. University of California Merced 9. World Wildlife Fund (WWF) 10. Global Food Security-support Analysis Data at 30m Project (GFSAD30)	1. ISRIC	1. EnvirometriX Ltd 2. FAO 3. GFSAD 1000	1. Oak Ridge National Laboratory Distributed Active Archive Center	1. EnvirometriX Ltd 2. Global Soil Partnership	1. Global Soil Partnership
Reputation	**Point of contact**	Google	ISRIC	UN-FAO With Google	NASA Earth Data	UN-FAO With Google	UN-FAO
Accuracy	**Use of Ancillary Data**	Yes	Yes	Yes	Yes	Yes	Yes

4.3.2. Accessibility Data Quality

The concept of accessibility data quality and its importance is recognized by data users. The data producers should ensure that data is readily available and usable for its intended purpose. To date, data users rely heavily on online access that makes it a more critical concern for data producers. Data needs to be easily accessible, usable online and downloadable for it to have a significant value to the data users. This category comprises the elements of accessibility and access security.

Table 4 shows the data access, possible product formats available for users, and the security of uploading own dataset to be used in the platform. For data access, only the ISRIC geospatial platform gives full access of functionalities to users without requiring a user account. Meanwhile, the 5 other platforms (GEE, EarthMap, DAAC, HiH and GloSIS) require users to have an account for full access of the platform.

Among these platforms, only GEE has the functionality of uploading users own dataset that makes it more flexible and adaptable to different data users. Meanwhile, EarthMap has the functionality of selecting the place of interest in watershed, national, regional, and continental scale. Both the FAO developed applications, HiH and GloSIS, give the user the ability to layout and focus on the intended area of interest and have an online sharing of map produced by users. This also allows users to make a story map using the available datasets. For ISRIC and DAAC, users can only visualize the data coverage of certain datasets but really cannot sparse it into smaller areas, when users need it. Sparsing data can be done outside the platform already.

Table 7. The accessibility component of EO web platforms.

Data Quality	DQ-Element	GEE	ISRIC	EARTHMAP	DAAC	FAO HIH	GLOSIS
Accessibility	**Data Access**	Need user account	Free	Need user account	Need user account	Need user account	Need user account
Accessibility	**Product Format**	GeoTIFF, TFRecord, Feature Collection as CSV, SHP, GeoJSON, KML, KMZ, Map Tiles	GeoTIFF, JSON, Geopackage, Shapezip, Text, CSV, XML, WMS, VRT (GDAL Virtual Format)	PNG, GeoTIFF, CSV, XCLS	Arc/Info ASCII Grid, Erdas Imagine, GeoTIFF, netcdf, NITF, XYZ	PNG, GeoTIFF, CSV, XCLS	GeoTIFF, WMS
Access Security	**Security of uploading own data**	Allows upload of own dataset without being	No function to upload your data.	No function to upload your data.	No function to upload your data.	No function to upload your data.	No function to upload your data.

For the product format, ISRIC can offer nine (9) different data types on downloaded dataset followed by GEE with eight (8) different data format available for download. However, GEE users should be knowledgeable in running some javascript or python to download a particular dataset. Users can download six (6) data types, four (4) data types for FAO-HiH and EarthMap while FAO-GloSIS which is still under development can only offer two (2) file formats. It can be noted that the common data type available across the platform is GeoTIFF, which users are more familiar with.

In terms of accessibility data quality, Google Earth Engine has more flexibility and adaptability to cater to different needs of users. Since this has a processing tool, users have more control over the type of dataset they want to use and limits on the needed data coverage and usage. It is worth noting that ISRIC is a platform that is open to the public as it can be operated and downloaded without any user registrations. Refer to Annexes 3 to 8 for the complete documentation of each platform's Accessibility Data Quality.

4.3.3. Representational Data Quality

The representational component of EO web platforms highlights the importance of clear and user-friendly data presentation. This encompasses the way datasets are structured and organized and considers the aspect of how easily the users can interpret the data and grasp its intended meaning. This section highlights the evaluation of data presentation in terms of the usability (ease of use), interpretability (well defined labels and available units), consistency (follows the same format throughout the platform and in accordance with EO data standards, understandability (ease of understanding) and reusability (ease of use for users).

In reference to the user's perspective, platform should be easy to use by novice even without knowledge on remote sensing. The ISRIC, DAAC, EarthMap, FAO-HiH and FAO-GloSIS have this platform capability as shown in Table 8. But for the dataset of GEE to be used extensively by users, requires users to have at least knowledge on programming, running and modifying the supplemental scripts available which can be a challenge for most users. At the same time, this functionality of GEE makes it more flexible and gives more leeway for users to change the dataset based on their area of interest and needs. In addition, it is only GEE among the evaluated platforms that can be used with Python or Javascript.

In terms of the available processing tools, most of the EO platforms (ISRIC, DAAC, FAO-HiH and FAO-GloSIS) only allow users to view and download datasets while EarthMap has the functionality of producing charts and graphs per selected area which can be done in the cloud. GEE, with its functionality, can do a lot more than just view and download data. It also allows users to manipulate and modify the available datasets from its catalog. Since EarthMap and GEE are powered by Google, they both have tools for geospatial analysis. This includes functions for time series analysis and obtain graphs and statistics to geospatial data.

All the platforms were found to have a user-friendly interface and easy dataset navigation. The User Interface of the applications has clear menus and navigation bars and consistent layout and designs. Consequently, it also has tooltips or “Help” guidelines functions. It was observed that all EO platforms offer a robust search engine that can be done through searching by keyword, location or region and for by date.

The scalability data quality element refers to the capability of the application to handle high volume of data which all the six (6) platforms are capable of, since those platforms are designed to handle big data and processing is done in cloud. This functionality of EO platform facilitates the challenge of using high-end computers to use EO derived data.

Table 8. The representational component of EO web platforms.

Data Quality	DQ Element	GEE	ISRIC	EARTHMAP	DAAC	HIH	GLOSIS
Usability	**User Technical Knowledge Requirement**	Requires basic knowledge of programming.	Familiarity with the platform	Familiarity with the platform	Familiarity with the platform.	Familiarity with the platform.	Familiarity with the platform
Usability	**Processing Tools**	Use built-in codes in javascript for data manipulation	Platform allows you to view and download data	Platform allows you to select the site, view and download data	Platform allows you to view and download data	Platform allows you to view and download data	Platform allows you to view and download data
Usability	**Programming Languages**	Python and Javascript	No	No	No	No	No
Usability	**Cloud Processing**	Yes. All functionalities are done in cloud.	No.	Yes, for the interactive graphs and charts.	No	No	No
Usability	**User Interface (UI)**	Yes. Easy to navigate and easy to find different datasets.	Yes. Easy to navigate and easy to find different datasets.	Yes. Easy to navigate and easy to find different datasets.	Yes. Easy to navigate and easy to find different datasets.	Yes. Easy to navigate and easy to find different datasets.	Yes. Easy to navigate and easy to find different datasets but difficult to find metadata.
Interpretability	**Visualization Tools**	Yes. It has Google Code Editor to view and manipulate data.	Yes. Allows visualizing the geotiff.	Yes. Allows visualizing the geotiff and graphs.	Yes.Allows to view dataset in kml and has	Yes. It has interface for data analysis, graphs and charts	Yes. Allows developing of own map with different datasets.

Data Quality	DQ Element	GEE	ISRIC	EARTHMAP	DAAC	HIH	GLOSIS
					Spatial Data Access Tool		
Interpretability	**Scalability**	Yes	Yes	Yes	Yes.	Yes.	Yes
Consistency	**Open Standards**	Yes	Yes	Yes	Yes	Yes	Yes
Understandability and Reusabilty	**User Support**	Yes	Yes	Yes	Yes	Yes	Yes
Understandability and Reusabilty	**Community**	Yes	Yes	Yes	Yes	Yes	Yes
Understandability and Reusabilty	**Product Manual**	Yes	Yes	Yes	Yes	Yes	Yes

In terms of consistency, understandability, and reusability data quality elements, all the six (6) platforms have consistent data formatting and units throughout. However, Google Earth Engine datasets have different methods and processing techniques used since it is produced by different organizations. This is also the same for EarthMap. The ISRIC, DAAC, FAO HiH and FAO GloSIS have similar data processing methods employed across its datasets. Hence, datasets from these platforms are easier to use for analysis.

All the six (6) platforms have available product documentation that can be used by users to refer to and understand how data is produced and along other data information. All platforms have the user support available and community that helps users for trouble shooting and openly asked questions on datasets which is actually helpful for further data processing or analysis that users intend to use.

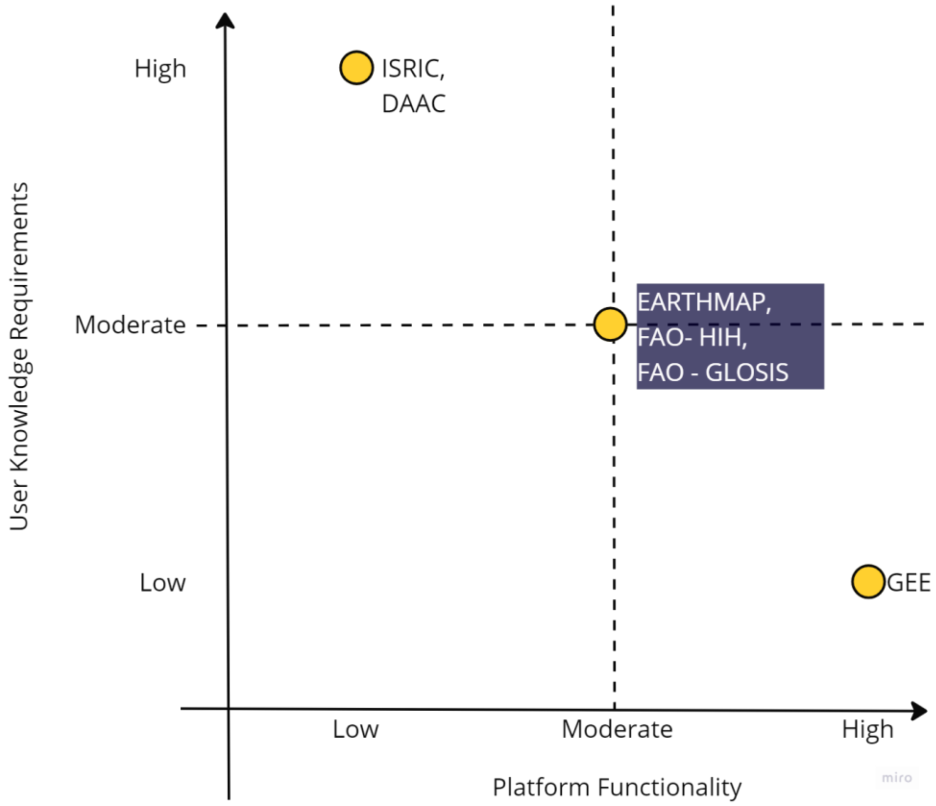

Figure 14. The rate of platform functionalities against the user knowledge requirements for data quality assessment.

Figure 14 illustrates the position of each platform in relation to the required knowledge for users and the platform functionality. Google Earth Engine as described earlier has higher functionalities because user can do more than just data viewing and download, it allows users to manipulate and do analysis tailored to user's needs. However, it is rated low on the user knowledge requirements because it requires the ability and knowledge of users on programming, even if it's basically just on running the script provided by GEE. The three (3) platforms (EarthMap, HiH and GloSIS) that are developed by FAO are in moderate level of user knowledge requirements and platform functionality, since it allows users to view, download, select the area of interest and can produce graphs and charts related to the selected area. EarthMap specifically offers the functionality on graphs and chart production while HiH and GloSIS provides users to create map story that is shareable to other users. On the other hand, even though ISRIC and DAAC have a high position in User Knowledge requirements, it also has the minimum functionalities it can offer, limited only to viewing and downloading of datasets. These are easiest to use but if the user needs to manipulate and focus on a certain area, another software should be used. It is a trade-off between learning a bit of programming or using another software to do analysis and produce maps. Refer to Annexes 3-8 for the detailed contextual data quality assessment of each platform in terms of soil health indicators.

4.3.4. Contextual Data Quality

The contextual data quality pertains to how well data serves its particular purpose. This purpose can be affected by the fitness of data to the user's need, coverage, date of data produced, granularity and accuracy. Different users have different contexts and needs on EO data, in this research, the sensors are important as it indicates the timeliness and completeness of data. Completeness in this research refers to whether there are datasets available for certain soil health indicators and covers a wide range of extent geographically. See Annexes 3 -8 for the detailed contextual data quality assessment for each indicator.

Looking at Figure 15, it indicates that the most common sensors used in production of EO datasets are MODIS, Landsat and DEM/DTM and its derivatives which are more extensively used in producing physical and chemical soil health indicators. This is understandable since Landsat is the longest-running EO satellite that started in 1972 while Moderate Resolution Imaging Spectroradiometer (MODIS) was onboarded in 1999. They are both the two most used sensors for a historical EO dataset. MODIS has been extensively used because it has a 1-2 day return period producing 250 meters to 1,000 meters spatial resolution. Meanwhile, Landsat has a better spatial resolution of 15 meters to 30 meters but with a 16-day return period. Most of the datasets available have 250 meters resolution even if it utilized both the Landsat and MODIS because the combination of two (2) different satellites into single dataset requires down sampling of Landsat to match with the coarser resolution of MODIS, reducing the details captured by Landsat. And since this research is focused on the global coverage of datasets, MODIS with moderate resolution with frequent coverage is more suitable especially for larger areas. On the other hand, Landsat with higher spatial resolution can be used for more detailed information on a smaller or specific location. It is also worth noting that Sentinel satellites came at a later date (2014), so it has only less usage in respect to soil health indicators.

With the fast development of earth observation technologies, radars and Lidars that penetrate the cloud covers are also explored in measurements of soil properties.

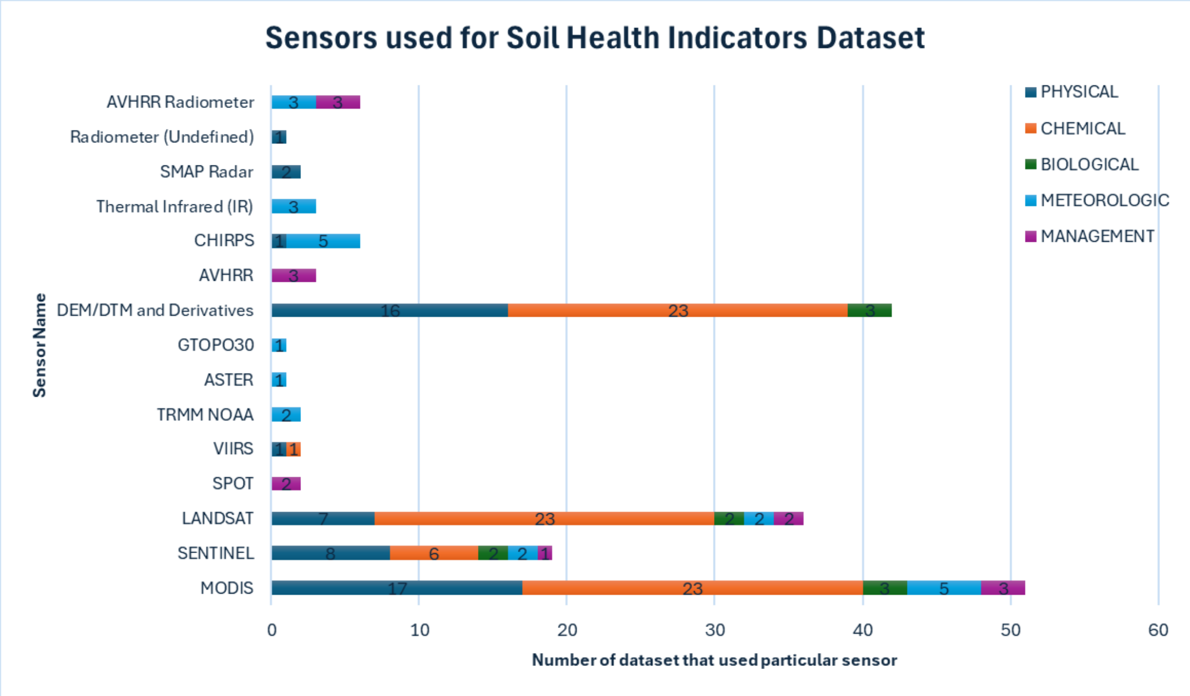

Figure 15. The available sensors are used in the production of soil health indicators.

Figure 16 illustrates the ranges of spatial resolution datasets available for the soil health indicators. It shows that for soil Physical Health indicators, the GEE has the best spatial resolution of 30 meters while DAAC has a very coarse dataset for about 100,000 meters. For the Chemical properties, GEE still stands out with the finer resolution while ISRIC is providing a very low resolution on chemical indicators, particularly only with Phosphorus and Cation Exchange Capacity (CEC) datasets. This kind of resolution is fitted for global scale analysis. For the biological properties and management factors, the spatial resolution available are in ranges from 30meters to 1 kilometer which is already significant for both smaller and national scale interpretation. While meteorological data is widely monitored, its spatial resolution can vary from 30 meters to 100,000 meters. However, this is often compensated for by the high temporal resolution, with datasets available in hourly or even 3-hour intervals. It must be noted that the 30-meter resolution datasets, are used for national and regional scale applications like the datasets produced for Africa, Australia, and the USA.

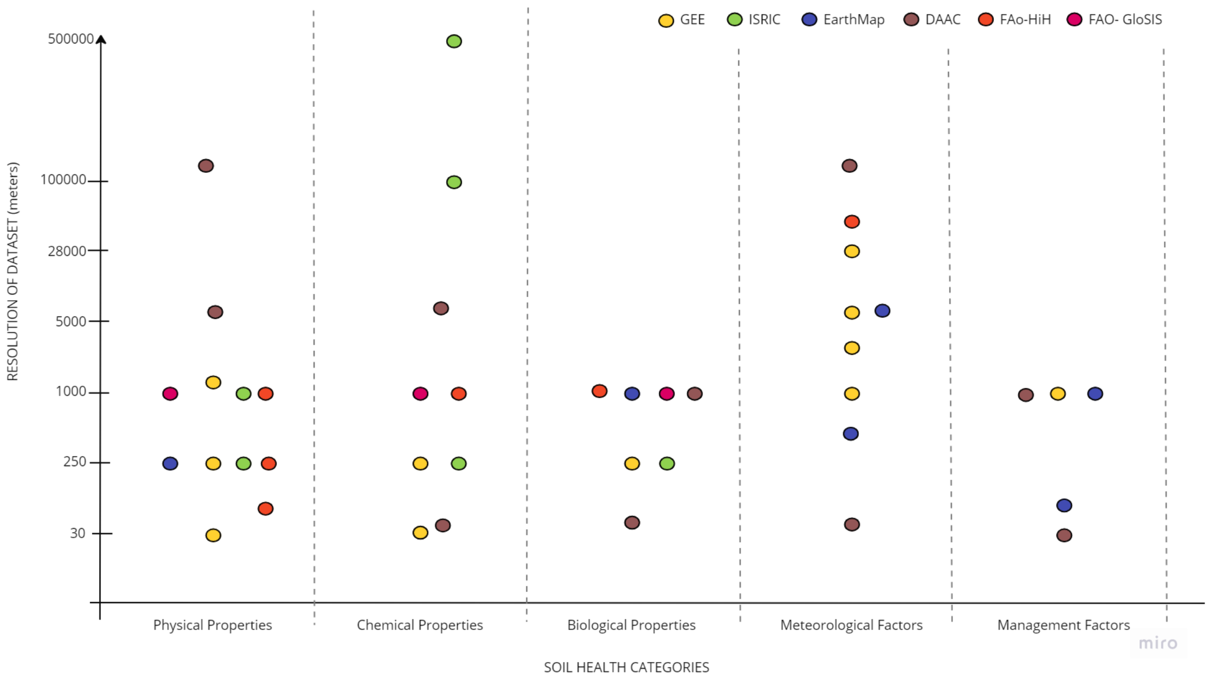

Figure 16. The spatial resolution of dataset available from platform categorized based on soil health.

Another element of the contextual data quality pertains to the timeliness which covers the frequency of dataset update and the temporal coverage. It was found out that there is only single dataset available for each indicator since 1950, thus call for the updating of EO datasets. However, with the use of the historical field data available, a machine learning algorithm and remote sensing images provided indicative or predicted soil property values for each indicator. This kind of dataset may not be the most efficient to use for monitoring since it didn't provide changes through time, but this can be used as benchmark data for mapping soil properties and soil health indicators to determine present conditions. Table 9 shows the available historical data used for the development of datasets. It is expected that the Physical and chemical indicators have data since 1950's because it is already recognized by scientists and farmers as relevant indicators for agricultural production. Biological indicators have field data starting 1981 to present as it's only in this era that people were aware of the biological factors affecting soil quality. The meteorological data started in the 1980's with great and fast improvements currently. Likewise, the agricultural monitoring of the irrigation availability in agricultural land started on the onset of the talks on smart agriculture.

Table 9. The historical data used in production of SH Indicator dataset.

	GEE	ISRIC	EARTH MAP	DAAC	HiH	GloSIS
Soil Texture	1950 to 2018	1950 to 2015	1950 to 2018	1950 to 1996	1950 to 2018	
Sand content (Sand-C)	1950 to 2018	1905 to 2016		1950 to 1996	1972 to 2004	1972 to 2004
Silt content (Silt-C)	2001 to 2017	1918 to 2013		1950 to 1996	1972 to 2004	1972 to 2004
Clay content (Clay-C)	2002 to 2017	1905 to 2016		1950 to 1996	1972 to 2004	1972 to 2004
Bulk Density (BD)	2003 to 2017	1950 to 2015		1950 to 1996	1972 to 2004	1972 to 2004
Available Water Capacity	1950 to 2018	1950 to 2015	1950 to 2015	1950 to 1996	1972 to 2004	1972 to 2004
Soil Moisture	2015 to 2024	1950 to 2015			1979 to near present	1972 to 2004
Hydrogen Potential (pH)	2005 to 2017	1951 to 2015		1950 to 1996	1972 to 2004	
Total nitrogen (TN)	2006 to 2017	1905 to 2016			1972 to 2004	
Potassium (K)	2006 to 2017	1980 to 2016				
Calcium (Ca)	2006 to 2017	1980 to 2016				
Magnesium (Mg)	2006 to 2017	1980 to 2016		1977 to 2012		
Phosphorus (P)	2006 to 2017	2001 to 2011				1972 to 2004
Soil Nutrient					1972 to 2004	1972 to 2004
Cation exchange capacity (CEC)	1950 to 2013	1980 to 2016			1972 to 2004	1971 to 2004
Soil Salinity		1918 to 2013	1972 to 2004	1950 to 1996	1971 to 2004	1972 to 2004
Soil Fertility	2001 to 2017	1986 to 2016		1950 to 1996	1972 to 2004	
Soil organic carbon (SOC)	1950 to 2018	1905 to 2016		1977 to 2012		

	GEE	ISRIC	EARTH MAP	DAAC	HiH	GloSIS
Soil microbial respiration (SMR)			1981 to Present	1995 to 2000	2002 to present	
Temperature (LST)	2000 to Present		2000 to 2021	1986 to 1995	1981 to near present	
Precipitation	1981 to Present			2008 to 2010	1981 to near present	
Evapotranspiration	1958 to 2023			1895 to 1993 (Annually)	1979 to near present	
Humidity	2000 to 2024					
Soil Erosion	2000		1981 to 2021	1859-08-01 to 2019-08-01 for modeling; 2017-04-07 to 2020-05-15 for imagery		
Cropland Irrigation	2010		2007 to 2012			

These Earth Observation (EO) platforms and soil data resources highlight the significant shift in the way soil is studied. The launch of GEE in 2001 marked a turning point in providing a powerful cloud-based platform for processing and analyzing large geospatial datasets, including those related to soil. The launch of ISRIC's SoilGrids platform in 2017, along with the ORNL-DAAC and FAO-GloSIS at the same year, further revolutionized soil properties data. And finally, the launch of FAO-HiH dataset in 2019 indicates continued advancements in the field of soil mapping. To date, FAO-HiH and GloSIS are still in development to improve the platform and expand datasets available to the public.

4.3.5. Summary of Data Quality

To determine possibility of using EO derived data in assessing agricultural soil health, visual illustration is made against data completeness to determine which platform has a rich dataset in respect to selected soil health indicators and determine the applicability based on the dataset's spatial resolution and coverage (global, regional and national), as well as the required user knowledge to use the platform.

Figure 9 demonstrates that generally, GEE stands out among the other platforms since it provides EO derived datasets in high resolution (30 meters maximum) and offers the highest number of soil health indicator datasets. It may mean that data from GEE with high resolution can be used for soil health assessment on a national scale. The downside of using GEE datasets is that data are processed and coming from different organizations, so if users intend to use it for further analysis, they should consider understanding and checking the methods used in producing the dataset. The compatibility of units used should also be considered before deciding on using the GEE datasets. Another user requirement in opting to utilize GEE dataset, is that user must have familiarity and knowledge on programming/scripts to extensively used the functionality of the platform including the uploading of own data, manipulating off-the-shelf data found in its catalog and do analysis, hence, making GEE more adaptive and flexible to the requirements of users.

Meanwhile, FAO-HiH platform is also a good source of EO data for soil health since it provides datasets to the four SH components except the management factors. It also has datasets with 30meters resolution, but most of its datasets are in 1000meters spatial resolution. This resolution might be more readily available and require less processing time, making it a practical choice for large-scale studies. FAO-HiH platform also has the functionality of making the maps in the platform and sharing it to other users. The downside of using this platform is that users cannot upload their own data nor select a specific area to load in the map interface. However, data downloaded from this platform can be processed in other geospatial applications like Quantum GIS or ArcGIS.

FAO-GloSIS, being the application developed solely for soil properties can be source of SH indicators but only with the physical and chemical properties of soil. The dataset is all in a spatial resolution of 1000-meter and has the same functionalities and limitations as with the FAO-HiH.

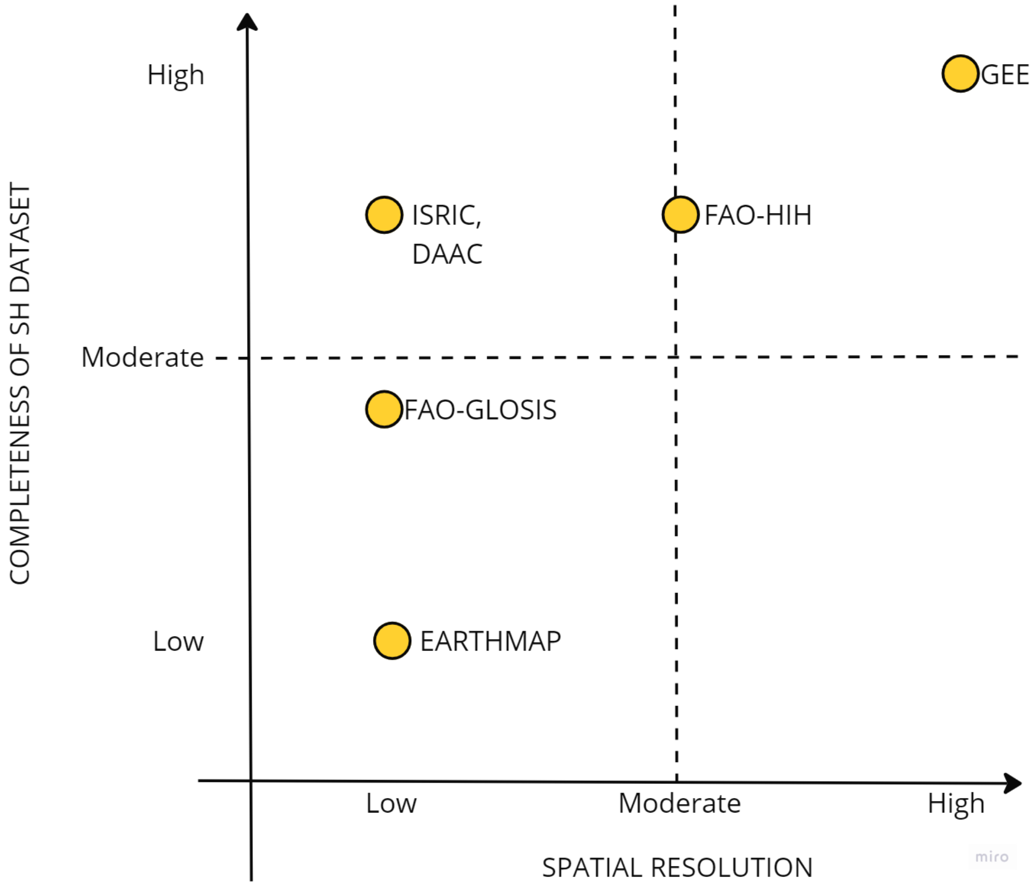

Figure 17. The rate of the completeness of dataset against the available spatial resolution.

Like the FAO-GloSIS, ISRIC is a platform developed mainly for global soil mapping and has most of its datasets focused on the biological, physical and chemical properties in 250 meters resolution except the phosphorus dataset which is in 5 million spatial resolution. ISRIC has also limited functionality, it allows the user to view and download datasets but with its simple functionality, it’s easier for users to navigate around the interface.

Likewise, the ORNL-DAAC has a high number of datasets fit for the SH indicators but has dataset with resolution that ranges from 1 degree (around 100,000 meters) that can only be fit for regional and global scale analysis. With only the functionality of viewing and downloading datasets, it has a more user-friendly and simpler interface.

Lastly, the EarthMap is an EO platform that is used for a wide range of environmental applications, it offers only a limited number of datasets in respect to the soil health indicators identified in this research. However, this has a good functionality of producing charts and graphs using a specified area of users. The spatial resolution of datasets ranges from 250 meters to around 6,000 meters which can be useful even on a national scale.

Overall, the six Earth Observation (EO) platforms have datasets from highly respected organizations like Google, NASA, ISRIC, and FAO. While these platforms have compiled data from diverse sources, they ensure adherence to established data quality standards set by entities like the Open Geospatial Consortium (OGC). Most datasets use machine learning algorithms and model predictions to fill data gaps, especially in areas lacking field data. Some platforms like GEE, FAO-HiH and FAO GloSIS offer readily available scripts and packages for users to understand data processing methods. This consistency guarantees data reliability and facilitates seamless integration between platforms.

Data accessibility is further enhanced by the availability of different downloadable formats across EO platforms, with GeoTIFF being the most universally supported. User-friendliness is another strength of these platforms and has powerful search engines that allow users to locate data using keywords, location parameters, or date ranges.

While a limitation exists in the form of infrequent historical updates (typically offering a single data since 1950), platforms used the historical field data and machine learning to produce maps with predicted values. This approach, combined with the cloud-based nature of these platforms, empowers users anywhere with an internet connection to access valuable soil health information. Additionally, the availability of pre-processed data allows even individuals without prior remote sensing expertise to gain insights into agricultural soil health conditions and have a better management to attain sustainable agriculture land use.

V. CASE STUDY

This chapter outlines the methodology used to assess the effectiveness of Earth Observation (EO) derived data in evaluating soil health in agricultural lands on a regional scale. The case study focuses on selected soil health indicators and data quality elements to identify potential gaps and challenges in using free and readily available EO derived data in Central Asia. This also highlights the importance of these data quality elements in ensuring the reliability and effectiveness of EO derived data for assessing soil health in agricultural lands.

5.1. Study Area – Central Asia

Millions of people in Central Asia depend on fertile soil for their livelihoods, prosperity, and overall well-being. This critical resource, however, faces a growing threat from desertification, land degradation, and drought (DLDD), all worsened by climate change. According to UNCCD data, around a third of Central Asia's land is already degraded, making it one of the worst-affected regions globally (UNCCD, n.d.)The situation is particularly concerning given that around 60% of the population in Central Asia relies on agriculture for their survival and well-being (FAO, n.d.-a). Healthy soil acts as the foundation for food security in the region, ensuring sufficient food production to feed its population and generate income (FAO, n.d.). Low agricultural productivity, however, remains a major hurdle for both economic development and food security in Central Asia (Khitakhunov, n.d.). Addressing these challenges requires a multi-faceted approach that prioritizes sustainable land management practices and promotes a more productive agricultural sector. However, achieving this requires detailed soil knowledge. The massive variability of soil properties across the region, along with the diverse symptoms of degradation, is necessary to identify beforehand. Only by considering this variability and implementing the most appropriate technological and management solutions for each specific area can Central Asia truly achieve sustainable soil health and a productive agricultural sector (Guggenberger et al. 2022).

The Central Asian countries - Kazakhstan, Kyrgyzstan, Tajikistan, Turkmenistan, and Uzbekistan, spans around 3.88 million square kilometers bordered by Afghanistan and Iran to the south, China to the east, and Russia to the west and north (Interstate Commission for Water Coordination, n.d.) as shown in Figure 18. This shows that Central Asia has a vast soil resource but unfortunately, this resource faces significant challenges like desertification and erosion. But with proper management, the soil can be restored and supported, adapting the concept of soil health. Prioritizing sustainable land management practices and adopting a soil health focus, Central Asia can not only restore degraded soils but use it to its full potential - securing the region's future (Guggenberger et al. 2022) on food security.

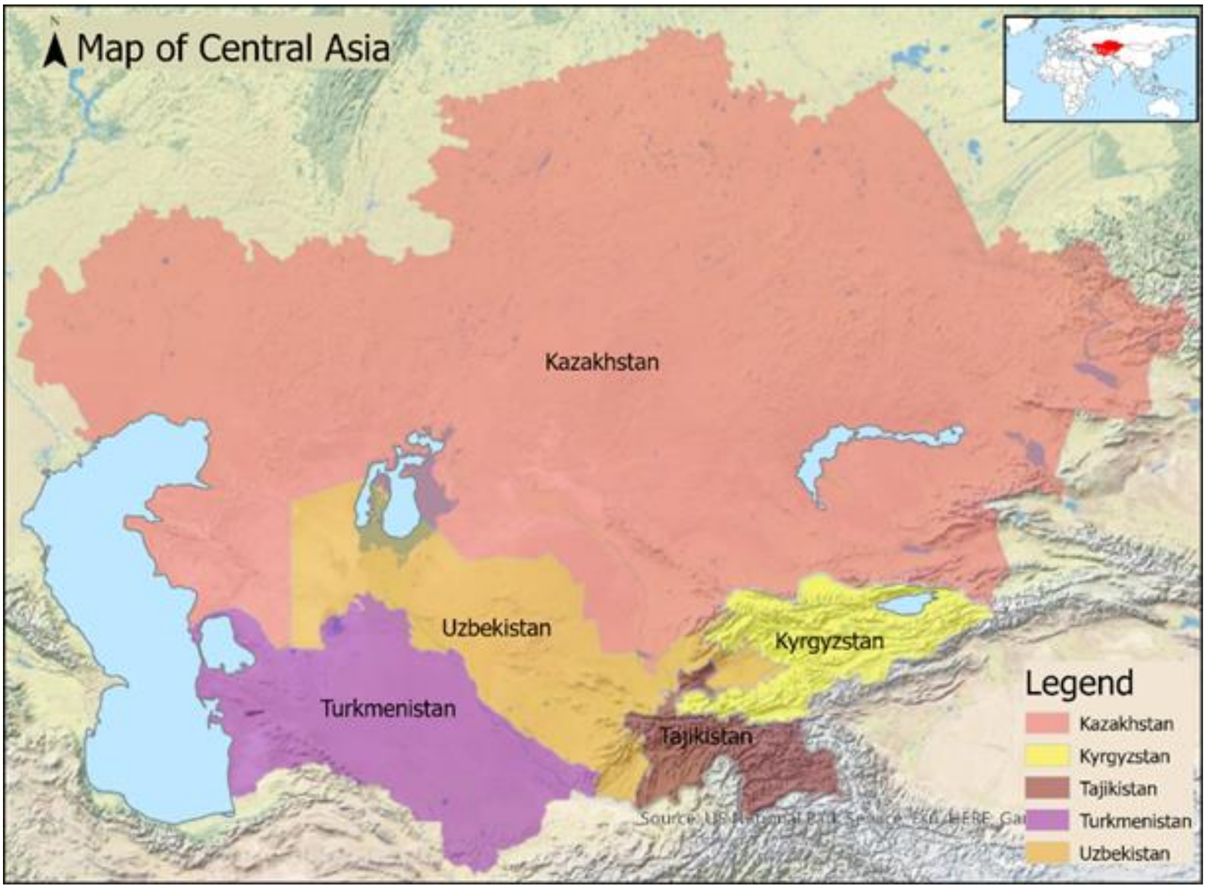

Figure 18. Central Asia Map.

With the identified soil problem in Central Asia region, it is timely to use the readily available EO derived data to assess the soil properties of agricultural soil. These soil properties can serve as indicators on the current agricultural land problems like salinization which is very much related with Cation Exchange Capacity (CEC) and influenced by factors such as pH, organic matter content, and clay texture. These properties can be reflected in maps provided by different EO platforms.

This case study also determines the free and available EO platforms that can be used for assessment of agricultural soil health in Central Asia and identify the needs and data gaps of the region in terms of earth observation derived data or geospatial technology, per se.

5.2. Central Asia Available EO Data Platforms for Soil Health

There are four (4) EO platforms identified that can be used for the assessment of soil health in the region of Central Asia. Two of which are focused in Central Asia region (GeoAgro by ICARDA and the Central Asia Caucasus Geoportal (CAC) that is powered by ESRI) and the other two platforms cover global scale but with resolutions in 250 meters, these are the SoilGrids by ISRIC and Google Earth Engine by Google which has the petabytes of spatial data available to the public. The three (3) platforms are described as follows, Google Earth Engine is described under the Results and Discussion Section 4.2.1:

A. GeoAgro

Geoinformatics Spatial Solutions for Integrated Agro-ecosystems (GeoAgro) is the geo-informatics platform of International Center for Agricultural Research in the Dry Areas (ICARDA) offering Geospatial Science, Technology and Application (GeSTA) solutions on food insecurity and environmental sustainability in dry regions (ICARDA, n.d.). It has datasets on a global, regional, national scale as well as dryland systems defined into North Africa and West Asia, Central Asia, South Asia, Eastern and Southern Africa, and West African Sahel and Dry Savannas. Focusing on themes such as Climate, Soils, Poverty, Agriculture, Suitability, Drought, Land Use/Land Cover and Water harvesting.

This geo-informatics portal is developed using open-source technology and database solutions as seen in Figure 19. The user-interface side uses PHP and MySQL for a simple and customizable experience. A system called "Geo-Store" holds the map metadata and allows preview of product. It is also designed for easy expansion through usage of DDD (Domain Driven Design) over MVC pattern (Model-View-Controller) by multiple developers. For data analysis, extraction and statistical calculation on the backend, it uses Python and the GDAL library. Since this platform is still in development, the next phase will integrate additional open-source tools like MapServer and OpenLayers for a complete solution (ICARDA, n.d.).

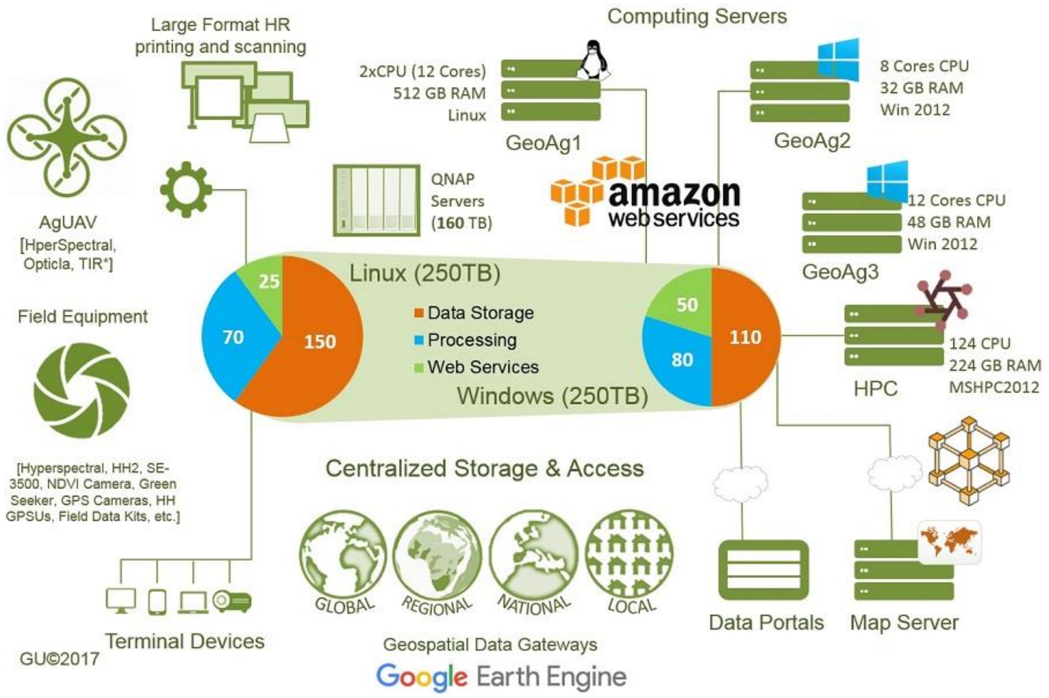

Figure 19.ICARDA Laboratory diagram. (Source: ICARDA)

B. Central Asia Caucasus Geoportal (CAC)

The Central Asia Caucasus Geoportal is the first free platform for Central Asia that allows existing GIS user groups to collaborate effectively and was launched in May 2024. Currently, there are many open source and open data platforms in the region but it's lacking the tools and user training needed to fully utilize advanced functionalities (ESRI 2024). As one of ESRI's commitments, this platform enables communities in Central Asia and Caucasus region to understand the issues using maps and data analytics. Since this geoportal is powered by ESRI, almost all functionalities of ArcGIS online like story maps, survey 123, instant apps, field maps, dashboards, etc are available.

The Central Asia and Caucasus Geoportal is not just a data repository but a powerful tool for addressing real-world processing by providing the geoportal to stakeholders like researchers and GIS professionals with industry-leading spatial analytics and authoritative data and use a geographic approach to critical issues. The CAC allows users to visualize datasets spatially to reveal crucial patterns and trends. Likewise, this geoportal is a platform for collaboration, enabling users to share data, and analyses to the public. This collaborative environment fosters better-informed decisions, leading to more sustainable and impactful solutions for the region.

This geoportal contains data in the categories of agriculture, climatology, meteorology and atmosphere, earth science, physical science and population.

C. Soil Grids

As mentioned in the Results and Discussion of this research, SoilGrids is developed and maintained by ISRIC - World Soil Information. Basically, this is a system that produced digital soil maps based on the global compilation of soil profile data (WoSIS) and environmental layers (Information Soils Grid, n.d.). SoilGrids relies on the large collection of soil profiles available to predict soil properties and processed consistently on a global scale. It is a data-driven system that uses global covariates and globally fitted models. However, it is also suggested to investigate the national soil survey available in the WoSIS database for more detailed information of specific regions.

5.3. Central Asia EO Data Quality Assessment

5.3.1. Intrinsic Data Quality

This Central Asia case study utilizes EO platforms from reputable organizations, including SoilGrids by ISRIC, Google Earth Engine by Google, the Central Asia and Caucasus Geoportal powered by ESRI, and GeoAgro developed by ICARDA. Consequently, the data providers from each platforms are from Environmetrix and ISRIC that provide 250-meter spatial resolution data. Other providers are the GCOM, UCSB/CHG, NASA, NOAA/NWS and NSIDC for the meteorological and management factors datasets that have lower resolution ranging from 5,000 meters to 11,000 meters but in hourly/daily time revisits.

Interestingly, the evaluation revealed that machine learning is the most common method for producing these datasets with most datasets utilizing MODIS data and elevation derivatives as key inputs.

Finally, it's important to highlight that all four platforms utilize ancillary data to generate the final datasets, further enriching the information available for the case study.

Refer to Annexes 9 to 12 for the detailed assessment of each platforms.

5.3.2. Accessibility Data Quality

Considering data accessibility and quality, both Google Earth Engine (GEE) and the Central Asia and Caucasus Geoportal (CAC) are free to use. However, they require user accounts to unlock full functionalities like downloading data and performing overlay analysis. This can be a barrier for some users who might prefer a more immediate, download-only approach.

On the other hand, GeoAgro and SoilGrids are completely free and open to use without any account creation. However, their functionality is limited solely to downloading pre-processed datasets. This offers convenience but limits users who want to perform further analysis requiring advanced tools.

As shown in Table 10, GEE, CAC, and SoilGrids offer various dataset formats suitable for further processing and analysis in external software. These may include formats like GeoTIFF, shapefiles, or VRT that allows users to choose the format that best suits their needs. In contrast, GeoAgro's output is limited to visual formats like JPEG, TIFF, or PDF. These formats are not ideal for advanced analysis as they lack the underlying data structure needed for manipulation. Users who want to analyze GeoAgro data would need to manually extract information from these visuals, which can be time-consuming and prone to human errors.

Only CAC and GEE offer functionalities for secure storage and use of personal data which can be beneficial for users working with sensitive information. However, it's important to note that both platforms require some knowledge of spatial data analysis to fully utilize their capabilities. Users who are unfamiliar with these concepts might need to invest additional time in learning the platform or seek external help for complex analysis.

Table 10. Accessibility Data Quality of each EO platforms.

	GEOAGRO	CAC	SOILGRIDS	GEE
Data Access	Free to download	Free but needs User Account.	Free to download	Free but needs User Account.
Product Format	JPEG, TIFF, PDF	API, KMZ, VRT format, ArcGIS web layer and Export map to PDF	WMS, WCS, WebDAV, VRT, OVR, GeoTIFF,CSV, Geopackage, Shapefile, GML	GeoTIFF, TFRecord, Feature Collection as CSV, SHP, GeoJSON, KML, KMZ and Map Tiles
Security of Data	No function to upload data.	Allows upload of own dataset to be used for analysis without sharing to public	No function to upload data.	Allows upload of own dataset to be used for analysis without sharing to public

Despite the differences in functionalities and data formats, all four platforms offer free and easy access to various datasets in formats catering to different user needs. Some platforms even provide built-in processing tools, further streamlining the workflow for those with the required expertise. The ideal platform choice depends on the user's specific needs – whether it requires maximum flexibility for analysis, GEE or CAC might be better, while for quick data access without analysis needs, GeoAgro or SoilGrids could be sufficient.

Refer to Annexes 9 to 12 for the detailed assessment of each platforms.

5.3.3. Representational Data Quality

As it relates to the global EO web platform assessment, GEE provides the most extensive functionalities. Users can upload their own data, manipulate, analyze, visualize, and download it using the code editor with JavaScript or Python. This flexibility offers powerful analysis capabilities but requires basic scripting knowledge, potentially hindering accessibility for users that are non-programmers.

As shown in Figure 21, the second platform with good functionalities is CAC. Powered by ESRI, CAC offers tools similar to ArcGIS Online, enabling users to create dashboards, map stories, online surveys, and web applications. This graphical user interface (GUI) eliminates coding needs but requires familiarity with the interface and steps involved. Noticeably, some map functionalities are currently under maintenance, that affects data access and visualization.

On the other hand, SoilGrids is a user-friendly platform, allowing users to easily extract datasets for further analysis, if needed, with minimal technical expertise. However, manipulating, overlaying, or conducting complex analysis needs additional software.

As observed, GeoAgro is designed for quick data access but has the least functionality. While users can directly click and download maps, these maps are in formats (JPEG, TIFF, PDF) unsuitable for advanced analysis. Additionally, GeoAgro lacks the “Search” button to easily look for available datasets. This can be time consuming and not practical for users as they need to go over the long list to find the dataset needed.

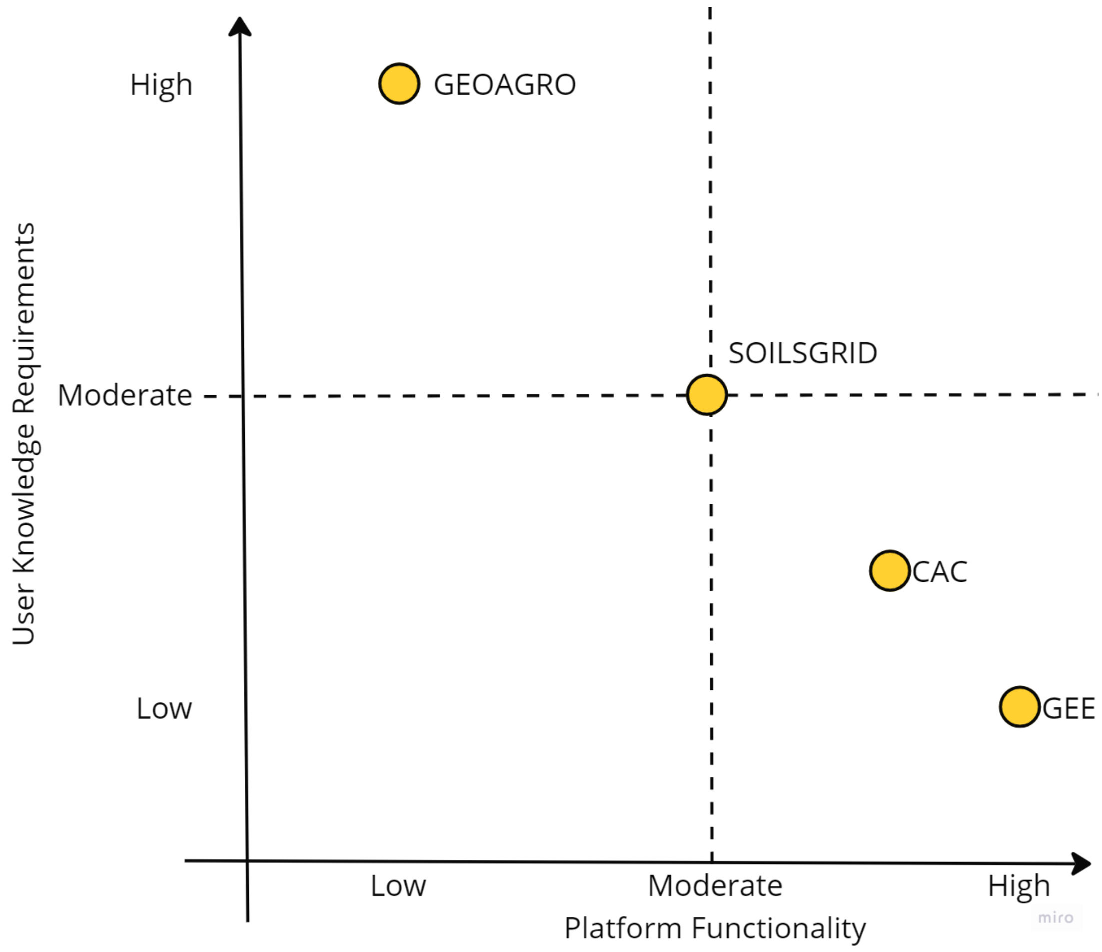

Figure 20. The user knowledge requirements and platform functionality of each EO platform.

As it is reflected in Table 11, all platforms use cloud storage for processing data. CAC and GEE perform analysis entirely in the cloud. Meanwhile, CAC even if it acts as a repository, the product manual is absent, potentially hindering analysis without understanding the data derivation methods. But both SoilGrids and GEE, as global data repositories, offer readily accessible user and product manuals.

In terms of user support, only GeoAgro lacks a community forum where users can directly seek help. While it claims user-friendliness, open communication with platform owners or data developers remains important for users to utilize the dataset and web platforms exhaustively.

Table 11. Representational Data Quality of each EO Platforms.

	GEOAGRO	CAC	SOILGRIDS	GEE
Processing Tools	No	Yes. Allows you to select your own area	No	Requires basic knowledge of programming.
User Technical Knowledge Requirement	Familiarity of interface	Familiarity with the interface	Familiarity with interface	Yes. Use built-in codes in javascript for data manipulation
Programming Languages	No	No	No	Python and Javascript
Cloud Processing	Yes	Yes	Yes	Yes
User Interface (UI)	Difficult to search specific datasets.	Yes. Easy to navigate and find datasets.	Yes. Easy to navigate and find datasets.	Yes. Easy to navigate and easy to find different datasets.
Visualization Tools	No. Only visualize datasets	Yes. Allows customization of maps.	No. Only visualize datasets	Yes. It has Google Code Editor to view and manipulate data.
Scalability	Yes	Yes	Yes	Yes
Open Standards	Follows FGDC Standard	Yes	Yes	Yes
User Support	Yes	Yes	Yes	Yes
Product User Guide/Manual	No	No	Yes	Yes
Community	None	Yes	Yes	Yes

Considering CAC's recent launch in May 2024, it's understandable that functionalities are under development, and complete data documentation is unavailable yet. However, its free and open access focused on the Central Asian region demonstrates its strong potential as source for researchers and people interested in the region.

This analysis highlights how diverse functionalities across platforms can translate to varying user knowledge demands. By presenting these differences of web platforms, this guide empowers users to select the platform that best suits their specific needs and technical skills.

Refer to Annexes 9 to 12 for the detailed assessment of each platforms.

5.3.4. Contextuality Data Quality

Figure 22 illustrates the available spatial resolution of each EO dataset for various soil health indicators. It shows that GeoAgro offers datasets for biological, physical, chemical, and meteorological factors within a 1,000-meter resolution. This scale provides regional details but might be insufficient for national or site-specific analysis.

As previously discussed, SoilGrids offers soil properties at a 250-meter resolution. This data can be a valuable supplement for national assessments, particularly in areas lacking soil surveys within the World Soil Information Service (WOSIS).

Similar to SoilGrids, GEE provides soil health indicator datasets for physical, chemical, and biological properties at 250-meter resolution. However, meteorological data ranges from 500 meters to a coarser 11,000-meter resolution. Even if the spatial resolution isn't ideal, these meteorological datasets have high temporal resolution of hourly updates.

The CAC platform uses the SoilGrids data for physical, chemical and biological soil properties with 250 meters spatial resolution. However, meteorological and management factors have a coarser resolution of around 28,000 meters but compensated by the 3-hourly temporal resolution. Overall, CAC provides good resolution datasets that is customized for Central Asian coverage.

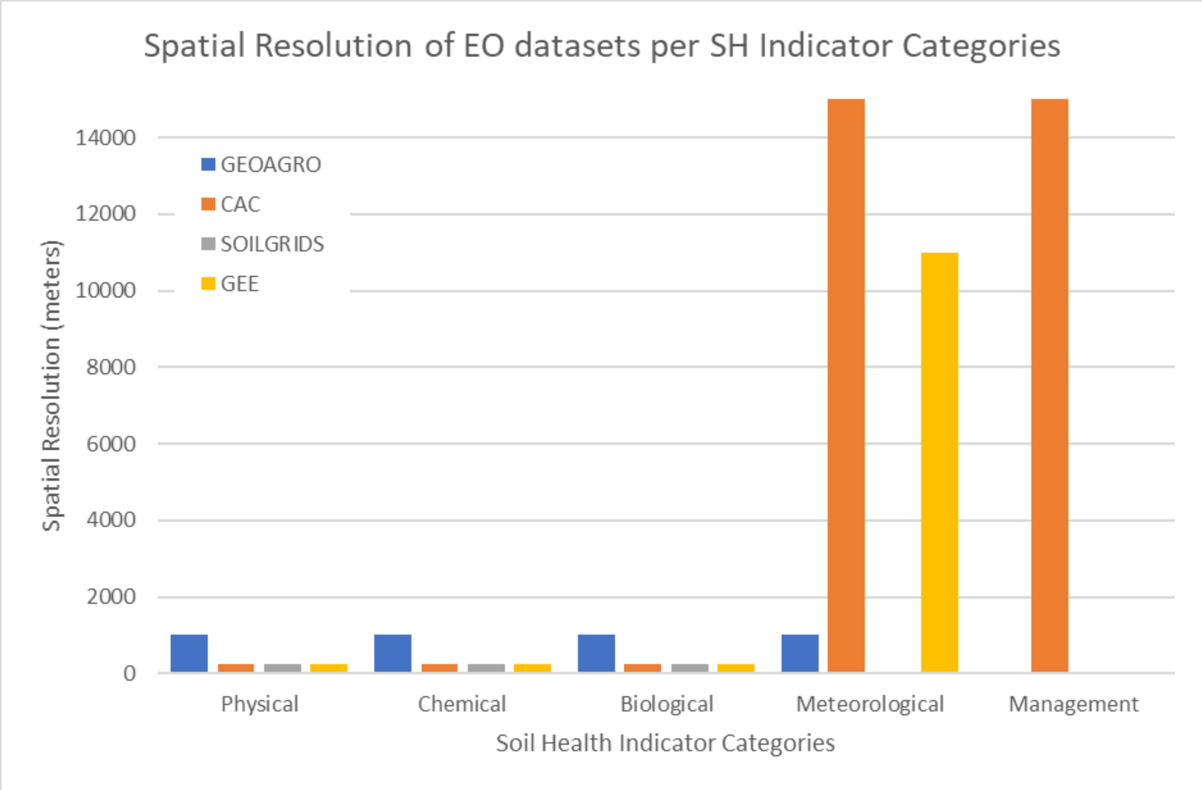

Figure 21. The spatial resolution of EO data against SH Indicator Categories.

Beyond spatial resolution, data update frequency is another critical factor for EO data soil health assessments. Currently, global soil datasets often lack frequent updates, with some indicators having only a single dataset since 1950. This highlights the need for new soil property data to effectively monitor soil health changes.

As mentioned, both CAC and SoilGrids datasets rely on ISRIC data, with training data for machine learning algorithms spanning 1950 to 2016. Conversely, GEE utilizes Environmetrix datasets with training data from 1950 to 2018. While these existing datasets might not be ideal for monitoring ongoing soil health changes, they can still serve as valuable baselines for future assessments.

In conclusion, this analysis emphasizes the importance of considering both spatial resolution and data update frequency when selecting EO datasets for soil health assessments. Different platforms offer varying levels of detail and temporal coverage, requiring users to choose the one that best aligns with their specific needs and desired level of monitoring.

Refer to Annexes 9 to 12 for the detailed assessment of each platforms.

5.3.5. Overall Data Quality

Figure 23 serves as a guide for selecting the optimal platform based on the balance between data completeness and platform functionality. Platforms positioned in the upper right quadrant indicate that it excels in both aspects. The CAC and GEE occupy the high range, offering comprehensive EO datasets (including physical, chemical and biological properties, meteorological and management factors) suitable for conducting detailed agricultural soil health assessments. Additionally, CAC and GEE provide functionalities for analysis and manipulation, making it ideal for users with diverse needs.

Meanwhile, SoilGrids falls into a “good balance” zone. Although its datasets focus lies primarily on physical, chemical, and biological soil properties, it offers high-quality data with a good user-friendly interface. This makes it a valuable resource for users prioritizing data credibility and ease of use over extensive functionality.

GeoAgro as it is situated in the lower quadrant, it presents a unique property. Despite its focus on Central Asia, the data output format (JPEG, TIFF, PDF) impedes its use for advanced analysis. While its user-friendliness might be appealing, increasing the data output type flexibility would significantly enhance its value, especially for researchers.

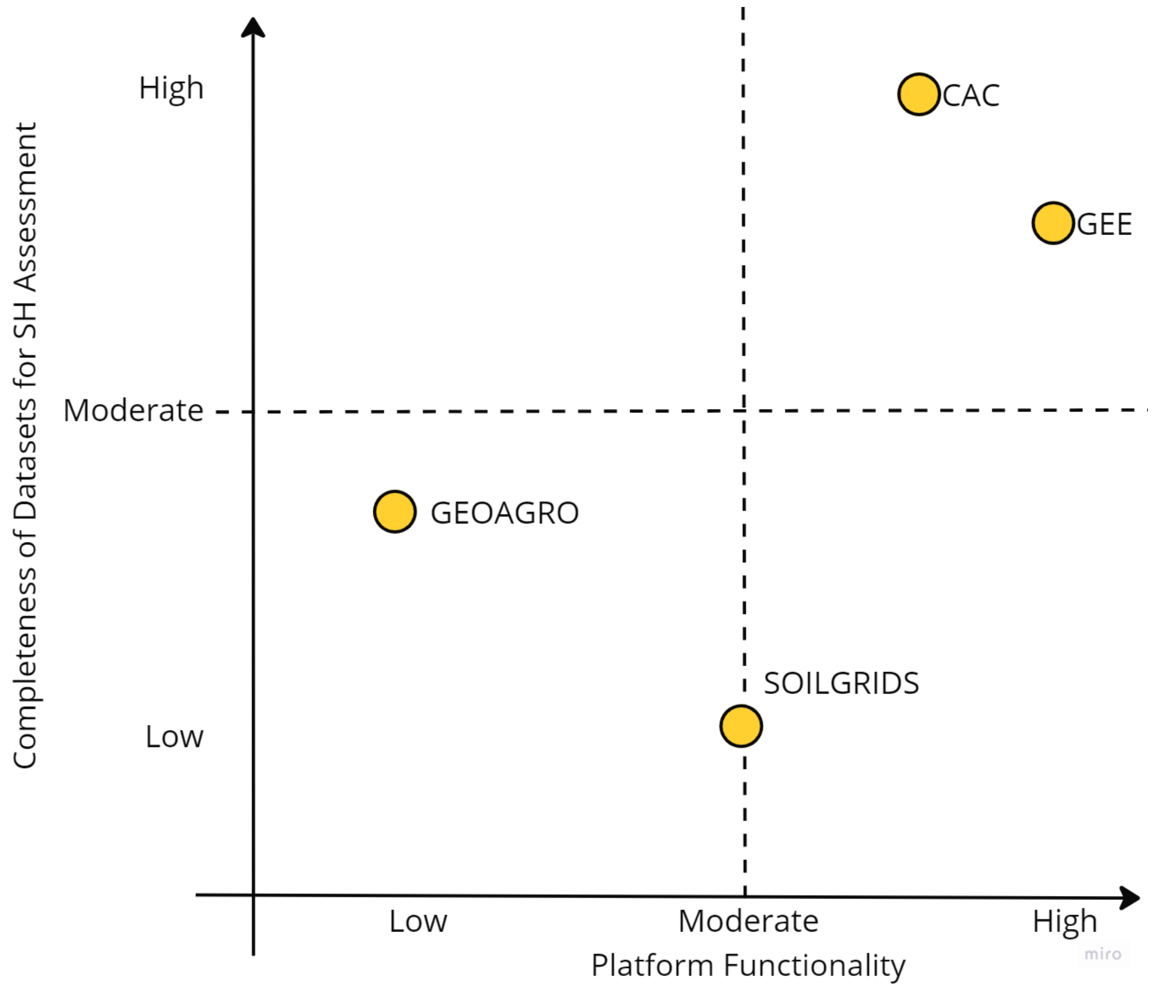

Figure 22. The positional rate of SH Indicator dataset against available platform functionality.

Hence, the optimal platform choice ultimately depends on the user's specific needs. But in this analysis, if the user wants an in-depth analysis requiring data on various soil health indicators (including meteorological and management factors), CAC or GEE are the best options. Likewise, if users intend to have data for advanced analysis and research purposes, platforms like CAC or GEE offer flexible output formats suitable for further manipulation and modeling. And if the primary concern is obtaining high-quality data on physical, chemical, and biological soil properties, and user-friendliness of platform is important, SoilGrids can be used.

Figure 24 highlights another important factor - the interplay between data completeness and user experience (ease of use). The ideal platform depends on the user's technical skills and intended use. The GeoAgro platform prioritizes user-friendliness. Downloading data in map format is straightforward. However, searching for specific datasets is burdensome due to the lack of filtering functionality and must be noted that data are in a coarser resolution of 1,000-meter. Despite these limitations, maps generated by GeoAgro are compatible with other datasets produced by this platform because it uses the same algorithms and scale. SoilGrids on the other hand, offers a user-friendly interface with good spatial resolution (250 meters) and easy data search capabilities. However, it focuses solely on soil properties data, limiting its application. The CAC has the most comprehensive data offering and a relatively user-friendly interface with clickable functionalities. Conversely, it currently lacks product manuals and thorough metadata for users, with some functionalities observed to be down or under maintenance during the research period. GEE is a tool that has high functionalities and vast dataset collection. However, it can only be operated using codes. Even though there are already pre-written available codes for specific datasets, users must still have knowledge on programming to maximize the full potential of the platform.

Selection of platform for soil health assessment considers trade-off between different data quality elements. The user-friendliness of the platform is one of the best basis for determining a suitable platform. If the user's priority and only data needs are basic soil data for visualization purposes, then GeoAgro or SoilGrids might be sufficient. These platforms offer easy access but have limitations in data scope. For in-depth analysis requiring comprehensive datasets that include meteorological and management factors, the Central Asia and Caucasus Geoportal (CAC) or Google Earth Engine (GEE) can be used despite the coding requirements for GEE or potential maintenance issues with CAC.

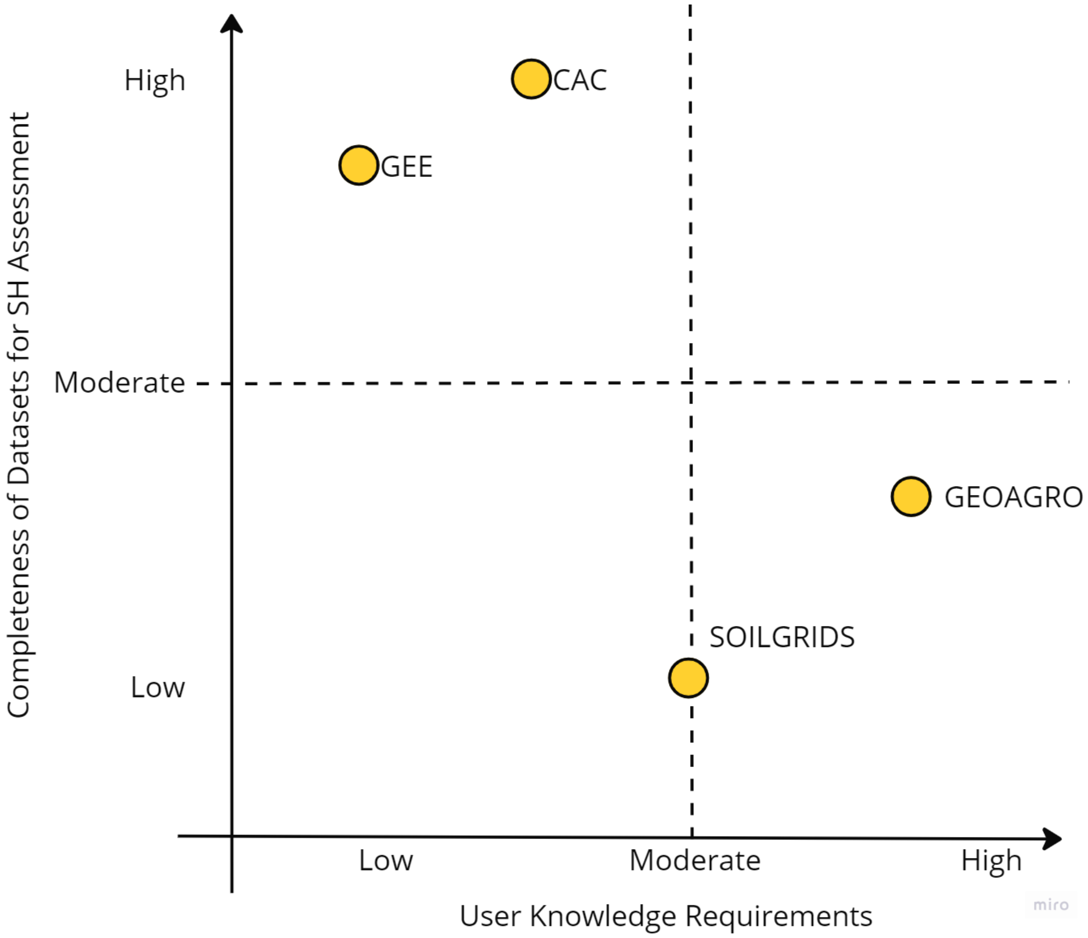

Figure 23.The positional rate of SH Indicator datasets against the user knowledge requirements.

Finally, if users want a balance between user-friendliness and data availability, SoilGrids could be a good platform, especially if the focus is primarily on soil properties. Understanding the user's technical skills and the complexity of intended analysis is crucial for selecting the platform that best suits the conduct of a successful soil health assessment.

By understanding the strengths and limitations of each platform, as highlighted by Figures 23 and 24, and this analysis, users can make an informed decision and select the tool that best facilitates the users specific soil health assessment needs.

VI. CONCLUSION

This research explored the potential of using readily available, free Earth Observation (EO) data for assessing soil health in attaining sustainable agriculture. It highlights the strengths and weaknesses of EO-derived data as a user-centric approach for evaluating agricultural soil health.

The research adopted a "Management Goal Approach" to identify soil health indicators that could be effectively assessed using freely available EO data. Since the goal was sustainable agriculture, the chosen indicators focused on physical, chemical, and biological soil properties. However, meteorological and management factors were also considered due to their impact on soil health. This approach provides a framework for users to select soil health indicators relevant to their specific needs.

The user's perspective approach was used to develop a framework for evaluating EO data quality and platform performance. This framework assessed intrinsic characteristics, context, representation, and accessibility of web platforms, resulting in the development of a comprehensive inventory of freely available EO data products suitable for soil health assessment (See Annex 3 to 12). This inventory characterizes data quality and platform capabilities, facilitating informed decision-making when choosing the most suitable data and platform for user needs. The assessment was conducted on both global and regional scales, with a specific focus on Central Asia.

The results revealed there's a limited number of EO platforms that offer free, readily available or "off-the-shelf" global datasets for soil properties. Only six out of twenty-seven platforms met this criteria that was evaluated for their data element qualities. Similarly, in the Central Asian case study, only four web platforms were found that provides open and free access for public users. Overall, the research identified potential limitations and biases in these freely available EO data products. These included limitations particularly in spatial resolution to which many datasets lacked high resolution, making it unsuitable for site-specific evaluations; and the limited temporal coverage where most datasets are presented only in a single year, even if historical data was available since 1950. Up-to-date data is highly important for both assessment and monitoring. Understanding these limitations is crucial for users to accurately interpret data, select appropriate platforms based on their needs, and avoid misinterpretations when assessing soil health.

By highlighting the potential of freely available EO data, this research can contribute to developing cost-effective methods for assessing soil health in agricultural lands across extensive areas. This can significantly advance sustainable agriculture for researchers, practitioners, and farm planners by providing insights into agricultural soil conditions over large areas. Overall, this research shows the potential of EO data and its practical application in soil health assessment, ultimately promoting data-driven decision making for sustainable agricultural land management.

VII. RECOMMENDATIONS

This research offers a valuable foundation for further exploration and research. To refine the user-centric approach, it is recommended to conduct surveys among various data-user groups, including researchers, EO professionals, NGOs, and government agencies, etc. This would provide invaluable insights into the specific needs and gaps experienced by different data user groups.

The evaluation of existing platforms revealed a small number of platforms with soil datasets, and only provide one-time data publications. To address this limitation, it's crucial to encourage organizations to develop time series data for soil properties, particularly focusing on chemical characteristics that are highly sensitive to agricultural management practices.

In the context of Central Asia, where desertification is a pressing concern, continued development of free and accessible EO platforms and tools focused on soil analysis is essential. These tools can empower decision-makers and governments to implement best practices for sustainable agriculture and combat desertification. Furthermore, incorporating time series soil mapping capabilities into EO systems would be a valuable addition for the region, as it would allow monitoring of soil health changes over time and inform more effective long-term strategies.

ANNEXES

Annex 1: EO Platforms Available

Platform	WAPOR	AQUAMAPS	Hand-in-Hand Geospatial Platform	Global Surface Water Explorer	GEOGLAM Crop Monitor
Website	https://data.apps.fao.org/wapor/?lang=en	fao.org/aquastat/en/geospatial-information/aquamaps/	https://www.fao.org/hih-geospatial-platform/en/	Global Surface Water Explorer (global-surface-water.appspot.com)	https://www.cropmonitor.org/
Description	FAO's portal to monitor Water Productivity through Open access of Remotely sensed derived data. The WaPOR project aims to assist partner countries in developing their capacity to monitor and improve water and land productivity in agriculture, both rainfed and irrigated, responding therefore to the challenges that are posed by the dwindling of fresh-water resources and the need to sustain agri-cultural production to ensure food security in the face of a changing climate. The first output of the project is the WaPOR database and portal, which provides open access to near-real time information on key land and water variables.	AQUAMAPS is AQUASTAT's online geospatial database on water and agriculture. Through a sophisticated web platform, regional and global spatial datasets on water resources and water management, produced by FAO and by external data providers, are made accessible.	FAO's open-access Hand in Hand (HIH) Geospatial Platform provides advanced information, including food security indicators and agricultural statistics, for more targeted agriculture interventions. The platform unlocks millions of data layers from different domains and sources to serve as the key enabling tool for FAO's HiH Initiative and serve digital agriculture experts, economists, government and non-government agencies, and other stakeholders working in the food and agriculture sector. The data has been sourced from FAO and other leading public data providers across the UN and NGOs, academia, private sector and space agencies, including key FAO flagship databases such as FAOSTAT data on food and agriculture for over 245 countries and territories from 1961 to the most recent year available.	The European Commission's Joint Research Centre developed this water dataset in the framework of the Copernicus Programme. This maps the location and temporal distribution of water surfaces at the global scale over the past 3.8 decades and provides statistics on the extent and change of those water surfaces. The dataset, produced from Landsat imagery, will support applications including water resource management, climate modelling, biodiversity conservation and food security.	The Crop Monitors were designed to provide open, timely, and science-driven information on crop conditions to support market transparency for the G20 Agricultural Market Information System (AMIS). The Crop Monitor for AMIS brings together over 40 partners from various monitoring systems, agencies, organizations, and universities to provide a consensus assessment of crop growing conditions, status, and agro-climatic factors that may impact global production, focusing on the major producing and trading countries for the four primary crops monitored by AMIS (wheat, maize, rice, and soybeans). Since its launch in 2013, the Crop Monitor for AMIS has become an internationally recognized source of information on global crop prospects, widely quoted by public and private agencies as well as top-tier media.
Available/ Used Sensors and Radars. Return Period	MODIS (1day), PROBA-V (2 days), Landsat (16 days), MSG (1day), TRMM (1day), GPM (1day), MERRA/GEOS (1day)	NASA's Shuttle Radar Topography Mission (SRTM), GTOPO30, HydroSHEDS and Hydro1K, ECMWF ERA5 data	Sentinel, Landsat, MODIS. Other datasets are derived from other organizations accompanied with the metadata.	Landsat 5 (16 days), Landsat 7 (16 days) and Landsat 8 (8-day combined repeat cycle, with each satellite having a 16-day repeat cycle offset by 8 days)	Landsat 8 (8-day combined repeat cycle, with each satellite having a 16-day repeat cycle offset by 8 days), Aqua, Terra (16days)
Spatial Resolution	100m, 250m, 30m	15 arc-seconds between 60 N and 60 S latitude, 5 arc/min, 19km	1km, 250m, 100m, 30m, 10m	100m, 60m, 30m	100m, 60m, 30m
Spatial Coverage	The continental-level data (250m) covers continental Africa and large parts of the Near East (L1). The national-level data (100m) covers 21 countries and four river basins (L2). The third level (30m) covers eight irrigation areas (L3).	Global, Regional, Some selected countries	68 countries in Africa, Asia, Europe, Latin America, and the Middle East	Global	Global
Temporal Coverage	Latest version was October 4, 2023 on the release of WAPOR 3.0. For Africa and the Near East in near real-time covering the period from 01-January-2009 to present	Monthly, Annually. Some datasets started on 1960s	Not mentioned. Dependent on the FAOSTAT API Updates.	Annually (1984 -2021)	1999 - Present (Near real-time)
Data Access	Downloadable and can do data analysis if you have account.	Downloadable if you have account .	Downloadable if you have account .	Downloadable and data is open for public.	Downloadable if you have account .
Ease of Use (low, Moderate, High)	High. Not much processing to be done by users. Directly select the area and then download.	High. Not much processing to be done by users. Directly select the area and then download.	High. Not much processing to be done by users. Directly select the area and then download.	Moderate. Not much processing to be done by users to get image of data. Directly select the grid for area of interest and then download tiff. But if needed dataset, need to run python script or download the filezilla client.	
Processing Abstraction (Cloud processing)	Off the shelf data powered by GEE	Off the shelf data.	Off the shelf data.	Off the shelf data powered by GEE	
Use of Ancillary Data for validation	Yes	Yes	Yes	Yes	
Output Type	Tiff, WMS, With QGIS Plug-ins, Some datasets are available in GEE public catalogue	WMS, WFS, Shapefile,	Tiff, WMS, CSV, Shapefile	TIFF, WMS, GEE Catalog, Dataset can be downloaded using FME or running the python script that can be found in github (https://github.com/mentaljam/download_water_data/tree/master/download_water_data)	JPEG, TIFF
Language of Platform	English, Spanish, French	English, Spanish, French	English, Spanish, French, Russian, Chinese	English	English
Platform Availability	Free and Open. Need to create user account.	Free and Open. Need to create user account.	Free and Open. Need to create user account.	Free and Open. No need for user account.	Free and Open. Need to create user account.
Platform Owner	FAO	FAO	FAO	JRC Europa	Group on Earth Observation
Platform Partners	IHE-Delft Institute for Water Education, and the International Water Management Institute (IWMI), University of Twente with funds from Ministry of Foreign Affairs Netherlands	WWF , International Water Management Institute (IWMI)	Data has also been sourced from World Meteorological Organization (WMO), the United Nations Environment Programme (UNEP), and the UN Geospatial Unit, FAO partners and public data providers across the UN and NGOs, private sector and space agencies.	USGS and NASA for the satellite images	Data provided by USGS. University of Maryland.
Availability of Process Documentation	Yes	Yes	Yes	Yes.	
Availability of User Guidelines	Yes	Yes	Yes	Yes.	

Platform	Copernicus Global Land Services (CGLS)	EarthExplorer (EE)	Sentinel Hub (EO Browser)	Global Forest Watch	Global Irrigated Area Mapping (GIAM)
Website	https://gbov.land.copernicus.eu/dataaccessLP/	EarthExplorer (usgs.gov)	Explore (sentinel-hub.com)	https://www.globalforestwatch.org/	Global Irrigated Area Mapping (GIAM) - Dataset - waterdata
Description	The Copernicus Land Monitoring Service provides geographical information on land cover and its changes, land use, ground motions, vegetation state, water cycle and Earth's surface energy variables to a broad range of users in Europe and across the World in the field of environmental terrestrial applications.	The EarthExplorer (EE) user interface is an online search, discovery, and ordering tool developed by the United States Geological Survey (USGS). EE supports the searching of satellite, aircraft, and other remote sensing inventories through interactive and textual-based query capabilities.	Sentinel Hub is a multi-spectral and multi-temporal big data satellite imagery service, capable of fully automated archiving, real-time processing and distribution of remote sensing data and related EO products. Users can use APIs to retrieve satellite data over their AOI and specific time range from full archives in a matter of seconds.	Global Forest Watch (GFW) is an online platform that provides data and tools for monitoring forests. By harnessing cutting-edge technology, GFW allows anyone to access near real-time information about where and how forests are changing around the world.	This site presents results of IWMI's first attempt to map global irrigated and rainfed croplands for the nominal year of 2000
Available/ Used Sensors and Radars. Return Period	PROBA-V(2 days) , Sentinel 2 (5days) , Sentinel 1 (12 days) , Sentinel 3 (1day), Landsat 5 (16 days), Landsat 7 (16 days), Landsat 8 (8 days), DEIMOS-2 (2days), Pleiades (5ays), GeoEye-1 (3days), SPOT 5, SPOT 6, SPOT 7, WorldView-2	Landsat Collection Landsat 1-3 (18 days) Landsat 4-6 (16 days) Landsat 7-8 (8 days), MODIS, Commercial Data Purchases (CDP) Imagery	Sentinel-1, Sentinel-2, Sentinel-3, Sentinel-5P, Landsat 1-5 MSS L1, Landsat 4-5 TM, Landsat 7 ETM+, Landsat 8-9, Harmonized Landsat Sentinel, MODIS, DEM, Copernicus Services, Proba-V, GIBS, Planet NICFI,	Landsat 7, VIIRS, MODIS, Global Ecosystem Dynamics Investigation (GEDI) lidar	AVHRR, SPOT 1, MODIS, GTOPO 30, CRU
Spatial Resolution	2.5m, 5m, 10m, 20m, 30m, 60m,100m, 250m,300m, 500m, 1km, 5km, 12.5km, 5m x 20m, 100m x 100m, 50km x 50km	0.5m, 10m, 15m, 30m, 60m,100m, 250m, 500m, 1km	5m, 10m, 15m, 30m, 60m,100m, 250m, 500m, 1km. Allows uploading of own images.	10m , 60m, 375m, 250m, 300m, 1km, 10km	10km, 1km, 500m, 30m
Spatial Coverage	Global, EEA and the UK, Europe, Selected Sites Globally	Global, Site specific	Global, Europe, Germany, Specific sites requested.	Global, Some selected countries	Basin, Regional, Global
Temporal Coverage	Daily, 10-days, Monthly, Annually	Daily, Monthly, Annually. From 1980 to present.	Daily, Monthly, Annually.	Near Real Time, Daily, Monthly	Year 2000
Data Access	Downloadable if you have account .	Downloadable if you have account .	Free download in 30 days.	Downloadable if you have account .	Downloadable.
Ease of Use (low, Moderate, High)	High. Directly select the area and then download.	High. Directly select the area and then download.	High. Directly select the area and then download.	High. Directly select the area and then download.	High. Directly select the area and then download.
Processing Abstraction (Cloud processing)	Off-the-shelf dataset	Available Off-the-shelf dataset and raw images.	Processing done in cloud.	Off-the-shelf dataset and cloud data analysis.	Off-the-shelf dataset
Use of Ancillary Data for validation	Yes		Yes	Yes	Yes
Output Type	GeoTIFF, Network Common Data Form (NetCDF), File Geodatabase (FGDB), Geography Markup Language (GML), Geopackage (GPKG), GeoJSON, Shapefile (SHP) ,Some datasets are available in GEE public catalogue	Geotiff, xml, txt	WMS, WFS, WCS, WMTS, GeoTiff	Statistics in csv or excel files. GeoTIFF, Shapefile, JPEG, and DBF formats	TIFF, TFW
Language of Platform	English	English	English	English, Chinese, French,	English
Platform Availability	Free and Open. Need to create user account.	Free and Open. Need to create user account.	Free download in 30 days.	Free and Open. Need to create user account.	Free and Open.
Platform Owner	Copernicus	US Geological Survey	Planet Labs	WRI	IWMI
Platform Partners	Implement jointly by European Environment Agency (EEA), Eionet Action Group on Land monitoring in Europe (EAGLE) and the DG Joint Research Centre of the European Commission	AmericaView, Boston University (BU), Committee on Earth Observation Satellites (CEOS), Global Environmental Remote Sensing (GERS) Laboratory, High Plains Regional Climate Center, University of Nebraska, Lincoln International Charter, Space and Major Disasters, NASA - Applied Sciences Program, NASA - Goddard Space	Supported by Amazon Web Services, ESA, Europen Union and Republic of Slovakia, European Union, EvoLand, Global Earth Monitor, AI4EO by ESA, LandSense Project, Odine, EO-VAS project, Perceptive Sentinel, Open EO, EnviroLENS, EOSC Hub, Opertus Mundi, Dione, Sen4CAP Project, NIVA Project	Founding Partners: Bobolink Foundation, Blue raster, Carto, Center for Global Development, Danida, ESRI, GEF, Google, Imazon, Ministry of Foreign Affairs of the Netherlands, Norwegian Ministry of Climate and	World Bank, World Bank Water Data, Global Water Security and Sanitation Partnership
Availability of Process Documentation	Yes	Yes	Yes	Yes	Yes
Availability of User Guidelines	Yes	Yes	Yes	Yes	Yes

Platform	Global Agro-Ecological Zones	Dynamic World	My Fire Watch	Global Mangrove Watch (GMW)	ISRIC
Website	https://gaez.fao.org/	https://dynamicworld.app/explore	https://myfirewatch.landgate.wa.gov.au/about.html	https://www.globalmangrovewatch.org/	https://www.isric.org/explore
Description	AEZ utilizes established land evaluation principles to assess natural resources, identifying resource limitations and opportunities based on plant characteristics, climatic, and soil requirements to evaluate crop suitability and production potentials under specific conditions. This framework aids in making informed planning decisions by managing agro-ecological constraints and determining optimal crop choices for productivity, sustainability, and resilience to climate variability. The AEZ methodology has contributed significantly to various United Nations Sustainable Development Goals (SDGs) by providing insights into agricultural production risks, irrigation water demand, and crop development options, aligning with global efforts to address food security and sustainable land use practices. The evolution of the AEZ assessments from the 1980s to the latest version, GAEZ v3, reflects advancements in computing and technology, enabling detailed global resource evaluations and supporting agro-ecological development on multiple scales to achieve SDGs effectively.	Dynamic World is a near realtime 10m resolution global land use land cover dataset, produced using deep learning, freely available and openly licensed. It is intended to be used as a data product for users to add custom rules with which to assign final class values, producing derivative land cover maps.	MyFireWatch is an online map-based tool that provides important information about hotspots to emergency service managers and the general public. FireWatch is a suite of fire monitoring products, services and solutions developed by Landgate's Imagery team. FireWatch uses satellite imagery to detect and report on hotspots as they are observed and monitor their subsequent effect on the land.	The Global Mangrove Watch is an online platform that provides the remote sensing data and tools for monitoring mangroves. The tool provides near real-time information on where and what changes there are to mangroves across the world, and highlights why they are valuable.	The International Soil Reference and Information Centre (ISRIC) acts as the global custodian of soil information with the mission of raising awareness and understanding of soil on global challenges. This is achieved by providing soil data and knowledge at various levels (global, national, sub-national) to promote sustainable soil and land management practices
Available/ Used Sensors and Radars. Return Period	Not explicitly mentioned.	Sentinel (2-5 days)	NOAA-18, NOAA-19, MetOp-B, MetOp-C, Terra, Aqua, SNPP, JPSS-1 (also known as NOAA-20), FY3-D	Combination of L-band radar (ALOS PALSAR) and optical (Landsat-5 and Landsat-7) satellite data, Sentinel-2, JERS-1 SAR, ALOS PALSAR and ALOS-2 PALSAR-2	MODIS land products, SRTM DEM derivatives, climatic images and global landform and lithology maps
Spatial Resolution	5 arc-minute, 30 arc-second	10m	30cm, 10-30m, 60m-2km	1 x 1 degree, 20m, 0.8 arc seconds	250 meters
Spatial Coverage	Basin, Country, Sub-national, Continental, Global	Global	Australia	Global	Global except antarctica
Temporal Coverage	1961 - 2100. Projections are made.	Started on 2015. Near-real tiem data.	2-4 days Updating. Near real time viewer.	1996, 2007, 2008, 2009, 2010, 2015, 2016, 2017, 2018, 2019, 2020	2020 Single Dataset with data used from 1905 to 2016
Data Access	Downloadable if you have account .	Downloadable if you have account .	Only for viewing. WA public sector agencies and Local Governments can download.	Downloadable	Downloadable
Ease of Use (low, Moderate, High)	High. Directly select the area and then download.	Low. Download the dataset in GEE. User must have basic knowledge on scripting.	High. Easy to navigate around the interphase.	High. Straight forward download	High. Straight forward download
Processing Abstraction (Cloud processing)	Off-the-shelf dataset	Off the shelf data powered by GEE	Cloud processing	Off the Shelf Data	Off the Shelf Data
Use of Ancillary Data for validation	Yes	Yes	Yes	Yes	Yes
Output Type	PNG, CSV, SHAPEFILE, TIFF	JSON	Not applicable for public users.	WMS, Shapefile	JSON, Geopackage, Shapezip, Text, CSV,XML
Language of Platform	Arabic, Chinese, English, French, Russian, Spanish	Arabic, Chinese, English, French,	English	English, French, Spanish	English
Platform Availability	Free and Open. Need to create user account.	Free and Open.	Free and Open.	Free and Open. No Need for User Account.	Free and Open. No Need for User Account.
Platform Owner	FAO	WRI	Landgate	Wetlands International	International Soil Reference and Information Centre (ISRIC) Data Hub
Platform Partners	International Institute for Applied Systems Analysis (IIASA)	Partnered with Google	Australian Government	With support from the Oak Foundation, DoB Ecology, Aberysthwyth university and soloEO, Global Mangrove Watch was initiated by The Nature Conservancy and Wetlands International, working with dozens of universities, NGOs and government agencies across the world.	ISRIC - World Soil Information is situated on the Wageningen University & Research campus in Wageningen, the Netherlands.
Availability of Process Documentation	Yes	Yes	No	Yes	Yes
Availability of User Guidelines	Yes	Yes	Yes	Yes	Yes

Platform	FAO GLOSIS	Global Dam Watch (GDW)	Earth Map (Open FORIS)	Google Earth Engine (GEE)	CEOS COVE Tool
Website	https://www.fao.org/global-soil-partnership/areas-of-work/soil-	http://globaldamwatch.org/	http://earthmap.org/	https://earthengine.google.com/platform/	https://ceos-cove.org/en/
Description	GloSIS, the Global Soil Information System of the FAO Global Soil Partnership is a central hub for global soil information and data created to provide easy access to dynamic soil resource information that is globally harmonized	Global Dam Watch is an international collaboration which aims to improve our understanding of the costs and benefits of dams to our world by providing open access data and tools focused on dams and reservoirs.	Earth Map, a collaborative effort between the Food and Agriculture Organization of the United Nations (FAO) and Google, is an innovative solution that aids countries, research institutions, and farmers in efficiently monitoring their land. This tool leverages Google Earth Engine's robust big data capabilities, allowing users to conduct sophisticated analyses of earth observation and climate data without requiring expertise in remote sensing or GIS. Earth Map is structured into thematic categories, offering visualization layers and statistical tools for a comprehensive range of datasets encompassing climate, vegetation, land degradation, water resources, forestry, and biodiversity, facilitating easy and in-depth land monitoring.	Google Earth Engine is a computing platform that allows users to run geospatial analysis on Google's infrastructure. There are several ways to interact with the platform. The Code Editor is a web-based IDE for writing and running scripts. The Explorer is a lightweight web app for exploring our data catalog and running simple analyses.	The Committee on Earth Observation Satellites (CEOS) Visualization Environment (COVE) is a comprehensive suite of tools developed by the NASA CEOS Systems Engineering Office (SEO) for analyzing satellite sensor coverage from over 100 Earth-observing satellites. The tools within COVE include the Acquisition Forecaster, enabling users to predict satellite imaging opportunities for specific regions; the Coverage Analyzer for historical satellite coverage analysis; the Revisits Calculator for estimating satellite coverage and revisits; the Coincident Calculator for determining instrument coincidences; the Data Browser for viewing satellite image archives; the Data Policy database for information on data policies; the Country Coverage tool for generating reports on over 70 countries; and Utilities for quick parameter estimations related to CEOS satellite missions.
Available/ Used Sensors and Radars. Return Period	MODIS, Landsat, DEM derivatives, climatic images and global landform and lithology maps	Not defined. Cited in the methodology about the use of Google Earth.	Synthetic Aperture Radar (SAR), Interferometric SAR (InSAR), Lidar (Light Detection and Ranging), Atmospheric Infrared Sounder (AIRS), Cloud Profiling Radar (CPR), Cloud-Aerosol Lidar with Orthogonal Polarization (CALIOP), Cloud-Aerosol Transport System (CATS), Cross-track Infrared Sounder (CrIS), Delay Doppler Mapping Instrument (DDMI), Differential Absorption LIDAR (DIAL), Digital Mapping System (DMS), Accumulation Radar (AR), Advanced Baseline Imager (ABI),Advanced Himawari Imager (AHI), Advanced Microwave Precipitation Radiometer (AMPR), Advanced Microwave Radiometer (AMR-2), Advanced Microwave Radiometer for Climate (AMR-C), Advanced Microwave Scanning Radiometer 2 (AMSR2), Advanced Microwave Scanning Radiometer for EOS (AMSR-E), Advanced Microwave Sounding Unit (AMSU), Advanced Spaceborne Thermal Emission and Reflection Radiometer (ASTER), Advanced Technology Microwave Sounder (ATMS), Advanced Topographic Laser Altimeter System (ATLAS), Advanced Very High Resolution Radiometer (AVHRR), Airborne Topographic Mapper (ATM), Airborne Visible InfraRed Imaging Spectrometer (AVIRIS), GeoEye-1, KOMPSAT-3, Landsat Collection, Sentinel Collection, GF-1 PMS	Synthetic Aperture Radar (SAR), Interferometric SAR (InSAR), Lidar (Light Detection and Ranging), Atmospheric Infrared Sounder (AIRS), Cloud Profiling Radar (CPR), Cloud-Aerosol Lidar with Orthogonal Polarization (CALIOP), Cloud-Aerosol Transport System (CATS), Cross-track Infrared Sounder (CrIS), Delay Doppler Mapping Instrument (DDMI), Differential Absorption LIDAR (DIAL), Digital Mapping System (DMS), Accumulation Radar (AR), Advanced Baseline Imager (ABI),Advanced Himawari Imager (AHI), Advanced Microwave Precipitation Radiometer (AMPR), Advanced Microwave Radiometer (AMR-2), Advanced Microwave Radiometer for Climate (AMR-C), Advanced Microwave Scanning Radiometer 2 (AMSR2), Advanced Microwave Scanning Radiometer for EOS (AMSR-E), Advanced Microwave Sounding Unit (AMSU), Advanced Spaceborne Thermal Emission and Reflection Radiometer (ASTER), Advanced Technology Microwave Sounder (ATMS), Advanced Topographic Laser Altimeter System (ATLAS), Advanced Very High Resolution Radiometer (AVHRR), Airborne Topographic Mapper (ATM), Airborne Visible InfraRed Imaging Spectrometer (AVIRIS), GeoEye-1, KOMPSAT-3, Landsat Collection, Sentinel Collection, GF-1 PMS	Landsat 5 TM, Landsat 7 ETM, Landsat 8 OLI/TIRS, Landsat 9 OLI/TIRS, Sentinel-1A, Sentinel-1B, Sentinel-2A MSI, Sentinel-2B MSI
Spatial Resolution	1000 meters	Not defined.	10m, 15m, 30m, 60m,100m, 250m, 500m, 1km	5m, 10m, 15m, 30m, 60m,100m, 250m, 500m, 1km. Allows uploading of own images.	30m, 60m, 100m
Spatial Coverage	Global	Global	Global, Regional, National, Site specific	Global, Regional, National, Site specific	Global
Temporal Coverage	Single dataset published on 2012	Not defined. Includes dam that appear after 1983	Daily, Monthly, Annually. Ranges from 1980 to present.	Daily, Monthly, Annually. Ranges from 1980 to present.	Landsat 5 (Mar 1964: Jun 2013), Landsat 7 (Apr 1999: now), Landsat 8 (Feb 2013: now), Landsat 9 (Sep 2021: now), Sentinel-1A (Apr 2017: now), Sentinel-1B (Apr 2016: Aug 2022)
Data Access	Downloadable if you have account .	Downloadable	Downloadable if you have account .	Downloadable using GEE account	Downloadable
Ease of Use (low, Moderate, High)	High. Straight forward download	High. Fill out the google form and directly redirected in the FTP File Folder. Data are in zip file.	High. Directly select the area and then download.	Low. User must have basic knowledge of programming in python or javascript.	High. Straight forward download
Processing Abstraction (Cloud processing)		Off the Shelf Data	Off the shelf data powered by GEE	Avaialable Off the shelf data. All processing done in cloud using the GEE code editor.	Cloud Processing
Use of Ancillary Data for validation	Yes	Yes	Yes	Yes	
Output Type	GeoTIFF, WMS	Shapefile with metadata in pdf	PNG, GeoTIFF, GeoTIFF QGIS Style, Chart, Statistical Chart Image, CSV	PNG,jpeg, GeoTIFF, CSV, txt	PNG, JSON, GeoTIFF, NetCDF and csv
Language of Platform	English	English	English, Spanish, French, Portugese, Macedonian Language	English, Deutsch, Spanish, French, Indonesia, Portugues-Brazil, Chinese, Japanese, Korean	English
Platform Availability	Free and Open. User account needed.	Free and Open. No Need for User Account.	Free and Open. Need to create user account.	Free and Open. Need to create user account.	Free and Open. User account needed.
Platform Owner	FAO	Owned and operated by a global consortium of organizations :	FAO	Google	Committee on Earth Observation Satellites (CEOS)
Platform Partners	International Network of Soil Information Institutions (INSII) is leading the development, with member countries contributing their data	Partners, Sponsors and Supporters: The Nature Conservancy, Conservation International, Amber International, World Resources Institute, UN Environment World, Conservation, Monitoring Center	Developed within FAO's Open Foris Initiative with the support of the Government of Germany through The International Climate Initiative (IKI) from the Federal Ministry of the Environment, Nature Conservation and Nuclear Safety		None.
Availability of Process Documentation	Yes	Yes	Yes	Yes	Yes
Availability of User Guidelines	Yes	Yes	Yes	Yes	Yes

Platform	RAPP	SEPAL (Open FORIS)	Open EO Platform	Open Data Cube (ODC)	EOSDIS Worldview	Trends.Earth
Website	http://map.geo-rapp.org/	https://sepal.io/	https://openeo.cloud/	https://www.opendatacube.org/	EOSDIS Worldview (nasa.gov)	Trends.Earth — Trends.Earth 2.1.8 documentation
Description	Geoglam RaPP Map is an online tool providing timeseries data on vegetation and environmental conditions. RAPP Map is the spatial data platform for the National Landcare Regional Partnerships Program. This online tool provides time-series data on vegetation and environmental conditions, allowing national and regional reporting of vegetation cover change.	SEPAL is part of the Open Foris suite of tools that enables users to quickly and efficiently query, process, and analyze satellite data using cloud-based supercomputing resources and modern geospatial data infrastructures like Google Earth Engine. The platform allows users to tailor products for local needs and produce sophisticated geospatial analyses, leveraging a combination of open-source software like GDAL, Jupyter, and R to access and process historical Landsat data as well as high-resolution imagery from the Copernicus program, empowering autonomous land monitoring through remotely sensed data.	openEO Platform is a new cloud processing and analytics environment built on top of openEO. It brings openEO to production and offers data access and data processing services to the EO community.	The Open Data Cube (ODC) is an Open Source Geospatial Data Management and Analysis Software project that helps you harness the power of Satellite data. At its core, the ODC is a set of Python libraries and PostgreSQL database that helps you work with geospatial raster data. The ODC seeks to increase the value and impact of global Earth observation satellite data by providing an open and freely accessible exploitation architecture. The ODC project seeks to foster a community to develop, sustain, and grow the technology and the breadth and depth of its applications for societal benefit.	This open source code app from NASA's EOSDIS provides the capability to interactively browse over 1000 global, full-resolution satellite imagery layers and then download the underlying data. Many of the imagery layers are updated daily and are available within three hours of observation - essentially showing the entire Earth as it looks "right now".	Trends.Earth is a free and open source tool to understand land change: the how and why behind changes on the ground. Trends.Earth allows users to draw on the best available information from across a range of sources - from globally available data to customized local maps. A broad range of users are applying Trends.Earth for projects ranging from planning and monitoring restoration efforts, to tracking urbanization, to developing official national reports for submission to the United Nations Convention to Combat Desertification (UNCCD).
Available/ Used Sensors and Radars. Return Period	MODIS, CHIRPS	Landsat, Sentinel, Planet, NICFI, Synthetic Aperture Radar (SAR), Interferometric SAR (InSAR), Lidar (Light Detection and Ranging), MODIS, AVHRR, Quickbird, Worldview	Agera, ALOS PALSAR, Landsat 8, Sentinel 1,Sentinel 2, Sentinel 3, Sentinel 5, PROBA-V, Worldview, SPOT/VEGETATION, Terra, TDX, TSX, VITO, Landsat 4, Landsat 5	Depends on data ingested to the software. But give you access with Landsat 4/5/7/8/9 , Sentinel-2/2A, CBERS-4/4A , MODIS, Harmonized Landsat-Sentinel (HLS) data, which combines Landsat 8 and Sentinel-2 imagery	MODIS (1day), Worldview (3-5 hours), NOAA-20/VIIRS, Suomi NPP/VIIRS	AVHRR/GIMMS, MOD13Q1-coll6, MERRA 2, ERA I, GPCP, GPCC V6, CHIRPS, PERSIANN-CDR, MOD16A2, ESA CCI Land Cover
Spatial Resolution	500m, 5km	0.5m, 5m, 10m, 15m, 30m, 60m,100m, 250m, 500m, 1km	2m, 3m, 10m, 30m, 90m, 300m, 1km	5m, 10m, 15m, 30m, 60m,100m, 250m, 500m, 1km. Allows uploading of own images.	250m	10m, 50m, 8km, 250m, 300m, 1km, 5km, 25km, 0.5° x 0.625°, 0.75° x 0.75°
Spatial Coverage	Australia, Global	Global, Regional, National, Site specific	Global, Europe, Selected sites	Global, Regional, National, Site specific	Global	Global
Temporal Coverage	8-days, Monthly	Daily, Monthly, Annually. Ranges from 1980 to present.	Daily, Monthly, Annually. Some datasets starts on 1984 to present	Daily, Monthly, Annually.	Daily	Some dataset started on 1980 to present
Data Access	Downloadable	Downloadable using GEE account	Offers a free 30-day trial with 1000 free credits. Paid account.	Need to download application.	Downloadable if you have account .	Downloadable.
Ease of Use (low, Moderate, High)	High. Directly select the area and then download.	Moderate. Run integrated workflows with modules- no need of coding experience required.	Low. Need knowledge on basic programming.	Low. User must have learn the software and its tools.	High. Select area and dataset then it will redirect you to NASA Earth Data School where to download data.	Moderate. Need to process data.
Processing Abstraction (Cloud processing)	Cloud Based Processing	Off the shelf data powered by GEE	Processing in Data Cube	Processing in the machine.	Images powered by GIBS	Processing in the machine.
Use of Ancillary Data for validation	Yes	Yes	Yes	Yes	Yes	Yes
Output Type	NetCDF and GeoTIFF	PNG,jpeg, GeoTIFF, CSV	CoverageJSON (CovJSON), CSV, GeoTiff, Jpeg, Json, Network Common Data Form (NetCDF), GeoParquet, Portable N	GeoTiff, Tiff, PNG, JPEG, NetCDF (Network Common Data Form)	JPEG, PNG, GeoTIFF, and KML	GeoTIFF
Language of Platform	English	English, Spanish, French	English	English	English	Arabic, Chinese, English, Spanish, French, Portuguese, Kiswahili
Platform Availability	Free and Open. No Need for User Account.	Free and Open. Need to create user account.	Offers a free 30-day trial with 1000 free credits to allow users to explore the	Free and Open.	Free and Open. Need to create user account.	Free and Open. Plugins to QGIS
Platform Owner	GEOGLAM	FAO	European Space Agency (ESA)	Open Data Cube	NASA	Conservation International
Platform Partners	It has been developed and is currently hosted by Data61. RAPP Map is supported by CSIRO, Department of Agriculture, Fisheries and Forestry Australian Bureau of Agricultural and Resource Economics and Sciences (ABARES). Through funding from the Department of Climate Change, Energy, the Environment and Water and Regional land partnership flagship project under Australian Government's National Landcare Program.	Funded by the Government of Norway from the Forestry Department of the Food and Agriculture Organization of the United Nations (FAO). Supported by Spatial Informatics Group (SIG), SERVIR, SilvaCarbon, KFW, NASA, JICA, JAXA, Google, Global Forest Observations Initiative (GFOI), Germany Federal Ministry for the Environment, Nature Conservation and Nuclear Safety, Formin Finland, ETH Zurich, ESA, European Commission	openEO Platform builds upon the EO cloud processing platforms managed by VITO, Sinergise, and EODC, and the platform management and software development experience of all partners	The Open Data Cube initiative is supported by six institutional partners: Geoscience Australia (GA), NASA / Committee on Earth Observation Satellite (CEOS), United States Geological Survey (USGS), Commonwealth Scientific and Industrial Research Organisation (CSIRO), Catapult Satellite Applications, and Analytical Mechanics Associates (AMA).	SSAI (Science Systems and Applications, Inc.) and ASRC (Arctic Slope Regional Corporation)	Produced with partnership to Lund University, and the National Aeronautics and Space Administration (NASA), with the support of the Global Environment Facility (GEF). It was further developed through a partnership with University of Bern, University of Colorado in partnership with USDA and USAID, University of California - Santa Barbara in partnership with University of North Carolina - Wilmington and Brown University with additional funding from the Global Environment Facility (GEF).
Availability of Process Documentation	Yes	Yes	Yes	Yes	Yes	Yes
Availability of User Guidelines	Yes	Yes	Yes	Yes	Yes	Yes

Annex 2: Name of EO Datasets for Soil Health Indicators

SH CATEGORIES		Physical Properties							Chemical Properties										Biological Properties		Meteorological data				Other Factors	
SH IDICATORS		Soil Texture	Sand content (Sand-C)	Silt content (Silt-C)	Clay content (Clay-C)	Bulk Density (BD)	Soil Moisture	Water Holding Capacity	Hydrogen Potential (pH)	Electrical Conductivity	Nitrogen (N)	Potassium (K)	Calcium (Ca)	Magnesium (Mg)	Phosphorus (P)	Soil Nutrient	Cation exchange capacity (CEC)	Soil Salinity	Soil organic carbon (SOC)	Soil microbial respiration (SMR)	Temperature (LST)	Precipitation	Evapotranspi-ration	Humidity	Soil Erosion	Cropland Irrigation
EO WEB P	GEE	OpenLand Map Soil Texture Class (USDA System)	OpenLand Map Sand Content	iSDAsoil USDA Texture Class	OpenLandMap Clay Content	OpenLandMap Soil Bulk Density	SPL4SMGP.007 SMAP L4 Global 3-hourly 9-km Surface and Root Zone Soil Moisture	OpenLandMap Soil Water Content at 33kPa (Field Capacity)	OpenLandMap Soil pH in H_2O		iSDAsoil Total Nitrogen	iSDAsoil Extractable Potassium	iSDAsoil Extractable Calcium	iSDAsoil Extractable Magnesium	iSDAsoil Extractable Phosphorus	iSDAsoil Fertility Capability Classification	SLGA: Soil and Landscape Grid of Australia (Soil Attributes)		OpenLandMap Soil Organic Carbon Content		MOD11A1.061 Terra Land Surface Temperature and Emissivity Daily Global 1km	CHIRPS Daily: Climate Hazards Group InfraRed Precipitation With Station Data (Version 2.0 Final)	TerraClimate: Monthly Climate and Climatic Water Balance for Global Terrestrial Surfaces, University of	GLDAS-2.1: Global Land Data Assimilation System	WWF HydroATLAS Basins Level 05	GFSAD1000: Cropland Extent 1km Multi-Study Crop Mask, Global Food-Support Analysis
	ISRIC	SoilGrids250m 2017-03 - Texture class (USDA system)	SoilGrids250m 2.0 - Sand content	WoSIS latest - Silt total	SoilGrids 250m 2.0 - Clay content	SoilGrids 250m 2017-03 - Bulk density (fine earth)	Africa SoilGrids - Root zone moisture content at wilting point aggregated at ERZD	SoilGrids250m 2017-03 - Derived available soil water capacity (volumetric fraction) with FC = pF 2.5	SoilGrids250m 2017-03 - Soil pH in H2O		SoilGrids 250m 2.0 - Total nitrogen	Africa SoilGrids nutrients - Extractable Potassium (K)	Africa SoilGrids nutrients - Extractable Calcium (Ca)	Africa SoilGrids nutrients - Extractable Magnesium (Mg)	Global distribution of soil phosphorus retention potential	Africa SoilGrids nutrients - Nutrient clusters based on fuzzy k-means	WoSIS latest - Effective cation exchange capacity - ISRIC	Global Soil Salinity Map	SoilGrids 250m 2.0 - Soil organic carbon content							
	FAO -GLOSIS		Topsoil Sand Fraction	Topsoil Silt Fraction	Topsoil Clay Fraction	Subsoil Reference Bulk Density		Available water storage capacity (Global)	Subsoil pH (H2O)							Soil Nutrient Availability from HWSD v1.2 (Global)	Topsoil CEC (CLAY)	Excess Salts from HWSD v1.2 (Global)	Global Soil Organic Carbon Map v1.5 (GSOC)							
	EARTH MAP	OpenLand Map Soil texture						Total Available Water											Global Soil Organic Carbon - GSOC			Prec (average) - CHIRPS	PET (average) - MODIS		Runoff (yearly) - ECMWF Land / CCI LC / SRTM DEM	Cropland - GFSAD1000
	ORNL - DAAC	Global Soil Texture and Derived Water-Holding Capacities (Webb et al.)	Global Soil Texture and Derived Water-Holding Capacities (Webb et al.)	Global Soil Texture and Derived Water-Holding Capacities (Webb et al.)	Global Soil Texture and Derived Water-Holding Capacities (Webb et al.)	Global Gridded Surfaces of Selected Soil Characteristics (IGBP-DIS)	Global Gridded Surfaces of Selected Soil Characteristics (IGBP-DIS)	Global Soil Texture and Derived Water-Holding Capacities (Webb et al.)			Global Gridded Surfaces of Selected Soil Characteristics (IGBP-DIS)				A Compilation of Global Soil Microbial Biomass Carbon, Nitrogen, and Phosphorus Data				Global Gridded Surfaces of Selected Soil Characteristics (IGBP-DIS)	A Compilation of Global Soil Microbial Biomass Carbon, Nitrogen, and Phosphorus Data	SAFARI 2000 AVHRR-derived Land Surface Temperature Maps, Africa, 1995-2000	ISLSCP II GLOBAL PRECIPITATION CLIMATOLOGY PROJECT VERSION 1, PENTAD PRECIPITATION	CMS: Evapotranspiration and Meteorology, Water-Limited Shrublands, Mexico, 2008-2010	VEMAP 2: U.S. ANNUAL CLIMATE (1895-1993): Mean Daily Irradiance	Topographic and Soil Carbon Reconstructions in Agricultural Fields, Iowa	GFSAD1KCM v001 Global Food Security Support Analysis Data (GFSAD) Crop Mask 2010
	FAO - HIH	OpenLand Map Soil texture	Topsoil Sand Fraction	Topsoil Silt Fraction	Topsoil Clay Fraction	Subsoil Reference Bulk Density	Relative root zone soil moisture - beta product (Global - Dekadal - 300m) - WaPOR v3	Available water storage capacity (Global)	Subsoil pH (H2O)		Subsoil pH (H2O)					Soil Nutrient Availability from HWSD v1.2 (Global)	Topsoil CEC (CLAY)	Excess Salts from HWSD v1.2 (Global)	Global Soil Organic Carbon Map v1.5 (GSOC)		Land Surface Temperature Day-Time (Global - L3, 8-Day - MYD11A2) MODIS Aqua	Precipitation (Global - Daily - Approximately 5km) - WaPOR v3	Reference Evapotranspiration (Global - Daily - Approximately 30km) - WaPOR v3	Relative humidity at 06h local time - AgERA5 (Global - Daily - ~10km)		

Annex 3: Google Earth Engine EO data for soil

Physical Property	INTRINSIC					CONTEXTUALITY					
	Name	Dataset Provider	Point of contact	Methods of Processing EO Data	Use of Ancillary Data	Sensor Used for production of dataset	Spatial Resolution	Temporal Coverage	Spatial Coverage	Temporal Resolution/Frequency	Year of EO launch
	What is the name of platform	Who is the source of data	Who is the responsible organization for the EO Platform	What type of method is used	Did it utilize additional/auxiliary and supplementary?	What are the sensors used to produced data?	What is the level of detail captured in an image by a sensor system (pixel size)?	Does it have historical data and the revisit frequency you require?	What is the geographical coverage of data?	What is the repeat/revisit time, e.g., time between successive observations of a given location	When did the EO platform become available and functional?
Soil Texture	OpenLandMap Soil Texture Class (USDA System)	EnvirometriX Ltd	Google	Predicted soil texture fractions using the soiltexture package in R	Yes	Not explicitly ment	250m	1950 to 2018	Global except Antarctica	Single dataset with soil texture class bands 0 cm - 200 cm with 6 standard depths	2001
Sand content (Sand-C)	OpenLandMap Sand Content	EnvirometriX Ltd	Google	Based on machine learning predictions (Random Forest and XGBoost) from global compilation of soil profiles and samples	Yes	Not explicitly ment	250m	1950 to 2018	Global except Antarctica	Single dataset with class bands 0 cm - 200 cm with 6 standard depths	2001
Silt content (Silt-C)	iSDAsoil USDA Texture Class	Innovative Solutions for Decision Agriculture Ltd. (iSDA)	Google	Used the multiscale ensemble machine learning approach combininf 2 different geospatial resolutions then application of 5 regression modelling algorithms (Ranger, XG Boost, Cubist, Deepnet, glmnet) and a training set of over 100,000 analyzed soil samples	Yes	MODIS, digital terrain model (DTM) derivatives, Sentinel-2 satellite and Landsat-7/8 cloud-free composite images	30m	2001 to 2017	Africa	Single dataset with soil depths of 0-20 cm and 20-50 cm	2001
Clay content (Clay-C)	OpenLandMap Clay Content	EnvirometriX Ltd	Google	Based on machine learning predictions (Random Forest and XGBoost) from global compilation of soil profiles and samples	Yes	Not explicitly ment	250m	2002 to 2017	Global except Antarctica	Single dataset with class bands 0 cm - 200 cm with 6 standard depths	2001
Bulk Density (BD)	OpenLandMap Soil Bulk Density	EnvirometriX Ltd	Google	Based on machine learning predictions (Random Forest and XGBoost) from global compilation of soil profiles and samples	Yes	Not explicitly ment	250m	2003 to 2017	Global except Antarctica	Single dataset with class bands 0 cm - 200 cm with 6 standard depths	2001
Available Water Capacity	OpenLandMap Soil Water Content at 33kPa (Field Capacity)	EnvirometriX Ltd. Training points are based on a global compilation of: USDA NCSS, AfSPDB, ISRIC WISE, EGRPR, SPADE, CanNPDB, UNSODA, SWIG, HYBRAS, HydroS	Google	Used spatial prediction (Random Forest, XGBoost). Available water capacity in mm (derived as a difference between field capacity and wilting point multiplied by layer thickness).	Yes	Modis, Landsat, Sentinel 2	250m	1950 to 2018	Global except Antarctica	Single dataset with class bands 0 cm - 200 cm with 6 standard depths	2002
Soil Moisture	SPL4SMGP.007 SMAP L4 Global 3-hourly 9-km Surface and Root Zone Soil Moisture	Google and NSIDC	Google	2 key processes using GEOS Catchment Land Surface and Microwave Radiative Transfer Model, GEOS Ensemble-Based Land Data Assimilation Algorithm	Yes	SMAP radar (active) and a radiometer (passive)	11000 meters	2015 to 2024	Global	3-hourly	2001

Physical Property	ACCESSIBILITY			REPRESENTATIONAL										
	Data Access	Product Format	Security	User Technical Knowledge Requirement	Processing Tools	Programming Languages	Cloud Processing	User Interface (UI)	Visualization Tools	Scalability	Open Standards	User Support	Community	Product User Guide/Manual
	Is the data free, subscription-based, or pay-per-use?	What are the EO data formats available for download?	Does the platform allows you to use your own data securely?	What is the minimum technical knowledge of user to access and download data?	Does the platform offer built-in tools for data analysis ?	Does it support your preferred programming language (e.g., Python, R) for custom analysis?	Does it leverage cloud computing for handling datasets?	Is the interface user-friendly and easy to navigate?	Does it offer tools for visualizing and exploring the data (e.g., generating maps, charts)?	Can the platform handle high volume of data (e.g., multiple downloads from global to community scale)?	Does it adhere to open data standards for interoperability with other platforms?	Does the platform offer adequate user support (e.g., video tutorials, case studies)?	Is there an active user community for knowledge sharing and troubleshooting?	Is there minimum set of documentation for users?
Soil Texture	Need user account	GeoTIFF, TFRecord, Feature Collection as CSV,SHP, GeoJSON,	Yes. Allows upload of own dataset without being	Requires basic knowledge of programming.	Yes. Use built-in codes in javascript for data manipulation	Python and Javascript	Yes	Yes. Easy to navigate and easy to find different datasets.	Yes. It has Google Code Editor to view and manipulate data.	Yes	Yes	Yes	Yes	Yes
Sand content (Sand-C)	Need user account	GeoTIFF, TFRecord, Feature Collection as CSV,SHP, GeoJSON, KML, KMZ and Map Tiles	Yes. Allows upload of own dataset without being	Requires basic knowledge of programming.	Yes. Use built-in codes in javascript for data manipulation	Python and Javascript	Yes	Yes. Easy to navigate and easy to find different datasets.	Yes. It has Google Code Editor to view and manipulate data.	Yes	Yes	Yes	Yes	Yes
Silt content (Silt-C)	Need user account	GeoTIFF, TFRecord, Feature Collection as CSV,SHP, GeoJSON, KML, KMZ and Map Tiles	Yes. Allows upload of own dataset without being	Requires basic knowledge of programming.	Yes. Use built-in codes in javascript for data manipulation	Python and Javascript	Yes	Yes. Easy to navigate and easy to find different datasets.	Yes. It has Google Code Editor to view and manipulate data.	Yes	Yes	Yes	Yes	Yes
Clay content (Clay-C)	Need user account	GeoTIFF, TFRecord, Feature Collection as CSV,SHP, GeoJSON, KML, KMZ and Map Tiles	Yes. Allows upload of own dataset without being	Requires basic knowledge of programming.	Yes. Use built-in codes in javascript for data manipulation	Python and Javascript	Yes	Yes. Easy to navigate and easy to find different datasets.	Yes. It has Google Code Editor to view and manipulate data.	Yes	Yes	Yes	Yes	Yes
Bulk Density (BD)	Need user account	GeoTIFF, TFRecord, Feature Collection as CSV,SHP, GeoJSON, KML, KMZ and Map Tiles	Yes. Allows upload of own dataset without being	Requires basic knowledge of programming.	Yes. Use built-in codes in javascript for data manipulation	Python and Javascript	Yes	Yes. Easy to navigate and easy to find different datasets.	Yes. It has Google Code Editor to view and manipulate data.	Yes	Yes	Yes	Yes	Yes
Available Water Capacity	Need user account	GeoTIFF, TFRecord, Feature Collection as CSV,SHP, GeoJSON, KML, KMZ and Map Tiles	Yes. Allows upload of own dataset without being	Requires basic knowledge of programming.	Yes. Use built-in codes in javascript for data manipulation	Python and Javascript	Yes	Yes. Easy to navigate and easy to find different datasets.	Yes. It has Google Code Editor to view and manipulate data.	Yes	Yes	Yes	Yes	Yes
Soil Moisture	Need user account	GeoTIFF, TFRecord, Feature Collection as CSV,SHP, GeoJSON, KML, KMZ and Map Tiles	Yes. Allows upload of own dataset without being	Requires basic knowledge of programming.	Yes. Use built-in codes in javascript for data manipulation	Python and Javascript	Yes	Yes. Easy to navigate and easy to find different datasets.	Yes. It has Google Code Editor to view and manipulate data.	Yes	Yes	Yes	Yes	Yes

Chemical Property	INTRINSIC					CONTEXTUALITY						ACCESSIBILITY			REPRESENTATIONAL										
	Name	**Dataset Provider**	**Point of contact**	**Methods of Processing EO Data**	**Use of Ancillary Data**	**Sensor Used for production of dataset**	**Spatial Resolution**	**Temporal Coverage**	**Spatial Coverage**	**Temporal Resolution/Frequency**	**Year of EO launch**	**Data Access**	**Product Format**	**Security**	**User Technical Knowledge Requirement**	**Processing Tools**	**Programming Languages**	**Cloud Processing**	**User Interface (UI)**	**Visualization Tools**	**Scalability**	**Open Standards**	**User Support**	**Community**	**Product User Guide/Manual**
	What is the name of platform	Who is the source of data	Who is the responsible organization for the EO Platform	What type of method is used	Did it utilize additional/auxiliary and supplementary?	What are the sensors used to produced data?	What is the level of detail captured in an image by a sensor system (pixel size)?	Does it have historical data and the revisit frequency you require?	What is the geographical coverage of data?	What is the repeat/revisit time, e.g., time between successive observations of a given location	When did the EO platform become available and functional?	Is the data free, subscription-based, or pay-per-use?	What are the EO data formats available for download?	Does the platform allows you to use your own data securely?	What is the minimum technical knowledge of user to access and download data?	Does the platform offer built-in tools for data analysis ?	Does it support your preferred programming language (e.g., Python, R) for custom analysis?	Does it leverage cloud computing for handling datasets?	Is the interface user-friendly and easy to navigate?	Does it offer tools for visualizing and exploring the data (e.g., generating maps, charts)?	Can the platform handle high volume of data (e.g., multiple downloads from global to community scale)?	Does it adhere to open data standards for interoperability with other platforms?	Does the platform offer adequate user support (e.g., video tutorials, case studies)?	Is there an active user community for knowledge sharing and troubleshooting?	Is there minimum set of documentation for users?
Hydrogen Potential (pH)	OpenLandMap Soil pH in H_2O	EnvirometriX Ltd	Google	Used spatial prediction (Random Forest, XGBoost).	Yes	Not explicitly ment	250m	2005 to 2017	Global except Antarctica	Single dataset with class bands 0 cm - 200 cm with 6 standard depths	2002	Need user account	GeoTIFF, TFRecord, Feature Collection as CSV,SHP, GeoJSON, KML, KMZ and Map Tiles	Yes. Allows upload of own dataset without being	Requires basic knowledge of programming.	Yes. Use built-in codes in javascript for data manipulation	Python and Javascript	Yes	Yes. Easy to navigate and easy to find different datasets.	Yes. It has Google Code Editor to view and manipulate data.	Yes	Yes	Yes	Yes	Yes
Total nitrogen (TN)	iSDAsoil Total Nitrogen	Innovative Solutions for Decision Agriculture Ltd. (iSDA	Google	Used the multiscale ensemble machine learning approach combininf 2 different geospatial resolutions then application of 5 regression modelling algorithms (Ranger, XG Boost, Cubist, Deepnet, glmnet) and a training set of over 100,000 analyzed soil samples	Yes	MODIS, digital terrain model (DTM) derivatives, Sentinel-2 satellite and Landsat-7/8 cloud-free composite images	250m	2006 to 2017	Africa	Single dataset with soil depths of 0-20 cm and 20-50 cm	2001	Need user account	GeoTIFF, TFRecord, Feature Collection as CSV,SHP, GeoJSON, KML, KMZ and Map Tiles	Yes. Allows upload of own dataset without being	Requires basic knowledge of programming.	Yes. Use built-in codes in javascript for data manipulation	Python and Javascript	Yes	Yes. Easy to navigate and easy to find different datasets.	Yes. It has Google Code Editor to view and manipulate data.	Yes	Yes	Yes	Yes	Yes
Potassium (K)	iSDAsoil Extractable Potassium	Innovative Solutions for Decision Agriculture Ltd. (iSDA	Google	Used the multiscale ensemble machine learning approach combininf 2 different geospatial resolutions then application of 5 regression modelling algorithms (Ranger, XG Boost, Cubist, Deepnet, glmnet) and a training set of over 100,000 analyzed soil samples	Yes	MODIS, digital terrain model (DTM) derivatives, Sentinel-2 satellite and Landsat-7/8 cloud-free composite images	250m	2006 to 2017	Africa	Single dataset with soil depths of 0-20 cm and 20-50 cm	2001	Need user account	GeoTIFF, TFRecord, Feature Collection as CSV,SHP, GeoJSON, KML, KMZ and Map Tiles	Yes. Allows upload of own dataset without being	Requires basic knowledge of programming.	Yes. Use built-in codes in javascript for data manipulation	Python and Javascript	Yes	Yes. Easy to navigate and easy to find different datasets.	Yes. It has Google Code Editor to view and manipulate data.	Yes	Yes	Yes	Yes	Yes
Calcium (Ca)	iSDAsoil Extractable Calcium	Innovative Solutions for Decision Agriculture Ltd. (iSDA	Google	Used the multiscale ensemble machine learning approach combininf 2 different geospatial resolutions then application of 5 regression modelling algorithms (Ranger, XG Boost, Cubist, Deepnet, glmnet) and a training set of over 100,000 analyzed soil samples	Yes	MODIS, digital terrain model (DTM) derivatives, Sentinel-2 satellite and Landsat-7/8 cloud-free composite images	250m	2006 to 2017	Africa	Single dataset with soil depths of 0-20 cm and 20-50 cm	2001	Need user account	GeoTIFF, TFRecord, Feature Collection as CSV,SHP, GeoJSON, KML, KMZ and Map Tiles	Yes. Allows upload of own dataset without being	Requires basic knowledge of programming.	Yes. Use built-in codes in javascript for data manipulation	Python and Javascript	Yes	Yes. Easy to navigate and easy to find different datasets.	Yes. It has Google Code Editor to view and manipulate data.	Yes	Yes	Yes	Yes	Yes
Magnesium (Mg)	iSDAsoil Extractable Magnesium	Innovative Solutions for Decision Agriculture Ltd. (iSDA	Google	Used the multiscale ensemble machine learning approach combininf 2 different geospatial resolutions then application of 5 regression modelling algorithms (Ranger, XG Boost, Cubist, Deepnet, glmnet) and a training set of over 100,000 analyzed soil samples	Yes	MODIS, digital terrain model (DTM) derivatives, Sentinel-2 satellite and Landsat-7/8 cloud-free composite images	250m	2006 to 2017	Africa	Single dataset with soil depths of 0-20 cm and 20-50 cm	2001	Need user account	GeoTIFF, TFRecord, Feature Collection as CSV,SHP, GeoJSON, KML, KMZ and Map Tiles	Yes. Allows upload of own dataset without being	Requires basic knowledge of programming.	Yes. Use built-in codes in javascript for data manipulation	Python and Javascript	Yes	Yes. Easy to navigate and easy to find different datasets.	Yes. It has Google Code Editor to view and manipulate data.	Yes	Yes	Yes	Yes	Yes
Phosphorus (P)	iSDAsoil Extractable Phosphorus	Innovative Solutions for Decision Agriculture Ltd. (iSDA)	Google	Used the multiscale ensemble machine learning approach combininf 2 different geospatial resolutions then application of 5 regression modelling algorithms (Ranger, XG Boost, Cubist, Deepnet, glmnet) and a training set of over 100,000 analyzed soil samples	Yes	MODIS, digital terrain model (DTM) derivatives, Sentinel-2 satellite and Landsat-7/8 cloud-free composite images	250m	2006 to 2017	Africa	Single dataset with soil depths of 0-20 cm and 20-50 cm	2001	Need user account	GeoTIFF, TFRecord, Feature Collection as CSV,SHP, GeoJSON, KML, KMZ and Map Tiles	Yes. Allows upload of own dataset without being	Requires basic knowledge of programming.	Yes. Use built-in codes in javascript for data manipulation	Python and Javascript	Yes	Yes. Easy to navigate and easy to find different datasets.	Yes. It has Google Code Editor to view and manipulate data.	Yes	Yes	Yes	Yes	Yes
Soil Nutrient																									
Cation exchange capacity (CEC)	SLGA: Soil and Landscape Grid of Australia (Soil Attributes)	Commonwealth Scientific and Industrial Research Organisation	Google	Used modelling (cubist algorithm) that describe spatial distribution of soil attributes using existing soil data and environmental covariates	Yes	vis-NIR spectra	250m	1950 to 2013	Australia	Single dataset with soil depths of 0-5cm, 5-15cm, 15-30cm, 30-60cm, 60-100cm, and 100-200cm. Consistent with the Gobal Soil Map.	2001	Need user account	GeoTIFF, TFRecord, Feature Collection as CSV,SHP, GeoJSON, KML, KMZ and Map Tiles	Yes. Allows upload of own dataset without being	Requires basic knowledge of programming.	Yes. Use built-in codes in javascript for data manipulation	Python and Javascript	Yes	Yes. Easy to navigate and easy to find different datasets.	Yes. It has Google Code Editor to view and manipulate data.	Yes	Yes	Yes	Yes	Yes
Soil Salinity																									
Soil Fertility	iSDAsoil Fertility Capability Classification	iSDA	Google	Applied 5 regression modelling algorithms: Random forest , Ranger package, XGBoost, Cubist regression models, Neural network algorithms, and GLM with Lasso or Elasticnet Regularization	Yes	MODIS, digital terrain model (DTM) derivatives, Sentinel-2 satellite and and Landsat-7/8 cloud-free composite images	30 meters	2001 to 2017	Africa	Single dataset with soil depths of 0-20 cm and 20-50 cm	2001	Need user account	GeoTIFF, TFRecord, Feature Collection as CSV,SHP, GeoJSON, KML, KMZ and Map Tiles	Yes. Allows upload of own dataset without being	Requires basic knowledge of programming.	Yes. Use built-in codes in javascript for data manipulation	Python and Javascript	Yes	Yes. Easy to navigate and easy to find different datasets.	Yes. It has Google Code Editor to view and manipulate data.	Yes	Yes	Yes	Yes	Yes

Soil Property	INTRINSIC					CONTEXTUALITY						ACCESSIBILITY			REPRESENTATIONAL										
	Name	**Dataset Provider**	**Point of contact**	**Methods of Processing EO Data**	**Use of Ancillary Data**	**Sensor Used for production of dataset**	**Spatial Resolution**	**Temporal Coverage**	**Spatial Coverage**	**Temporal Resolution/Frequency**	**Year of EO launch**	**Data Access**	**Product Format**	**Security**	**User Technical Knowledge Requirement**	**Processing Tools**	**Programming Languages**	**Cloud Processing**	**User Interface (UI)**	**Visualization Tools**	**Scalability**	**Open Standards**	**User Support**	**Community**	**Product User Guide/Manual**
	What is the name of platform	Who is the source of data	Who is the responsible organization for the EO Platform	What type of method is used	Did it utilize additional/auxiliary and supplementary?	What are the sensors used to produced data?	What is the level of detail captured in an image by a sensor system (pixel size)?	Does it have historical data and the revisit frequency you require?	What is the geographical coverage of data?	What is the repeat/revisit time, e.g., time between successive observations of a given location	When did the EO platform become available and functional?	Is the data free, subscription-based, or pay-per-use?	What are the EO data formats available for download?	Does the platform allows you to use your own data securely?	What is the minimum technical knowledge of user to access and download data?	Does the platform offer built-in tools for data analysis ?	Does it support your preferred programming language (e.g., Python, R) for custom analysis?	Does it leverage cloud computing for handling datasets?	Is the interface user-friendly and easy to navigate?	Does it offer tools for visualizing and exploring the data (e.g., generating maps, charts)?	Can the platform handle high volume of data (e.g., multiple downloads from global to community scale)?	Does it adhere to open data standards for interoperability with other platforms?	Does the platform offer adequate user support (e.g., video tutorials, case studies)?	Is there an active user community for knowledge sharing and troubleshooting?	Is there minimum set of documentation for users?
Soil organic carbon (SOC)	OpenLandMap Soil Organic Carbon Content	Commonwealth Scientific and Industrial Research Organisation	Google	Used spatial prediction (Random Forest, XGBoost).	Yes	Not explicitly mentioned	250m	1950 to 2018	Global	Single dataset with soil depths of 0-20 cm and 20-50 cm	2001	Need user account	GeoTIFF, TFRecord, Feature Collection as CSV,SHP, GeoJSON, KML, KMZ and Map Tiles	Yes. Allows upload of own dataset without being	Requires basic knowledge of programming.	Yes. Use built-in codes in javascript for data manipulation	Python and Javascript	Yes	Yes. Easy to navigate and easy to find different datasets.	Yes. It has Google Code Editor to view and manipulate data.	Yes	Yes	Yes	Yes	Yes
Soil microbial respiration (SMR)																									Yes
Temperature (LST)	MOD11A1.061 Terra Land Surface Temperature and Emissivity Daily Global 1km	NASA LP DAAC at the U	Google	Algorithm based from Wan and Dozier,1989	Yes	MODIS, ASTER, Atmospheric I	1km	2000 to Present	Global	Daily data	2001	Need user account	GeoTIFF, TFRecord, Feature Collection as CSV,SHP, GeoJSON, KML, KMZ and Map Tiles	Yes. Allows upload of own dataset without being	Requires basic knowledge of programming.	Yes. Use built-in codes in javascript for data manipulation	Python and Javascript	Yes	Yes. Easy to navigate and easy to find different datasets.	Yes. It has Google Code Editor to view and manipulate data.	Yes	Yes	Yes	Yes	Yes
Precipitation	CHIRPS Daily: Climate Hazards Group InfraRed Precipitation With Station Data (Version 2.0 Final)	Climate Hazards Center UC Santa Barbara	Google	Generated through a two-part process. Firstly, IR Precipitation (IRP) pentad rainfall estimates are derived from satellite data by calculating the percentage of time during the pentad that the IR observations indicate cold cloud tops (<235° K). Then the station data is integrated with the CHIRP data to create the final product, CHIRPS.	Yes	CHPClim, Quasi-global geostationary thermal infrared (IR) satellite observations from two NOAA sources, Tropical Rainfall Measuring Mission (TRMM) 3B42 product from NASA, Atmospheric model rainfall fields from the NOAA Climate Forecast System	5566 meters	1981 to Present	Global	Daily data	2001	Need user account	GeoTIFF, TFRecord, Feature Collection as CSV,SHP, GeoJSON, KML, KMZ and Map Tiles	Yes. Allows upload of own dataset without being	Requires basic knowledge of programming.	Yes. Use built-in codes in javascript for data manipulation	Python and Javascript	Yes	Yes. Easy to navigate and easy to find different datasets.	Yes. It has Google Code Editor to view and manipulate data.	Yes	Yes	Yes	Yes	Yes
Evapotranspiration	TerraClimate: Monthly Climate and Climatic Water Balance for Global Terrestrial Surfaces, University of Idaho	University of California Merced	Google	Combination of high-resolution WorldClim climatology with time-varying CRU Ts4.0 and JRA55 data using climatically aided interpolation, creating a high-resolution dataset with a broader temporal record.	Yes	CRU Ts4.0, WorldClim dataset (uses SRTM), JRA55	4638.3 meters	1958 to 2023	Global	Monthly Data	2001	Need user account	GeoTIFF, TFRecord, Feature Collection as CSV,SHP, GeoJSON, KML, KMZ and Map Tiles	Yes. Allows upload of own dataset without being	Requires basic knowledge of programming.	Yes. Use built-in codes in javascript for data manipulation	Python and Javascript	Yes	Yes. Easy to navigate and easy to find different datasets.	Yes. It has Google Code Editor to view and manipulate data.	Yes	Yes	Yes	Yes	Yes
Humidity	GLDAS-2.1: Global Land Data Assimilation System	NASA GES DISC at NASA Goddard Space Flight Center	Google	Combination of model and observation data from 2000 to the present with an "open-loop" products without data assimilation. Featuring enhanced models forced by a mix of GDAS, disaggregated GPCP, and AGRMET radiation datasets	Yes	MODIS, GTOPO30	27830 meters	2000 to 2024	Global	Daily Data	2001	Need user account	GeoTIFF, TFRecord, Feature Collection as CSV,SHP, GeoJSON, KML, KMZ and Map Tiles	Yes. Allows upload of own dataset without being	Requires basic knowledge of programming.	Yes. Use built-in codes in javascript for data manipulation	Python and Javascript	Yes	Yes. Easy to navigate and easy to find different datasets.	Yes. It has Google Code Editor to view and manipulate data.	Yes	Yes	Yes	Yes	Yes
Soil Erosion	WWF HydroATLAS Basins Level 05	WWF	Google	Used iterative neighborhood analysis to fill no-data voids	Yes	SRTM, DTED	16 arc-second resolu	2000	South America	Single dataset	2001	Need user account	GeoTIFF, TFRecord, Feature Collection as CSV,SHP, GeoJSON, KML, KMZ and Map Tiles	Yes. Allows upload of own dataset without being	Requires basic knowledge of programming.	Yes. Use built-in codes in javascript for data manipulation	Python and Javascript	Yes	Yes. Easy to navigate and easy to find different datasets.	Yes. It has Google Code Editor to view and manipulate data.	Yes	Yes	Yes	Yes	Yes
Cropland Irrigation	GFSAD1000: Cropland Extent 1km Multi-Study Crop Mask, Global Food-Support Analysis Data	Global Food Security-support Analysis Data at 30m Project (GFSAD30)	Google	Used of Automated Cropland Classification Algorithm (ACCA)	Yes	Landsat, Sentinel, MODIS and AVHRR	1000 meters	2010	Global	Single dataset derived from data on 2007 to 2012	2001	Need user account	GeoTIFF, TFRecord, Feature Collection as CSV,SHP, GeoJSON, KML, KMZ and Map Tiles	Yes. Allows upload of own dataset without being	Requires basic knowledge of programming.	Yes. Use built-in codes in javascript for data manipulation	Python and Javascript	Yes	Yes. Easy to navigate and easy to find different datasets.	Yes. It has Google Code Editor to view and manipulate data.	Yes	Yes	Yes	Yes	Yes

Annex 4: ISRIC EO data for soil

Physiccal Property	INTRINSIC					CONTEXTUALITY						ACCESSIBILITY			REPRESENTATIONAL										
	Name	Dataset Provider	Point of contact	Methods of Processing EO Data	Use of Ancillary Data	Sensor Used for production of dataset	Spatial Resolution	Temporal Coverage	Spatial Coverage	Temporal Resolution/Frequency	Year of EO launch	Data Access	Product Format	Security	User Technical Knowledge Requirement	Processing Tools	Programming Languages	Cloud Processing	User Interface (UI)	Visualization Tools	Scalability	Open Standards	User Support	Community	Product User Guide/Manual
	What is the name of platform	Who is the source of data	Who is the responsible organization for the EO Platform	What type of method is used	Did it utilize additional/auxiliary and supplementary?	What are the sensors used to produced data?	What is the level of detail captured in an image by a sensor system (pixel size)?	Does it have historical data and the revisit frequency you require?	What is the geographical coverage of data?	What is the repeat/revisit time, e.g., time between successive observations of a given location	When did the EO platform become available and functional?	Is the data free, subscription-based, or pay-per-use?	What are the EO data formats available for download?	Does the platform allows you to use your own data securely?	What is the minimum technical knowledge of user to access and download data?	Does the platform offer built-in tools for data analysis ?	Does it support your preferred programming language (e.g., Python, R) for custom analysis?	Does it leverage cloud computing for handling datasets?	Is the interface user-friendly and easy to navigate?	Does it offer tools for visualizing and exploring the data (e.g., generating maps, charts)?	Can the platform handle high volume of data (e.g., multiple downloads from global to community scale)?	Does it adhere to open data standards for interoperability with other platforms?	Does the platform offer adequate user support (e.g., video tutorials, case studies)?	Is there an active user community for knowledge sharing and troubleshooting?	Is there minimum set of documentation for users?
Soil Texture	SoilGrids250m 2017-03 - Texture class (USDA system)	ISRIc	ISRIc	Ensemble of machine learning methods — random forest and gradient boosting and/or multinomial logistic regression — as implemented in the R packages ranger, xgboost, nnet and caret	Yes	MODIS land products, SRTM DEM derivatives, climatic images and global landform and lithology maps	250m	1950 to 2015	Global except Antarctica	Single dataset published on 2017	2017	Free	GeoTIFF	No function to upload your data.	Familiarity with the platform	No. Platform allows you to view and download data	No	No	Yes. Easy to navigate and easy to find different datasets.	Yes. Only for visualizing the geotiff.	Not applicable	Yes	Yes	Yes	Yes
Sand content (Sand-C)	SoilGrids250m 2.0 - Sand content	ISRIC	ISRIC	Ensemble of machine learning methods — random forest and gradient boosting and/or multinomial logistic regression — as implemented in the R packages ranger, xgboost, nnet and caret	Yes	MODIS land products, SRTM DEM derivatives, climatic images and global landform and lithology maps	250m	1905 to 2016	Global except Antarctica	Single dataset published on 2020	2017	Free	GeoTIFF	No function to upload your data.	Familiarity with the platform	No. Platform allows you to view and download data	No	No	Yes. Easy to navigate and easy to find different datasets.	Yes. Only for visualizing the geotiff.	Not applicable	Yes	Yes	Yes	Yes
Silt content (Silt-C)	WoSIS latest - Silt total	ISRIC	ISRIC	Ensemble of machine learning methods — random forest and gradient boosting and/or multinomial logistic regression — as implemented in the R packages ranger, xgboost, nnet and caret	Yes	Not used. Based from global compilation of soil profile data and environmental layers	250m	1918 to 2013	Global	Single dataset published on 2021	2017	Free	JSON, Geopackage, Shapezip, Text, CSV,XML	No function to upload your data.	Familiarity with the platform	No. Platform allows you to view and download data	No	No	Yes. Easy to navigate and easy to find different datasets.	Yes. Only for visualizing the geotiff.	Not applicable	Yes	Yes	Yes	Yes
Clay content (Clay-C)	SoilGrids250m 2.0 - Clay content	ISRIC	ISRIC	Ensemble of machine learning methods — random forest and gradient boosting and/or multinomial logistic regression — as implemented in the R packages ranger, xgboost, nnet and caret	Yes	MODIS land products, SRTM DEM derivatives, climatic images and global landform and lithology maps	250m	1905 to 2016	Global except Antarctica	Single dataset published on 2020	2017	Free	WMS,VRT (GDAL Virtual Format)	No function to upload your data.	Familiarity with the platform	No. Platform allows you to view and download data	No	No	Yes. Easy to navigate and easy to find different datasets.	Yes. Only for visualizing the geotiff.	Not applicable	Yes	Yes	Yes	Yes
Bulk Density (BD)	SoilGrids250m 2017-03 - Bulk density (fine earth)	ISRIC	ISRIC	Ensemble of machine learning methods — random forest and gradient boosting and/or multinomial logistic regression — as implemented in the R packages ranger, xgboost, nnet and caret	Yes	MODIS land products, SRTM DEM derivatives, climatic images and global landform and lithology maps	250m	1950 to 2015	Global except Antarctica	Single dataset published on 2017	2017	Free	GeoTIFF	No function to upload your data.	Familiarity with the platform	No. Platform allows you to view and download data	No	No	Yes. Easy to navigate and easy to find different datasets.	Yes. Only for visualizing the geotiff.	Not applicable	Yes	Yes	Yes	Yes
Available Water Capacity	SoilGrids250m 2017-03 - Derived available soil water capacity (volumetric fraction) with FC = pF 2.5	ISRIC	ISRIC	Ensemble of machine learning methods — random forest and gradient boosting and/or multinomial logistic regression — as implemented in the R packages ranger, xgboost, nnet and caret	Yes	MODIS land products, SRTM DEM derivatives, climatic images and global landform and lithology maps	250m	1950 to 2015	Global except Antarctica	Single dataset published on 2017	2017	Free	GeoTIFF	No function to upload your data.	Familiarity with the platform	No. Platform allows you to view and download data	No	No	Yes. Easy to navigate and easy to find different datasets.	Yes. Only for visualizing the geotiff.	Not applicable	Yes	Yes	Yes	Yes
Soil Moisture	**Africa SoilGrids - Root zone moisture content at wilting point aggregated at ERZD**	ISRIC	ISRIC	From compiles georeferenced dataset of 28,000 soil profiles for SSA and application of DSM techniques. Developing and parameterizing pedotransfer functions, rules, and criteria.	Yes	Not explicitly mentioned	1000 meters	1950 to 2015	Global	Single dataset published on 2017	2017	Free	GeoTiff	No function to upload your data.	Familiarity with the platform	No. Platform allows you to view and download data	No	No	Yes. Easy to navigate and easy to find different datasets.	Yes. Only for visualizing the geotiff.	Not applicable	Yes	Yes	Yes	Yes

Chemical Property	INTRINSIC					CONTEXTUALITY						ACCESSIBILITY			REPRESENTATIONAL										
	Name	Dataset Provider	Point of contact	Methods of Processing EO Data	Use of Ancillary Data	Sensor Used for production of dataset	Spatial Resolution	Temporal Coverage	Spatial Coverage	Temporal Resolution/Frequency	Year of EO launch	Data Access	Product Format	Security	User Technical Knowledge Requirement	Processing Tools	Programming Languages	Cloud Processing	User Interface (UI)	Visualization Tools	Scalability	Open Standards	User Support	Community	Product User Guide/Manual
	What is the name of platform	Who is the source of data	Who is the responsible organization for the EO Platform	What type of method is used	Did it utilize additional/auxiliary and supplementary?	What are the sensors used to produced data?	What is the level of detail captured in an image by a sensor system (pixel size)?	Does it have historical data and the revisit frequency you require?	What is the geographical coverage of data?	What is the repeat/revisit time, e.g., time between successive observations of a given location	When did the EO platform become available and functional?	Is the data free, subscription-based, or pay-per-use?	What are the EO data formats available for download?	Does the platform allows you to use your own data securely?	What is the minimum technical knowledge of user to access and download data?	Does the platform offer built-in tools for data analysis ?	Does it support your preferred programming language (e.g., Python, R) for custom analysis?	Does it leverage cloud computing for handling datasets?	Is the interface user-friendly and easy to navigate?	Does it offer tools for visualizing and exploring the data (e.g., generating maps, charts)?	Can the platform handle high volume of data (e.g., multiple downloads from global to community scale)?	Does it adhere to open data standards for interoperability with other platforms?	Does the platform offer adequate user support (e.g., video tutorials, case studies)?	Is there an active user community for knowledge sharing and troubleshooting?	Is there minimum set of documentation for users?
Hydrogen Potential (pH)	SoilGrids250m 2017-03 - Soil pH in H2O	ISRIC	ISRIC	Ensemble of machine learning methods — random forest and gradient boosting and/or multinomial logistic regression — as implemented in the R packages ranger, xgboost, nnet and caret	Yes	MODIS land products, SRTM DEM derivatives, climatic images and global landform and lithology maps	250m	1951 to 2015	Global except Antarctica	Single dataset published on 2018	2017	Free	GeoTIFF	No function to upload your data.	Familiarity with the platform	No. Platform allows you to view and download data	No	No	Yes. Easy to navigate and easy to find different datasets.	Yes. Only for visualizing the geotiff.	Not applicable	Yes	Yes	Yes	Yes
Total nitrogen (TN)	SoilGrids250m 2.0 - Total nitrogen	ISRIC	ISRIC	Ensemble of machine learning methods — random forest and gradient boosting and/or multinomial logistic regression — as implemented in the R packages ranger, xgboost, nnet and caret	Yes	MODIS land products, SRTM DEM derivatives, climatic images and global landform and lithology maps	250m	1905 to 2016	Global except Antarctica	Single dataset published on 2019	2017	Free	WMS, VRT (GDAL Virtual Format)	No function to upload your data.	Familiarity with the platform	No. Platform allows you to view and download data	No	No	Yes. Easy to navigate and easy to find different datasets.	Yes. Only for visualizing the geotiff.	Not applicable	Yes	Yes	Yes	Yes
Potassium (K)	Africa SoilGrids nutrients - Extractable Potassium (K)	ISRIC	ISRIC	Model fitting and prediction were undertaken using an ensemble of two Machine Learning algorithms (MLA): ranger (random forest) and xgboost (Gradient Boosting Tree), as implemented in the R environment for statistical computing.	Yes	MODIS, Landsat, DEM derived surfaces, Global Surface Water dynamics images, Land cover map of the world at 300 m resolution	250m	1980 to 2016	Africa	Single dataset published on 2016	2017	Free	GeoTiff	No function to upload your data.	Familiarity with the platform	No. Platform allows you to view and download data	No	No	Yes. Easy to navigate and easy to find different datasets.	Yes. Only for visualizing the geotiff.	Not applicable	Yes	Yes	Yes	Yes
Calcium (Ca)	Africa SoilGrids nutrients - Extractable Calcium (Ca)	ISRIC	ISRIC	Model fitting and prediction were undertaken using an ensemble of two Machine Learning algorithms (MLA): ranger (random forest) and xgboost (Gradient Boosting Tree), as implemented in the R environment for statistical computing.	Yes	MODIS, Landsat, DEM derived surfaces, Global Surface Water dynamics images, Land cover map of the world at 300 m resolution	250 meters	1980 to 2016	Africa	Single dataset published on 2016	2017	Free	GeoTiff	No function to upload your data.	Familiarity with the platform	No. Platform allows you to view and download data	No	No	Yes. Easy to navigate and easy to find different datasets.	Yes. Only for visualizing the geotiff.	Not applicable	Yes	Yes	Yes	Yes
Magnesium (Mg)	Africa SoilGrids nutrients - Extractable Magnesium (Mg)	ISRIC	ISRIC	Model fitting and prediction were undertaken using an ensemble of two Machine Learning algorithms (MLA): ranger (random forest) and xgboost (Gradient Boosting Tree), as implemented in the R environment for statistical computing.	Yes	MODIS, Landsat, DEM derived surfaces, Global Surface Water dynamics images, Land cover map of the world at 300 m resolution	250 meters	1980 to 2016	Africa	Single dataset published on 2016	2017	Free	GeoTiff	No function to upload your data.	Familiarity with the platform	No. Platform allows you to view and download data	No	No	Yes. Easy to navigate and easy to find different datasets.	Yes. Only for visualizing the geotiff.	Not applicable	Yes	Yes	Yes	Yes
Phosphorus (P)	Global distribution of soil phosphorus retention potential	ISRIC	ISRIC	Model fitting and prediction were undertaken using an ensemble of two Machine Learning algorithms (MLA): ranger (random forest) and xgboost (Gradient Boosting Tree), as implemented in the R environment for statistical computing.	Yes	Not explicitly mentioned	5000000 meters	2001 to 2011	Global	Single dataset published on 2011	2017	Free	Shapefile	No function to upload your data.	Familiarity with the platform	No. Platform allows you to view and download data	No	No	Yes. Easy to navigate and easy to find different datasets.	Yes. Only for visualizing the thumbnail image.	Not applicable	Yes	Yes	Yes	Yes
Soil Nutrient	Africa SoilGrids nutrients - Nutrient clusters based on fuzzy k-means	ISRIC	ISRIC	Generated by training machine learning models with data from 59,000+ soil samples and a diverse set of covariates, including remote sensing data, landform, lithologic, and land cover layers, using ensemble methods like random forest and gradient boosting in R packages ranger and xgboost.	Yes	MODIS, Landsat, DEM derived surfaces, Global Surface Water dynamics images, Land cover map of the world at 300 m resolution	250 meters	1980 to 2016	Africa	Single dataset published on 2016	2017	Free	GeoTiff	No function to upload your data.	Familiarity with the platform	No. Platform allows you to view and download data	No	No	Yes. Easy to navigate and easy to find different datasets.	Yes. Only for visualizing the geotiff.	Not applicable	Yes	Yes	Yes	Yes
Cation exchange capacity (CEC)	WoSIS latest - Effective cation exchange capacity - ISRIC	ISRIC	ISRIC	Ensemble of machine learning methods — random forest and gradient boosting and/or multinomial logistic regression — as implemented in the R packages ranger, xgboost, nnet and caret	Yes	Not used. Based from global compilation of soil profile data and environmental layers	100000	1918 to 2013	Global	Single dataset published on 2020	2017	Free	Shapefile	No function to upload your data.	Familiarity with the platform	No. Platform allows you to view and download data	No	No	Yes. Easy to navigate and easy to find different datasets.	Yes. Only for visualizing the geotiff.	Not applicable	Yes	Yes	Yes	Yes
Soil Salinity	Global Soil Salinity Map	ISRIC	ISRIC	Used random forest classifier that was trained using seven soil properties maps, thermal infrared imagery and the ECe point data from the WoSIS database	Yes	USGS Landsat 5 Surface Reflectance Tier 1 collection and USGS Landsat 8 Surface Reflectance Tier 1 collection	250 meters	1986 to 2016	Global	Years 1986, 1992, 2000, 2002, 2005, 2009 and 2016	2017	Free	GeoTiff	No function to upload your data.	Familiarity with the platform	No. Platform allows you to view and download data	No	No	Yes. Easy to navigate and easy to find different datasets.	Yes. Only for visualizing the geotiff.	Not applicable	Yes	Yes	Yes	Yes
Soil Fertility																									

Soil Property	INTRINSIC					CONTEXTUALITY						ACCESSIBILITY			REPRESENTATIONAL										
	Name	**Dataset Provider**	**Point of contact**	**Methods of Processing EO Data**	**Use of Ancillary Data**	**Sensor Used for production of dataset**	**Spatial Resolution**	**Temporal Coverage**	**Spatial Coverage**	**Temporal Resolution/Frequency**	**Year of EO launch**	**Data Access**	**Product Format**	**Security**	**User Technical Knowledge Requirement**	**Processing Tools**	**Programming Languages**	**Cloud Processing**	**User Interface (UI)**	**Visualization Tools**	**Scalability**	**Open Standards**	**User Support**	**Community**	**Product User Guide/Manual**
	What is the name of platform	Who is the source of data	Who is the responsible organization for the EO Platform	What type of method is used	Did it utilize additional/auxiliary and supplementary?	What are the sensors used to produced data?	What is the level of detail captured in an image by a sensor system (pixel size)?	Does it have historical data and the revisit frequency you require?	What is the geographical coverage of data?	What is the repeat/revisit time, e.g., time between successive observations of a given location	When did the EO platform become available and functional?	Is the data free, subscription-based, or pay-per-use?	What are the EO data formats available for download?	Does the platform allows you to use your own data securely?	What is the minimum technical knowledge of user to access and download data?	Does the platform offer built-in tools for data analysis ?	Does it support your preferred programming language (e.g., Python, R) for custom analysis?	Does it leverage cloud computing for handling datasets?	Is the interface user-friendly and easy to navigate?	Does it offer tools for visualizing and exploring the data (e.g., generating maps, charts)?	Can the platform handle high volume of data (e.g., multiple downloads from global to community scale)?	Does it adhere to open data standards for interoperability with other platforms?	Does the platform offer adequate user support (e.g., video tutorials, case studies)?	Is there an active user community for knowledge sharing and troubleshooting?	Is there minimum set of documentation for users?
Soil organic carbon (SOC)	SoilGrids250m 2.0 - Soil organic carbon content	ISRIC	ISRIC	Used Quantile Random Forest digital soil mapping models trained on a global dataset of soil profiles and environmental covariates	Yes	MODIS land products, SRTM DEM derivatives, climatic images and global landform and lithology maps	250 meters	1905 to 2016	Global	Single dataset published on 2020	2017	Free	WMS, VRT (GDAL Virtual Format)	No function to upload your data.	Familiarity with the platform	No. Platform allows you to view and download data	No	No	Yes. Easy to navigate and easy to find different datasets.	Yes. Only for visualizing the geotiff.	Not applicable	Yes	Yes	Yes	Yes
Soil microbial respiration (SMR)																									
Temperature (LST)																									
Precipitation																									
Evapotranspiration																									
Humidity																									
Soil Erosion																									
Cropland Irrigation																									

Annex 5: Earth Map EO data for soil

Soil Property	INTRINSIC					CONTEXTUALITY						ACCESSIBILITY			REPRESENTATIONAL										
	Name	**Dataset Provider**	**Point of contact**	**Methods of Processing EO Data**	**Use of Ancillary Data**	**Sensor Used for production of dataset**	**Spatial Resolution**	**Temporal Coverage**	**Spatial Coverage**	**Temporal Resolution/Frequency**	**Year of EO launch**	**Data Access**	**Product Format**	**Security**	**User Technical Knowledge Requirement**	**Processing Tools**	**Programming Languages**	**Cloud Processing**	**User Interface (UI)**	**Visualization Tools**	**Scalability**	**Open Standards**	**User Support**	**Community**	**Product User Guide/Manual**
	What is the name of platform	Who is the source of data	Who is the responsible organization for the EO Platform	What type of method is used	Did it utilize additional/auxiliary and supplementary?	What are the sensors used to produced data?	What is the level of detail captured in an image by a sensor system (pixel size)?	Does it have historical data and the revisit frequency you require?	What is the geographical coverage of data?	What is the repeat/revisit time, e.g., time between successive observations of a given location	When did the EO platform become available and functional?	Is the data free, subscription-based, or pay-per-use?	What are the EO data formats available for download?	Does the platform allows you to use your own data securely?	What is the minimum technical knowledge of user to access and download data?	Does the platform offer built-in tools for data analysis ?	Does it support your preferred programming language (e.g., Python, R) for custom analysis?	Does it leverage cloud computing for handling datasets?	Is the interface user-friendly and easy to navigate?	Does it offer tools for visualizing and exploring the data (e.g., generating maps, charts)?	Can the platform handle high volume of data (e.g., multiple downloads from global to community scale)?	Does it adhere to open data standards for interoperability with other platforms?	Does the platform offer adequate user support (e.g., video tutorials, case studies)?	Is there an active user community for knowledge sharing and troubleshooting?	Is there minimum set of documentation for users?
Soil Texture	OpenLandMap Soil texture	EnvirometriX Ltd	UN-FAO with Google	Predicted soil texture fractions using the soiltexture package in R	Yes	Not explicitly mentioned	250m	1950 to 2018	Global except Antarctica	Single dataset with soil texture class bands 0 cm - 200 cm with 6 standard depths		Need user account	PNG, GeoTIFF, CSV,XCLS	Yes. Not possible to add your own Area of Interest	Yes. Use built-in codes in javascript for data manipulation	Familiarity with the platform.	Not applicable	Yes	Yes. Easy to navigate and easy to find different datasets.	Yes. It has interface for data analysis, graphs and charts	Yes. It is powered by Google Earth Engine	Yes	Yes	Yes	Yes
Sand content (Sand-C)																									
Silt content (Silt-C)																									
Clay content (Clay-C)																									
Bulk Density (BD)																									
Available Water Capacity	Total Available Water	FAO	UN-FAO with Google	Determining the Total Available Water (TAW) in the root zone. This calculation considers the difference between the water content at field capacity and the wilting point.	Yes	Not explicitly mentioned	250 meters		Global	Single dataset published on 2012		Need user account	PNG, GeoTIFF, CSV,XCLS	Yes. Not possible to add your own Area of Interest	Yes. Use built-in codes in javascript for data manipulation	Familiarity with the platform.	Not applicable	Yes	Yes. Easy to navigate and easy to find different datasets.	Yes. It has interface for data analysis, graphs and charts	Yes. It is powered by Google Earth Engine	Yes	Yes	Yes	Yes

Soil Property	INTRINSIC					CONTEXTUALITY						ACCESSIBILITY			REPRESENTATIONAL										
	Name	Dataset Provider	Point of contact	Methods of Processing EO Data	Use of Ancillary Data	Sensor Used for production of dataset	Spatial Resolution	Temporal Coverage	Spatial Coverage	Temporal Resolution/Frequency	Year of EO launch	Data Access	Product Format	Security	User Technical Knowledge Requirement	Processing Tools	Programming Languages	Cloud Processing	User Interface (UI)	Visualization Tools	Scalability	Open Standards	User Support	Community	Product User Guide/Manual
	What is the name of platform	Who is the source of data	Who is the responsible organization for the EO Platform	What type of method is used	Did it utilize additional/auxiliary and supplementary?	What are the sensors used to produced data?	What is the level of detail captured in an image by a sensor system (pixel size)?	Does it have historical data and the revisit frequency you require?	What is the geographical coverage of data?	What is the repeat/revisit time, e.g., time between successive observations of a given location	When did the EO platform become available and functional?	Is the data free, subscription-based, or pay-per-use?	What are the EO data formats available for download?	Does the platform allows you to use your own data securely?	What is the minimum technical knowledge of user to access and download data?	Does the platform offer built-in tools for data analysis ?	Does it support your preferred programming language (e.g., Python, R) for custom analysis?	Does it leverage cloud computing for handling datasets?	Is the interface user-friendly and easy to navigate?	Does it offer tools for visualizing and exploring the data (e.g., generating maps, charts)?	Can the platform handle high volume of data (e.g., multiple downloads from global to community scale)?	Does it adhere to open data standards for interoperability with other platforms?	Does the platform offer adequate user support (e.g., video tutorials, case studies)?	Is there an active user community for knowledge sharing and troubleshooting?	Is there minimum set of documentation for users?
Hydrogen Potential (pH)																									
Total nitrogen (TN)																									
Potassium (K)																									
Calcium (Ca)																									
Magnesium (Mg)																									
Phosphorus (P)																									
Soil Nutrient																									
Cation exchange capacity (CEC)																									
Soil Salinity																									
Soil organic carbon (SOC)	Global Soil Organic Carbon - GSOC	FAO	UN-FAO with Google	Use nearest neighbor resampling, bilinear interpolation to resample to 30 arc-second grid then mosaicked using nearest neighbor resampling. NoData values at national borders were filled using GDAL gapfilling with inverse distance weighting and a four direction conic search within a 5-km buffer mask, and a global mask of water bodies was applied.	Yes	MODIS, Landsat, DEM derivatives, climatic images and global landform and lithology maps	1000 meters	1972 to 2004 (Based from the data used)	Global	Single dataset published on 2012		Need user account	PNG, GeoTIFF, CSV,XCLS	Yes. Not possible to add your own Area of Interest	Yes. Use built-in codes in javascript for data manipulation	Familiarity with the platform.	Not applicable	Yes	Yes. Easy to navigate and easy to find different datasets.	Yes. It has interface for data analysis, graphs and charts	Yes. It is powered by Google Earth Engine	Yes	Yes	Yes	Yes
Soil microbial respiration (SMR)																									
Temperature (LST)																									
Precipitation	Prec (average) - CHIRPS	FAO	UN-FAO with Google	Generated through a two-part process. Firstly, IR Precipitation (IRP) pentad rainfall estimates are derived from satellite data by calculating the percentage of time during the pentad that the IR observations indicate cold cloud tops (<235° K). Then the station data is integrated with the CHIRP data to create the final product, CHIRPS.	Yes	CHPClim, Quasi-global geostationary thermal infrared (IR) satellite observations from two NOAA sources, Tropical Rainfall Measuring Mission (TRMM) 3B42 product from NASA, Atmospheric model rainfall fields from the NOAA Climate Forecast System	5566 meters	1981 to Present	Global	Daily data		Need user account	PNG, GeoTIFF, CSV,XCLS	Yes. Not possible to add your own Area of Interest	Yes. Use built-in codes in javascript for data manipulation	Familiarity with the platform.	Not applicable	Yes	Yes. Easy to navigate and easy to find different datasets.	Yes. It has interface for data analysis, graphs and charts	Yes. It is powered by Google Earth Engine	Yes	Yes	Yes	Yes
Evapotranspiration	PET (average) - MODIS	FAO	UN-FAO with Google	Based on the logic of the Penman-Monteith equation, which includes inputs of daily meteorological reanalysis data along with MODIS remotely sensed data products such as vegetation property dynamics, albedo, and land cover.	Yes	MODIS	500 meters	2000 to 2021	Global	Single dataset published on 2018		Need user account	PNG, GeoTIFF, CSV,XCLS	Yes. Not possible to add your own Area of Interest	Yes. Use built-in codes in javascript for data manipulation	Familiarity with the platform.	Not applicable	Yes	Yes. Easy to navigate and easy to find different datasets.	Yes. It has interface for data analysis, graphs and charts	Yes. It is powered by Google Earth Engine	Yes	Yes	Yes	Yes
Humidity																									
Soil Moisture																									
Soil Fertility																									
Soil Erosion	**Runoff (yearly) - ECMWF Land / CCI LC / SRTM DEM**	GFSAD 1000	UN-FAO with Google	Calculate runoff using the CN and precipitation data	Yes	derived from 2019 CCI LandCover 300 m/SoilGrids Clay	90 meters	1981 to 2021	Global	Annually (1981 to 2024)		Need user account	PNG, GeoTIFF, CSV,XCLS	Yes. Not possible to add your own Area of Interest	Yes. Use built-in codes in javascript for data manipulation	Familiarity with the platform.	Not applicable	Yes	Yes. Easy to navigate and easy to find different datasets.	Yes. It has interface for data analysis, graphs and charts	Yes. It is powered by Google Earth Engine	Yes	Yes	Yes	Yes
Cropland Irrigation	Cropland - GFSAD1000	GFSAD 1000	UN-FAO with Google	overlaying the five dominant crops of the world from studies by Ramankutty et al. (2008), Monfreda et al. (2008), and Portman et al. (2009) over a global irrigated and rainfed cropland area map derived from remote sensing data by the International Water Management Institute (IWMI; Thenkabail et al., 2009a, 2009b, 2011, Biradar et al., 2009).	Yes	MODIS, SPOT, AVHRR	1000 meters	2007 to 2012	Global	Single dataset published on 2010		Need user account	PNG, GeoTIFF, CSV,XCLS	Yes. Not possible to add your own Area of Interest	Yes. Use built-in codes in javascript for data manipulation	Familiarity with the platform.	Not applicable	Yes	Yes. Easy to navigate and easy to find different datasets.	Yes. It has interface for data analysis, graphs and charts	Yes. It is powered by Google Earth Engine	Yes	Yes	Yes	Yes

Annex 6: ORNL-DAAC EO data for soil

Soil Property	INTRINSIC					CONTEXTUALITY						ACCESSIBILITY			REPRESENTATIONAL										
	Name	Dataset Provider	Point of contact	Methods of Processing EO Data	Use of Ancillary Data	Sensor Used for production of dataset	Spatial Resolution	Temporal Coverage	Spatial Coverage	Temporal Resolution/Frequency	Year of EO launch	Data Access	Product Format	Security	User Technical Knowledge Requirement	Processing Tools	Programming Languages	Cloud Processing	User Interface (UI)	Visualization Tools	Scalability	Open Standards	User Support	Community	Product User Guide/Manual
	What is the name of platform	Who is the source of data	Who is the responsible organization for the EO Platform	What type of method is used	Did it utilize additional/auxiliary and supplementary?	What are the sensors used to produced data?	What is the level of detail captured in an image by a sensor system (pixel size)?	Does it have historical data and the revisit frequency you require?	What is the geographical coverage of data?	What is the repeat/revisit time, e.g., time between successive observations of a given location	When did the EO platform become available and functional?	Is the data free, subscription-based, or pay-per-use?	What are the EO data formats available for download?	Does the platform allows you to use your own data securely?	What is the minimum technical knowledge of user to access and download data?	Does the platform offer built-in tools for data analysis ?	Does it support your preferred programming language (e.g., Python, R) for custom analysis?	Does it leverage cloud computing for handling datasets?	Is the interface user-friendly and easy to navigate?	Does it offer tools for visualizing and exploring the data (e.g., generating maps, charts)?	Can the platform handle high volume of data (e.g., multiple downloads from global to community scale)?	Does it adhere to open data standards for interoperability with other platforms?	Does the platform offer adequate user support (e.g., video tutorials, case studies)?	Is there an active user community for knowledge sharing and troubleshooting?	Is there minimum set of documentation for users?
Soil Texture	Global Soil Texture and Derived Water-Holding Capacities (Webb et al.)	Oak Ridge National Laboratory Distributed Active Archive Center	NASA Earth Data	Harmonizing soil water characteristics curves (SWCCs) from various sites to create a global soil hydraulic properties (GSHP) database, with parameters estimated using the van Genuchten (vG) SWCC model and pedotransfer functions (PTFs) to estimate missing information	Yes	Not explicitly mentioned	1 degree	Single dataset published on 2000	Global except Antarctica	1950 to 1996	2017	Need user account	Arc/Info ASCII Grid, Erdas Imagine, GeoTIFF, netcdf, NITF, XYZ	Yes. Not possible to add your own Area of Interest	Familiarity with the platform.	No. Only allow view and download of dataset.	Not applicable	No	Yes. Easy to navigate and easy to find different datasets.	Yes.Allows to view dataset in kml and has Spatial Data Access Tool	Yes.	Yes	Yes	Yes	Yes
Sand content (Sand-C)	Global Soil Texture and Derived Water-Holding Capacities (Webb et al.)	Oak Ridge National Laboratory Distributed Active Archive Center	NASA Earth Data	Harmonizing soil water characteristics curves (SWCCs) from various sites to create a global soil hydraulic properties (GSHP) database, with parameters estimated using the van Genuchten (vG) SWCC model and pedotransfer functions (PTFs) to estimate missing information	Yes	Not explicitly mentioned	1 degree	Single dataset published on 2000	Global except Antarctica	1950 to 1996	2017	Need user account	Arc/Info ASCII Grid, Erdas Imagine, GeoTIFF, netcdf, NITF, XYZ	Yes. Not possible to add your own Area of Interest	Familiarity with the platform.	No. Only allow view and download of dataset.	Not applicable	No	Yes. Easy to navigate and easy to find different datasets.	Yes.Allows to view dataset in kml and has Spatial Data Access Tool	Yes.	Yes	Yes	Yes	Yes
Silt content (Silt-C)	Global Soil Texture and Derived Water-Holding Capacities (Webb et al.)	Oak Ridge National Laboratory Distributed Active Archive Center	NASA Earth Data	Harmonizing soil water characteristics curves (SWCCs) from various sites to create a global soil hydraulic properties (GSHP) database, with parameters estimated using the van Genuchten (vG) SWCC model and pedotransfer functions (PTFs) to estimate missing information	Yes	Not explicitly mentioned	1 degree	Single dataset published on 2000	Global except Antarctica	1950 to 1996	2017	Need user account	Arc/Info ASCII Grid, Erdas Imagine, GeoTIFF, netcdf, NITF, XYZ	Yes. Not possible to add your own Area of Interest	Familiarity with the platform.	No. Only allow view and download of dataset.	Not applicable	No	Yes. Easy to navigate and easy to find different datasets.	Yes.Allows to view dataset in kml and has Spatial Data Access Tool	Yes.	Yes	Yes	Yes	Yes
Clay content (Clay-C)	Global Soil Texture and Derived Water-Holding Capacities (Webb et al.)	Oak Ridge National Laboratory Distributed Active Archive Center	NASA Earth Data	Harmonizing soil water characteristics curves (SWCCs) from various sites to create a global soil hydraulic properties (GSHP) database, with parameters estimated using the van Genuchten (vG) SWCC model and pedotransfer functions (PTFs) to estimate missing information	Yes	Not explicitly mentioned	1 degree	Single dataset published on 2000	Global except Antarctica	1950 to 1996	2017	Need user account	Arc/Info ASCII Grid, Erdas Imagine, GeoTIFF, netcdf, NITF, XYZ	Yes. Not possible to add your own Area of Interest	Familiarity with the platform.	No. Only allow view and download of dataset.	Not applicable	No	Yes. Easy to navigate and easy to find different datasets.	Yes.Allows to view dataset in kml and has Spatial Data Access Tool	Yes.	Yes	Yes	Yes	Yes
Bulk Density (BD)	Global Gridded Surfaces of Selected Soil Characteristics (IGBP-DIS)	Oak Ridge National Laboratory Distributed Active Archive Center	NASA Earth Data	Employs a statistical bootstrapping technique to associate pedon data from the Global Pedon Database with the FAO/UNESCO Digital Soil Map of the World.	Yes	Not explicitly mentioned	5x5 arc-minutes	Single dataset published on 2000	Global	1950 to 1996	2017	Need user account	Arc/Info ASCII Grid, Erdas Imagine, GeoTIFF, netcdf, NITF, XYZ	Yes. Not possible to add your own Area of Interest	Familiarity with the platform.	No. Only allow view and download of dataset.	Not applicable	No	Yes. Easy to navigate and easy to find different datasets.	Yes.Allows to view dataset in kml and has Spatial Data Access Tool	Yes.	Yes	Yes	Yes	Yes
Available Water Capacity	Global Soil Texture and Derived Water-Holding Capacities (Webb et al.)	Oak Ridge National Laboratory Distributed Active Archive Center	NASA Earth Data	Harmonizing soil water characteristics curves (SWCCs) from various sites to create a global soil hydraulic properties (GSHP) database, with parameters estimated using the van Genuchten (vG) SWCC model and pedotransfer functions (PTFs) to estimate missing information	Yes	Not explicitly mentioned	1 degree	Single dataset published on 2000	Global	1950 to 1996	2017	Need user account	Arc/Info ASCII Grid, Erdas Imagine, GeoTIFF, netcdf, NITF, XYZ	Yes. Not possible to add your own Area of Interest	Familiarity with the platform.	No. Only allow view and download of dataset.	Not applicable	No	Yes. Easy to navigate and easy to find different datasets.	Yes.Allows to view dataset in kml and has Spatial Data Access Tool	Yes.	Yes	Yes	Yes	Yes
Hydrogen Potential (pH)																									
Total nitrogen (TN)	Global Gridded Surfaces of Selected Soil Characteristics (IGBP-DIS)	Oak Ridge National Laboratory Distributed Active Archive Center	NASA Earth Data	Employs a statistical bootstrapping technique to associate pedon data from the Global Pedon Database with the FAO/UNESCO Digital Soil Map of the World.	Yes	Not explicitly mentioned	5x5 arc-minutes	Single dataset published on 2000	Global	1950 to 1996	2017	Need user account	Arc/Info ASCII Grid, Erdas Imagine, GeoTIFF, netcdf, NITF, XYZ	Yes. Not possible to add your own Area of Interest	Familiarity with the platform.	No. Only allow view and download of dataset.	Not applicable	No	Yes. Easy to navigate and easy to find different datasets.	Yes.Allows to view dataset in kml and has Spatial Data Access Tool	Yes.	Yes	Yes	Yes	Yes
Potassium (K)																									
Calcium (Ca)																									
Magnesium (Mg)																									

Soil Property	INTRINSIC					CONTEXTUALITY						ACCESSIBILITY			REPRESENTATIONAL										
	Name	**Dataset Provider**	**Point of contact**	**Methods of Processing EO Data**	**Use of Ancillary Data**	**Sensor Used for production of dataset**	**Spatial Resolution**	**Temporal Coverage**	**Spatial Coverage**	**Temporal Resolution/Frequency**	**Year of EO launch**	**Data Access**	**Product Format**	**Security**	**User Technical Knowledge Requirement**	**Processing Tools**	**Programming Languages**	**Cloud Processing**	**User Interface (UI)**	**Visualization Tools**	**Scalability**	**Open Standards**	**User Support**	**Community**	**Product User Guide/Manual**
	What is the name of platform	Who is the source of data	Who is the responsible organization for the EO Platform	What type of method is used	Did it utilize additional/auxiliary and supplementary?	What are the sensors used to produced data?	What is the level of detail captured in an image by a sensor system (pixel size)?	Does it have historical data and the revisit frequency you require?	What is the geographical coverage of data?	What is the repeat/revisit time, e.g., time between successive observations of a given location	When did the EO platform become available and functional?	Is the data free, subscription-based, or pay-per-use?	What are the EO data formats available for download?	Does the platform allows you to use your own data securely?	What is the minimum technical knowledge of user to access and download data?	Does the platform offer built-in tools for data analysis ?	Does it support your preferred programming language (e.g., Python, R) for custom analysis?	Does it leverage cloud computing for handling datasets?	Is the interface user-friendly and easy to navigate?	Does it offer tools for visualizing and exploring the data (e.g., generating maps, charts)?	Can the platform handle high volume of data (e.g., multiple downloads from global to community scale)?	Does it adhere to open data standards for interoperability with other platforms?	Does the platform offer adequate user support (e.g., video tutorials, case studies)?	Is there an active user community for knowledge sharing and troubleshooting?	Is there minimum set of documentation for users?
Phosphorus (P)	A Compilation of Global Soil Microbial Biomass Carbon, Nitrogen, and Phosphorus Data	Oak Ridge National Laboratory Distributed Active Archive Center	NASA Earth Data	Used multiple linear regression to obtain correlations between soil microbial biomass or soil element concentrations and long term climate variables done using the R 2.12.3	Yes	Not explicitly mentioned	0.05-degree by 0.5-degree	Single dataset published on 2015	Global	1977 to 2012	2017	Need user account	Arc/Info ASCII Grid, Erdas Imagine, GeoTIFF, netcdf, NITF, XYZ	Yes. Not possible to add your own Area of Interest	Familiarity with the platform.	No. Only allow view and download of dataset.	Not applicable	No	Yes. Easy to navigate and easy to find different datasets.	Yes.Allows to view dataset in kml and has Spatial Data Access Tool	Yes.	Yes	Yes	Yes	Yes
Soil Nutrient																									
Cation exchange capacity (CEC)																									
Soil Salinity																									
Soil Moisture	Global Gridded Surfaces of Selected Soil Characteristics (IGBP-DIS)	Oak Ridge National Laboratory Distributed Active Archive Center	NASA Earth Data	Employs a statistical bootstrapping technique to associate pedon data from the Global Pedon Database with the FAO/UNESCO Digital Soil Map of the World.	Yes	Not explicitly mentioned	5x5 arc-minutes	Single dataset published on 2000	Global	1950 to 1996	2017	Need user account	Arc/Info ASCII Grid, Erdas Imagine, GeoTIFF, netcdf, NITF, XYZ	Yes. Not possible to add your own Area of Interest	Familiarity with the platform.	No. Only allow view and download of dataset.	Not applicable	No	Yes. Easy to navigate and easy to find different datasets.	Yes.Allows to view dataset in kml and has Spatial Data Access Tool	Yes.	Yes	Yes	Yes	Yes
Soil organic carbon (SOC)	Global Gridded Surfaces of Selected Soil Characteristics (IGBP-DIS)	Oak Ridge National Laboratory Distributed Active Archive Center	NASA Earth Data	Employs a statistical bootstrapping technique to associate pedon data from the Global Pedon Database with the FAO/UNESCO Digital Soil Map of the World.	Yes	Not explicitly mentioned	5x5 arc-minutes	Single dataset published on 2000	Global	1950 to 1996	2017	Need user account	Arc/Info ASCII Grid, Erdas Imagine, GeoTIFF, netcdf, NITF, XYZ	Yes. Not possible to add your own Area of Interest	Familiarity with the platform.	No. Only allow view and download of dataset.	Not applicable	No	Yes. Easy to navigate and easy to find different datasets.	Yes.Allows to view dataset in kml and has Spatial Data Access Tool	Yes.	Yes	Yes	Yes	Yes
Soil microbial biomass	A Compilation of Global Soil Microbial Biomass Carbon, Nitrogen, and Phosphorus Data	Oak Ridge National Laboratory Distributed Active Archive Center	NASA Earth Data	Used multiple linear regression to obtain correlations between soil microbial biomass or soil element concentrations and long term climate variables done using the R 2.12.3	Yes	Utilized global maps of vegetation distribution, soil properties, and long-term climate data for spatial extrapolation to estimate soil microbial biomass storage in terrestrial ecosystems	0.05-degree by 0.5-degree	Single dataset published on 2015	Global	1977 to 2012	2017	Need user account	Arc/Info ASCII Grid, Erdas Imagine, GeoTIFF, netcdf, NITF, XYZ	Yes. Not possible to add your own Area of Interest	Familiarity with the platform.	No. Only allow view and download of dataset.	Not applicable	No	Yes. Easy to navigate and easy to find different datasets.	Yes.Allows to view dataset in kml and has Spatial Data Access Tool	Yes.	Yes	Yes	Yes	Yes
Temperature (LST)	SAFARI 2000 AVHRR-derived Land Surface Temperature Maps, Africa, 1995-2000	Oak Ridge National Laboratory Distributed Active Archive Center	NASA Earth Data	The AVHRR GAC data is projected into Albers Equal Area then cloud filtering using the CLAVR-1 algorithm, LST values are estimated using a split-window technique that exploits the differential absorption of thermal infrared signal in AVHRR bands 4 and 5, surface emissivity is generated by combining a land cover classification map, FAO soil map of Africa, and maps of tree, herbaceous and bare soil percent cover, and collateral products like cloud mask, time-of-scan, latitude, longitude and land/water mask are also generated	Yes	AVHRR, GAC, NOAA	8 km	Single dataset published on 2006	Africa	1995 to 2000	2017	Need user account	Arc/Info ASCII Grid, Erdas Imagine, GeoTIFF, netcdf, NITF, XYZ	Yes. Not possible to add your own Area of Interest	Familiarity with the platform.	No. Only allow view and download of dataset.	Not applicable	No	Yes. Easy to navigate and easy to find different datasets.	Yes.Allows to view dataset in kml and has Spatial Data Access Tool	Yes.	Yes	Yes	Yes	Yes
Precipitation	ISLSCP II GLOBAL PRECIPITATION CLIMATOLOGY PROJECT VERSION 1, PENTAD PRECIPITATION	Oak Ridge National Laboratory Distributed Active Archive Center	NASA Earth Data	Used 3-step process: combining satellite estimates linearly using Maximum Likelihood Estimation with weighting coefficients inversely proportional to individual error variance, blending this output with gauge data through the Reynolds method to remove bias, and adjusting the blended analysis with GPCP monthly precipitation analysis to ensure consistency in accumulation between pentad and monthly analyses.	Yes	Satellite observations of IR, OLR, MSU and SSM/I	1 degree and 2.5 degrees in both latitude and longitude	5-days	Global	1986 to 1995	2017	Need user account	Arc/Info ASCII Grid, Erdas Imagine, GeoTIFF, netcdf, NITF, XYZ	Yes. Not possible to add your own Area of Interest	Familiarity with the platform.	No. Only allow view and download of dataset.	Not applicable	No	Yes. Easy to navigate and easy to find different datasets.	Yes.Allows to view dataset in kml and has Spatial Data Access Tool	Yes.	Yes	Yes	Yes	Yes
Evapotranspiration	CMS: Evapotranspiration and Meteorology, Water-Limited Shrublands, Mexico, 2008-2010	Oak Ridge National Laboratory Distributed Active Archive Center	NASA Earth Data	Modified the widely used Penman-Monteith equation to incorporate SWC to model ET in water-limited regions (Sun et al, 2013)	Yes	Not explicitly mentioned	Point data	Daily (Published data started on 2016)	Mexico	2008 to 2010	2017	Need user account	Arc/Info ASCII Grid, Erdas Imagine, GeoTIFF, netcdf, NITF, XYZ	Yes. Not possible to add your own Area of Interest	Familiarity with the platform.	No. Only allow view and download of dataset.	Not applicable	No	Yes. Easy to navigate and easy to find different datasets.	Yes.Allows to view dataset in kml and has Spatial Data Access Tool	Yes.	Yes	Yes	Yes	Yes
Humidity	VEMAP 2: U.S. ANNUAL CLIMATE (1895-1993): Mean Daily Irradiance	Oak Ridge National Laboratory Distributed Active Archive Center	NASA Earth Data	a set of selected biogeochemical and coupled biogeochemical-biogeographical models from 1895 to 1993 to compare model responses to historical time series and current ecosystem biogeochemistry, then running the same models on projected 1994 to 2100 data to compare ecological responses to transient scenarios of climate and atmospheric CO2 change	Yes	Not explicitly mentioned	0.5 x 0.5 degrees		US	1895 to 1993 (Annually)	2017	Need user account	Arc/Info ASCII Grid, Erdas Imagine, GeoTIFF, netcdf, NITF, XYZ	Yes. Not possible to add your own Area of Interest	Familiarity with the platform.	No. Only allow view and download of dataset.	Not applicable	No	Yes. Easy to navigate and easy to find different datasets.	Yes.Allows to view dataset in kml and has Spatial Data Access Tool	Yes.	Yes	Yes	Yes	Yes
Soil Fertility																									
Soil Erosion	Topographic and Soil Carbon Reconstructions in Agricultural Fields, Iowa	Oak Ridge National Laboratory Distributed Active Archive Center	NASA Earth Data	Used a hillslope diffusion model	Yes	Not explicitly mentioned	3-m raster cells and point locations	Single dataset published on 2022	Iowa State	1859-08-01 to 2019-08-01 for modeling; 2017-04-07 to 2020-05-15 for imagery	2017	Need user account	Arc/Info ASCII Grid, Erdas Imagine, GeoTIFF, netcdf, NITF, XYZ	Yes. Not possible to add your own Area of Interest	Familiarity with the platform.	No. Only allow view and download of dataset.	Not applicable	No	Yes. Easy to navigate and easy to find different datasets.	Yes.Allows to view dataset in kml and has Spatial Data Access Tool	Yes.	Yes	Yes	Yes	Yes
Cropland Irrigation	GFSAD1KCM v001 Global Food Security Support Analysis Data (GFSAD) Crop Mask 2010	Oak Ridge National Laboratory Distributed Active Archive Center	NASA Earth Data	Disaggregated a five-class global cropland extent map derived from four major studies: Thenkabail et al. (2009a, 2011), Pittman et al. (2010), Yu et al. (2013), and Friedl et al. (2010	Yes	Landsat, AVHRR, SPOT and MODIS	1km	Single dataset published on 2016	Global		2017	Need user account	GeoTIFF	Yes. Not possible to add your own Area of Interest	Familiarity with the platform.	No. Only allow view and download of dataset.	Not applicable	No	Yes. Easy to navigate and easy to find different datasets.	Yes.Allows to view dataset in kml and has Spatial Data Access Tool	Yes.	Yes	Yes	Yes	Yes

Annex 7: FAO-HiH EO data for soil

Soil Property	INTRINSIC					CONTEXTUALITY						ACCESSIBILITY			REPRESENTATIONAL										
	Name	Dataset Provider	Point of contact	Methods of Processing EO Data	Use of Ancillary Data	Sensor Used for production of dataset	Spatial Resolution	Temporal Coverage	Spatial Coverage	Temporal Resolution/Frequency	Year of EO launch	Data Access	Product Format	Security	User Technical Knowledge Requirement	Processing Tools	Programming Languages	Cloud Processing	User Interface (UI)	Visualization Tools	Scalability	Open Standards	User Support	Community	Product User Guide/Manual
	What is the name of platform	Who is the source of data	Who is the responsible organization for the EO Platform	What type of method is used	Did it utilize additional/auxiliary and supplementary?	What are the sensors used to produced data?	What is the level of detail captured in an image by a sensor system (pixel size)?	Does it have historical data and the revisit frequency you require?	What is the geographical coverage of data?	What is the repeat/revisit time, e.g., time between successive observations of a given location	When did the EO platform become available and functional?	Is the data free, subscription-based, or pay-per-use?	What are the EO data formats available for download?	Does the platform allows you to use your own data securely?	What is the minimum technical knowledge of user to access and download data?	Does the platform offer built-in tools for data analysis ?	Does it support your preferred programming language (e.g., Python, R) for custom analysis?	Does it leverage cloud computing for handling datasets?	Is the interface user-friendly and easy to navigate?	Does it offer tools for visualizing and exploring the data (e.g., generating maps, charts)?	Can the platform handle high volume of data (e.g., multiple downloads from global to community scale)?	Does it adhere to open data standards for interoperability with other platforms?	Does the platform offer adequate user support (e.g., video tutorials, case studies)?	Is there an active user community for knowledge sharing and troubleshooting?	Is there minimum set of documentation for users?
Soil Texture	OpenLandMap Soil texture	EnvirometriX Ltd	UN-FAO with Google	Predicted soil texture fractions using the soiltexture package in R	Yes	Not explicitly mentioned	250m	Single dataset with soil texture class bands 0 cm - 200 cm with 6 standard depths	Global except Antarctica	1950 to 2018	2019	Need user account	PNG, GeoTIFF, CSV,XCLS	Yes. Not possible to add your own Area of Interest	Familiarity with the platform.	No. Only allow view and download of dataset.	Not applicable	No	Yes. Easy to navigate and easy to find different datasets.	Yes. It has interface for data analysis, graphs and charts	Yes. It is powered by Google Earth Engine	Yes	Yes	Yes	Yes
Sand content (Sand-C)	Topsoil Sand Fraction	Global Soil Partnership	UN-FAO	Use nearest neighbor resampling, bilinear interpolation to resample to 30 arc-second grid then mosaicked using nearest neighbor resampling. NoData values at national borders were filled using GDAL gapfilling with inverse distance weighting and a four direction conic search within a 5-km buffer mask, and a global mask of water bodies was applied.	Yes	MODIS, Landsat, DEM derivatives, climatic images and global landform and lithology maps	1000 meters	Single dataset published on 2012	Global	1972 to 2004 (Based from the data used)	2019	Need user account	GeoTIFF, WMS	Yes. Not possible to add your own Area of Interest	Familiarity with the platform	No. Only allow view and download of dataset.	No	No	Yes. Easy to navigate and easy to find different datasets but difficult to find metadata.	Yes. Allows developing of own map with different datasets.	Yes	Yes	Yes	Yes	Yes
Silt content (Silt-C)	Topsoil Silt Fraction	Global Soil Partnership	UN-FAO	Use nearest neighbor resampling, bilinear interpolation to resample to 30 arc-second grid then mosaicked using nearest neighbor resampling. NoData values at national borders were filled using GDAL gapfilling with inverse distance weighting and a four direction conic search within a 5-km buffer mask, and a global mask of water bodies was applied.	Yes	MODIS, Landsat, DEM derivatives, climatic images and global landform and lithology maps	1000 meters	Single dataset published on 2012	Global	1972 to 2004 (Based from the data used)	2019	Need user account	GeoTIFF, WMS	Yes. Not possible to add your own Area of Interest	Familiarity with the platform	No. Only allow view and download of dataset.	No	No	Yes. Easy to navigate and easy to find different datasets but difficult to find metadata.	Yes. Allows developing of own map with different datasets.	Yes	Yes	Yes	Yes	Yes
Clay content (Clay-C)	Topsoil Clay Fraction	Global Soil Partnership	UN-FAO	Use nearest neighbor resampling, bilinear interpolation to resample to 30 arc-second grid then mosaicked using nearest neighbor resampling. NoData values at national borders were filled using GDAL gapfilling with inverse distance weighting and a four direction conic search within a 5-km buffer mask, and a global mask of water bodies was applied.	Yes	MODIS, Landsat, DEM derivatives, climatic images and global landform and lithology maps	1000 meters	Single dataset published on 2012	Global	1972 to 2004 (Based from the data used)	2019	Need user account	GeoTIFF, WMS	Yes. Not possible to add your own Area of Interest	Familiarity with the platform	No. Only allow view and download of dataset.	No	No	Yes. Easy to navigate and easy to find different datasets but difficult to find metadata.	Yes. Allows developing of own map with different datasets.	Yes	Yes	Yes	Yes	Yes
Bulk Density (BD)	Subsoil Reference Bulk Density	Global Soil Partnership	UN-FAO	Use nearest neighbor resampling, bilinear interpolation to resample to 30 arc-second grid then mosaicked using nearest neighbor resampling. NoData values at national borders were filled using GDAL gapfilling with inverse distance weighting and a four direction conic search within a 5-km buffer mask, and a global mask of water bodies was applied.	Yes	MODIS, Landsat, DEM derivatives, climatic images and global landform and lithology maps	1000 meters	Single dataset published on 2012	Global	1972 to 2004 (Based from the data used)	2019	Need user account	GeoTIFF, WMS	Yes. Not possible to add your own Area of Interest	Familiarity with the platform	No. Only allow view and download of dataset.	No	No	Yes. Easy to navigate and easy to find different datasets but difficult to find metadata.	Yes. Allows developing of own map with different datasets.	Yes	Yes	Yes	Yes	Yes
Available Water Capacity	Available water storage capacity (Global)	Global Soil Partnership	UN-FAO	Use nearest neighbor resampling, bilinear interpolation to resample to 30 arc-second grid then mosaicked using nearest neighbor resampling. NoData values at national borders were filled using GDAL gapfilling with inverse distance weighting and a four direction conic search within a 5-km buffer mask, and a global mask of water bodies was applied.	Yes	MODIS, Landsat, DEM derivatives, climatic images and global landform and lithology maps	1000 meters	Single dataset published on 2012	Global	1972 to 2004 (Based from the data used)	2019	Need user account	GeoTIFF, WMS	Yes. Not possible to add your own Area of Interest	Familiarity with the platform	No. Only allow view and download of dataset.	No	No	Yes. Easy to navigate and easy to find different datasets but difficult to find metadata.	Yes. Allows developing of own map with different datasets.	Yes	Yes	Yes	Yes	Yes
Soil Moisture	Relative root zone soil moisture - beta product (Global - Dekadal - 300m) - WaPOR v3	FAO Water Productivity Open-access portal (WaPOR)	UN-FAO	applying grid and variable-specific regression equations trained on ECMWF's HRES model at the same resolution	Yes	CHIRPS Precipitation, Copernicus DEM, (Ag)ERA5 Meteorological Data, GEOS-5 Meteorological Data, IMERG Precipitation, Landsat satellites, MODIS sensors, MSG satellites, Sentinel-2 satellites, VIIRS sensors and WorldCover Land Cover.	100 meters	5-days published on 2023	Global	1979 to near present	2019	Need user account	GeoTIFF, WMS	Yes. Not possible to add your own Area of Interest	Familiarity with the platform	No. Only allow view and download of dataset.	No	No	Yes. Easy to navigate and easy to find different datasets but difficult to find metadata.	Yes. Allows developing of own map with different datasets.	Yes	Yes	Yes	Yes	Yes
Hydrogen Potential (pH)	Subsoil pH (H_2O)	Global Soil Partnership	UN-FAO	Use nearest neighbor resampling, bilinear interpolation to resample to 30 arc-second grid then mosaicked using nearest neighbor resampling. NoData values at national borders were filled using GDAL gapfilling with inverse distance weighting and a four direction conic search within a 5-km buffer mask, and a global mask of water bodies was applied.	Yes	MODIS, Landsat, DEM derivatives, climatic images and global landform and lithology maps	1000 meters	Single dataset published on 2012	Global	1972 to 2004 (Based from the data used)	2019	Need user account	GeoTIFF, WMS	Yes. Not possible to add your own Area of Interest	Familiarity with the platform	No. Only allow view and download of dataset.	No	No	Yes. Easy to navigate and easy to find different datasets but difficult to find metadata.	Yes. Allows developing of own map with different datasets.	Yes	Yes	Yes	Yes	Yes
Total nitrogen (TN)	Subsoil pH (H_2O)	Global Soil Partnership	UN-FAO	Use nearest neighbor resampling, bilinear interpolation to resample to 30 arc-second grid then mosaicked using nearest neighbor resampling. NoData values at national borders were filled using GDAL gapfilling with inverse distance weighting and a four direction conic search within a 5-km buffer mask, and a global mask of water bodies was applied.	Yes	MODIS, Landsat, DEM derivatives, climatic images and global landform and lithology maps	1000 meters	Single dataset published on 2012	Global	1972 to 2004 (Based from the data used)	2019	Need user account	GeoTIFF, WMS	Yes. Not possible to add your own Area of Interest	Familiarity with the platform	No. Only allow view and download of dataset.	No	No	Yes. Easy to navigate and easy to find different datasets but difficult to find metadata.	Yes. Allows developing of own map with different datasets.	Yes	Yes	Yes	Yes	Yes
Potassium (K)																									
Calcium (Ca)																									
Magnesium (Mg)																									
Phosphorus (P)																									
Cropland Irrigation	GFSAD1KCM v001 Global Food Security Support Analysis Data (GFSAD) Crop Mask 2010	Oak Ridge National Laboratory Distributed Active Archive Center	NASA Earth Data	Disaggregated a five-class global cropland extent map derived from four major studies: Thenkabail et al. (2009a, 2011), Pittman et al. (2010), Yu et al. (2013), and Friedl et al. (2010	Yes	Landsat, AVHRR, SPOT and MODIS	1km	Single dataset published on 2016	Global		2017	Need user account	GeoTIFF	Yes. Not possible to add your own Area of Interest	Familiarity with the platform.	No. Only allow view and download of dataset.	Not applicable	No	Yes. Easy to navigate and easy to find different datasets.	Yes.Allows to view dataset in kml and has Spatial Data Access Tool	Yes.	Yes	Yes	Yes	Yes

Soil Property	INTRINSIC					CONTEXTUALITY						ACCESSIBILITY			REPRESENTATIONAL										
	Name	**Dataset Provider**	**Point of contact**	**Methods of Processing EO Data**	**Use of Ancillary Data**	**Sensor Used for production of dataset**	**Spatial Resolution**	**Temporal Coverage**	**Spatial Coverage**	**Temporal Resolution/Frequency**	**Year of EO launch**	**Data Access**	**Product Format**	**Security**	**User Technical Knowledge Requirement**	**Processing Tools**	**Programming Languages**	**Cloud Processing**	**User Interface (UI)**	**Visualization Tools**	**Scalability**	**Open Standards**	**User Support**	**Community**	**Product User Guide/Manual**
	What is the name of platform	Who is the source of data	Who is the responsible organization for the EO Platform	What type of method is used	Did it utilize additional/auxiliary and supplementary?	What are the sensors used to produced data?	What is the level of detail captured in an image by a sensor system (pixel size)?	Does it have historical data and the revisit frequency you require?	What is the geographical coverage of data?	What is the repeat/revisit time, e.g., time between successive observations of a given location	When did the EO platform become available and functional?	Is the data free, subscription-based, or pay-per-use?	What are the EO data formats available for download?	Does the platform allows you to use your own data securely?	What is the minimum technical knowledge of user to access and download data?	Does the platform offer built-in tools for data analysis ?	Does it support your preferred programming language (e.g., Python, R) for custom analysis?	Does it leverage cloud computing for handling datasets?	Is the interface user-friendly and easy to navigate?	Does it offer tools for visualizing and exploring the data (e.g., generating maps, charts)?	Can the platform handle high volume of data (e.g., multiple downloads from global to community scale)?	Does it adhere to open data standards for interoperability with other platforms?	Does the platform offer adequate user support (e.g., video tutorials, case studies)?	Is there an active user community for knowledge sharing and troubleshooting?	Is there minimum set of documentation for users?
Soil Nutrient	Soil Nutrient Availability from HWSD v1.2 (Global)	Global Soil Partnership	UN-FAO	Use nearest neighbor resampling, bilinear interpolation to resample to 30 arc-second grid then mosaicked using nearest neighbor resampling. NoData values at national borders were filled using GDAL gapfilling with inverse distance weighting and a four direction conic search within a 5-km buffer mask, and a global mask of water bodies was applied.	Yes	MODIS, Landsat, DEM derivatives, climatic images and global landform and lithology maps	1000 meters	Single dataset published on 2012	Global	1972 to 2004 (Based from the data used)	2019	Need user account	GeoTIFF, WMS	Yes. Not possible to add your own Area of Interest	Familiarity with the platform	No. Only allow view and download of dataset.	No	No	Yes. Easy to navigate and easy to find different datasets but difficult to find metadata.	Yes. Allows developing of own map with different datasets.	Yes	Yes	Yes	Yes	Yes
Cation exchange capacity (CEC)	Topsoil CEC (CLAY)	Global Soil Partnership	UN-FAO	Use nearest neighbor resampling, bilinear interpolation to resample to 30 arc-second grid then mosaicked using nearest neighbor resampling. NoData values at national borders were filled using GDAL gapfilling with inverse distance weighting and a four direction conic search within a 5-km buffer mask, and a global mask of water bodies was applied.	Yes	MODIS, Landsat, DEM derivatives, climatic images and global landform and lithology maps	1000 meters	Single dataset published on 2012	Global	1972 to 2004 (Based from the data used)	2019	Need user account	GeoTIFF, WMS	Yes. Not possible to add your own Area of Interest	Familiarity with the platform	No. Only allow view and download of dataset.	No	No	Yes. Easy to navigate and easy to find different datasets but difficult to find metadata.	Yes. Allows developing of own map with different datasets.	Yes	Yes	Yes	Yes	Yes
Soil Salinity	Excess Salts from HWSD v1.2 (Global)	Global Soil Partnership	UN-FAO	Use nearest neighbor resampling, bilinear interpolation to resample to 30 arc-second grid then mosaicked using nearest neighbor resampling. NoData values at national borders were filled using GDAL gapfilling with inverse distance weighting and a four direction conic search within a 5-km buffer mask, and a global mask of water bodies was applied.	Yes	MODIS, Landsat, DEM derivatives, climatic images and global landform and lithology maps	1000 meters	Single dataset published on 2012	Global	1971 to 2004 (Based from the data used)	2019	Need user account	GeoTIFF, WMS	Yes. Not possible to add your own Area of Interest	Familiarity with the platform	No. Only allow view and download of dataset.	No	No	Yes. Easy to navigate and easy to find different datasets but difficult to find metadata.	Yes. Allows developing of own map with different datasets.	Yes	Yes	Yes	Yes	Yes
Soil organic carbon (SOC)	Global Soil Organic Carbon Map v1.5 (GSOC)	Global Soil Partnership	UN-FAO	Use nearest neighbor resampling, bilinear interpolation to resample to 30 arc-second grid then mosaicked using nearest neighbor resampling. NoData values at national borders were filled using GDAL gapfilling with inverse distance weighting and a four direction conic search within a 5-km buffer mask, and a global mask of water bodies was applied.	Yes	MODIS, Landsat, DEM derivatives, climatic images and global landform and lithology maps	1000 meters	Single dataset published on 2012	Global	1972 to 2004 (Based from the data used)	2019	Need user account	GeoTIFF, WMS	Yes. Not possible to add your own Area of Interest	Familiarity with the platform	No. Only allow view and download of dataset.	No	No	Yes. Easy to navigate and easy to find different datasets but difficult to find metadata.	Yes. Allows developing of own map with different datasets.	Yes	Yes	Yes	Yes	Yes
Soil microbial respiration (SMR)											2019			Yes. Not possible to add your own Area of Interest											
Temperature (LST)	Land Surface Temperature Day-Time (Global - L3, 8-Day - MYD11A2) MODIS Aqua	USGS-NASA	UN-FAO	Algorithm based from Wan and Dozier,1989	Yes	MODIS,	1000 meters	8-days	Global	2002 to present	2019	Need user account	GeoTIFF, WMS	Yes. Not possible to add your own Area of Interest	Familiarity with the platform	No. Only allow view and download of dataset.	No	No	Yes. Easy to navigate and easy to find different datasets but difficult to find metadata.	Yes. Allows developing of own map with different datasets.	Yes	Yes	Yes	Yes	Yes
Precipitation	Precipitation (Global - Daily - Approximately 5km) - WaPOR v3	FAO Water Productivity Open-access portal (WaPOR)	UN-FAO	Generated through a two-part process. Firstly, IR Precipitation (IRP) pentad rainfall estimates are derived from satellite data by calculating the percentage of time during the pentad that the IR observations indicate cold cloud tops (<235° K). Then the station data is integrated with the CHIRP data to create the final product, CHIRPS.	Yes	CHIRPS (Climate Hazards Group InfraRed Precipitation with Station) quasi-global rainfall dataset	1000 meters	Daily published on 2024	Global	1981 to near present	2019	Need user account	GeoTIFF, WMS	Yes. Not possible to add your own Area of Interest	Familiarity with the platform	No. Only allow view and download of dataset.	No	No	Yes. Easy to navigate and easy to find different datasets but difficult to find metadata.	Yes. Allows developing of own map with different datasets.	Yes	Yes	Yes	Yes	Yes
Evapotranspiration	Reference Evapotranspiration (Global - Daily - Approximately 30km) - WaPOR v3	FAO Water Productivity Open-access portal (WaPOR)	UN-FAO	Derived using the Penman-Monteith equation, with the distinction that most of the variables are predefined	Yes	CHIRPS Precipitation, Copernicus DEM, (Ag)ERA5 Meteorological Data, GEOS-5 Meteorological Data, IMERG Precipitation, Landsat satellites, MODIS sensors, MSG satellites, Sentinel-2 satellites, VIIRS sensors and WorldCover Land Cover.	30 km	Daily published on 2024	Global	1981 to near present	2019	Need user account	GeoTIFF, WMS	Yes. Not possible to add your own Area of Interest	Familiarity with the platform	No. Only allow view and download of dataset.	No	No	Yes. Easy to navigate and easy to find different datasets but difficult to find metadata.	Yes. Allows developing of own map with different datasets.	Yes	Yes	Yes	Yes	Yes
Humidity	Relative humidity at 06h local time - AgERA5 (Global - Daily - ~10km)	Copernicus Climate Change Service	UN-FAO	Application of grid and variable-specific regression equations to the ERA5 dataset interpolated at a 0.1° grid, with the equations trained on ECMWF's operational high-resolution atmospheric model (HRES) at a 0.1° resolution.	Yes	CHIRPS Precipitation, Copernicus DEM, (Ag)ERA5 Meteorological Data, GEOS-5 Meteorological Data, IMERG Precipitation, Landsat satellites, MODIS sensors, MSG satellites, Sentinel-2 satellites, VIIRS sensors and WorldCover Land Cover.	0.1° x 0.1°	Daily published on 2021	Global	1979 to near present	2019	Need user account	GeoTIFF, WMS	Yes. Not possible to add your own Area of Interest	Familiarity with the platform	No. Only allow view and download of dataset.	No	No	Yes. Easy to navigate and easy to find different datasets but difficult to find metadata.	Yes. Allows developing of own map with different datasets.	Yes	Yes	Yes	Yes	Yes
Soil Fertility																									
Soil Erosion																									
Cropland Irrigation																									
Phosphorus (P)																									
Cropland Irrigation	GFSAD1KCM v001 Global Food Security Support Analysis Data (GFSAD) Crop Mask 2010	Oak Ridge National Laboratory Distributed Active Archive Center	NASA Earth Data	Disaggregated a five-class global cropland extent map derived from four major studies: Thenkabail et al. (2009a, 2011), Pittman et al. (2010), Yu et al. (2013), and Friedl et al. (2010	Yes	Landsat, AVHRR, SPOT and MODIS	1km	Single dataset published on 2016	Global		2017	Need user account	GeoTIFF	Yes. Not possible to add your own Area of Interest	Familiarity with the platform.	No. Only allow view and download of dataset.	Not applicable	No	Yes. Easy to navigate and easy to find different datasets.	Yes.Allows to view dataset in kml and has Spatial Data Access Tool	Yes.	Yes	Yes	Yes	Yes

Annex 8: FAO-GloSIS EO data for soil

Soil Property	INTRINSIC					CONTEXTUALITY						ACCESSIBILITY			REPRESENTATIONAL										
	Name	Dataset Provider	Point of contact	Methods of Processing EO Data	Use of Ancillary Data	Sensor Used for production of dataset	Spatial Resolution	Temporal Coverage	Spatial Coverage	Temporal Resolution/Frequency	Year of EO launch	Data Access	Product Format	Security	User Technical Knowledge Requirement	Processing Tools	Programming Languages	Cloud Processing	User Interface (UI)	Visualization Tools	Scalability	Open Standards	User Support	Community	Product User Guide/Manual
	What is the name of platform	Who is the source of data	Who is the responsible organization for the EO Platform	What type of method is used	Did it utilize additional/auxiliary and supplementary?	What are the sensors used to produced data?	What is the level of detail captured in an image by a sensor system (pixel size)?	Does it have historical data and the revisit frequency you require?	What is the geographical coverage of data?	What is the repeat/revisit time, e.g., time between successive observations of a given location	When did the EO platform become available and functional?	Is the data free, subscription-based, or pay-per-use?	What are the EO data formats available for download?	Does the platform allows you to use your own data securely?	What is the minimum technical knowledge of user to access and download data?	Does the platform offer built-in tools for data analysis ?	Does it support your preferred programming language (e.g., Python, R) for custom analysis?	Does it leverage cloud computing for handling datasets?	Is the interface user-friendly and easy to navigate?	Does it offer tools for visualizing and exploring the data (e.g., generating maps, charts)?	Can the platform handle high volume of data (e.g., multiple downloads from global to community scale)?	Does it adhere to open data standards for interoperability with other platforms?	Does the platform offer adequate user support (e.g., video tutorials, case studies)?	Is there an active user community for knowledge sharing and troubleshooting?	Is there minimum set of documentation for users?
Soil Texture																									
Sand content (Sand-C)	Topsoil Sand Fraction	Global Soil Partnership	UN-FAO	Use nearest neighbor resampling, bilinear interpolation to resample to 30 arc-second grid then mosaicked using nearest neighbor resampling. NoData values at national borders were filled using GDAL gapfilling with inverse distance weighting and a four direction conic search within a 5-km buffer mask, and a global mask of water bodies was applied.	Yes	MODIS, Landsat, DEM derivatives, climatic images and global landform and lithology maps	1000 meters	Single dataset published on 2012	Global	1972 to 2004 (Based from the data used)	2017	Need user account	GeoTIFF, WMS	Yes	Familiarity with the platform	No. Only allow view and download of dataset.	No	No	Yes. Easy to navigate and easy to find different datasets but difficult to find metadata.	Yes. Allows developing of own map with different datasets.	Yes	Yes	Yes	Yes	Yes
Silt content (Silt-C)	Topsoil Silt Fraction	Global Soil Partnership	UN-FAO	Use nearest neighbor resampling, bilinear interpolation to resample to 30 arc-second grid then mosaicked using nearest neighbor resampling. NoData values at national borders were filled using GDAL gapfilling with inverse distance weighting and a four direction conic search within a 5-km buffer mask, and a global mask of water bodies was applied.	Yes	MODIS, Landsat, DEM derivatives, climatic images and global landform and lithology maps	1000 meters	Single dataset published on 2012	Global	1972 to 2004 (Based from the data used)	2017	Need user account	GeoTIFF, WMS	Yes	Familiarity with the platform	No. Only allow view and download of dataset.	No	No	Yes. Easy to navigate and easy to find different datasets but difficult to find metadata.	Yes. Allows developing of own map with different datasets.	Yes	Yes	Yes	Yes	Yes
Clay content (Clay-C)	Topsoil Clay Fraction	Global Soil Partnership	UN-FAO	Use nearest neighbor resampling, bilinear interpolation to resample to 30 arc-second grid then mosaicked using nearest neighbor resampling. NoData values at national borders were filled using GDAL gapfilling with inverse distance weighting and a four direction conic search within a 5-km buffer mask, and a global mask of water bodies was applied.	Yes	MODIS, Landsat, DEM derivatives, climatic images and global landform and lithology maps	1000 meters	Single dataset published on 2012	Global	1972 to 2004 (Based from the data used)	2017	Need user account	GeoTIFF, WMS	Yes	Familiarity with the platform	No. Only allow view and download of dataset.	No	No	Yes. Easy to navigate and easy to find different datasets but difficult to find metadata.	Yes. Allows developing of own map with different datasets.	Yes	Yes	Yes	Yes	Yes
Bulk Density (BD)	Subsoil Reference Bulk Density	Global Soil Partnership	UN-FAO	Use nearest neighbor resampling, bilinear interpolation to resample to 30 arc-second grid then mosaicked using nearest neighbor resampling. NoData values at national borders were filled using GDAL gapfilling with inverse distance weighting and a four direction conic search within a 5-km buffer mask, and a global mask of water bodies was applied.	Yes	MODIS, Landsat, DEM derivatives, climatic images and global landform and lithology maps	1000 meters	Single dataset published on 2012	Global	1972 to 2004 (Based from the data used)	2017	Need user account	GeoTIFF, WMS	Yes	Familiarity with the platform	No. Only allow view and download of dataset.	No	No	Yes. Easy to navigate and easy to find different datasets but difficult to find metadata.	Yes. Allows developing of own map with different datasets.	Yes	Yes	Yes	Yes	Yes
Available Water Capacity	Available water storage capacity (Global)	Global Soil Partnership	UN-FAO	Use nearest neighbor resampling, to resample to 30 arc-second grid then mosaicked using nearest neighbor resampling. NoData values at national borders were filled using GDAL gapfilling with inverse distance weighting and a four direction conic search within a 5-km buffer mask, and a global mask of water bodies was applied.	Yes	MODIS, Landsat, DEM derivatives, climatic images and global landform and lithology maps	1000 meters	Single dataset published on 2012	Global	1972 to 2004 (Based from the data used)	2017	Need user account	GeoTIFF, WMS	Yes	Familiarity with the platform	No. Only allow view and download of dataset.	No	No	Yes. Easy to navigate and easy to find different datasets but difficult to find metadata.	Yes. Allows developing of own map with different datasets.	Yes	Yes	Yes	Yes	Yes
Hydrogen Potential (pH)	Subsoil pH (H2O)	Global Soil Partnership	UN-FAO	Use nearest neighbor resampling, bilinear interpolation to resample to 30 arc-second grid then mosaicked using nearest neighbor resampling. NoData values at national borders were filled using GDAL gapfilling with inverse distance weighting and a four direction conic search within a 5-km buffer mask, and a global mask of water bodies was applied.	Yes	MODIS, Landsat, DEM derivatives, climatic images and global landform and lithology maps	1000 meters	Single dataset published on 2012	Global	1972 to 2004 (Based from the data used)	2017	Need user account	GeoTIFF, WMS	Yes	Familiarity with the platform	No. Only allow view and download of dataset.	No	No	Yes. Easy to navigate and easy to find different datasets but difficult to find metadata.	Yes. Allows developing of own map with different datasets.	Yes	Yes	Yes	Yes	Yes
Total nitrogen (TN)																									
Potassium (K)																									
Calcium (Ca)																									
Magnesium (Mg)																									
Phosphorus (P)																									

Soil Property	INTRINSIC					CONTEXTUALITY						ACCESSIBILITY			REPRESENTATIONAL										
	Name	Dataset Provider	Point of contact	Methods of Processing EO Data	Use of Ancillary Data	Sensor Used for production of dataset	Spatial Resolution	Temporal Coverage	Spatial Coverage	Temporal Resolution/Frequency	Year of EO launch	Data Access	Product Format	Security	User Technical Knowledge Requirement	Processing Tools	Programming Languages	Cloud Processing	User Interface (UI)	Visualization Tools	Scalability	Open Standards	User Support	Community	Product User Guide/Manual
	What is the name of platform	Who is the source of data	Who is the responsible organization for the EO Platform	What type of method is used	Did it utilize additional/auxiliary and supplementary?	What are the sensors used to produced data?	What is the level of detail captured in an image by a sensor system (pixel size)?	Does it have historical data and the revisit frequency you require?	What is the geographical coverage of data?	What is the repeat/revisit time, e.g., time between successive observations of a given location	When did the EO platform become available and functional?	Is the data free, subscription-based, or pay-per-use?	What are the EO data formats available for download?	Does the platform allows you to use your own data securely?	What is the minimum technical knowledge of user to access and download data?	Does the platform offer built-in tools for data analysis ?	Does it support your preferred programming language (e.g., Python, R) for custom analysis?	Does it leverage cloud computing for handling datasets?	Is the interface user-friendly and easy to navigate?	Does it offer tools for visualizing and exploring the data (e.g., generating maps, charts)?	Can the platform handle high volume of data (e.g., multiple downloads from global to community scale)?	Does it adhere to open data standards for interoperability with other platforms?	Does the platform offer adequate user support (e.g., video tutorials, case studies)?	Is there an active user community for knowledge sharing and troubleshooting?	Is there minimum set of documentation for users?
Soil Nutrient	Soil Nutrient Availability from HWSD v1.2 (Global)	Global Soil Partnership	UN-FAO	Use nearest neighbor resampling, bilinear interpolation to resample to 30 arc-second grid then mosaicked using nearest neighbor resampling. NoData values at national borders were filled using GDAL gapfilling with inverse distance weighting and a four direction conic search within a 5-km buffer mask, and a global mask of water bodies was applied.	Yes	MODIS, Landsat, DEM derivatives, climatic images and global landform and lithology maps	1000 meters	Single dataset published on 2012	Global	1972 to 2004 (Based from the data used)	2017	Need user account	GeoTIFF, WMS	Yes	Familiarity with the platform	No. Only allow view and download of dataset.	No	No	Yes. Easy to navigate and easy to find different datasets but difficult to find metadata.	Yes. Allows developing of own map with different datasets.	Yes	Yes	Yes	Yes	Yes
Cation exchange capacity (CEC)	Topsoil CEC (CLAY)	Global Soil Partnership	UN-FAO	Use nearest neighbor resampling, bilinear interpolation to resample to 30 arc-second grid then mosaicked using nearest neighbor resampling. NoData values at national borders were filled using GDAL gapfilling with inverse distance weighting and a four direction conic search within a 5-km buffer mask, and a global mask of water bodies was applied.	Yes	MODIS, Landsat, DEM derivatives, climatic images and global landform and lithology maps	1000 meters	Single dataset published on 2012	Global	1972 to 2004 (Based from the data used)	2017	Need user account	GeoTIFF, WMS	Yes	Familiarity with the platform	No. Only allow view and download of dataset.	No	No	Yes. Easy to navigate and easy to find different datasets but difficult to find metadata.	Yes. Allows developing of own map with different datasets.	Yes	Yes	Yes	Yes	Yes
Soil Salinity	Excess Salts from HWSD v1.2 (Global)	Global Soil Partnership	UN-FAO	Use nearest neighbor resampling, bilinear interpolation to resample to 30 arc-second grid then mosaicked using nearest neighbor resampling. NoData values at national borders were filled using GDAL gapfilling with inverse distance weighting and a four direction conic search within a 5-km buffer mask, and a global mask of water bodies was applied.	Yes	MODIS, Landsat, DEM derivatives, climatic images and global landform and lithology maps	1000 meters	Single dataset published on 2012	Global	1971 to 2004 (Based from the data used)	2017	Need user account	GeoTIFF, WMS	Yes	Familiarity with the platform	No. Only allow view and download of dataset.	No	No	Yes. Easy to navigate and easy to find different datasets but difficult to find metadata.	Yes. Allows developing of own map with different datasets.	Yes	Yes	Yes	Yes	Yes
Soil organic carbon (SOC)	Global Soil Organic Carbon Map v1.5 (GSOC)	Global Soil Partnership	UN-FAO	Use nearest neighbor resampling, bilinear interpolation to resample to 30 arc-second grid then mosaicked using nearest neighbor resampling. NoData values at national borders were filled using GDAL gapfilling with inverse distance weighting and a four direction conic search within a 5-km buffer mask, and a global mask of water bodies was applied.	Yes	MODIS, Landsat, DEM derivatives, climatic images and global landform and lithology maps	1000 meters	Single dataset published on 2012	Global	1972 to 2004 (Based from the data used)	2017	Need user account	GeoTIFF, WMS	Yes	Familiarity with the platform	No. Only allow view and download of dataset.	No	No	Yes. Easy to navigate and easy to find different datasets but difficult to find metadata.	Yes. Allows developing of own map with different datasets.	Yes	Yes	Yes	Yes	Yes
Soil microbial respiration (SMR)																									
Temperature (LST)																									
Precipitation																									
Evapotranspiration																									
Humidity																									
Soil Moisture																									
Soil Fertility																									
Soil Erosion																									
Cropland Irrigation																									

Annex 9: GeoAgro EO data for Central Asia soil

Soil Property	INTRINSIC					CONTEXTUALITY						ACCESSIBILITY			REPRESENTATIONAL										
	Name	Dataset Provider	Point of contact	Methods of Processing EO Data	Use of Ancillary Data	Sensor Used for production of dataset	Spatial Resolution	Temporal Coverage	Spatial Coverage	Temporal Resolution/Frequency	Year of EO launch	Data Access	Product Format	Security	User Technical Knowledge Requirement	Processing Tools	Programming Languages	Cloud Processing	User Interface (UI)	Visualization Tools	Scalability	Open Standards	User Support	Community	Product User Guide/Manual
	What is the name of platform	Who is the source of data	Who is the responsible organization for the EO Platform	What type of method is used	Did it utilize additional/auxiliary and supplementary?	What are the sensors used to produced data?	What is the level of detail captured in an image by a sensor system (pixel size)?	Does it have historical data and the revisit frequency you require?	What is the geographical coverage of data?	What is the repeat/revisit time, e.g., time between successive observations of a given location	When did the EO platform become available and functional?	Is the data free, subscription-based, or pay-per-use?	What are the EO data formats available for download?	Does the platform allows you to use your own data securely?	What is the minimum technical knowledge of user to access and download data?	Does the platform offer built-in tools for data analysis ?	Does it support your preferred programming language (e.g., Python, R) for custom analysis?	Does it leverage cloud computing for handling datasets?	Is the interface user-friendly and easy to navigate?	Does it offer tools for visualizing and exploring the data (e.g., generating maps, charts)?	Can the platform handle high volume of data (e.g., multiple downloads from global to community scale)?	Does it adhere to open data standards for interoperability with other platforms?	Does the platform offer adequate user support (e.g., video tutorials, case studies)?	Is there an active user community for knowledge sharing and troubleshooting?	Is there minimum set of documentation for users?
Soil Texture	Coarse-textured Soils in CWANA and Eurasia	ICARDA	ICARDA	Not explicitly mentioned	Yes. GTOPO30	Not explicitly mentioned	1000 meters	Not explicitly mentioned	Central and West Asia - North Africa and Eurasia regions	Not explicitly mentioned	2008	Free to download	JPEG, PDF	Not applicable	No	Familiarity of interface	No	Yes	Difficult to search specific datasets.	Yes	Yes	Follows FGDC Standard	Yes	No	None
Sand content (Sand-C)	Central Asia: sandy soils	ICARDA	ICARDA	Not explicitly mentioned	Yes	Not explicitly mentioned	1000 meters	1975 or earlier	Central Asia	Not explicitly mentioned	2008	Free to download	JPEG, TIFF	Not applicable	No	Familiarity of interface	No	Yes	Difficult to search specific datasets.	Yes	Yes	Follows FGDC Standard	Yes	No	None
Silt content (Silt-C)																									
Clay content (Clay-C)	Central Asia: soils with fine texture	ICARDA	ICARDA	Not explicitly mentioned	Yes	Not explicitly mentioned	1000 meters	1975 or earlier	Central Asia	Not explicitly mentioned	2008	Free to download	JPEG, TIFF	Not applicable	No	Familiarity of interface	No	Yes	Difficult to search specific datasets.	Yes	Yes	Follows FGDC Standard	Yes	No	None
Bulk Density (BD)																									
Available Water Capacity	Wet Soils in CWANA and Eurasia	ICARDA	ICARDA	Not explicitly mentioned	Yes. GTOPO30	Not explicitly mentioned	1000 meters	Not explicitly mentioned	Central and West Asia - North Africa and Eurasia regions	Not explicitly mentioned	2008	Free to download	JPEG, PDF	Not applicable	No	Familiarity of interface	No	Yes	Difficult to search specific datasets.	Yes	Yes	Follows FGDC Standard	Yes	No	None
Hydrogen Potential (pH)																									
Total nitrogen (TN)																									
Potassium (K)																									
Calcium (Ca)																									
Magnesium (Mg)																									
Phosphorus (P)	High Risk of Phosphorous Fixation Soils in CWANA and Eurasia	ICARDA	ICARDA	Not explicitly mentioned	Yes. GTOPO30	Not explicitly mentioned	1000 meters	Not explicitly mentioned	Central and West Asia - North Africa and Eurasia regions	Not explicitly mentioned	2008	Free to download	JPEG, PDF	Not applicable	No	Familiarity of interface	No	Yes	Difficult to search specific datasets.	Yes	Yes	Follows FGDC Standard	Yes	No	None

Soil Property	INTRINSIC					CONTEXTUALITY						ACCESSIBILITY			REPRESENTATIONAL										
	Name	**Dataset Provider**	**Point of contact**	**Methods of Processing EO Data**	**Use of Ancillary Data**	**Sensor Used for production of dataset**	**Spatial Resolution**	**Temporal Coverage**	**Spatial Coverage**	**Temporal Resolution/Frequency**	**Year of EO launch**	**Data Access**	**Product Format**	**Security**	**User Technical Knowledge Requirement**	**Processing Tools**	**Programming Languages**	**Cloud Processing**	**User Interface (UI)**	**Visualization Tools**	**Scalability**	**Open Standards**	**User Support**	**Community**	**Product User Guide/Manual**
	What is the name of platform	Who is the source of data	Who is the responsible organization for the EO Platform	What type of method is used	Did it utilize additional/auxiliary and supplementary?	What are the sensors used to produced data?	What is the level of detail captured in an image by a sensor system (pixel size)?	Does it have historical data and the revisit frequency you require?	What is the geographical coverage of data?	What is the repeat/revisit time, e.g., time between successive observations of a given location	When did the EO platform become available and functional?	Is the data free, subscription-based, or pay-per-use?	What are the EO data formats available for download?	Does the platform allows you to use your own data securely?	What is the minimum technical knowledge of user to access and download data?	Does the platform offer built-in tools for data analysis ?	Does it support your preferred programming language (e.g., Python, R) for custom analysis?	Does it leverage cloud computing for handling datasets?	Is the interface user-friendly and easy to navigate?	Does it offer tools for visualizing and exploring the data (e.g., generating maps, charts)?	Can the platform handle high volume of data (e.g., multiple downloads from global to community scale)?	Does it adhere to open data standards for interoperability with other platforms?	Does the platform offer adequate user support (e.g., video tutorials, case studies)?	Is there an active user community for knowledge sharing and troubleshooting?	Is there minimum set of documentation for users?
Soil Nutrient																									
Cation exchange capacity (CEC)																									
Soil Salinity	Saline Soils in CWANA and Eurasia	ICARDA	ICARDA	Not explicitly mentioned	Yes. GTOPO30	Not explicitly mentioned	1000 meters	Not explicitly mentioned	Central and West Asia - North Africa and Eurasia regions	Not explicitly mentioned	2008	Free to download	JPEG, PDF	Not applicable	No	Familiarity of interface	No	Yes	Difficult to search specific datasets.	Yes	Yes	Follows FGDC Standard	Yes	No	None
Soil organic carbon (SOC)	Soils with High Organic Matter Content in CWANA and Eurasia	ICARDA	ICARDA	Not explicitly mentioned	Yes. GTOPO30	Not explicitly mentioned	1000 meters	Not explicitly mentioned	Central and West Asia - North Africa and Eurasia regions	Not explicitly mentioned	2008	Free to download	JPEG, PDF	Not applicable	No	Familiarity of interface	No	Yes	Difficult to search specific datasets.	Yes	Yes	Follows FGDC Standard	Yes	No	None
Soil microbial respiration (SMR)																									
Temperature (LST)	Central Asia and Xingjiang Province (China) Annual mean temperature 2011-2040 A2 Scenario	ICARDA	ICARDA	Not explicitly mentioned	Yes.	Not explicitly mentioned	1000 meters	2011 to 2040. Annually	Central Asia and Xingjiang Province (China)	Not explicitly mentioned	2010	Free to download	JPEG, TIFF	Not applicable	No	Familiarity of interface	No	Yes	Difficult to search specific datasets.	Yes	Yes	Follows FGDC Standard	Yes	No	None
Precipitation																									
Evapotranspiration	Central Asia and Xingjiang Province (China) Annual Potential Evapo Transpiration 2011-2040 A1b Scenario	ICARDA	ICARDA	Not explicitly mentioned	Yes.	Not explicitly mentioned	1000 meters	2011 to 2040. Annually	Central Asia and Xingjiang Province (China)	Not explicitly mentioned	2010	Free to download	JPEG, TIFF	Not applicable	No	Familiarity of interface	No	Yes	Difficult to search specific datasets.	Yes	Yes	Follows FGDC Standard	Yes	No	None
Humidity																									
Soil Moisture	Moisture-limited Growing Period in CWANA and Eurasia	ICARDA	ICARDA	The climatic moisture-limited growing period is calculated by means of a model developed by the Food and Agriculture Organization of the United Nations (FAO, 1978)	Yes. GTOPO30	Not explicitly mentioned	1000 meters	Not explicitly mentioned	Central and West Asia - North Africa and Eurasia regions	Not explicitly mentioned	2008	Free to download	JPEG, PDF	Not applicable	No	Familiarity of interface	No	Yes	Difficult to search specific datasets.	Yes	Yes	Follows FGDC Standard	Yes	No	None
Soil Fertility																									
Soil Erosion																									
Cropland Irrigation																									

Annex 10: CAC EO data for Central Asia soil

Soil Property	INTRINSIC					CONTEXTUALITY						ACCESSIBILITY			REPRESENTATIONAL										
	Name	**Dataset Provider**	**Point of contact**	**Methods of Processing EO Data**	**Use of Ancillary Data**	**Sensor Used for production of dataset**	**Spatial Resolution**	**Temporal Coverage**	**Spatial Coverage**	**Temporal Resolution/Frequency**	**Year of EO launch**	**Data Access**	**Product Format**	**Security**	**User Technical Knowledge Requirement**	**Processing Tools**	**Programming Languages**	**Cloud Processing**	**User Interface (UI)**	**Visualization Tools**	**Scalability**	**Open Standards**	**User Support**	**Community**	**Product User Guide/Manual**
	What is the name of platform	Who is the source of data	Who is the responsible organization for the EO Platform	What type of method is used	Did it utilize additional/auxiliary and supplementary?	What are the sensors used to produced data?	What is the level of detail captured in an image by a sensor system (pixel size)?	Does it have historical data and the revisit frequency you require?	What is the geographical coverage of data?	What is the repeat/revisit time, e.g., time between successive observations of a given location	When did the EO platform become available and functional?	Is the data free, subscription-based, or pay-per-use?	What are the EO data formats available for download?	Does the platform allows you to use your own data securely?	What is the minimum technical knowledge of user to access and download data?	Does the platform offer built-in tools for data analysis ?	Does it support your preferred programming language (e.g., Python, R) for custom analysis?	Does it leverage cloud computing for handling datasets?	Is the interface user-friendly and easy to navigate?	Does it offer tools for visualizing and exploring the data (e.g., generating maps, charts)?	Can the platform handle high volume of data (e.g., multiple downloads from global to community scale)?	Does it adhere to open data standards for interoperability with other platforms?	Does the platform offer adequate user support (e.g., video tutorials, case studies)?	Is there an active user community for knowledge sharing and troubleshooting?	Is there minimum set of documentation for users?
Soil Texture	World Soils Harmonized World Soil Database - Texture (Mature Support)	Not mentioned	Central Asia and Caucasus GeoPortal Powered by ESRI	Not mentioned	Not mentioned	Lacks information on dataset	Not mentioned	Lacks information on dataset	Central Asia	Lacks information on dataset	2024	Free but needs User Account.	API, KMZ, VRT format , ArcGIS web layer and Export map to PDF	Yes	Yes. Allows you to select your own area	Familiarity with the interface	No	Yes	Yes	Yes. Allows customization of maps.	Yes	Yes	Yes	No	Yes
Sand content (Sand-C)	CAC Soils 250m Percent Sand	Soilgrids.org (ISRIC) released in May 2020	Central Asia and Caucasus GeoPortal Powered by ESRI	Not mentioned	Not mentioned	Lacks information on dataset	250 m	Lacks information on dataset	Central Asia	Lacks information on dataset	2024	Free but needs User Account.	API, KMZ, VRT format , ArcGIS web layer and Export map to PDF	Yes	Yes. Allows you to select your own area	Familiarity with the interface	No	Yes	Yes	Yes. Allows customization of maps.	Yes	Yes	Yes	No	Yes
Silt content (Silt-C)	CAC Soils 250m Percent Silt	Soilgrids.org (ISRIC) released in May 2020	Central Asia and Caucasus GeoPortal Powered by ESRI	Not mentioned	Not mentioned	Lacks information on dataset	250 m	Lacks information on dataset	Central Asia	Lacks information on dataset	2024	Free but needs User Account.	API, KMZ, VRT format , ArcGIS web layer and Export map to PDF	Yes	Yes. Allows you to select your own area	Familiarity with the interface	No	Yes	Yes	Yes. Allows customization of maps.	Yes	Yes	Yes	No	Yes
Clay content (Clay-C)	CAC Soils 250m Percent Clay	Soilgrids.org (ISRIC) released in May 2020	Central Asia and Caucasus GeoPortal Powered by ESRI	Not mentioned	Not mentioned	Lacks information on dataset	250 m	Lacks information on dataset	Central Asia	Lacks information on dataset	2024	Free but needs User Account.	API, KMZ, VRT format , ArcGIS web layer and Export map to PDF	Yes	Yes. Allows you to select your own area	Familiarity with the interface	No	Yes	Yes	Yes. Allows customization of maps.	Yes	Yes	Yes	No	Yes
Bulk Density (BD)	CAC Soils 250m Bulk Density	Soilgrids.org (ISRIC) released in May 2020	Central Asia and Caucasus GeoPortal Powered by ESRI	Not mentioned	Not mentioned	Lacks information on dataset	Not mentioned	Lacks information on dataset	Central Asia	Lacks information on dataset	2024	Free but needs User Account.	API, KMZ, VRT format , ArcGIS web layer and Export map to PDF	Yes	Yes. Allows you to select your own area	Familiarity with the interface	No	Yes	Yes	Yes. Allows customization of maps.	Yes	Yes	Yes	No	Yes
Available Water Capacity																									
Hydrogen Potential (pH)																									
Total nitrogen (TN)	CAC Soils 250m Nitrogen	Soilgrids.org (ISRIC) released in May 2020	Central Asia and Caucasus GeoPortal Powered by ESRI	Central Asia and Caucasus GeoPortal	Not mentioned	Lacks information on dataset	250 m	Lacks information on dataset	Central Asia	Lacks information on dataset	2024	Free but needs User Account.	API, KMZ, VRT format , ArcGIS web layer and Export map to PDF	Yes	Yes. Allows you to select your own area	Familiarity with the interface	No	Yes	Yes	Yes. Allows customization of maps.	Yes	Yes	Yes	No	Yes
Potassium (K)																									
Calcium (Ca)																									
Magnesium (Mg)																									
Phosphorus (P)																									
Soil Nutrient																									

Soil Property	INTRINSIC					CONTEXTUALITY						ACCESSIBILITY			REPRESENTATIONAL										
	Name	**Dataset Provider**	**Point of contact**	**Methods of Processing EO Data**	**Use of Ancillary Data**	**Sensor Used for production of dataset**	**Spatial Resolution**	**Temporal Coverage**	**Spatial Coverage**	**Temporal Resolution/Frequency**	**Year of EO launch**	**Data Access**	**Product Format**	**Security**	**User Technical Knowledge Requirement**	**Processing Tools**	**Programming Languages**	**Cloud Processing**	**User Interface (UI)**	**Visualization Tools**	**Scalability**	**Open Standards**	**User Support**	**Community**	**Product User Guide/Manual**
	What is the name of platform	Who is the source of data	Who is the responsible organization for the EO Platform	What type of method is used	Did it utilize additional/auxiliary and supplementary?	What are the sensors used to produced data?	What is the level of detail captured in an image by a sensor system (pixel size)?	Does it have historical data and the revisit frequency you require?	What is the geographical coverage of data?	What is the repeat/revisit time, e.g., time between successive observations of a given location	When did the EO platform become available and functional?	Is the data free, subscription-based, or pay-per-use?	What are the EO data formats available for download?	Does the platform allows you to use your own data securely?	What is the minimum technical knowledge of user to access and download data?	Does the platform offer built-in tools for data analysis ?	Does it support your preferred programming language (e.g., Python, R) for custom analysis?	Does it leverage cloud computing for handling datasets?	Is the interface user-friendly and easy to navigate?	Does it offer tools for visualizing and exploring the data (e.g., generating maps, charts)?	Can the platform handle high volume of data (e.g., multiple downloads from global to community scale)?	Does it adhere to open data standards for interoperability with other platforms?	Does the platform offer adequate user support (e.g., video tutorials, case studies)?	Is there an active user community for knowledge sharing and troubleshooting?	Is there minimum set of documentation for users?
Cation exchange capacity (CEC)	Central Asia and Caucasus Soil Atlas | CAC Soils 250m Cation Exchange Capacity	Soilgrids.org (ISRIC) released in May 2020	Central Asia and Caucasus GeoPortal Powered by ESRI	Central Asia and Caucasus GeoPortal	Not mentioned	Lacks information on dataset	250 m	Lacks information on dataset	Central Asia	Lacks information on dataset	2024	Free but needs User Account.	API, KMZ, VRT format , ArcGIS web layer and Export map to PDF	Yes	Yes. Allows you to select your own area	Familiarity with the interface	No	Yes	Yes	Yes. Allows customization of maps.	Yes	Yes	Yes	No	Yes
Soil Salinity																									
Soil organic carbon (SOC)	CAC Soils 250m Organic Carbon Stocks	Soilgrids.org (ISRIC) released in May 2020	Central Asia and Caucasus GeoPortal Powered by ESRI	Central Asia and Caucasus GeoPortal	Not mentioned	Lacks information on dataset	250 m	Lacks information on dataset	Central Asia	Lacks information on dataset	2024	Free but needs User Account.	API, KMZ, VRT format , ArcGIS web layer and Export map to PDF	Yes	Yes. Allows you to select your own area	Familiarity with the interface	No	Yes	Yes	Yes. Allows customization of maps.	Yes	Yes	Yes	No	Yes
Soil microbial respiration (SMR)																									
Temperature (LST)																									
Precipitation	Early IMERG Precipitation Rate	NOAA	Central Asia and Caucasus GeoPortal Powered by ESRI	Created from hourly METAR station data provided from NOAA and contains approximately 11 weather variables for each location	No	METAR	30 minutes	2019 to present	Central Asia	Hourly	2024	Free but needs User Account.	API, VRT format , ArcGIS web layer and Export map to PDF	Yes	Yes. Allows you to select your own area	Familiarity with the interface	No	Yes	Yes	Yes. Allows customization of maps.	Yes	Yes	Yes	No	Yes
Evapotranspiration	GLDAS Evapotranspiration 2000 - Present	NASA	Central Asia and Caucasus GeoPortal Powered by ESRI	Calculated by NASA using the Noah land surface model, run at 0.25 degree spatial resolution using satellite and ground-based observational data from the Global Land Data Assimilation System (GLDAS-2.1). The model is run with 3-hourly time steps and aggregated into monthly averages.	Yes	Lacks information on dataset	28 km	2000 to present	Global	3-hourly	2024	Free but needs User Account.	API, KMZ, VRT format , ArcGIS web layer and Export map to PDF	Yes	Yes. Allows you to select your own area	Familiarity with the interface	No	Yes	Yes	Yes. Allows customization of maps.	Yes	Yes	Yes	No	Yes
Humidity																									
Soil Moisture																									
Soil Fertility																									
Soil Erosion	GLDAS Runoff 2000 - Present	NASA	Central Asia and Caucasus GeoPortal Powered by ESRI	Calculated by NASA using the Noah land surface model, run at 0.25 degree spatial resolution using satellite and ground-based observational data from the Global Land Data Assimilation System (GLDAS-2.1). The model is run with 3-hourly time steps and aggregated into monthly averages.	Yes	Lacks information on dataset	28 km	2000 to present	Global	3-hourly	2024	Free but needs User Account.	API, KMZ, VRT format , ArcGIS web layer and Export map to PDF	Yes	Yes. Allows you to select your own area	Familiarity with the interface	No	Yes	Yes	Yes. Allows customization of maps.	Yes	Yes	Yes	No	Yes
Cropland Irrigation																									

Annex 11: SoilGrids EO data for Central Asia soil

Soil Property	INTRINSIC					CONTEXTUALITY						ACCESSIBILITY			REPRESENTATIONAL										
	Name	**Dataset Provider**	**Point of contact**	**Methods of Processing EO Data**	**Use of Ancillary Data**	**Sensor Used for production of dataset**	**Spatial Resolution**	**Temporal Coverage**	**Spatial Coverage**	**Temporal Resolution/Frequency**	**Year of EO launch**	**Data Access**	**Product Format**	**Security**	**User Technical Knowledge Requirement**	**Processing Tools**	**Programming Languages**	**Cloud Processing**	**User Interface (UI)**	**Visualization Tools**	**Scalability**	**Open Standards**	**User Support**	**Community**	**Product User Guide/Manual**
	What is the name of platform	Who is the source of data	Who is the responsible organization for the EO Platform	What type of method is used	Did it utilize additional/auxiliary and supplementary?	What are the sensors used to produced data?	What is the level of detail captured in an image by a sensor system (pixel size)?	Does it have historical data and the revisit frequency you require?	What is the geographical coverage of data?	What is the repeat/revisit time, e.g., time between successive observations of a given location	When did the EO platform become available and functional?	Is the data free, subscription-based, or pay-per-use?	What are the EO data formats available for download?	Does the platform allows you to use your own data securely?	What is the minimum technical knowledge of user to access and download data?	Does the platform offer built-in tools for data analysis ?	Does it support your preferred programming language (e.g., Python, R) for custom analysis?	Does it leverage cloud computing for handling datasets?	Is the interface user-friendly and easy to navigate?	Does it offer tools for visualizing and exploring the data (e.g., generating maps, charts)?	Can the platform handle high volume of data (e.g., multiple downloads from global to community scale)?	Does it adhere to open data standards for interoperability with other platforms?	Does the platform offer adequate user support (e.g., video tutorials, case studies)?	Is there an active user community for knowledge sharing and troubleshooting?	Is there minimum set of documentation for users?
Soil Texture																									
Sand content (Sand-C)	Sand in g/kg	ISRIC	SoilGrids	Uses quantile regression forests with the soil property data from ISRIC World Soil Information Service (WOSIS)	Yes	MODIS land products, SRTM DEM derivatives, climatic images and global landform and lithology maps	250 meters	Based on long-term temporal signatures of the soil surface derived from 15 years of MODIS images	Global	2016	2013	Free to download	WMS, WCS, WebDAV, VRT, OVR, GeoTIFF	Yes	No	Familiarity with interface	No	Yes	Yes. Easy to navigate and find datasets.	No. Only visualize datasets	Yes	Yes	Yes	Yes	Yes
Silt content (Silt-C)	Silt in g/kg	ISRIC	SoilGrids	Uses quantile regression forests with the soil property data from ISRIC World Soil Information Service (WOSIS)	Yes	MODIS land products, SRTM DEM derivatives, climatic images and global landform and lithology maps	250 meters	Based on long-term temporal signatures of the soil surface derived from 15 years of MODIS images	Global	2016	2013	Free to download	WMS, WCS, WebDAV, VRT, OVR, GeoTIFF	Yes	No	Familiarity with interface	No	Yes	Yes. Easy to navigate and find datasets.	No. Only visualize datasets	Yes	Yes	Yes	Yes	Yes
Clay content (Clay-C)	Clay content in g/kg	ISRIC	SoilGrids	Uses quantile regression forests with the soil property data from ISRIC World Soil Information Service (WOSIS)	Yes	MODIS land products, SRTM DEM derivatives, climatic images and global landform and lithology maps	250 meters	Based on long-term temporal signatures of the soil surface derived from 15 years of MODIS images	Global	2016	2013	Free to download	WMS, WCS, WebDAV, VRT, OVR, GeoTIFF	Yes	No	Familiarity with interface	No	Yes	Yes. Easy to navigate and find datasets.	No. Only visualize datasets	Yes	Yes	Yes	Yes	Yes
Bulk Density (BD)	Bulk density in cg/cm³	ISRIC	SoilGrids	Uses quantile regression forests with the soil property data from ISRIC World Soil Information Service (WOSIS)	Yes	MODIS land products, SRTM DEM derivatives, climatic images and global landform and lithology maps	250 meters	Based on long-term temporal signatures of the soil surface derived from 15 years of MODIS images	Global	2016	2013	Free to download	WMS, WCS, WebDAV, VRT, OVR, GeoTIFF	Yes	No	Familiarity with interface	No	Yes	Yes. Easy to navigate and find datasets.	No. Only visualize datasets	Yes	Yes	Yes	Yes	Yes
Available Water Capacity	Vol. water content at -10 kPain (10−2 cm3 cm−3)*10	ISRIC	SoilGrids	Uses quantile regression forests with the soil property data from ISRIC World Soil Information Service (WOSIS)	Yes	MODIS land products, SRTM DEM derivatives, climatic images and global landform and lithology maps	250 meters	Based on long-term temporal signatures of the soil surface derived from 15 years of MODIS images	Global	2016	2013	Free to download	WMS, WCS, WebDAV, VRT, OVR, GeoTIFF	Yes	No	Familiarity with interface	No	Yes	Yes. Easy to navigate and find datasets.	No. Only visualize datasets	Yes	Yes	Yes	Yes	Yes
Hydrogen Potential (pH)	pH waterin pH*10	ISRIC	SoilGrids	Uses quantile regression forests with the soil property data from ISRIC World Soil Information Service (WOSIS)	Yes	MODIS land products, SRTM DEM derivatives, climatic images and global landform and lithology maps	250 meters	Based on long-term temporal signatures of the soil surface derived from 15 years of MODIS images	Global	2016	2013	Free to download	WMS, WCS, WebDAV, VRT, OVR, GeoTIFF	Yes	No	Familiarity with interface	No	Yes	Yes. Easy to navigate and find datasets.	No. Only visualize datasets	Yes	Yes	Yes	Yes	Yes
Total nitrogen (TN)	Nitrogen in cg/kg	ISRIC	SoilGrids	Uses quantile regression forests with the soil property data from ISRIC World Soil Information Service (WOSIS)	Yes	MODIS land products, SRTM DEM derivatives, climatic images and global landform and lithology maps	250 meters	Based on long-term temporal signatures of the soil surface derived from 15 years of MODIS images	Global	2016	2013	Free to download	WMS, WCS, WebDAV, VRT, OVR, GeoTIFF	Yes	No	Familiarity with interface	No	Yes	Yes. Easy to navigate and find datasets.	No. Only visualize datasets	Yes	Yes	Yes	Yes	Yes
Potassium (K)																									
Calcium (Ca)																									
Magnesium (Mg)																									
Phosphorus (P)																									
Soil Nutrient																									
Cation exchange capacity (CEC)	Cation exchange capacity (at ph 7)in mmol(c)/kg	ISRIC	SoilGrids	Uses quantile regression forests with the soil property data from ISRIC World Soil Information Service (WOSIS)	Yes	MODIS land products, SRTM DEM derivatives, climatic images and global landform and lithology maps	250 meters	Based on long-term temporal signatures of the soil surface derived from 15 years of MODIS images	Global	2016	2013	Free to download	WMS, WCS, WebDAV, VRT, OVR, GeoTIFF	Yes	No	Familiarity with interface	No	Yes	Yes. Easy to navigate and find datasets.	No. Only visualize datasets	Yes	Yes	Yes	Yes	Yes
Soil Salinity																									
Soil organic carbon (SOC)	Soil organic carbon in dg/kg	ISRIC	SoilGrids	Uses quantile regression forests with the soil property data from ISRIC World Soil Information Service (WOSIS)	Yes	MODIS land products, SRTM DEM derivatives, climatic images and global landform and lithology maps	250 meters	Based on long-term temporal signatures of the soil surface derived from 15 years of MODIS images	Global	2016	2013	Free to download	WMS, WCS, WebDAV, VRT, OVR, GeoTIFF	Yes	No	Familiarity with interface	No	Yes	Yes. Easy to navigate and find datasets.	No. Only visualize datasets	Yes	Yes	Yes	Yes	Yes
Soil microbial respiration (SMR)																									
Temperature (LST)																									
Precipitation																									
Evapotranspiration																									
Humidity																									
Soil Moisture																									
Soil Fertility																									
Soil Erosion																									
Cropland Irrigation																									

Annex 12: GEE EO data for Central Asia soil

oil Property	INTRINSIC					CONTEXTUALITY						ACCESSIBILITY			REPRESENTATIONAL										
	Name	Dataset Provider	Point of contact	Methods of Processing EO Data	Use of Ancillary Data	Sensor Used for production of dataset	Spatial Resolution	Temporal Coverage	Spatial Coverage	Temporal Resolution/Frequency	Year of EO launch	Data Access	Product Format	Security	User Technical Knowledge Requirement	Processing Tools	Programming Languages	Cloud Processing	User Interface (UI)	Visualization Tools	Scalability	Open Standards	User Support	Community	Product User Guide/Manual
	What is the name of platform	Who is the source of data	Who is the responsible organization for the EO Platform	What type of method is used	Did it utilize additional/auxiliary and supplementary?	What are the sensors used to produced data?	What is the level of detail captured in an image by a sensor system (pixel size)?	Does it have historical data and the revisit frequency you require?	What is the geographical coverage of data?	What is the repeat/revisit time, e.g., time between successive observations of a given location	When did the EO platform become available and functional?	Is the data free, subscription-based, or pay-per-use?	What are the EO data formats available for download?	Does the platform allows you to use your own data securely?	What is the minimum technical knowledge of user to access and download data?	Does the platform offer built-in tools for data analysis ?	Does it support your preferred programming language (e.g., Python, R) for custom analysis?	Does it leverage cloud computing for handling datasets?	Is the interface user-friendly and easy to navigate?	Does it offer tools for visualizing and exploring the data (e.g., generating maps, charts)?	Can the platform handle high volume of data (e.g., multiple downloads from global to community scale)?	Does it adhere to open data standards for interoperability with other platforms?	Does the platform offer adequate user support (e.g., video tutorials, case studies)?	Is there an active user community for knowledge sharing and troubleshooting?	Is there minimum set of documentation for users?
ꞩoil Texture	OpenLandMap Soil Texture Class (USDA System)	EnvirometriX Ltd	Google Earth Engine	Derived from predicted soil texture fractions using the soiltexture package in R	Yes	Historic land use maps HYDE data set, MODIS land products, Global Precipitation Measurement Integrated Multi-satellitE Retrievals for GPM (IMERG) Global landscape degradation degree map 1992-2015	250 meters	1950 to 2018	Global	2018	2001	Need user account	GeoTIFF, TFRecord, Feature Collection as CSV,SHP, GeoJSON, KML, KMZ and Map Tiles	Yes. Allows upload of own dataset without being	Requires basic knowledge of programming.	Yes. Use built-in codes in javascript for data manipulation	Python and Javascript	Yes	Yes. Easy to navigate and easy to find different datasets.	Yes. It has Google Code Editor to view and manipulate data.	Yes	Yes	Yes	Yes	Yes
Sand content (Sand-C)	OpenLandMap Sand Content	EnvirometriX Ltd	Google Earth Engine	Derived from predicted soil texture fractions using the soiltexture package in R	Yes	Historic land use maps HYDE data set, MODIS land products, Global Precipitation Measurement Integrated Multi-satellitE Retrievals for GPM (IMERG) Global landscape degradation degree map 1992-2015	250 meters	1950 to 2018	Global	2018	2001	Need user account	GeoTIFF, TFRecord, Feature Collection as CSV,SHP, GeoJSON, KML, KMZ and Map Tiles	Yes. Allows upload of own dataset without being	Requires basic knowledge of programming.	Yes. Use built-in codes in javascript for data manipulation	Python and Javascript	Yes	Yes. Easy to navigate and easy to find different datasets.	Yes. It has Google Code Editor to view and manipulate data.	Yes	Yes	Yes	Yes	Yes
Silt content (Silt-C)																									
Clay content (Clay-C)	OpenLandMap Clay Content	EnvirometriX Ltd	Google Earth Engine	Derived from predicted soil texture fractions using the soiltexture package in R	Yes	Historic land use maps HYDE data set, MODIS land products, Global Precipitation Measurement Integrated Multi-satellitE Retrievals for GPM (IMERG) Global landscape degradation degree map 1992-2015	250 meters	1950 to 2018	Global	2018	2001	Need user account	GeoTIFF, TFRecord, Feature Collection as CSV,SHP, GeoJSON, KML, KMZ and Map Tiles	Yes. Allows upload of own dataset without being	Requires basic knowledge of programming.	Yes. Use built-in codes in javascript for data manipulation	Python and Javascript	Yes	Yes. Easy to navigate and easy to find different datasets.	Yes. It has Google Code Editor to view and manipulate data.	Yes	Yes	Yes	Yes	Yes
Bulk Density (BD)	OpenLandMap Soil Bulk Density	EnvirometriX Ltd	Google Earth Engine	Derived from predicted soil texture fractions using the soiltexture package in R	Yes	Historic land use maps HYDE data set, MODIS land products, Global Precipitation Measurement Integrated Multi-satellitE Retrievals for GPM (IMERG) Global landscape degradation degree map 1992-2015	250 meters	1950 to 2018	Global	2018	2001	Need user account	GeoTIFF, TFRecord, Feature Collection as CSV,SHP, GeoJSON, KML, KMZ and Map Tiles	Yes. Allows upload of own dataset without being	Requires basic knowledge of programming.	Yes. Use built-in codes in javascript for data manipulation	Python and Javascript	Yes	Yes. Easy to navigate and easy to find different datasets.	Yes. It has Google Code Editor to view and manipulate data.	Yes	Yes	Yes	Yes	Yes
Available Water Capacity																									
Hydrogen Potential (pH)	OpenLandMap Soil pH in H2O	EnvirometriX Ltd	Google Earth Engine	Derived from predicted soil texture fractions using the soiltexture package in R	Yes	Historic land use maps HYDE data set, MODIS land products, Global Precipitation Measurement Integrated Multi-satellitE Retrievals for GPM (IMERG) Global landscape degradation degree map 1992-2015	250 meters	1950 to 2018	Global	2018	2001	Need user account	GeoTIFF, TFRecord, Feature Collection as CSV,SHP, GeoJSON, KML, KMZ and Map Tiles	Yes. Allows upload of own dataset without being	Requires basic knowledge of programming.	Yes. Use built-in codes in javascript for data manipulation	Python and Javascript	Yes	Yes. Easy to navigate and easy to find different datasets.	Yes. It has Google Code Editor to view and manipulate data.	Yes	Yes	Yes	Yes	Yes
Total nitrogen (TN)																									
ꞌotassium (K)																									
ꞌalcium (Ca)																									
Magnesium (Mg)																									
ꞌhosphorus (P)																									

Soil Property	INTRINSIC					CONTEXTUALITY						ACCESSIBILITY			REPRESENTATIONAL										
	Name	Dataset Provider	Point of contact	Methods of Processing EO Data	Use of Ancillary Data	Sensor Used for production of dataset	Spatial Resolution	Temporal Coverage	Spatial Coverage	Temporal Resolution/Frequency	Year of EO launch	Data Access	Product Format	Security	User Technical Knowledge Requirement	Processing Tools	Programming Languages	Cloud Processing	User Interface (UI)	Visualization Tools	Scalability	Open Standards	User Support	Community	Product User Guide/Manual
	What is the name of platform	Who is the source of data	Who is the responsible organization for the EO Platform	What type of method is used	Did it utilize additional/auxiliary and supplementary?	What are the sensors used to produced data?	What is the level of detail captured in an image by a sensor system (pixel size)?	Does it have historical data and the revisit frequency you require?	What is the geographical coverage of data?	What is the repeat/revisit time, e.g., time between successive observations of a given location	When did the EO platform become available and functional?	Is the data free, subscription-based, or pay-per-use?	What are the EO data formats available for download?	Does the platform allows you to use your own data securely?	What is the minimum technical knowledge of user to access and download data?	Does the platform offer built-in tools for data analysis ?	Does it support your preferred programming language (e.g., Python, R) for custom analysis?	Does it leverage cloud computing for handling datasets?	Is the interface user-friendly and easy to navigate?	Does it offer tools for visualizing and exploring the data (e.g., generating maps, charts)?	Can the platform handle high volume of data (e.g., multiple downloads from global to community scale)?	Does it adhere to open data standards for interoperability with other platforms?	Does the platform offer adequate user support (e.g., video tutorials, case studies)?	Is there an active user community for knowledge sharing and troubleshooting?	Is there minimum set of documentation for users?
Soil Nutrient																									
Cation exchange capacity (CEC)																									
Soil Salinity																									
Soil organic carbon (SOC)	OpenLandMap Soil Organic Carbon Content	EnvirometriX Ltd	Google Earth Engine	Derived from predicted soil texture fractions using the soiltexture package in R	Yes	Historic land use maps HYDE data set, MODIS land products, Global Precipitation Measurement Integrated Multi-satellitE Retrievals for GPM (IMERG) Global landscape degradation degree map 1992-2015	250 meters	1950 to 2018	Global	2018	2001	Need user account	GeoTIFF, TFRecord, Feature Collection as CSV,SHP, GeoJSON, KML, KMZ and Map Tiles	Yes. Allows upload of own dataset without being	Requires basic knowledge of programming.	Yes. Use built-in codes in javascript for data manipulation	Python and Javascript	Yes	Yes. Easy to navigate and easy to find different datasets.	Yes. It has Google Code Editor to view and manipulate data.	Yes	Yes	Yes	Yes	Yes
Soil microbial respiration (SMR)																									
Temperature (LST)	GCOM-C/SGLI L3 Land Surface Temperature (V3)	Global Change Observation Mission (GCOM)	Google Earth Engine	Using the SGLI LST Split window algorithm	Yes	Synthetic Aperture Radar (SAR) Global Land Observations (SGLI) Level 1 Terrain Observation (LTO) Atmospheric Correction (AQ), Sentinel-2 Global Land Infrared CLoud Free Gaps and Quality, AVHRR, MODIS, ASTER LST	4638.3 meters	2021 to present. 3-4 days	Global	2020	2001	Need user account	GeoTIFF, TFRecord, Feature Collection as CSV,SHP, GeoJSON, KML, KMZ and Map Tiles	Yes. Allows upload of own dataset without being	Requires basic knowledge of programming.	Yes. Use built-in codes in javascript for data manipulation	Python and Javascript	Yes	Yes. Easy to navigate and easy to find different datasets.	Yes. It has Google Code Editor to view and manipulate data.	Yes	Yes	Yes	Yes	Yes
Precipitation	CHIRPS Daily: Climate Hazards Group InfraRed Precipitation With Station Data (Version 2.0 Final)	UCSB/CHG	Google Earth Engine	Generated through a two-part process. Firstly, IR Precipitation (IRP) pentad rainfall estimates are derived from satellite data by calculating the percentage of time during the pentad that the IR observations indicate cold cloud tops (<235° K). Then the station data is integrated with the CHIRP data to create the final product, CHIRPS.	Yes	CHPClim, Quasi-global geostationary thermal infrared (IR) satellite observations from two NOAA sources, Tropical Rainfall Measuring Mission (TRMM) 3B42 product from NASA, Atmospheric model rainfall fields from the NOAA Climate Forecast System	5566 meters	1981 to prese	Global	2015	2001	Need user account	GeoTIFF, TFRecord, Feature Collection as CSV,SHP, GeoJSON, KML, KMZ and Map Tiles	Yes. Allows upload of own dataset without being	Requires basic knowledge of programming.	Yes. Use built-in codes in javascript for data manipulation	Python and Javascript	Yes	Yes. Easy to navigate and easy to find different datasets.	Yes. It has Google Code Editor to view and manipulate data.	Yes	Yes	Yes	Yes	Yes
Evapotranspiration	MOD16A2.061: Terra Net Evapotranspiration 8-Day Global 500m	NASA LP DAAC at the USGS EROS Center	Google Earth Engine	MOD16 algorithm uses the Penman-Monteith equation to calculate evapotranspiration. It incorporates daily meteorological data and MODIS data on vegetation properties, albedo, and land cover to make these calculations	Yes	MODIS Terra and Aqua satellites	500 meters	2001 to prresent. 5 to 6 days.	Global	2019	2001	Need user account	GeoTIFF, TFRecord, Feature Collection as CSV,SHP, GeoJSON, KML, KMZ and Map Tiles	Yes. Allows upload of own dataset without being	Requires basic knowledge of programming.	Yes. Use built-in codes in javascript for data manipulation	Python and Javascript	Yes	Yes. Easy to navigate and easy to find different datasets.	Yes. It has Google Code Editor to view and manipulate data.	Yes	Yes	Yes	Yes	Yes
Humidity	RTMA: Real-Time Mesoscale Analysis	NOAA/NWS	Google Earth Engine	Direct Records from RTMA	Yes	MODIS Terra and Aqua satellites	2500 meters	2011 to present. Hourly	Global	2018	2001	Need user account	GeoTIFF, TFRecord, Feature Collection as CSV,SHP, GeoJSON, KML, KMZ and Map Tiles	Yes. Allows upload of own dataset without being	Requires basic knowledge of programming.	Yes. Use built-in codes in javascript for data manipulation	Python and Javascript	Yes	Yes. Easy to navigate and easy to find different datasets.	Yes. It has Google Code Editor to view and manipulate data.	Yes	Yes	Yes	Yes	Yes
Soil Moisture	SPL4SMGP.007 SMAP L4 Global 3-hourly 9-km Surface and Root Zone Soil Moisture	Google and NSIDC	Google Earth Engine	SMAP L4 soil moisture is based on land model simulations alone, without the concomitant assimilation of SMAP brightness temperature observations. Significant SMAP instrument outages	Yes	NOAA Climate Prediction Center "Unified" (CPCU) global, 0.5 degree, daily, gauge-based precipitation data, NASA Integrated Multi-satellitE Retrievals for the Global Precipitation Measurement mission (IMERG) quasi-global, • Goddard Earth Observing System (GEOS) Forward Processing (FP) global, 0.25-degree, SMAP L1C Radiometer	11000 meters	2015 to present. 3-hourly	Global	2022	2001	Need user account	GeoTIFF, TFRecord, Feature Collection as CSV,SHP, GeoJSON, KML, KMZ and Map Tiles	Yes. Allows upload of own dataset without being	Requires basic knowledge of programming.	Yes. Use built-in codes in javascript for data manipulation	Python and Javascript	Yes	Yes. Easy to navigate and easy to find different datasets.	Yes. It has Google Code Editor to view and manipulate data.	Yes	Yes	Yes	Yes	Yes
Soil Fertility																									
Soil Erosion																									

Reference List

Abdulraheem, Mukhtar Iderawumi, Wei Zhang, Shixin Li, Ata Jahangir Moshayedi, Aitazaz A. Farooque, and Jiandong Hu. 2023. “Advancement of Remote Sensing for Soil Measurements and Applications: A Comprehensive Review.” *Sustainability* 15 (21): 15444. https://doi.org/10.3390/su152115444.

Andrewsi, Susan S, and C Ronald Carroll. 2024. “Designing a Soil Quality Assessment Tool for Sustainable Agroecosystem Management.” *Ecological Applications*.

Andries, Ana, Stephen Morse, Richard J. Murphy, Jim Lynch, and Emma R. Woolliams. 2022. “Using Data from Earth Observation to Support Sustainable Development Indicators: An Analysis of the Literature and Challenges for the Future.” *Sustainability* 14 (3): 1191. https://doi.org/10.3390/su14031191.

APEX Publishers. n.d. “Soil Properties: Physical, Chemical, and Biological.” Accessed May 23, 2024. https://www.organic-crop-production.com/organic_crop_production/soil_management_organic_farms/soil_properties.htm.

Aravind. 2021. “Earth Observation Hype Cycle: 2023 Edition.” June 7, 2021. https://newsletter.terrawatchspace.com/p/earth-observation-hype-cycle-2023.

Awais, Muhammad, Syed Muhammad Zaigham Abbas Naqvi, Hao Zhang, Linze Li, Wei Zhang, Fuad A. Awwad, Emad A. A. Ismail, M. Ijaz Khan, Vijaya Raghavan, and Jiandong Hu. 2023. “AI and Machine Learning for Soil Analysis: An Assessment of Sustainable Agricultural Practices.” *Bioresources and Bioprocessing* 10 (1): 90. https://doi.org/10.1186/s40643-023-00710-y.

Ben-Dor, Eyal, J. Irons, and G.F. Epema. 1999. *Soil Reflectance*.

Bünemann, Else K., Giulia Bongiorno, Zhanguo Bai, Rachel E. Creamer, Gerlinde De Deyn, Ron De Goede, Luuk Fleskens, et al. 2018. “Soil Quality – A Critical Review.” *Soil Biology and Biochemistry* 120 (May):105–25. https://doi.org/10.1016/j.soilbio.2018.01.030.

Cappiello, Cinzia, Chiara Francalanci, and Barbara Pernici. 2004. “Data Quality Assessment from the User’s Perspective.” In *Proceedings of the 2004 International Workshop on Information Quality in Information Systems*, 68–73. Paris France: ACM. https://doi.org/10.1145/1012453.1012465.

Cardoso, Elke Jurandy Bran Nogueira, Rafael Leandro Figueiredo Vasconcellos, Daniel Bini, Marina Yumi Horta Miyauchi, Cristiane Alcantara Dos Santos, Paulo Roger Lopes Alves, Alessandra Monteiro De Paula, André Shigueyoshi Nakatani, Jamil De Moraes Pereira, and Marco Antonio Nogueira. 2013. “Soil Health: Looking for Suitable Indicators. What Should Be Considered to Assess the Effects of Use and Management on Soil Health?” *Scientia Agricola* 70 (4): 274–89. https://doi.org/10.1590/S0103-90162013000400009.

Diaz-Gonzalez, Freddy A., Jose Vuelvas, Carlos A. Correa, Victoria E. Vallejo, and D. Patino. 2022. "Machine Learning and Remote Sensing Techniques Applied to Estimate Soil Indicators – Review." *Ecological Indicators* 135 (February):108517. https://doi.org/10.1016/j.ecolind.2021.108517.

Doran, John W., and Michael R. Zeiss. 2000. "Soil Health and Sustainability: Managing the Biotic Component of Soil Quality." *Applied Soil Ecology* 15 (1): 3–11. https://doi.org/10.1016/S0929-1393(00)00067-6.

Doran, J.W., M. Sarrantonio, and M.A. Liebig. 1996. "Soil Health and Sustainability." In *Advances in Agronomy*, 56:1–54. Elsevier. https://doi.org/10.1016/S0065-2113(08)60178-9.

Dwyer, John L., David P. Roy, Brian Sauer, Calli B. Jenkerson, Hankui K. Zhang, and Leo Lymburner. 2018. "Analysis Ready Data: Enabling Analysis of the Landsat Archive." *Remote Sensing* 10 (9): 1363. https://doi.org/10.3390/rs10091363.

EarthMap. n.d. "Earth Map." Accessed May 6, 2024. https://help.earthmap.org/.

Eisele, Andreas, Ian Lau, Robert Hewson, Dan Carter, Buddy Wheaton, Cindy Ong, Thomas

John Cudahy, Sabine Chabrillat, and Hermann Kaufmann. 2012. "Applicability of the Thermal Infrared Spectral Region for the Prediction of Soil Properties Across Semi-Arid Agricultural Landscapes." *Remote Sensing* 4 (11): 3265–86. https://doi.org/10.3390/rs4113265.

ESRI. 2024. "Esri Launches Central Asia and Caucasus Geoportal." *Esri* (blog). May 28, 2024. https://www.esri.com/about/newsroom/announcements/esri-launches-central-asia-and-caucasus-geoportal/.

FAO. 2015. "AGROECOLOGY TO REVERSE SOIL Degradation and Achieve Food Security." *EL SALVADOR*.

———. n.d.-a. "FAO." Accessed May 13, 2024. https://data.apps.fao.org/map/catalog/srv/eng/catalog.search#/metadata/ec81aff5-2b00-4534-a780-78178288d21d.

———. n.d.-b. "GLOSIS | Global Soil Partnership | Food and Agriculture Organization of the United Nations." Accessed June 3, 2024. https://www.fao.org/global-soil-partnership/areas-of-work/soil-information-and-data/glosis/en/.

Gasperi, Jerome, Frédéric Houbie, Andrew Woolf, and Steven Smolders. 2016. "OGC Earth Observation Metadata Profile of Observations & Measurements." Implementation Standard. Open Geospatial Consortium. June 9, 2016. https://docs.ogc.org/is/10-157r4/10-157r4.html#11.

Gomes, Vitor, Gilberto Queiroz, and Karine Ferreira. 2020. "An Overview of Platforms for Big Earth Observation Data Management and Analysis." *Remote Sensing* 12 (8): 1253. https://doi.org/10.3390/rs12081253.

"Google Earth Engine." n.d. Accessed June 13, 2024. https://earthengine.google.com.

Guggenberger, Georg, Olga Shibistova, Peter Liebelt, Hafiz Muminjanov, Amir Kassam, and Theodor Friedrich. 2022. "Knowledge-Based Soil Management for a Sustainable Ecological and Economic Development of Central Asia.” Application/pdf. GFZ German Research Centre for Geosciences. https://doi.org/10.48440/GCA.2022.005.

Holmes, Chris. 2018. “Analysis Ready Data Defined.” *Planet Stories* (blog). April 26, 2018. https://medium.com/planet-stories/analysis-ready-data-defined-5694f6f48815.

“Hand-in-Hand Geospatial Platform | Food and Agriculture Organization of the United Nations.” n.d. Accessed June 27, 2024. https://www.fao.org/hih-geospatial-platform/en.

Huete, A R. n.d. “REMOTE SENSING FOR ENVIRONMENTAL MONITORING.”
ICARDA. n.d. “Geoinformatics Spatial Solutions for Integrated Agro-Ecosystems.” Accessed June 11, 2024. http://geoagro.icarda.org/en/.

Idowu, Omololu J., Harold M. Van Es, George S. Abawi, David W. Wolfe, Judith I. Ball, Beth K. Gugino, Bianca N. Moebius, Robert R. Schindelbeck, and Ali V. Bilgili. 2008. “Farmer-Oriented Assessment of Soil Quality Using Field, Laboratory, and VNIR Spectroscopy Methods.” *Plant and Soil* 307 (1–2): 243–53. https://doi.org/10.1007/s11104-007-9521-0.

Information, ISRIC-World Soil. n.d. “SoilGrids Web Portal.” SoilGrids Web Portal. Accessed June 11, 2024. https://soilgrids.org.

Interstate Commission for Water Coordination. n.d. “Central Asia: Background Information.” Accessed June 7, 2024. http://www.icwc-aral.uz/general.htm.

“ISRIC — World Soil Information.” n.d. ISRIC — World Soil Information. Accessed June 13, 2024. https://www.isric.org.

Joksic, Dusan, and Branislav Bajat. 2004. “Elements of Spatial Data Quality as Information Technology Support for Sustainable Development Planning.” *Spatium*, no. 11, 77–83. https://doi.org/10.2298/SPAT0411077J.

Karagiannis, Akis. 2023. “The Challenge of Analysis Ready Data in Earth Observation.” April 19, 2023. https://www.spectralreflectance.space/p/the-challenge-of-analysis-ready-data-in-earth-observation-d978ea1df97.

Karlen, D.L., N.S. Eash, and P.W. Unger. 1992. “Soil and Crop Management Effects on Soil Quality Indicators.” *American Journal of Alternative Agriculture* 7 (1–2): 48–55. https://doi.org/10.1017/S0889189300004458.

Karlen, Douglas L, Susan S Andrews, Brian J Wienhold, and Ted M Zobeck. 2008. “Soil Quality Assessment: Past, Present and Future.”

Karlen, Douglas L., Craig A. Ditzler, and Susan S. Andrews. 2003. “Soil Quality: Why and How?” *Geoderma* 114 (3–4): 145–56. https://doi.org/10.1016/S0016-7061(03)00039-9.

Khanal, Sami, John Fulton, and Scott Shearer. 2017. “An Overview of Current and Potential Applications of Thermal Remote Sensing in Precision Agriculture.” *Computers and Electronics in Agriculture* 139 (June):22–32. https://doi.org/10.1016/j.compag.2017.05.001.

Khitakhunov, Azimzhan. n.d. “Agricultural Potential of Central Asian Countries.” Accessed June 7, 2024. https://www.eurasian-research.org/publication/agricultural-potential-of-central-asian-countries/.

Kibblewhite, M.G, K Ritz, and M.J Swift. 2008. “Soil Health in Agricultural Systems.” *Philosophical Transactions of the Royal Society B: Biological Sciences* 363 (1492): 685–701. https://doi.org/10.1098/rstb.2007.2178.

Kogut, Peter. 2023. “Soil Degradation | UNDRR.” Https://Eos.Com/Blog/Soil-Degradation/. June 7, 2023. http://www.undrr.org/understanding-disaster-risk/terminology/hips/en0005.

Kopittke, Peter M., Neal W. Menzies, Peng Wang, Brigid A. McKenna, and Enzo Lombi. 2019. “Soil and the Intensification of Agriculture for Global Food Security.” *Environment International* 132 (November):105078. https://doi.org/10.1016/j.envint.2019.105078.

Lehmann, Johannes, Deborah A. Bossio, Ingrid Kögel-Knabner, and Matthias C. Rillig. 2020. “The Concept and Future Prospects of Soil Health.” *Nature Reviews. Earth & Environment* 1 (10): 544–53. https://doi.org/10.1038/s43017-020-0080-8.

“Master Report - ARD for Africa - May 2022.Pdf.” n.d. Accessed May 3, 2024. https://www.digitalearthafrica.org/sites/default/files/file-uploads/Master%20Report%20-%20ARD%20for%20Africa%20-%20May%202022.pdf.

Morales, Carmen, Alfonso Sanchez-Paus Díaz, Daniel Dionisio, Laura Guarnieri, Giulio Marchi, Danae Maniatis, and Danilo Mollicone. 2023. “Earth Map: A Novel Tool for Fast Performance of Advanced Land Monitoring and Climate Assessment.” *Journal of Remote Sensing* 3 (January):0003. https://doi.org/10.34133/remotesensing.0003.

“ORNL DAAC for Biogeochemical Dynamics.” n.d. Accessed June 27, 2024. https://daac.ornl.gov/.

Parikh, Sanja. 2012a. “Soil: The Foundation of Agriculture | Learn Science at Scitable.” Https://Www.Nature.Com/Scitable/Knowledge/Library/Soil-the-Foundation-of-Agriculture-84224268/. 2012. https://www.nature.com/scitable/knowledge/library/soil-the-foundation-of-agriculture-84224268/.

Parr, J.F., R.I. Papendick, S.B. Hornick, and R.E. Meyer. 1992. “Soil Quality: Attributes and Relationship to Alternative and Sustainable Agriculture.” *American Journal of Alternative Agriculture* 7 (1–2): 5–11. https://doi.org/10.1017/S0889189300004367.

Prakash, Bhanu, and Shashi Kumar. 2022. “Chapter 14 - Emerging Techniques of Polarimetric Interferometric Synthetic Aperture Radar for Scattering-Based Characterization.” In *Radar Remote Sensing*, edited by Prashant K. Srivastava, Dileep Kumar Gupta, Tanvir Islam, Dawei Han, and Rajendra Prasad, 259–85. Earth Observation. Elsevier. https://doi.org/10.1016/B978-0-12-823457-0.00014-8.

Rinot, Oshri, Guy J. Levy, Yosef Steinberger, Tal Svoray, and Gil Eshel. 2019. “Soil Health Assessment: A Critical Review of Current Methodologies and a Proposed New Approach.” *Science of The Total Environment* 648 (January):1484–91. https://doi.org/10.1016/j.scitotenv.2018.08.259.

Safanelli, José L., José A.M. Demattê, Sabine Chabrillat, Raul R. Poppiel, Rodnei Rizzo, André C. Dotto, Nélida E.Q. Silvero, et al. 2021. "Leveraging the Application of Earth Observation Data for Mapping Cropland Soils in Brazil." *Geoderma* 396 (August):115042. https://doi.org/10.1016/j.geoderma.2021.115042.

"Satellite Missions Catalogue - eoPortal." n.d. Accessed May 27, 2024. https://www.eoportal.org/satellite-missions?Mission+type=EO.

Sweta, Leonard Ondongo, and Wietske Bijker. 2013. "Methodology for Assessing the Usability of Earth Observation-Based Data for Disaster Management." *Natural Hazards* 65 (1): 167–99. https://doi.org/10.1007/s11069-012-0351-x.

Toor, Gurpal S., Yun-Ya Yang, Srabani Das, Suzanne Dorsey, and Gary Felton. 2021. "Soil Health in Agricultural Ecosystems: Current Status and Future Perspectives." In *Advances in Agronomy*, 168:157–201. Elsevier. https://doi.org/10.1016/bs.agron.2021.02.004.

Tziolas, Nikolaos, Nikolaos Tsakiridis, Sabine Chabrillat, José A. M. Demattê, Eyal Ben-Dor, Asa Gholizadeh, George Zalidis, and Bas Van Wesemael. 2021. "Earth Observation Data-Driven Cropland Soil Monitoring: A Review." *Remote Sensing* 13 (21): 4439. https://doi.org/10.3390/rs13214439.

UNCCD. n.d. "Healthy Soil for a Food-Secure Future Becomes Focus of World Soil Day in Kazakhstan." UNCCD. Accessed June 6, 2024. https://www.unccd.int/news-stories/stories/healthy-soil-food-secure-future-becomes-focus-world-soil-day-kazakhstan.

US Department of Agriculture. 2024. "Soil Health | Natural Resources Conservation Service." April 29, 2024. https://www.nrcs.usda.gov/conservation-basics/natural-resource-concerns/soils/soil-health.

Viloria, Kimberly. 2023. "Remote Sensing for Impact Evaluation of Agriculture and Natural Resource Management Research: Guidelines for Use in One CGIAR."

Wand, Yair, and Richard Y. Wang. 1996. "Anchoring Data Quality Dimensions in Ontological Foundations." *Communications of the ACM* 39 (11): 86–95. https://doi.org/10.1145/240455.240479.

Wang, Richard Y., and Diane M. Strong. 1996. "Beyond Accuracy: What Data Quality Means to Data Consumers." *Journal of Management Information Systems* 12 (4): 5–33.

Wuepper, David, Pasquale Borrelli, Panos Panagos, Thomas Lauber, Thomas Crowther, Amy Thomas, and David A. Robinson. 2021. "A 'Debt' Based Approach to Land Degradation as an Indicator of Global Change." *Global Change Biology* 27 (21): 5407–10. https://doi.org/10.1111/gcb.15830.