Integration of a wireless sensor network and a participatory soil monitoring system for smallholder agriculture

Candidate: Amsale Zelalem

Promoters: Dr. Javier Morales Guarin, Dr. Rolf A. de By, Dr. Yaregal Assabie

PhD proposal submitted to:
Department of Geo-Information Processing, ITC

February 20, 2019

UNIVERSITY OF TWENTE.

Abstract

Agriculture is one of the pillars to the social, economical and political well-being of Ethiopia. It is a source of food, raw materials, employment and much more. Smallholders are predominant in this economic sector and mostly practice small-scale production. Several natural and human-induced obstacles challenge the performance of farm lands in the country. Weather, farm inputs, farm management practices and soil are among the significant parameters hindering farm productivity of smallholders. Extensive cropping and grazing practices have also adversely affected the country's soil health. These have caused poor water holding capacity, nutrient depletion, reduced soil depth and poor productivity. Soil is an important factor for improved farm yields and clearly farm-specific soil information is needed for planning any intervention. The spatial and temporal variability of soil attributes needs to be understood, both before and during cropping season, to help farmers take appropriate measures. The absence of up-to-date information about these attributes creates a knowledge gap with stakeholders that is essential to tackle. Technology and participatory research can be ways through which these problems are addressed. A Wireless Sensor Network-based Internet of Things (IoT) system together with digital citizen science is proposed in this work to create a robust, scalable and affordable field-level soil data collection and analysis infrastructure. Crop and soil characterization at the farm plot level can play a significant role in informed decisions in sustainable agricultural practices and hence reduce the severe prevalence of food insecurity. A knowledge base with a continuous and (near) real-time data flow coming directly from the field could fuel such logical decisions based on facts and assist in improved yield production; hence reduced food insecurity. New and useful information such as suitability of the land or optimal yield estimations can be inferred from such system. This work envisages to construct a knowledge base of acquired data and apply Artificial Neural Network (ANN) algorithm to model the inter-play between farmlands and crop yield. With farmer's direct involvement in the data collection and analysis, natural resource management and best farm practices are also aimed for in the long run. Such inclusive and participatory platforms also help to create sense of ownership among farmers, while a substantial amount of data is collected that can document changes at the farm level. This in turn enable farmers to monitor and adjust to changing soil and environmental conditions.

Keywords:IoT, WSN, citizen science, knowledge base, agriculture expert system

Contents

Chapter 1

Sustainable development and agriculture

1.1 Preamble

This chapter provides a general background of food security and information challenges faced by smallholders in its context. It then discusses the problem caused by information gap and how the proposed work aims to contribute towards filling the gap. Specific objectives and research questions this work attempts to address are also discussed in this chapter. Lastly, anticipated impact of the work and expected outcomes are discussed.

1.2 General perspectives

According to the UN, food security is a condition in which all people, at all times, have physical, social and economic access to sufficient, safe and nutritious food that meets their dietary needs and food preferences for an active and healthy life [1]. The Second UN Sustainable Development Goal (SDG 2), "Zero Hunger," aims to achieve food security by 2030 through promoting sustainable agricultural practices around the globe [2]. Consequently, productive agricultural practices are given utmost importance [3]. The SDG2 emphasizes agricultural transformation and rural empowerment as critical agents to the envisioned change [2]. Mainly because over 70 % of the world's poor people, where severe hunger exists, are rurally living and also heavily rely on agriculture [2]. Some progress has been observed towards realizing this goal, though it has not been consistent. The unprecedented number of natural and human–induced crises is the reason and the challenges have become of great concern [2, 4]. Drought, climate change, natural resource degradation, and conflicts are some of these factors for the fragile food security of the world [3]. According to [4], an increase in the number of people suffering from hunger has even been observed over the past three years. In 2016, 108 million people were reported as food-insecure from countries affected by conflicts and crisis; this number was 80 million in 2015 [5]. Millions of people are also affected by environmental and natural resource limitations and changes.

Africa is one of the regions where food insecurity has worsened and realization of the SDG2 seems to lag behind with almost half the population falling victim [6]. In Sub-Saharan Africa (SSA), the situation has escalated and more than 23 % of society is estimated to have suffered from severe food shortage in the past years [4]. Ethiopia, as a country in SSA, is no exception and faces difficulty to maintain sustainable food security for its population.

Agriculture is an important pillar for Ethiopia's economy. With more than 40 % contribution to the country's gross domestic products (GDP), about 85 % of employment is also in agriculture [7, 8]. Around 70 % of raw material requirements of local industries are also supplied by agricultural products within the country [9]. Ethiopian agriculture is smallholder-dominated and rain-fed with two crop seasons: Meher (main and long rain season from June to September) and Belg (shorter season from February to May). Despite its significant role in the well-being of the country, agriculture in Ethiopia is marked with frequent low productivity. High climate variability and limited advanced farming practices highly attribute to this situation. Drought and high rainfall intensities frequently result in crop failure and severe soil erosion. These incidents in turn cause land degradation and soil infertility, which are major threats to agriculture and food security [7]. Overall, rainfall patterns, soil health and land degradation, climate change, population growth and poor infrastructure have direct impact on the performance of agriculture. For instance, the El Niño-induced drought that occurred in 2015–16, resulted in a serious food insecurity problem and economic instability [10]. Agricultural production is estimated to decrease by 25 % during such droughts and this number runs to 75 % in some ecologically vulnerable areas [11].

Particularly, northern Ethiopia is known for poor agricultural productivity because of high rates of soil erosion as a result of erratic rainfall and land mismanagement [4]. The soil's physical base and nutrient content is severely damaged and most provinces have poor natural vegetation cover as well [12]. This has frequently caused reduced crop production and has challenged meeting minimum food needs [13]. A recent report [14] also indicates endangered food security reporting that expected harvest was 40 % below average. Food security, thus, remains a challenge to many households in Ethiopia and in particular to those of rural areas.

At the national front, a recent World Bank report stated more than 7 million people required food assistance in 2018 and 3.6 million people who received food aid in the past will also need food assistance until end of the year [15]. Such incidents are life-threatening for households with limited off-farm income generation means. Figure 1.1 shows recent food insecurity status of the country and specifically its prevalence in the Amhara region.

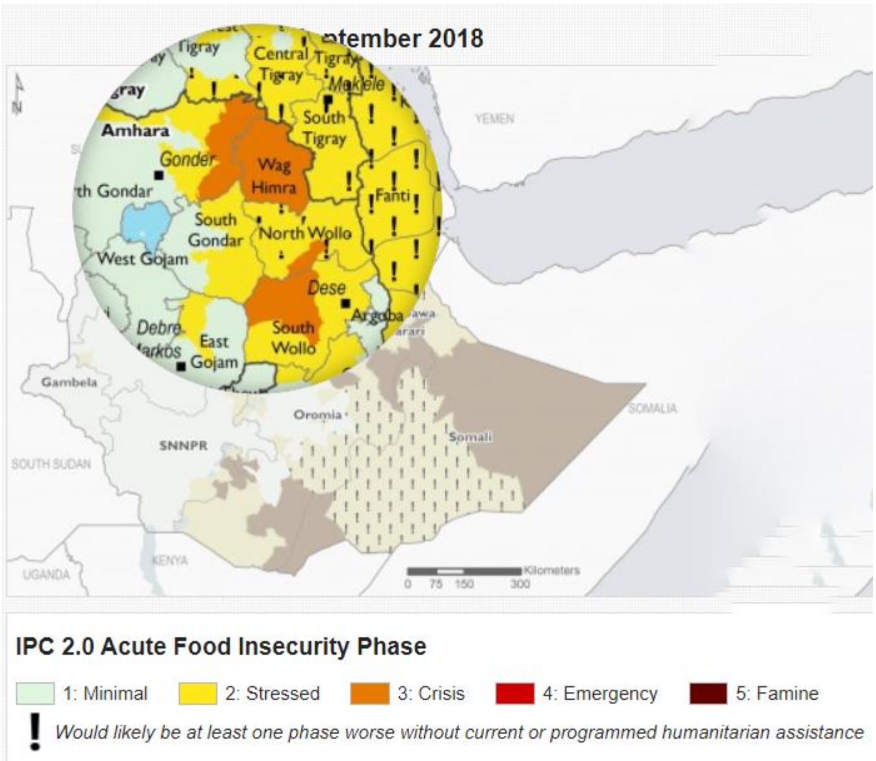


Figure 1.1: *Food Insecurity Status of Amhara Region* Source: [14]

1.3 Challenges

Agricultural production by small households makes a significant contribution to the world's food production while paradoxically at the same time these households are characterized as poor and food-insecure. Especially in the developing world, above 60 % of food supply comes from such small farms [4]. Accordingly, improving the productivity and income of smallholders and rural communities has been identified as a key target to eradicate food insecurity [2]. Most smallholder farmers own a small portion of land to cultivate, less than 2 ha on average, and mostly rely on natural biophysical conditions to crop growth [16]. In Ethiopia, for instance, the production is mainly covered by smallholder farmers which produce about 95 % of crops [8, 17]. Though the country has promising potential for agricultural development, only about 13 % of the land is cultivated with an average farmland size of 1.02 ha per household [18, 19]. Little attention has been given to these farmers and with current worldwide intervention trends, it appears unlikely that the SDG 2 will be achieved in many parts of the world by 2030, and this includes Ethiopia [16]. Thus, it is urgent to accelerate and scale up actions to address these challenges while maintaining a balance between natural resource conservation and adequate, healthy food production [4]. Building a resilient food system with the limited natural resources and unpredictable climatic conditions, thus, puts high pressure on current agriculture; food production systems are highly interlinked societal systems that require attention of coherence [20]. Increased investment in improved agriculture: climate-resilient, inclusive, environmentally sustainable and provide enough to feed the world, is urgently needed [21, 22].

Many smallholder farmers practice a mix of crop growing and livestock farming. Farms are diverse in agro-ecology, farming practices and management. Thus, top-down, one-for-all solutions do not seem to bring the change anticipated. A robust and bottom-up approach for sustainable food production is required that considers this heterogeneity [23]. A number of challenges must be addressed based on the local demands: sustainable agricultural practices, land and soil restoration, improved water usage and improved seed. This calls for collaboration among government, international organizations, civil society and research institutes to create an enabling environment and policy for smallholders to leverage on [23]. Better understanding of local contexts and real scenarios of farmers are important for effective execution of interventions. Agriculture is a dynamic and complex process that needs a holistic and integrated approach. Land, soil, water and other farm inputs are necessary and scarce inputs to farming. Sustained use and management of these resources is imperative to obtain improved yield; modern technology is seen as an enabler for such requirements [24]. Whether in the form of new agricultural practices, improved breeds or improved crops, technology is believed to significantly improve farm yields, reduce farm waste and risks [24]. Better and improved access to enabling technologies is, thus, a common interest among farmers throughout the world [23].

On the other hand, it is often the case that farmers show resistance and are hesitant to adopt technology-assisted farming [25, 26]. Especially in SSA, adoption of farming technology lags behind significantly [27]. Such resistance also plays a considerable role in the low yields of the region [28]. Age, policy, education, risk, start-up cost, lack of awareness and poor access to information communication technology (ICT) and non-participatory research efforts are some of the reasons to the resistance [25, 28]. In a study by [29], lack of appropriate information was given as one reason for weak adoption of farm technology. Proper access to reliable and authentic information on farming technology is, thus, an important requirement to remove the adoption barrier and also empower smallholders with better environmentally-friendly farming practices. Agricultural extension and advisory systems can be used as bridges between technology and farmers.

Agricultural extension services deliver useful information to farmers [30]. They aim to promote agricultural productivity, increase food security, improve livelihoods of farmers and facilitate economic growth through agriculture [31]. In addition, they can also serve as policy instruments and mobilize farmers for necessary behavioral changes [32]. In Ethiopia, the agricultural extension service dates back to 1953 and is known to be one of the densest systems in the world [32]. Training farmers, communicating and demonstrating improved farming techniques, offering market information and advisory services to farmers are the responsibility of the extension service. Nonetheless, the service has been stagnant and its impact has not been as effective as it should be [32]. Lack of precise, detailed and specific agricultural information, lack of adequate skills and knowledge on systematic data collection, representation, and dissemination, absence of localized user-friendly technology that present data in usable form are mentioned in this context [32]. It is imperative to assist these services by production of comprehensive data on farm inputs, livelihoods, natural resource usage, environmental and biophysical attributes. Such data is best disaggregated to community or even farm level. Through the provision of up-to-date agricultural information, extension services can be equipped with required knowledge and contextualized best practices, to be shared with farmers and assist in improved farm yields [32]. However, most agriculture data is variable in space and time. This calls for substantial and frequent maintenance to produce real-time and authentic information that helps farmers to make informed decisions. The whole process can be resource-intensive and sometimes difficult: data shall be collected from remote and inaccessible farmlands; access to extension service providers is limited. These and other socio-economic factors demand for cost-effective innovation that addresses the limitations while producing rich data to fill the information gap. An automated platform that allows proper and timely production and dissemination of information to involved stakeholders, especially farmers, can help make sustainable farming decisions.

1.4 Problem statement

Three attributes contribute to food security; adequate availability, adequate access and appropriate utilization of food [33]. With increased farm yield and better productivity, the first two elements can be achieved. Increase in farm yield, in turn, needs better understanding of local and farm level contexts of climate, soil, crop, and yield potentials. Accurate information about all is, thus, needed unceasingly. Nonetheless, in Africa and notably in Ethiopia, such information is found mostly in fragments and often does not quite fit with household levels, both in scale and content-wise. For most of the regions in the country, soil, crop and other agricultural data at household level are hard, if not impossible, to find. Consequently, intuitive decisions and traditional farm practices persist causing over-exploitation of nature yet poor production.

Agricultural extension workers also face challenges with the absence of down-scaled information. In Ethiopia and most developing countries, in general, agricultural extension services are offered to farmers through physical meet ups with agents. Addressing all farmlands is difficult and one-on-one agent-farmer advisories are impossible. In fact, the most common practice is a once-in-awhile field visit by an agent and meeting with groups of farmers or farmers' visit to agent mostly when critical advisory is needed. It is also unlikely for an agent to know every details of crop yield impacting factors. Thus, limited knowledge of an area is found, unless site-specific information from which adequate inference can be made exists. The recommendation by [34] indicates the severity of the information gap and the urgency for such down-scaled systems, particularly to the Amhara region.

The Amhara region is mostly characterized by rugged mountains, plateaus, valleys, and gorges. Agricultural practice is mostly mountainous with a slope gradient of 5 to 45 %. The region frequently experiences short violent rainfalls that damage the soil significantly, with a reported annual soil loss rate of 250 mm/year [34]. Extensive cropping and grazing practices also have adverse effect on the soil health. These have caused poor water holding capacity, nutrient depletion, reduced soil depth and poor productivity. The prevalence of such problems have caused low organic matter content and nutrients deficiency of the soil, which is a challenge faced in sustained food security and agricultural growth of the region. According to [35], productivity of the country's soil is 40 % lesser than the global average resulting in decline of per-capital food production. Irregular and erratic rainfall also results in frequent soil water stress and loss before used by crops. Soil is an important factor for improved farm yields and clearly farm-specific soil information is needed for planning any intervention. The spatial and temporal variability of soil attributes needs to be understood, both before and during cropping season, to help farmers take appropriate measures.

All the aforementioned challenges call for robust, efficient, site-specific, and agro-ecological information presentation schemes. Every year, the Ethiopia Central Statistics Agency (CSA) collects and disseminates agricultural data under the name Agricultural Sample Survey (AgSS) [36]. This data consists of agricultural production information for the country at region and zone level. However, for smallholder-dominated farming practices, such data cannot be of much use and certainly not at farm level. Further disaggregation to smaller administrative levels, and even to farm level, is thus, important to bring data to decisions. Access to the much-needed information for improved farm productivity and crop yield is expected to allow tackling the aforementioned challenges.

1.5 Hypothesis

Based on the analysis above, the proposed work is based on these hypothesis:

- In Ethiopia, farm field soil data is unavailable or is found only in fragments

- Even when found, these data are of poor quality, maybe outdated and not representative for smallholder farm fields
- Laboratory-based soil chemical analysis is unaffordable, complex, and time-consuming for use at farm level
- Stakeholders of the agri-chain lack adequate and up-to-date information and knowledge about farm level yield affecting parameters
- Crop yield heavily depends on soil properties, particularly on moisture and macro-nutrient levels, and better understanding of these is a cornerstone to any solution
- Dynamic modeling of soil macro-nutrient and moisture level against crop performance helps to better understand the inter-play between farm and crop, which in turn assists in site-specific logical decision makings

1.6 General objective

The general objective of this project is to create a platform for improved information flow from multiple sources to assist crop yield predictions and projections by implementing a (near) real-time, robust, usable, rapidly deployable and affordable soil data collection and analysis tool. An Internet of Things (IoT) Wireless Sensor Network (WSN) architecture on a free spectrum communication channel, will be designed and deployed to acquire soil moisture and temperature. Optimal sensor deployment schemes will be assessed to scale up and optimize data outcomes. This work also aims to design participatory citizen science model and use cost-efficient soil testing kits to determine the total nitrogen, phosphorus and potassium (NPK) nutrient levels of a soil. The data will then be mapped to spatial information through a digital camera and further processed. The research opts to engage the farmer community through an easy-to-use, localized data acquisition and analysis tool. These data, together with satellite provided data and experts knowledge will be used to construct a knowledge base from which usable information that supports agricultural advisory is extracted. A machine learning algorithm will be adapted to design and construct a new knowledge on the inter-dependency among the obtained data and farm productivity.

Specific objectives with their corresponding research questions are as follows:

1. Specific objective 1: Review and study the usability and impact of existing soil data sources for better farm productivity at farm-field level
   - What are the existing sources for soil data? How viable, representative and applicable are these data?
   - What is the relationship between crop types and soil; as advised and as found?
   - Which information gap challenges the agri-chain stakeholders?
2. Specific objective 2: Design and deploy a WSN-assisted IoT Architecture for farm field soil data collection
   - How can farm level in-situ soil data be collected using in-field sensors and how efficient is such a scheme?
   - How can we design robust, affordable and scalable spatio-temporal soil data collection infrastructure in the area of work?
   - How can collected data be efficiently transmitted to a central system, autonomously and in real-time?
   - What are the soil moisture and temperature levels of a farm field and how variable are they in space and time?
3. Specific objective 3: Explore and design participatory field-level soil macro-nutrient collection and analysis platforms and quantify the magnitude of such nutrients.
   - How can farmer communities engage in site-specific soil analysis works?
   - What participatory and user-friendly soil nutrient data collection system can be designed?

- How can we implement accurate and cost effective soil macro-nutrient analysis at farm field?
- What is the macro-nutrient (NPK) level of a farm field and its spatial-temporal variability?
- How can volunteer citizen participation in soil data collection be sustained ?

4. Specific objective 4: Design a knowledge base and expert system to support agricultural advisory services
   - How can data from heterogeneous sources be integrated to design farm-level agricultural knowledge base?
   - What is the spatial-temporal correlation between crop yield, soil moisture and macro-nutrient level?
   - What is the suitable soil water and nutrient level for the dominant crops growing in the study area?
   - What are the agronomic, management practices and climatic factors that affect farm productivity?
   - How suitable is the existing soil for crops growing in the area?
5. Specific objective 5: Conduct empirical analysis and evaluate the performance of the developed system
   - How efficient is the data collection infrastructure deployed in this system?
   - How accurate and usable is the information obtained through the designed inference engine?
   - What can be done to improve the work?

1.7 Research outputs

The output of the whole work is a PhD dissertation with possible articles as chapters. Each objective is anticipated to correspond with at least one scientific article output as follows:

- Assessment and Evaluation of existing soil data sources for Ethiopia at farm-level : limitations, gaps and possible improvements.
- Towards a robust IoT implementation for sustainable farming practices in rural Ethiopia
- Citizen Science for participatory soil data collection: opportunities, challenges and the way forward
- Empirical evaluation and performance assessment of WSN in remote out-door setup
- Artificial neural network algorithm for farm-field soil suitability and crop yield estimations based on soil moisture and macro-nutrient level.
- Integration of IoT, participatory approach and Remote Sensing (RS): robust and scalable primary data collection tool for agriculture

Two MSc research topics are also expected as outputs of this research effort, which will be conducted in parallel while complementing this work. These are draft ideas and can be re-defined as the research progresses.

- Design and implementation of active participatory data collection and dissemination platform for soil data
- Farm field delineation of smallholder farmlands using remote sensing and GIS
- Spatial correlation of farm fields and farm management practices

1.8 Research impact

Crop yield has strong correlation with socio-economic, environmental and biophysical situations. These attributes are variable in space and time and shall be addressed accordingly, if better yield is required. Agricultural decision support systems, hence, need to entertain all these for producing accurate and helpful information to all stakeholders of the agri-chain. The physical and chemical properties of soil, crop, environment, and topological nature of the area need to be available for such systems to succeed. With proper enabling tools, farm level in-situ data of the aforementioned parameters can be obtained and better findings can be formulated. This work envisions filling such requirements through a comprehensive tool development for yield affecting data collection and presentation. With proper integration of in-field and remotely sensed data and appropriate yield simulation technique, accurate and timely information on soil status, crop needs and their impact on crop yield will be available. This outcome offers rich primary data on land status of smallholders, which is important for better farming practices and improved nutritional value food production. The outcome of this work is believed to bring farmers, extension services, policy makers, collaborators and vendors to a common table and work on demand-driven crop yield improving interventions.

High-level stakeholders can make informed decisions at low-level and design agricultural strategies, considering the smallholders, which is a policy gap this work opts to fill. Farmers will also get better information on what and when to plant and anticipate the investment return or they can be assisted in their quest for what, how much and when to exercise interventions on their farm. As such, the research outcome is believed to equip farmers with better information on their farmlands and how to optimize what they have to increase their yield while cutting unnecessary chemical costs and also taking care of nature. This is the socio-economic aspect this work hopes to contribute. Extension services will also benefit from this work by obtaining farm-level information; making it easier to formulate and offer specific farming advisories. Moreover, pairing the information provided by this work with their existing knowledge can enrich their expertise and close the knowledge gap.

With the abundant and detailed data anticipated, this work can also provide a basis for future research on natural resource management and climate-resilient subsistence agriculture practices in the region. It is also to the researchers' belief that with the participatory and engaging platform, farmers will positively contribute to the work while attaining better understanding of their environment. This, in turn, enhances the adoption rate of further technological interventions pertaining to the research outcome. The proposed work strongly believes in inspiring younger generations of the community in science and technology researches and hope in creating a strong design thinking skills among the participants; where real problem solving innovations should come from. Through the youth participation, better environmental awareness is created with long-term impact on natural resource conservation while creating tech-enabled future farmers. The supposed positive influence the youth could make on their family and the community as a whole in terms of better farming practices is also believed to be strong. Assisting the community on the sustained food security journey through technology and engagement is the vision of this work.

1.9 Summary

Soil is an important attribute that affects productivity of farm-fields. In countries like Ethiopia, where agriculture has a significant economical role, it is imperative to monitor the status of the soil regularly to make logical decisions regarding farming practices, input use and resource management, among others. Smallholder farmers need expert knowledge and advices on such practices; stakeholders such as agriculture extension agents, governmental and non-governmental aid organizations also require these information to plan and base their contribution on. However, the absence of up-to-date soil data creates considerable knowledge gap and hinders such actions contributing to the poor productivity of farm-lands and thus considerable prevalence of food insecurity exists through out the country. Providing robust, scalable and efficient timely data on soil is thus required to fill the gap, which is what this research work aims to achieve.

Chapter 2

Related Work

2.1 Farm-level soil data collection and analysis

Real time, precise, and affordable data collection and dissemination is vital in increasing agricultural productivity. Agriculture-in particular crop production is a cross-sectoral process that depends on data coming from various sources: environment, soil, crop, and other spatio-temporal attributes. These data are variant in space and time, and data needs to be obtained, if possible, at the atomic level. On the other hand, in-situ data collection at the farm level is challenging, if not impossible. Conventional soil sampling techniques are resource intensive, time-consuming and allow only limited spatial coverage. Detailed, timely, and sufficient site-specific data usually cannot be obtained. Moreover, agricultural activities are mostly carried out in rural areas and often sites are inaccessible, especially in developing countries, making it difficult to acquire what is needed at the desired quality. Information communication technology (ICT) is identified as a potential enabler for efficient data acquisition to fill the gap. With the advent of technology and powerful computing infrastructures, it has become possible to obtain, analyze, and generate sound agricultural information in a cost-effective and timely manner. Vast and ever-growing amounts of data about environmental and climate conditions, soil physical and chemical properties, and others are being captured and produced by automated data collection systems. Soil is one of the most important components of agricultural production and has a dominant effect on farm productivity, both in terms of quantity and quality. Automated soil data collection techniques can be generalized into three classes: remote sensing (RS), proximal sensing [37], and in-field sensing [38].

The large spatial and temporal coverage, reduced cost, and recent advances in high-resolution imagery, has made RS and GIS technologies indispensable to the capture and analysis of soil and other farm-related data [39]. RS tools capture soil data remotely by analysing the electromagnetic radiation reflectance of the earth's surface in the visible or infrared (IR) and near-infrared (NIR) wavebands [37]. This reflectance varies due to the varying properties of the soil, and is used to both distinguish soils and to draw inferences about soil characteristics. By using optical and microwave RS tools, data on soil moisture, mineralogy, texture, and micro-nutrients are collected and analyzed. A recent work by [40] shows how Sentinel 1 and 2 as well as MODIS satellite data is used to monitor soil moisture and surface temperature.The potential of RS to sustainable farming practices was also demonstrated in [41]. According to the authors, information about agricultural water consumption, soil moisture level and other farm-related data are successfully captured in (near) real-time by Sentinel satellite. More detailed insight on the role RS and GIS play in agriculture can be found in [37, 39, 42]. However, most of the soil data obtained through remote sensing is qualitative and further processing is required to obtain discrete data. Moreover, the measurement space between the soil and the sensors means that findings are at the coarse level. The interference of noise and unwanted signals also affect the quality of data obtained through RS.

A closer look at the soil is achieved using proximal sensors that capture data at a distance of less than approximately 2 m above the soil surface [43]. This is achieved through applying high-tech and advanced observation and analysis tools. Electromagnetic induction, Soil Electrical Conductivity, Induced Polarization Measurements, Magnetic sensors, Ground Penetrating Radar(GPR), Gamma-Ray Spectrometry are techniques used in this context [37, 44]. In a paper by [45], Hyper-spectral Imaging(HSI) technology is used to classify soil types and to measure the total nitrogen in a soil. Images of soil samples were captured using near-infrared HSI and processed in the lab for further investigation. A review of proximal sensors is presented in [46]. The authors reviewed on-the-go soil Nitrogen, Phosphorous, and Potassium (NPK) data collection sensors and confirmed the applicability of optical sensors for efficient soil data collection. A review of both RS and proximal sensors technology in soil data collection and associated challenges was presented in [47]. Most studies indicate the high potential of such proximal sensing for better observation of soil properties both at the in-farm and laboratory scales [48]. However, the limited spatial coverage and high costs hinder full usability of proximal sensors: a price most farmers and specially of developing countries cannot afford to pay.

The use of small, inexpensive, and efficient sensing devices for in-field data collection has become significant in recent years [49]. The Internet of Things (IoT) and wireless sensors network (WSN) interconnects objects capable of gathering and exchanging information. Through this technology, it has become possible to capture more accurate data affordably–from the spot automatically to the palms of our hands. The IoT has gained great momentum both from the research and industry since first proposed by K. Ashton [50]. It is expected to grow even more in the near future: around 26 billion things will be connected to the Internet with a possible market value of \$450 billion [51, 52]. In agriculture, IoT is creating remarkable opportunities for farm improvements: farmers can monitor their farms almost in real time without necessarily being at the farm; application and use of resources has become more efficient, based on the accurate measurements obtained, leading to smart farming or precision agriculture. It is anticipated that the IoT will further push the future of farming to an even higher level [49]. Today, several large IoT-based agricultural projects are seen across the world: the European Union's Food and Farm 2020 project, Kansas water preserving through sensors, Bangladesh's new sensor technology and the NanoGanesh are worth mentioning [53]. Researchers are also investigating and designing affordable farm improvement mechanisms through using the IoT. For example, a wireless underground network system for the continuous monitoring of soil water contents is proposed by [54]. The work presents a WSN architecture for collecting soil moisture data by deploying sensors underground; implying that the monitoring of soil information over large spatial scales and in (near) real-time is possible at reasonable cost. Research works in [55–60] all show the effective utilization of WSN for farm soil data collection. According to a review by [61], out of the 72 papers reviewed dealing with sensors used in agriculture, 27% of the works concern soil data collection, as shown in Figure 2.1 below, showing the potential of the technology for field-level and underground soil data collection.

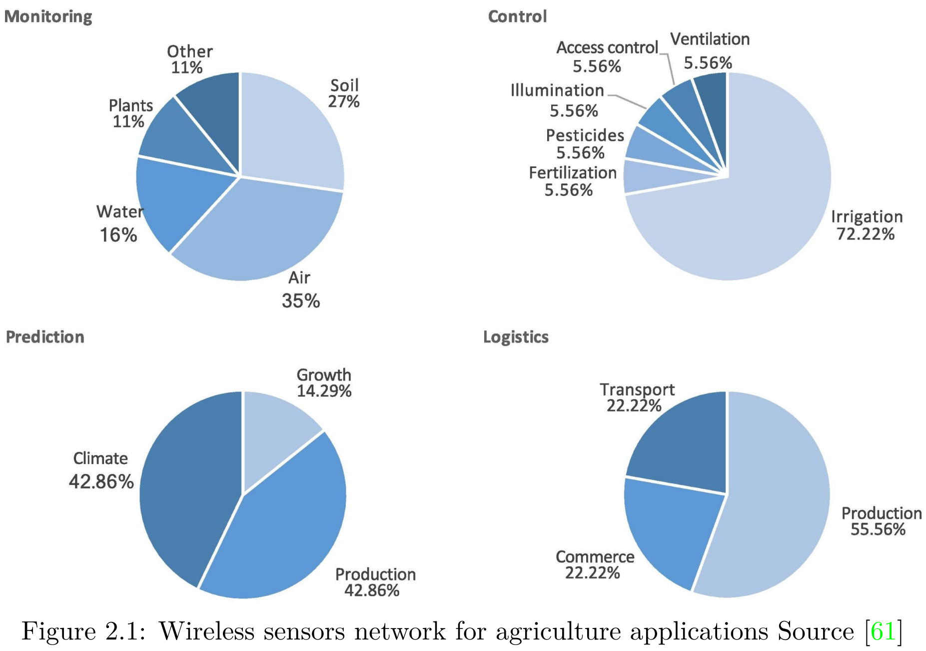


Figure 2.1: Wireless sensors network for agriculture applications Source [61]

In addition to identifying the physical properties of a soil, it is imperative to also know the fertility status and the soil's nutrient content. The nutrient identification and analysis is even more resource intensive than the collection. Common soil nutrient analysis is performed in highly-equipped soil analysis laboratories; mostly unaffordable for low-income countries such as Ethiopia [62]. Consequently, soil samples are often sent to remote laboratories and it takes too long for the analysis result to be returned. Even when such countries have suitable laboratories, findings are mostly unreliable for making further decisions due to the lack of experts and inaccurate readings [63], and are expensive for smallholder farmers to use. Therefore, on the spot, affordable and timely soil testing techniques are required. Remote sensing, satellite imagery and digital cameras have all proved their worth in identifying the nutrient content of a soil; mainly through canopy reflectance and imaging spectroscopy [45, 64–67]. A simpler and farmers-engaging soil nutrient analysis approach will involve the use of a colorimeter and test kits. Soil test kits can be used to determine the Nitrogen, Phosphorous and Potassium (NPK) level of a soil, by observing the color change of a sample soil when mixed with chemicals, without requiring complex laboratory tools or professionals. Kits are small and are simple to carry and operate in the field while they are also inexpensive, making them economical and timely. The usability, accuracy and precision of soil test kits for measuring soil properties was evaluated and positive findings were reported in [68]. They compared the data obtained using the test kit to that from standard laboratory analysis and reported the results to be comparable. The authors recommend repeated test kit measurements to obtain even better results.

In research by [69], a soil test kit is used with colorimetry to determine the NPK and PH content of a soil sample in Philippines. According to the authors, a color sensor device equipped with a light-to-digital converter was designed and tested. Soil directly taken from the farm is immersed in the test kit, which is then used by the color sensor to determine the NPK level. A promising result was reported in the finding and the authors concluded that the method is comparable to human readings. In another work by [70], a polyvinyl alcohol based hydrogel test kit was developed to detect pre- and post-blast trinitrotoluene (TNT). A digital image colorimetric (DIC) is used together with the test kit and the colorimetric products from the test kit are captured using the existing digital camera in a smart phone. The report highlighted that rapid quantitative and accurate analysis of TNT was achieved by using the test kit in combination with DIC. A review of three commercially available soil testing kits to assess their efficiency at determining soil NPK levels was conducted in [71]. LaMotte, Luster Leaf, and Rapitest are the reviewed kits. The review recognized the effectiveness of the kits but indicated notable performance variations among the kits. It revealed LaMotte's data to be more reliable than the others; most closely correlated with the test results of Oklahoma State University's Soil, Water, and Forage Analytical Laboratory (SWFAL). The work also reported the poor K level indication obtained from all sensors compared to the results found using standard laboratory procedures. Another work by [72] investigated the accuracy of five commercial soil testing kits: La Motte Soil Test Kit, Rapitest, Quick Soiltest, Nitty-Gritty, and Soil Kit. The authors reported 94% and 92% of accuracy for La Motte and Rapitest, respectively, but recommend Rapitest for its ease of use and of interpretation. The authors also argue soil chemical data can be obtained economically and in-time by using accurate test kits. However, all the above mentioned soil data collection and analysis tools will not offer data to the required extent, if used separately. There is a need to integrate and take advantage of the existing technological innovations and build a robust soil data acquisition system at reasonable cost.

2.2 Digital citizen science in agriculture

Public participation has a long and distinguished history in agriculture related research [73, 74] and has several advantages: (1) strong synergies among researchers and farmers; (2) mutual understanding of local contexts and solutions to address those; (3) usability and practicability of research findings; (4) improved understandings of climate, environment and nature resilient production; (5) sense of trust and ownership on the part of farmers; and (6) large amounts of accurate data [73]. Promoting equal partnership and a shared vision among stakeholders is crucial for a successful implementation of agricultural innovations. Research needs to integrate local knowledge from farmers, and farmers must also both feel and see the impact and contribution they are providing to the research and safety of the environment. Most important is also contribution they make to better farming practices, leading to higher productivity and sustained living incomes. Moreover, with a strong heterogeneity of socioeconomic requirements and environmental conditions, methodologies that engage farmers are gaining interest [74].

However, the trend of participatory agricultural research usually involve limited numbers of farmers, living close to the research facility, trained and recruited by researchers. Scalability is a concern as additional investment for training and larger farmer group organizations are needed [75]. Suitable citizen science methodologies to effectively engage many participants and robustly collect data from farms is vital. Citizen science is a collaborative approach to scientific research through public involvement as partners rather than just as users [76]. It is a situation in which citizens contribute to research through voluntarily collecting and analyzing scientific data [77]. Research utilizing citizen science benefits from large ground truth data to base the investigation on and produce feasible solutions. Digital citizen science has two requisites essential for success: access to digital communication or technology and the motivation of participants [78]. These attributes call for two tasks: development of efficient systems to collect, analyze, store and disseminate large datasets, and to both recruit and sustain volunteer participants [79].

A study by [80, 81] conducted in three different countries Ethiopia, India and Honduras assessed the usability of digital citizen science for agriculture. The study indicated the high potential for citizen science at the smallholder farmers' level and the high likelihood of farmers using their mobile phones as a communication medium. Two points are also raised in the work: (1) development of easy to use mobile applications for the data collection and submission, since the illiteracy rate among smallholder farmers of the studied area is high, and (2) use of gamification to motivate and keep participants engaged. Gamification is a new approach introduced as an enabler to digital citizen science in which game design elements are used in non-game contexts with the intention of motivating volunteer participation and thereby improve user experiences [82]. In that study, it was also reported that the majority of the participants had an intrinsic motivation to be involved in citizen science. According to Self-Determination Theory (SDT), intrinsic motivation is when an individual is interested in doing things with no pre-conditions but a mere satisfaction of performing it all by themselves; developing self-confidence, competence and efficiency as well as being connected with others [82, 83]. Thus, the researchers recommended gamification to sustain active participation of farmers.

ANONYMOUS project is an international project of 16 research institutes which implements digital citizen science and assesses the usability of gamification strategy for a collaborative agricultural knowledge construction[84]. The project aimed to develop ICT solutions and enable agri-chain stakeholders to better deal with risks and uncertain situations. The research runs a simulation based on historical agricultural data found in Wikipedia to evaluate the impact of gamification on the participants. According to the paper, the work had been evaluated using the revision history of 4,690 articles, including discussions and related articles extracted between April 2001 and April 2018. The authors found out that the number of actions continued to grow over time. In another study by [85], estimation of sesame yields in Ethiopia using digital citizen science and remote sensing was conducted. The research emphasized the benefits such harmonization offers to conduct crop yield gap analysis at a field level with reasonable accuracy. According to the research, integration of citizen science into sensor-based farm data collection creates a platform for capturing large datasets about biophysical, geo-location, management, farm/er characteristics, farm level incidents and socio-economic factors in real time. The work used farmers' collected crop phenology data to select images within the boundaries of a growing season per field and extracted vegetation indices values. The extracted values were then used by the researchers on an empirical model to estimate sesame yield and the researchers were able to observe improved model accuracy; highlighting the opportunity for using the crowd-sourced data from farm as a calibration to remote sensing based agricultural models. The information derived from remote sensing can also be used to validate crowd-sourced information. The study had also included an assessment on what drives farmers to participate. Agronomic advice was expected by most in return for their active and continued participation. Thus, the paper recommended that farm-level decision support systems be sent to farmers through an SMS or other mobile app platforms for their farm-level management. The research also touched on the need for sustaining such citizen science schemes for a long period through a strategic collaboration and partnership with potential stakeholders: public, private and NGO organizations working towards shared goals. In another attempt, development of a smartphone application for collecting soil depth, texture, pH and organic-matter content has also proved to be successful [86]. The paper discuss MySoil, an open project that engages citizens through a free mobile application used to view soil maps of the United Kingdom (UK). According to the authors, the application also allows users to upload photos and description of soils in their locality. The app has attracted more than 2 million web hits and 12,500 dedicated users since its launch in June 2012. The Open Air Laboratories (OPAL) soil and earthworm survey is another citizen science initiative in the UK that is designed to engage citizens in providing environmental information, including about soil [87]. The work aimed to achieve two objectives: involving the public in soil monitoring and building knowledge and commitment to look after the soil and inform policy makers on the soil state of the country. That survey is reported to collect more primary soil information from citizens than any other. The Global Learning Observations to Benefit the Environment (GLOBE) is another project aimed at protecting the environment through citizen science by specifically targeting youth participants. The work engages primary and secondary school students worldwide to collect and send data to a common database from school weather stations. Students are encouraged to study soils through structured data collection with defined protocols and standardized tools such as field fertility kits, soil thermometers and soil moisture sensors [88]. Use of citizen science as aid to the digital soil mapping (DSM) of classes and properties of the soil is also proposed in [89].

2.3 Deep learning and big data analysis for in-season crop yield estimation

With such enabling technologies in place, vast and growing amounts of farm data is being produced, both in-field and remotely. However, such data are of little value unless processed and presented in a usable form to farmers. This calls for an efficient data management and analysis platforms that seamlessly integrate the collected heterogeneous data. The ability to manage, process and interpret data has become more important than ever. Efficient data analytics and mapping technologies, such as GIS, Artificial Intelligence and machine learning, are vital to achieve these functionality [90]. New information about yield improvements, anticipated risks and mitigation, efficient resource utilization and better market strategies can be inferred [91]. Precision agriculture, which refers to small-scale information-based optimization of inputs to make agricultural processes more productive, relies heavily on the acquisition of accurate data and analysis methodologies to supply the needs of optimum production [92]. As agriculture is resource intensive, and resources are growing increasingly scarce, practitioners are more driven to precision agriculture to increase efficiency from existing resources while also cutting costs significantly [93]. Key elements enabling precision agriculture are ICT and other technological tools such as sensors, drones, satellites and autonomous vehicles [93]. In a paper by [94], existing big data projects in agriculture are reviewed, and the IoT and big data are presented as enablers of precision agriculture. According to the authors, by using IoT and big data, real-time data collection and real-time forecasting can be implemented that can assist farm management and operations, and thus change both the scope and organization of farming. theoretical framework for structuring and analysing data-intensive cases in agro-environmental research was proposed in [95]. That paper selected use cases from three European projects to showcase opportunities and challenges for big data in facilitating decision support in the context of agriculture. A survey of agriculture crop recommender systems is presented in [96]. The paper reviews machine learning and data mining approaches applied to farm management and precision agriculture. A crop recommendation system based on soil characteristics and using a machine learning algorithm is presented in [97]. The authors propose multiple machine learning algorithms to analyze soil and crop data in order to predict yield. An expert system (ES) using the IoT for pest, weed and fertilizer management of cotton fields in Pakistan was presented in [98]. The research used sensors for soil moisture and temperature capture and transmission in real-time, which is then used by the ES to make decisions and send results to farmers' mobile phones. According to the paper, the system offered information on diseases diagnosis, pest and weed attacks, pesticide recommendations as well as disease predictions for the cotton farm investigated.

In [99], a model for smart agriculture using the IoT is presented. A sensor board with temperature, relative humidity, light intensity, barometric pressure and proximity sensing and web cameras to capture images of crops are used in that work. The authors used a Wireless IP Network Gateway (WINGZ) to interconnect the sensor devices to a central server and a decision support system incorporating crop monitoring and alert service was implemented. The outcome of this analysis is delivered to users through web and mobile application interfaces. Agrisys is another precision agriculture system proposed in [59] which utilizes several soil and environmental sensors and analyze the resulting big data using a software tool called LabVIEW. With a controller component, the proposed system is also reported to act dynamically and respond according to the various inputs collected.

Crop yield estimations within a cropping season are of high importance for a number of food security related and other agricultural decisions [1]. Governmental and non-governmental bodies, business enterprises, industries, farmers and populations, as a whole, will benefit from accurate and timely farm yield predictions. In particular, for countries such as Ethiopia with an agriculture dependent economy, timely yield estimation significantly impacts the economical well-being of the country: it gives suitable warnings to decision makers on potential gain or loss in crop yields and helps plan for timely actions; alerts farmers to exploit other income generating options or alternative marketing strategies. Aid organizations, farm input supplies, agricultural trades and insurances, agriculture extension advisories all require the best possible crop yield predictions [100]. Farming practices, soil, inputs and climate affect farm yield, which, by themselves are very complex. Therefore, modeling is used to create a simplified representation of these elements and understand their relationship to the crop yield. A model is a collection of equations and procedures conceptually representing behaviors of a system to better understand the system and make improvements [101]. Crop yield models can be generally classified as empirical, statistical, sample-based or analytical [102]. Empirical models use direct descriptions of observed data or evidence from subjective sources. Statistical models use quantitative descriptions or regressions of the mechanisms and processes that cause the observed behavior of the system. Sampling models use actual measurements such as whole plot harvest, crop cut, and farmers' estimates. analytical models are simulation models that create a virtual representation of crop-weather-soil interactions and generate yield forecasts [101, 103]. Several review papers are available describing crop models and crop yield estimation models; comparisons, opportunities and challenges [104–107]. Crop simulation models incorporating environmental, biophysical and crop-specific parameters are reported to be accurate and versatile in terms of describing the nature of a crop as a function of the others [101]. The World Food Study (WOFOST), MARS Crop Yield Forecasting System (MCYFS), Cropping Systems Simulation Model (CropSyst), STICS, Environmental Policy Integrated Climate (EPIC), Decision Support System for AgroTechnology (DSSAT), and Agricultural Production Systems Simulator (APSIM) are some of the most common and widely used models for yield simulation [108–110]. Several research projects have been conducted using these models for estimating crop yields of farms on both small and large scales [111, 112], These models are generally termed parametric or process-oriented models as they tend to be closed to new situations that may occur after the model has been trained. The non-parametric models, on the other hand, have an open architecture in which every variable that happens to have an impact on the yield will be considered over the growing season[113]. These models describe the general conditions under which crops are grown and make an assumption that similar yields can be obtained under those conditions in the future. The models vary from simple to very complex expert systems that involve artificial intelligent and can infer the relationship between environmental conditions and crop yield such as by applying Normalized Difference Vegetation Indices (NDVI) or other temporal profiles of vegetation indices [114]. In both modeling techniques, data coming from various sources is required as inputs. The general data flow to yield forecasting systems is shown in figure 2.2 below.

[1] Throughout this paper, the terms prediction, estimation and forecast are used interchangeably

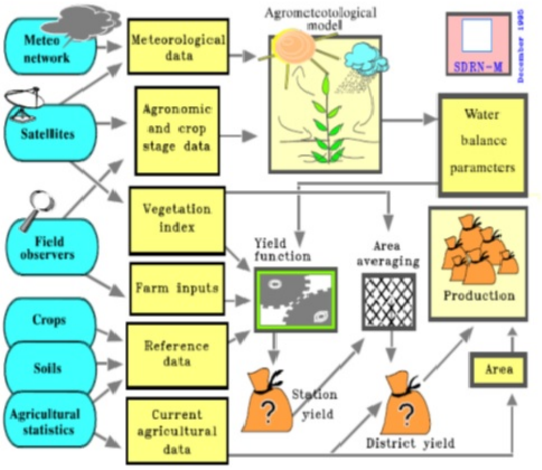


Figure 2.2: Data Flow in Yield Forecasting models Source: [115]

With wide spatial coverage, timely and open data access, remote sensing and satellite technologies are becoming particularly suited to inputs for yield forecasting models. Monitoring Agriculture with Remote Sensing (MARS) is a good example of such integration in which high spatial resolution satellite data is used for yield forecasting throughout Europe [110]. Remote sensing has greatly evolved in recent decades and become valuable for crop models [100, 116–119]. The yield prediction models depend on efficient tools and methodologies to generate reliable outcomes from the available data. A more recent advance in yield prediction models is also the use of algorithms such as artificial intelligence (AI) for data processing and inference making. Machine learning (ML), neural network (NN) and other AI approaches are used in various crop yield forecasting works and reported to be accurate. A convolutional neural network (CNN) and long-short-term-memory (LSTM) NN algorithms were used on a sequence of remotely sensed images to forecast county-level soya bean yield in the United States [120]. Once the models had been trained, using the imagery inputs, a Gaussian Process Modeling, which is a non-parametric probabilistic model, was run over the models to integrate spatio-temporal dependencies between data points. According to the authors, the models showed weak performance during the season's start but better progress was observed as crops grew and more information was gathered and fed to the models. The authors claim the performance of these models outweighs other existing systems for county level yield prediction. In another work by [121], the Convolutional Architecture for Fast Feature Embedding (Caffe) deep learning model is used for Corn yield prediction in Illinois. The authors used Enhanced Vegetation Index (EVI) from MODIS, and climate data including monthly max and min temperature, potential evaporation, and pressure as inputs to the model. The authors evaluated the work using ten cross-fold validation and found a correlation coefficient of 0.810. The paper claims the advantage deep learning and remote sensing brings for better yield estimation, particularly for areas for which data is poor.

The Maximum Entropy (MaxEnt) algorithm, which estimates the likelihood of the occurrence of a species based on presence-only data, is used in a work by [122] for agricultural crop suitability mapping of two crops in Thailand. Three variant input datasets were used in this paper: (1) a socio-demographic survey of households; (2) remote sensing land use and land cover classification using an assembled Landsat TM; and (3) time-series image and ground data from Geodetic Ground Control points. The authors reported that the MaxEnt model has great potential for crop suitability modeling , given careful sample size selection and distribution. According to the authors, the model's performance was highly affected by the sample size and distribution and produced varying results for the three different input datasets. With suitable model validation data, independent data and by ensuring effective distribution of sample data, further improved results can be obtained.

2.4 Research gaps and conclusion

Of course, the performance of any algorithm vary significantly based on the quality of input datasets. Sometimes, satellite imagery and remotely sensed data are coarse, especially for heterogeneous and small-scale areas such as smallholder farms in most developing countries. These farms are often small, exhibit variability in both planting and harvesting times and differ in farm management practices. These and other makes it difficult to even determine field boundaries from remote. The farms can be over-represented or captured data be coarse unless high-resolution tools are used. Specifically, the chemical soil properties of a farm need to be collected below the top surface which is difficult and inaccurate if only remote sensing tools are used. There is a need to complement these data sources with in-field data collection and actual observations. On the other hand, conventional in-situ data collection on such farm levels is also challenging. Technological innovations such as WSNs and croud-sourcing tools can be used to solve the paradox. The use of in-field sensors data in crop models has been experimented with in [123]. The work deployed hand-held Leaf Area Index (LAI) measuring instruments, soil moisture and precipitation sensors for data collection. In that work, the data was then integrated in the LINTUL-3 crop model to forecast potato crop yield for two experimental fields in the south of the Netherlands. The work used two years' worth of data 2010–2011 using the 2010 sensors data for model calibration and obtained yield simulation for 2011 with an R2 of 0.82. According to the author, the estimation error was reduced by implementing an assimilation on the LAI using various datasets including the in-situ LAI taken. The paper concludes by emphasizing the importance of sensors' provided data, especially for the calibration and validation of yield simulation models to improve accuracy. A well-structured digital infrastructure is recommended by the authors to support in-field spatial variations. The sensor network enables local and (near) real-time observations and monitoring of a farm while remote sensing offers larger scale data which might be expensive to obtain using in-field sensors. The IoT and remote sensing technology can offer accurate, fast and large parcel-level yield impact data collection and analysis tools at low cost. The combination of remotely sensed and in-field collected data is believed to create a strong backbone to obtain usable new information on which further farm-level agricultural advisories and decision support systems can rely. With such technological tools in place, more sustainable crop production practices can also be fostered so reducing the environmental impacts of agriculture and food insecurity risks. eleaf from the Netherlands and GEOSYS from France are good examples of platforms that use both remote sensing and in-field sensors for precision agriculture.

Chapter 3

Methodology

3.1 Preamble

The purpose of this chapter is to give a general explanation of the research methods used for the study. The chapter will discuss the research approach and design. Subsequently, conceptual framework and system architecture are presented, followed by methodological discussion per objective. Finally, a discussion about study area, sample size and sampling strategy is presented.

3.2 Proposed research approach

The proposed work aims to continuously monitor the spatial and temporal dynamicity of soil moisture, temperature and macro-nutrients and derive new information on how specific crop reacts to any variations of the behaviors. The overall system is broken down into components shown in Figure 3.1.

As can be seen from the figure, three core modules are defined. WSN and Participatory soil data infrastructure components to acquire *in-situ* soil data, in (near)real-time, and Application module that integrates and analyzes the information acquired and synthesize usable data for decision support in agriculture. Each module is discussed in detail in section 3.5, section 3.6 and section 3.7 below. To accomplish these functionality, a design science approach is adopted as a research methodology throughout the life cycle of this work. Design science is described as: "the design and investigation of artifacts in context" [124]. This approach emphasizes in understanding and improving a problem context by designing set of possible solutions and artifacts that interact with the context to bring the desired outcome. The design science approach has been identified suitable for sciences of the artificial such as computer science and information systems whose main focus is to synthesize useful correlations between nature and external environments [125]. The outcomes of such methodology are: constructs, methods, models, algorithms or even improved theories and new knowledge. We thus aim to produce constructs, models and tools for studying the correlation between soil properties and farm productivity and generate new and usable information to actors in agricultural value chains: farmers, agriculture extension agents, aid organizations, policy makers and other government agencies. Specifically, continual observations of aforementioned soil properties over a cropping season at farm-fields to analyze the yield variability is conducted. The in-field observations (interactions) through efficient data collection infrastructures (artifact) creates better understanding of the impact of soil in crop yield of farms(construct). Prior knowledge is studied and explored to understand the problem domain and design appropriate solutions. In doing so, the work also generates new knowledge and findings that are usable for further investigations and yield improvement actions. There are four pillars of actions to a design research: problem investigation and implementation evaluation, system design and specifications, designed system assessment and validation, and system implementation and deployment. The proposed work, thus, follows this approach to answer the formulated research questions, as shown in Figure 3.2. Conceptual framework and system architecture of the proposed system are also shown in Figure 3.3 and Figure 3.4, respectively.

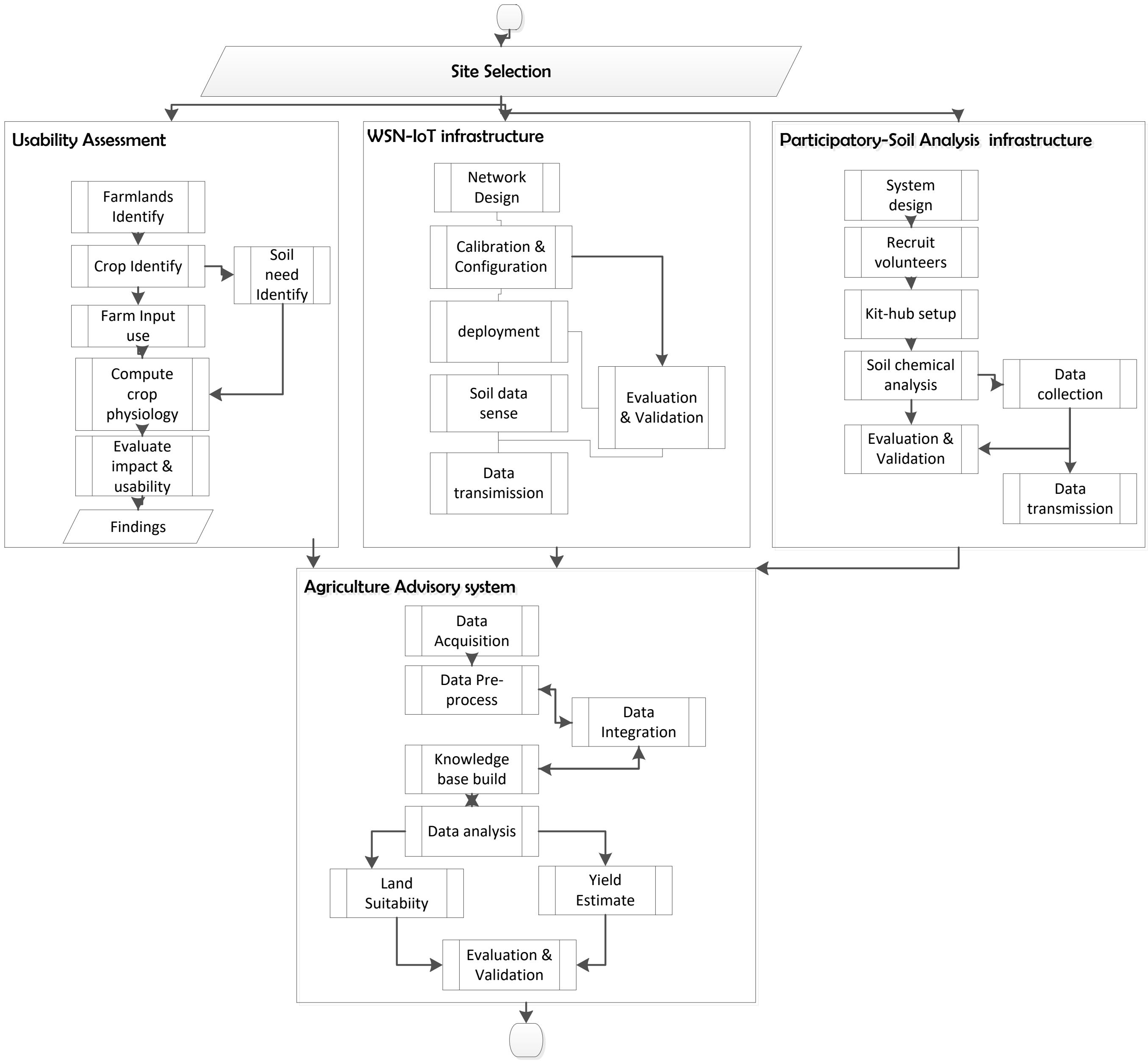


Figure 3.1: Components of the proposed work

3.3 Tools and methods

Both empirical and quantitative research approaches are used to answer the research questions defined in this work and also to evaluate the behaviors of the artifacts designed. Quantitative and descriptive statistical models are used to analyze and describe the inter-relationship between the variables studied and to evaluate existing interventions. A hierarchical linear model and analysis of variance model will be used to analyze such correlations. A clustered sampling approach is used to select sample population in which random sampling will be used for actual sample points. In doing so, four data collection tools are used: reviewing, surveying, crowd-sourcing, and observing.

Review : Literature, previous works and experts are consulted to determine the important yield impact criteria. Moreover, result validation and reliability testing will be conducted using knowledgeable resources. Crop and yield data from local agricultural offices, national and global repositories are also used, as we find

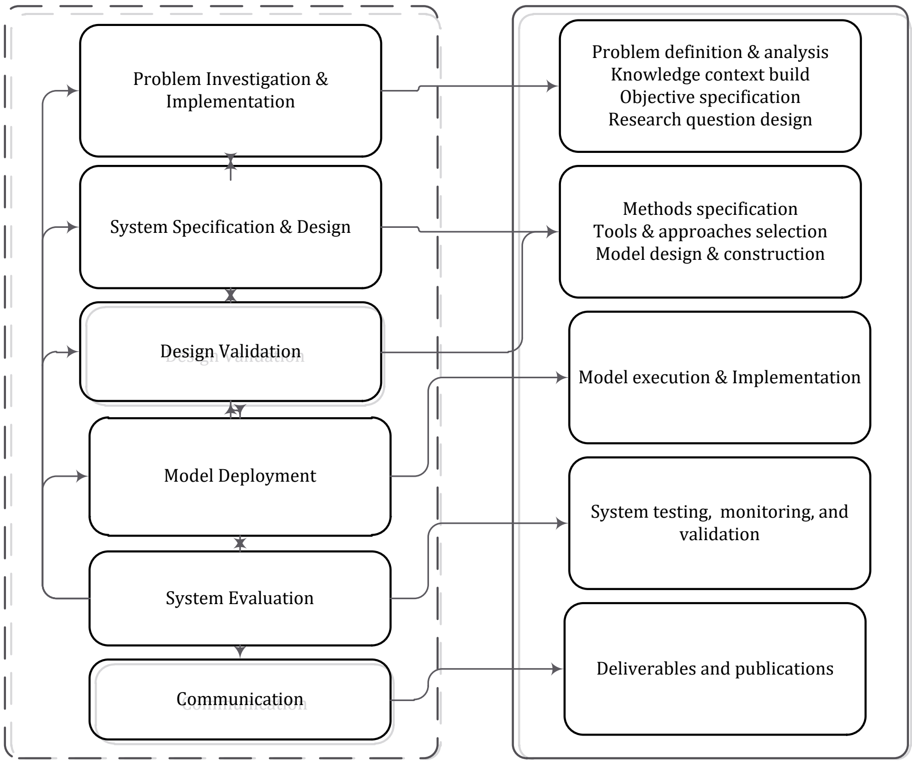


Figure 3.2: Action pillars of the proposed work

them.

Survey : A sample-based survey approach is used to collect data regarding farm management practices, cropping patterns and farm input use. Survey on past years' yield, as well as anticipated yield for the study duration, is also acquired to determine the performance of farmlands and fertilizer applications. Closed questionnaires are thought to be suitable for this purpose as they are less time-consuming, avoiding participants' possible resistance and are also less prone to subjectivity. Moreover, with an exhaustive and clear design, it allows to obtain accurate and concise findings. An informed consent form is prepared and is used before we use this method.

Crowd-sourcing : soil's macro-nutrient data is collected through volunteer citizens and using efficient, mobile and affordable soil data analysis kits. With this approach, data is collected in real-time and continually which is crucial for the temporal analysis of soil properties. It also gives the opportunity for local community to be involved in research works that targets their area and creates awareness about local conditions.

Observatory : WSN technology is chosen for the *in-situ* data acquisition of physical soil properties. Soil moisture and temperature is obtained at various soil depths. This approach is preferred mainly because of the infrastructural constraint of rural areas in obtaining representative farm level and accurate data. It also produces up-to-date information, which helps to monitor the soil, remotely, in real-time and spatially distributed.

Since the aim of this work is to assess and monitor smallholder farmlands, proper identification of permanent crop fields and exploration of dominant crop types growing in the farmlands is critical. We believe that, the Sen2Agri system is suitable for these purposes. It is a free and open source stand-alone system developed in an effort to support agricultural transformations and accurate information productions using remote sensing [126]. It is an initiative from European Space Agency (ESA) and captures and processes data automatically from Sentinel-2 (S2)and Landsat-8 (L8) satellites over an area of interest [127]. The system then generates outputs to serve various functionality, including land cover and crop types identification, which is then used for further analysis like Vegetation Indices (VIs). Detailed technical and research information about Sen2Agri can be accessed from [128, 129]. Monthly composites of cloud-free surface reflectance, cropland masks, crop type maps and periodic vegetation status maps, for the study area, will be obtained through this system. With 10–20 m spatial and 7 days temporal resolutions input data, the Sen2Agri system will allow to closely and timely monitor farmlands throughout the growing season. Thus, for the area of interest this work has targeted, the above-mentioned products will be produced over the duration of the work. In the following sections, we discuss in detail the methodologies used for each objective.

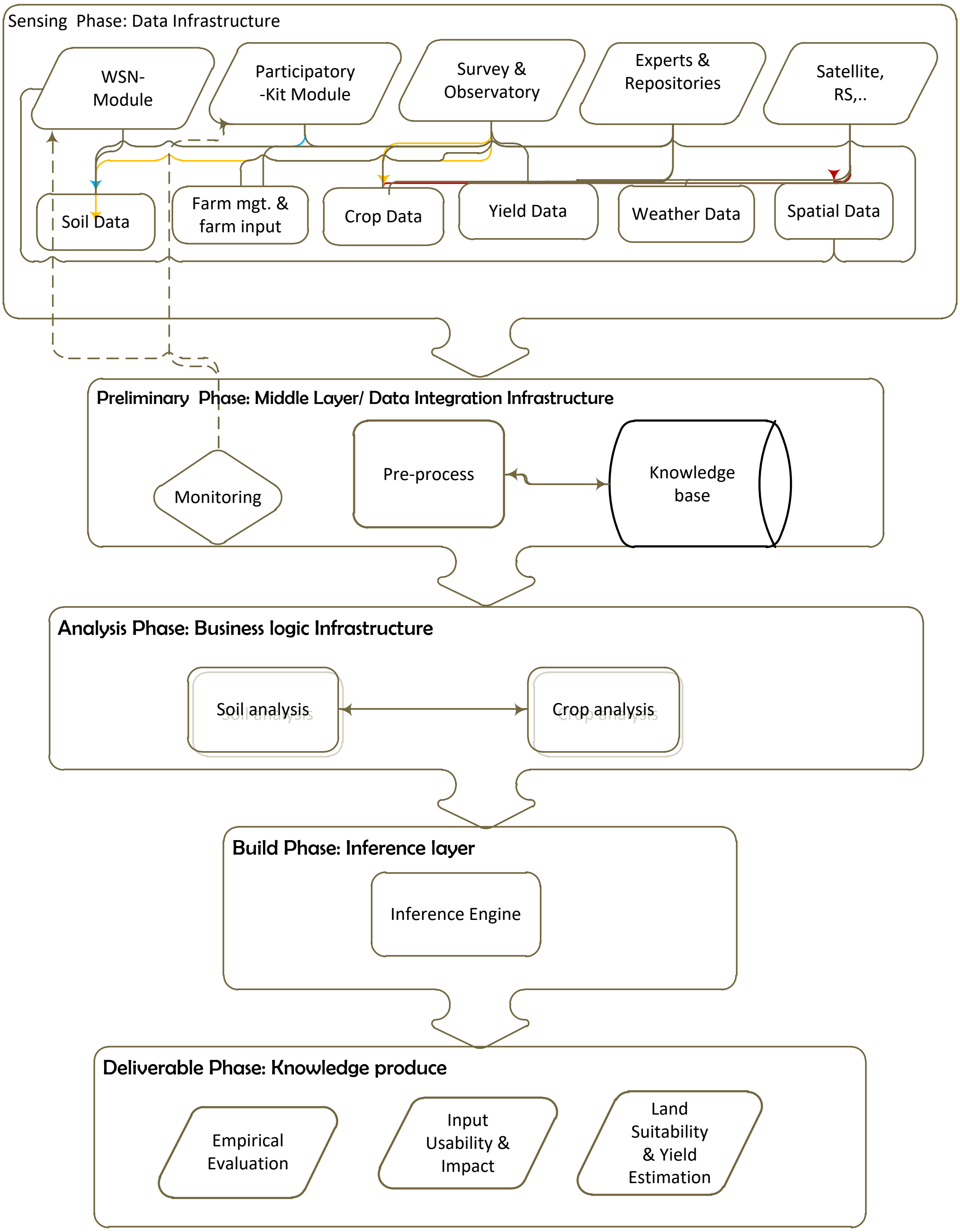


Figure 3.3: Conceptual diagram of the proposed work

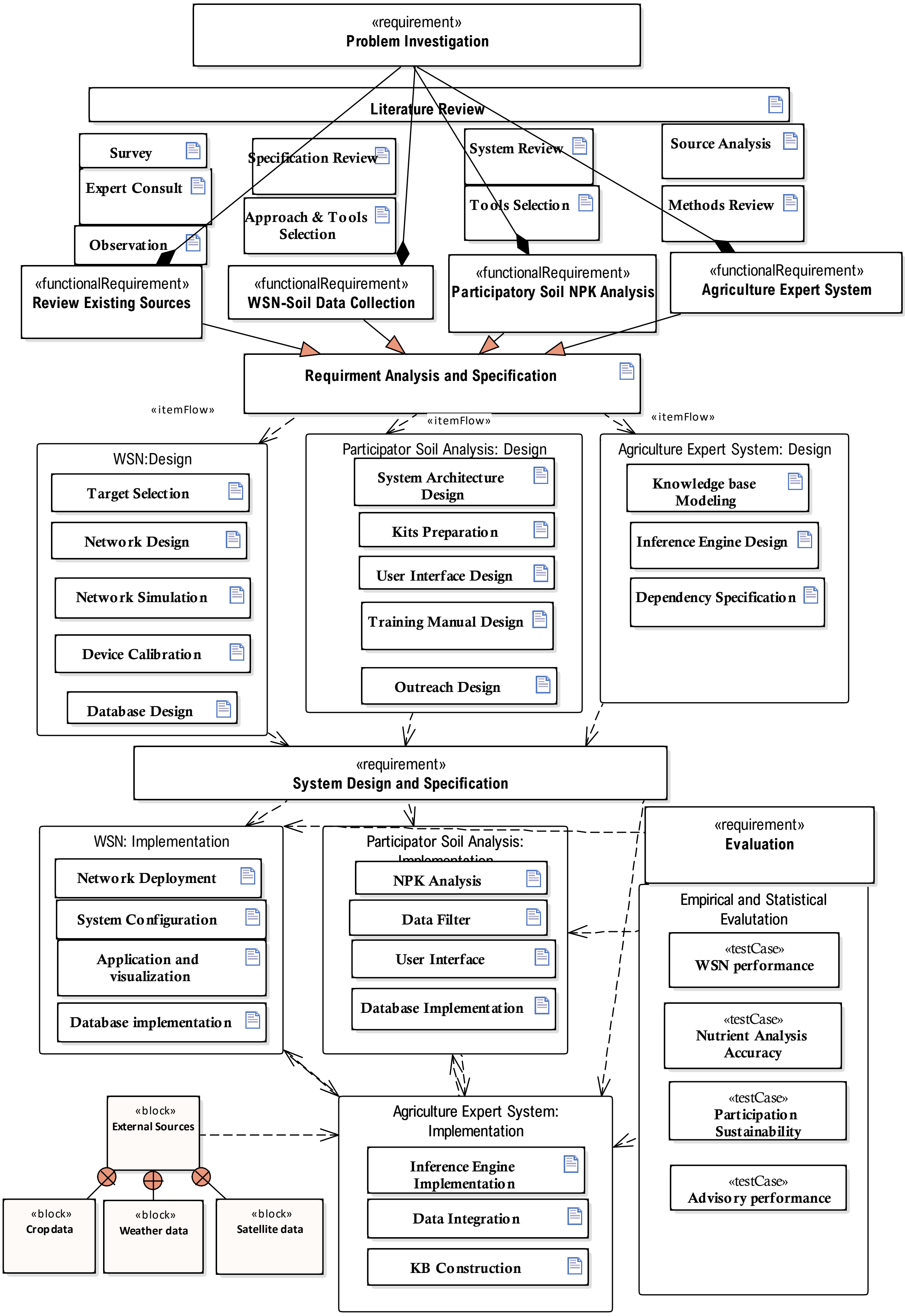


Figure 3.4: System architecture diagram of the proposed work

3.4 Assessment and evaluation of existing soil data sources

3.4.1 Concepts

Soil is a vital input to crop agriculture. It dictates the quality and productivity of a farmland. The soil also determines the investment a farmer needs to put in order to improve productivity and obtain better yield. Improved understanding of the soil is, thus, required to plan, manage and use properly. However, such insight requires abundant and timely data about the soil, which is found scarcely, especially in most developing countries. Even when found, it is often out-dated and at coarse level, either at national or regional level.

For Ethiopia, for instance, the most recent soil data map, EthioSIS, is from 2012 [130]. EthioSIS is the first ever attempt with country level mobilization of efforts and resources to create detailed soil fertility map [131]. This data, comprises physical and chemical soil properties, collected at woreda (district) level offering a generalized representation of soil profile. The soil map is mainly to inform fertilizer blending by a recommendation system, observing the nutrient levels of studied soils. In line with this, a fertilizer recommendation atlas was prepared for selected regions of the country, which informs about soil fertility status and possible nutrient combinations needed to be used as fertilizers. Accordingly, small fertilizer blending enterprises have been established to produce the recommended fertilizers. However, with the significant spatio-temporal variability of soil, the small farm-field details are over-looked. Thus, the more than 13 million smallholder farmers of the country have no or minimal information about their soil.

As part of the EthioSIS task, Agricultural Transformation Agency (ATA) produced soil fertility status and fertilizer recommendation atlas for 134 woredas of the Amhara region in August 2016 [130]. A total of 12,500 composite samples were collected by the team using a randomized and gridded sampling strategies. Surface soil and sub soil depths (0–20 cm and 50 cm respectively) were targeted for the sample collection, with 80 % taken from annual and perennial croplands and 20 % collected from potential croplands. According to [130], sample preparation and physical soil analysis is conducted at regional and national soil laboratories, but nutrient content of the soil was determined at Yara International Soil Testing Laboratory, London. The work reported the use of four prediction and classification algorithms for detailed soil nutrient prediction: Random Forest (RF), Deep Learning Neural Network (DNN), Gradient Boosting Model (GBM) and an Ensemble model. The prediction also used the K-fold, Repeated k-fold and Leave One out Cross Validation algorithms to tune the parameters for the selected classification algorithms and to estimate accuracy of the classifications. The performance of the developed prediction model is also reported to be tested and evaluated, so far with promising results. As a showcase, Table 3.1 presents a summarized fertilizer recommendation information for woredas of the Beshilo basin, collected from the region's atlas. The fertilizer recommendation is subsequently down-scaled to kebele level of each woreda to be used accordingly.

Farming practices are rather diverse between smallholder farms with frequent crop rotations and other intuitive fertility revival methods. Thus, it is important to also consider such practices and crop-specific requirements to recommend any intervention, which is not seen in the aforementioned work. Accordingly, it is imperative to assess the usability and impact of the fertilizer recommendation system and the methods used to produce it. Thus, the EthioSIS soil profile will be reviewed in this work to analyze the spatial and temporal distribution on fine-grain resolutions. Specifically, usability and impact of this data on smallholder farmlands found in selected woredas of the Amhara region is assessed. Hypothesis-driven, review-based research approach is used to explore this source and understand its usability, the gaps, limitations and possible ways of improvements that can be attempted. This data source is given more emphasis because it is the most recent data collected, and it has the better spatial resolution. Below is a discussion of the research approach with the tools and methods to be used.

Table 3.1: Fertilizers recommendation for woredas of Beshilo watershed

Woreda Name	Recommended fertilizer blends
Lay Gayint	NPSB (dominant), NPSZnB, NPSZnBCu, NPSBCu
Tach Gayint	NPSB (dominant), NPSZnB, very small NPSZnBCu and NPSBCu
Simada	NPSB (dominant), NPSZnB, very small NPS and NPSBCu
Meket	NPSB (dominant), NPSZnB
Wadla	NPSB (dominant), NPSZnB
Dawunt Delanta	NPSB (dominant), NPSZnB and NPSBCu
Guba Lafto	NPSB (dominant), NPSZnB, very small NPSZn, NPS and NPSBCu
Mekdela	NPSB (dominant), NPSZnB, very small NPSZnBCu
Tenta	NPSB (dominant), very small NPSZnB and NPSZnBCu
Kuta Ber	NPSB (dominant) and very small NPSZnB
Ambasel	NPSB (dominant), small NPSZnB and very small NPS
Tehulederie	NPSB (dominant), very small NPSZnB, NPS and NPSZn
Dessie Zuria	NPSB (dominant), small NPSZnB and very small NPSZnBCu and NPS
Legambo	NPSB (dominant), NPSZnBCu, NPSZnB and very small NPSBCu
Sayint	NPSB (dominant), small NPSZnB and NPSZnBCu

3.4.2 Operationalization: Approach and workflow

The farm-field soil fertility treatments of selected sites will be evaluated considering the crops planted and the status of the soil observed so far. Empirical evaluation to determine the between-fields as well as between-crops variations on the acceptance of the recommended blended fertilizer application is also conducted. Changes in crop production and yield of the study area is studied to further assess the impact of the blended fertilizer. Using temporal Sentinel-2 imagery products, dominant crop types growing in the selected area and various vegetation indices (VIs) will be computed in support of this activity. In related effort, a survey on dominantly cultivated crop types, fertilizer use and farm yield variations will be conducted. Farmers, local agriculture offices, NGOs and aid organizations working in the area and local and national agricultural data sources are targeted for this purpose. The favorable soil requirement of the identified crops is also reviewed and studied in detail. Finally, the correlation between crop yields of farmlands and the fertilizer application is investigated through hierarchical linear model (HLM) evaluation approach. This approach is selected due to the clustered-random sampling scheme used in this work, as presented in 3.9 and the fact that crop growth depends on multiple parameters. The crop types, landscape and fertilizer use variations are taken as independent variables which affects the yield or the dependent variable. We further investigate the significance of farmlands soil moisture level on the crop yields and fertilizer absorption capabilities of crops. The overall activities used to realize this objective are given as follows.

- Identify land mask and dominant crop types of the study area, possibly over two growing seasons: 2019 and 2020
- Collect *in-situ* data about the crops, yield variability and fertilizer use practices
- Produce seasonal VIs and compute the vegetation condition index (VCI) for selected crop types
- Assess the soil need variability of the selected crops
- Assess responses of farmlands and crops to the blended fertilizer applications using HLM:
- Identify data gaps

The conceptual framework of the proposed approach is presented in Figure 3.5.

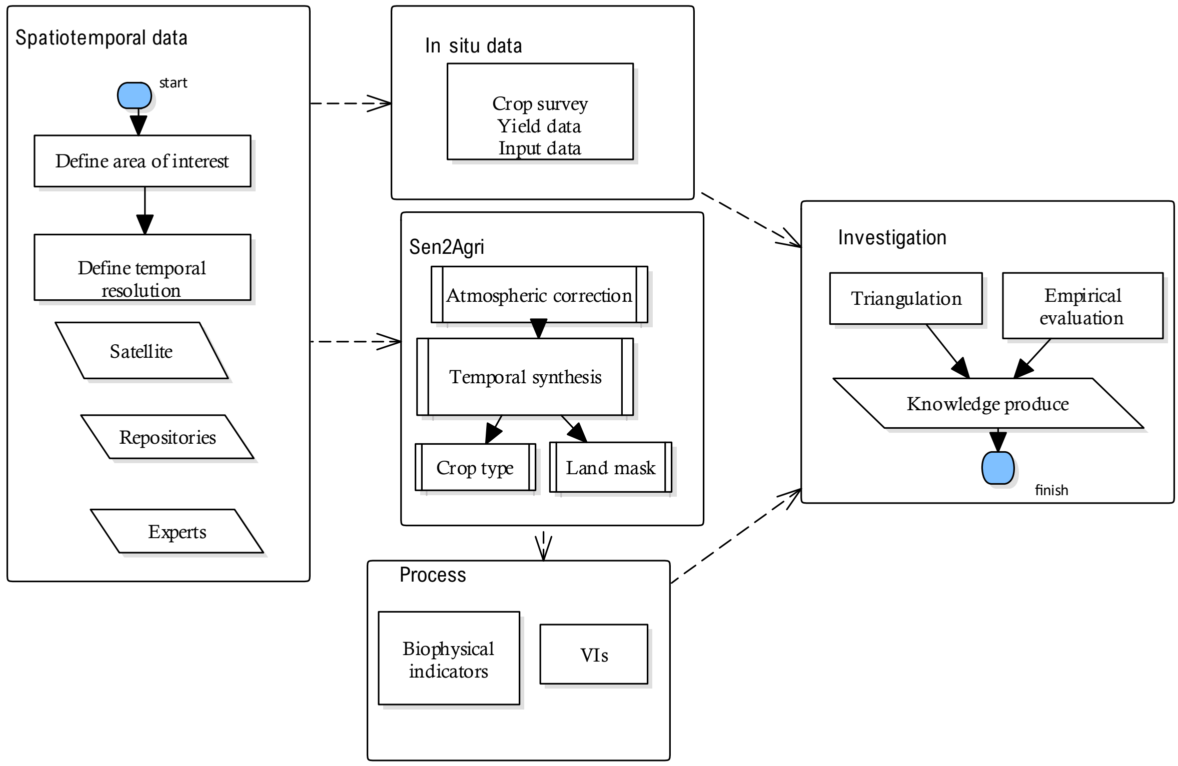


Figure 3.5: Conceptual diagram: assessment on usability of existing soil profile

3.5 Wireless sensors network-based IoT system for soil data collection

3.5.1 Concepts

IoT can be described as a platform where (physical and virtual) objects can be interconnected to generate and exchange relevant and valuable information to make logical decisions. Objects can be humans, hardware, software and data. In recent years, IoT's popularity has escalated mainly due to its considerable capacity to sense a wider environment more precisely and create accurate information that enable us understand our surroundings more. According to a study by [132, 133], more than 80 % of IoT projects address or have the potential to address most of the SDG's [132]. IoT's potential use in agriculture is also ranked among the top four application areas [134] with a wide possibility for the technology to even contribute to food security. Three basic components making IoT are: sensing or perceiving, communication or networking, and data manipulating or application.

Sensing Component : this component consists of sensor or actuator objects that are responsible to perceive or take action in the environment they are deployed in. Objects can be of different internal properties or architectures and can come from different vendors. In a specific instance, few to hundreds and even thousands of objects can be deployed.

Networking Component : this is the component which allows objects to exchange information with each other or to external entities. It defines set of communication protocols to allow information flow and ensure successful data transmission.

Application Component : is the component which defines services or actions to be taken on the data collected by objects. The acquired data can be used in various domains: from simple room monitoring to a complex industry or medical applications. This component also defines and manages the behaviors of actuator objects.

Flexible and layered architecture is required to embrace the heterogeneity and growing number of objects and communication protocols used with the advanced application requirements. The International Telecommunication Unit (ITU) presents a four-layered IoT architecture [135], as shown and summarized in Figure 3.6. Short discussion of each layer is presented below.

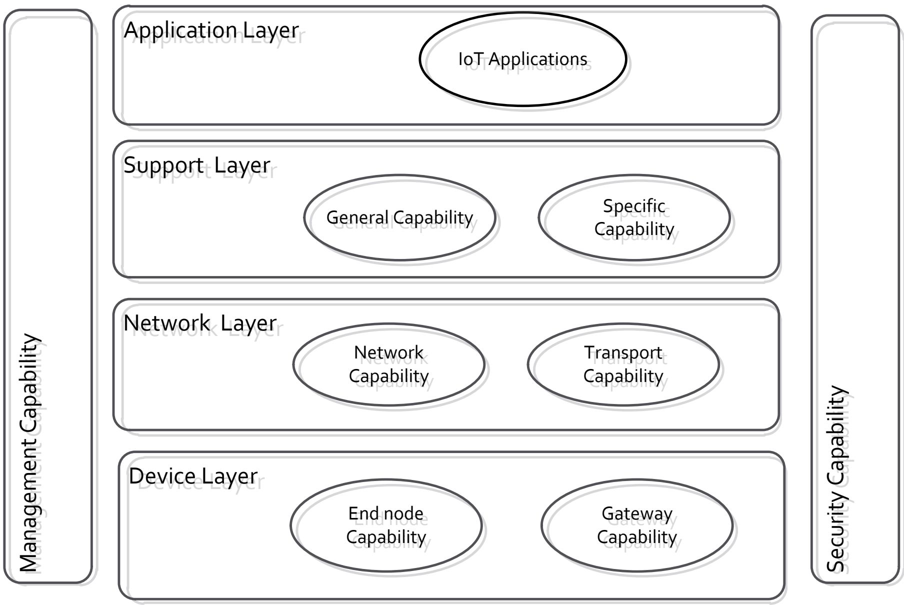


Figure 3.6: General IoT Architecture (Source: [135]

1. Device layer: A device is either an end node, which captures and transmits data locally, or a gateway that buffers end nodes data and transfers it globally. End nodes are capable of:

   - Direct interaction with the communication network: gather, upload and obtain information from the communication network without intermediate component
   - Indirect interaction: contact the communication network through a negotiator and transmit gathered data
   - Ad-hoc interaction: establish communication with other end nodes and devices
   - Sleep-n-wake: power utilization efficiency without compromising data collection and transmission

   Gateway capabilities are:

   - Multiple interface support: Support both wired and wireless communications
   - Protocol conversion: integrate heterogeneous device-level and network-level communication protocols

2. Network Layer: is the backbone of the IoT communication and serves as the brain of the whole system. It uses wired, wireless or satellite communication to interconnect the lower level objects to the upper layer services and applications. The network layer defines and manages the how, when and whereabouts of data acquisition and transfer from nodes to applications. Alternatives of communication technologies and protocols exist that are used by this layer. A brief discussion on the wireless protocols is presented in section 3.5.2. Reliable communication, switching and routing as well as protocol translations are important tasks of the network layer, in addition to Authentication, authorization and accounting (AAA) functions.

3. Service and Application support layer: is the layer that maps low-level objects to higher-level services. It handles data processing and storage requirements of applications and services. It also provides data abstraction service to allow system developers work on generic solution developments.
4. Application and business layer: is the high level component that manages end-user interactions to the IoT system through query processing. Smart city, smart home, precision agriculture are few examples of the application layers. The layer offers the functionality of high-level analysis of the data received from the objects and generates business models, graphs and flow charts.

3.5.2 Wireless communication protocols

One of the significant break-throughs of IoT is the advancement of the wireless communication technology. Radio frequency identifiers (RFID), short-range wireless communications such as Bluetooth and ad-hoc wireless sensor networks are examples of such. Table 3.2 summarizes the wireless technology used in IoT communication networks. These technologies have low operational power and operate for a considerable amount of time. However, most have short-range span and are commonly suitable for applications that address limited spatial coverage such as across buildings or rooms [136]. The Cellular wireless networks are other categories of wireless communication protocols, which can address the wide range limitation. But, as IoT keeps growing and more objects become connected, network traffic management and signal generation becomes a concern for the usability of these protocols, as it is, and requires further design improvements [137]. A rather recent communication technology, which aims to fill in these gaps, is the Low-Power Wide Area Network (LPWAN).

Table 3.2: *IoT enabling wireless communication technology* Source: [135, 138]

Wireless network	WLAN	WMAN	WPAN	WWAN	LPWWAN
Wireless technology	WiFi	WiMAX	LR-WPAN	Cellular	LoRa
Standard	IEEE 802.11	IEEE 802.16	IEEE 802.15.4	2G, 3G, 4G	LoRaWAN
Operating frequency	5–60 GHz	2–6 GHz	868/919 MHz, 2.4 GHz	2.4 GHz, 865 MHz, 2.4 GHz	868/900 MHz
Date rate	1 Mbps–6.75 Gbps	1 Mbps–1 Gbps, 50–100 Mbps	40–250 Kbps, 1–24 Mbps	50 Kbps–1 Gbps	0.3–50 Kbps
Transmission range	20–100 m	$<$ 10 km	8– 20 m	Entire cellular area	$>$ 30 km
Power consumption	High	Medium	Medium-Low	Medium	Very Low
Cost	High	High	Low	Medium	High
Operating life	Years	Hours	up to 2 yrs	Hours	10–20 yrs

LPWAN is a collection of various low-power, long-range wireless communication protocols with typical area coverage up to 100 Km, using gateways as a relay [139]. The downside of such technology is its limited data rate and thus suitable only for services with infrequent and small data exchange requirements, and is consequently expected to connect a wide domain of IoT applications. Parking, agriculture and environment monitoring are some typical applications benefiting from LPWAN. The technology is either proprietary or open standard and runs on both licensed and unlicensed radio frequency (RF) spectrum. Some of the well-known LPWAN technologies are:

SigFox Is a forerunner in the LPWAN technology and released in 2009. SigFox is a proprietary platform and runs only in specific regions where the network operator exists. It uses the ultra-narrow-band (UNB) channel with a distance coverage of up to 50 Km and transmits data over the unlicensed spectrum with frequency of 915 MHz in US and 868 MHz in Europe [140]. It has weak bi-directional communication and suitable only for one way communications. SigFox can be an appropriate choice for applications with small (12 bytes) and infrequent data transmission requirements [141].

Weightless Is a true open standard narrow-band LPWAN protocol in the Sub-GHz unlicensed spectrum with possible support to licensed spectrum as well. It has an adaptive data rate to maximize battery life but a limited distance coverage of less than three and five Kms in urban and rural environments, respectively [142].

Long Range Wide Area Network (LoRaWAN) Is a technology with wider spatial coverage and operates only in the unlicensed spectrum at various carrier bandwidth. It defines different implementation designs and supports bi-directional communication through which nodes can receive and send data, as required [143, 144].

Symphony Link Is a standard with guaranteed message receipt. It supports dynamic data rate and has no communication frequency limits, enabling nodes to transfer data whenever required. Another advantage of Symphony link is its fair and balanced link budget for all nodes in the network, regardless of their distance from a gateway,thus, suitable for large area coverage communication [144].

Table 3.3 presents a summary of the technology.

Table 3.3: LPWAN Communication Protocols

Name	*LoRaWAN*	*SigFox*	*Weightless*	*Symphony Link*
Source	Alliance	Proprietary	Alliance	Alliance
Node	LoRa	SigFox		LoRa
Security	AES, MCI	AES, CMAC	AES	AES
Frequency	433/868/915 MHz	868/962 MHz	400 MHz–1 GHz	150 MHz–1 GHz
Transfer rate	300 bps–50 Kbps	Upload: < 300 Kbps, Download: 8 bpd	100 bps–10 Mbps	Adaptive
Packet size	User defined	12 bpd	10–20 bpm	256 bpd
Range	Rural:15–50 Km, Urban: 2–5 Km	Rural: 30–50 Km, Urban: 3–10 Km	up to 10 Km	Flexible
Modulation	CSS	BPSK	DBSPK	FH, AF
Status	Deployment	Deployment	Introduction	Introduction
Duty cycle	1 %	1 %	1 %	NA

3.5.3 Wireless sensor network in agriculture

Different implications of IoT exist for developed and developing countries. Therefore, it is important to develop an appropriate implementation strategy for each context and according to what exists on the ground. This includes connectivity technology, cost implications, technical opportunities or limitations, cultural contexts and the intended impacts. For remote and rural areas, where infrastructure exists but it is commonly of poor quality, IoT deployment is rather difficult. More cost-effective, robust and power-efficient solutions such as Wireless Sensor Network (WSN) are of importance and needs to be integrated into the IoT [145]. Essentially, the LPWAN communications contribute to improving both IoT device connections to the Internet and the overall efficiency of the IoT application operation in such resource-constrained environments.

A WSN is an IoT architecture for a wireless network of spatially scattered nodes with embedded sensors for closely monitoring physical and environmental conditions [146, 147]. It supports up to hundreds or even thousands of nodes in a single setup which can communicate among themselves or the gateway, through wireless communication. The gateway serves as a sink for nodes-transmitted data and transfers the data to a network server. A typical node in a WSN is shown in Figure 3.7. As can be seen, a node has a radio transceiver with an internal antenna for communication, a micro-controller board which interfaces with sensors, and a power unit which is mostly a battery. a node can also have extra elements like power generator, mobilizer and a location finder to add external power source, location knowledge and mobility feature [148].

micro-controller Is the component with sensing and processing units. The sensing unit in turn consists of an analog signal sensor and analog-to-digital converter (ADC), for analog-to-digital or vice versa conversion. The processing unit comprises a processor and storage element for processing and storing data and program.

Transceiver Connects the node to a network through a radio frequency (RF) and internal antenna. It converts data to radio waves for the data to successfully propagate and reach to the destination. The radio spans mostly short-range of about 100 meters and low data-rate, about 10–100 Kbps. Better range is obtained through additional external antenna.

power Unit Is the energy source to the node and ensures prolonging network lifetime at the cost of lower throughput. Rechargeable batteries, capacitors and solar energy are common sources of power. In a typical outdoor setup, end node operates on batteries while a gateway node requires a direct supply.

A WSN node can be in one of these three modes during the life-cycle of the network: sleep, active and idle. Sleep is when the node is doing nothing and saving power; active, when it is engaged in data transmission (uplink) and idle is when the node is receiving data from others (downlink) [149].

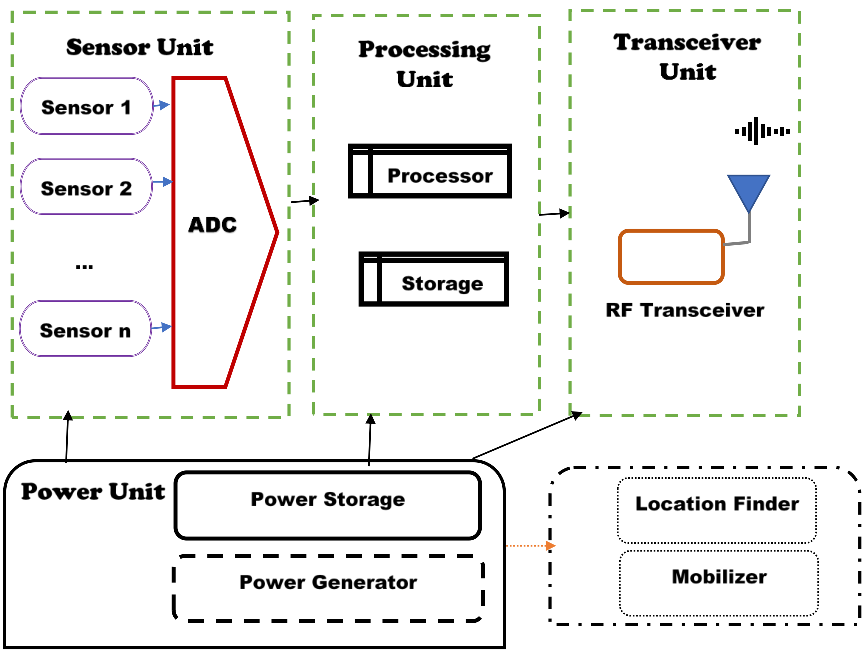


Figure 3.7: Typical end node in WSN

The WSN is receiving substantial attention in agricultural research works indicating the usability of the technology for such requirements including crop management, soil monitoring, irrigation management and precision agriculture, in general []. However, large-scale implementation of the WSN for outdoor investigations are very limited, which pinpoints the need for both empirical and engineering research works in this aspect. Based on our findings, most works are experimental, either in a controlled environment such as gardens or greenhouses and with few number of sensor nodes implemented through simulations. On the other hand, LPWAN technology promises a lot to offer, which are all to the advantage of large-scale data acquisition and processing schemes, implying more is needed to be done to fill the gap. The fact that most of the works targeted developed countries also adds to the extended work required on the usability of such technology to create fair and equal access to all. Only few works exist on the use of the WSN technology in the developing countries, particularly none for SSA, despite the fact these countries are constrained with agriculture-related data and have also poor infrastructure for such data acquisition. It is to our strong belief that the LPWAN technology can play a significant role to fill this infrastructure gap and contribute towards improved livelihoods of a society where land degradation and food insecurity is prevalent. Accordingly, this work, to best of our knowledge, is the first attempt to set up a large-scale WSN, using LoRaWAN protocol, to monitor soil properties at farm fields in remote areas in a developing country. A WSN is advantageous for such cases and fits well to rural areas as it has minimal infrastructure requirement such as power supply and Internet connection and also less human involvement.

3.5.4 Operationalization: Approach and workflow

A WSN-assisted IoT architecture is proposed for farm-field soil data collection to assess the spatial and temporal variability of soil. In particular, this work aims to deploy a LoRaWAN-based WSN for soil moisture and temperature monitoring. A general overview and workflow of the proposed work are shown in Figure 3.8 and Figure 3.9 as follows.

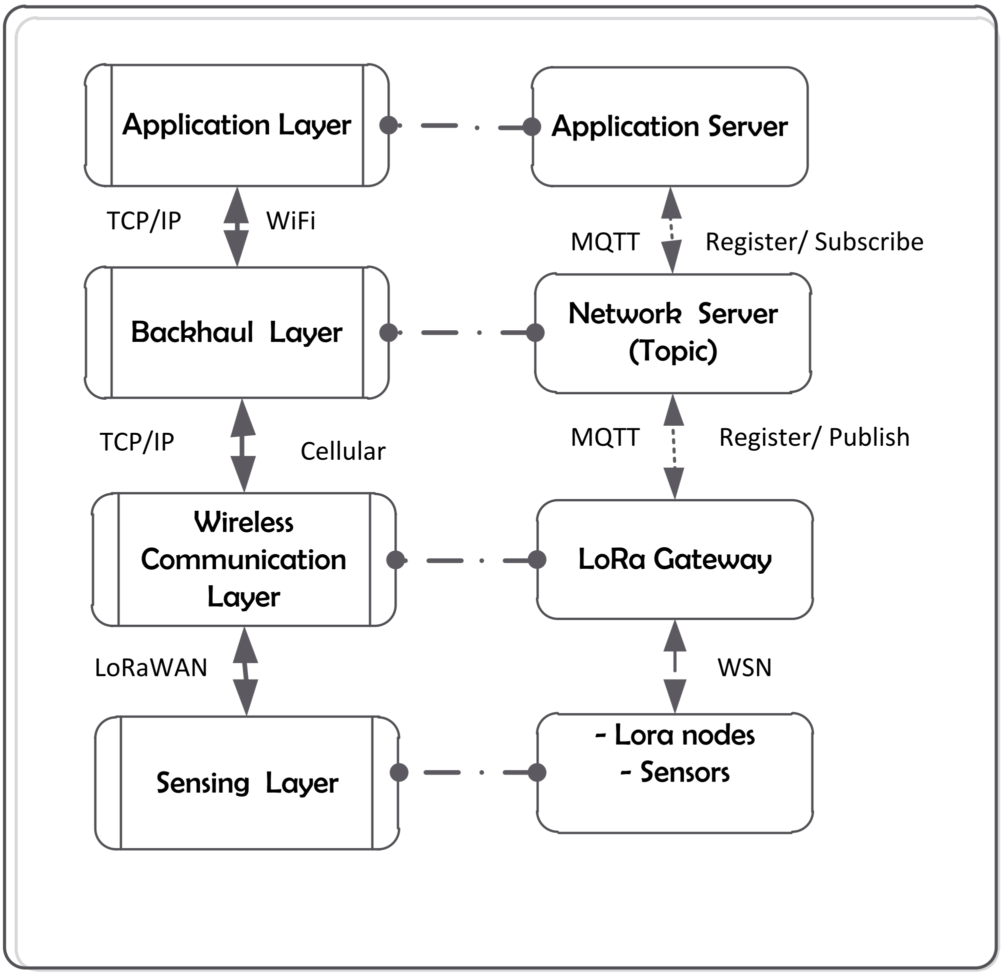


Figure 3.8: System architecture: WSN-based IoT soil monitoring system

As can be seen from Figure 3.8, the proposed system has four sub-components: Sensing layer, wireless communication layer, back-haul layer, and application layer.

1. Sensing Layer: LoRa nodes with off-the-shelf soil moisture and soil temperature sensors are setup across farmlands. The sensors are placed at depth variations of

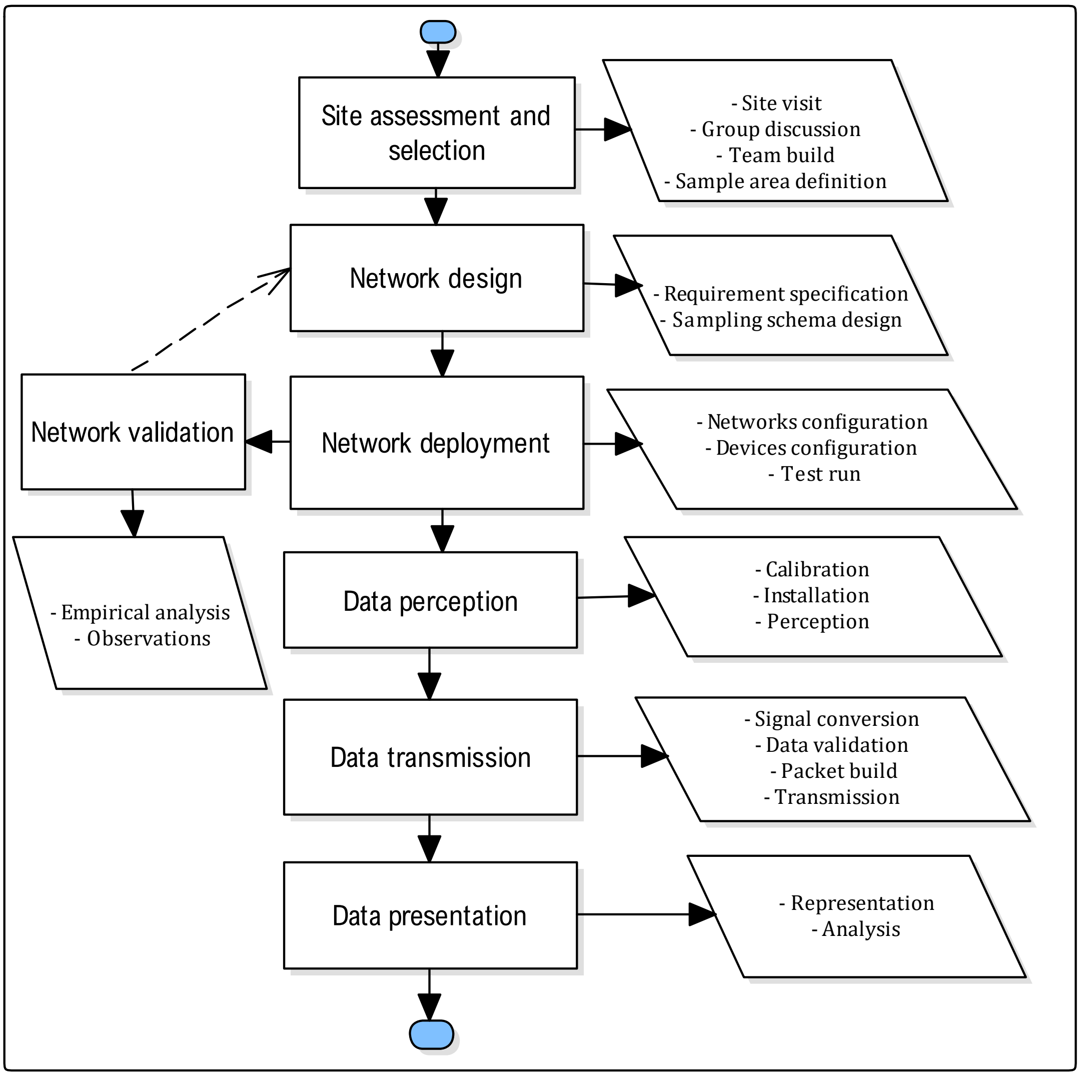


Figure 3.9: Workflow diagram: WSN-based IoT soil monitoring system

20–80 cm below the ground surface. The micro-controller of the node converts the sensor data to digital form, controls the frequency of readings by the sensor and time-stamp for transmission to the gateways. The location of each node is captured through a separate GPS device during installation. This enables proper identification of the spatial-temporal dynamics of the soil. To validate and cross check the quality of sensor data, laboratory soil tests will be conducted at some intervals. As soil exhibits variation in space and time, a well-calibrated sensor that smooths such variability and obtains accurate measurements of water content will be utilized. Sensors will be calibrated with site-specific soil samples, validate their measurements with laboratory obtained findings and perform proper adjustments prior to deployment.

Volumetric water content (VWC) and water potential (WP) are the two important factors to describe the moisture level of a soil. VWC measures the amount of water found in some volume of soil and can be useful in fertigation management, soil health monitoring and water balance studies [150]. WP, on the other hand, indicates the energy state of the water in the soil and can be of importance for monitoring water movement, precipitation level or water availability for plantation in the soil [150]. In this work, Decagon 5 TM sensors will be used to measure the VWC of a soil and determine its moisture level. The sensor also has a thermistor in thermal contact with the sensor tips to read the soil temperature. These sensors use dielectric permittivity measurement, through electromagnetic field computation, to determine the VWC of a soil. A 70 MHz oscillating wave is supplied to the sensor tips that charges based on the dielectric amount in the soil. The stored charge is taken as proportional to soil dielectric permittivity value and generated as output. Further details about the Decagon 5TM sensor can be found on the manual [151]. The raw dielectric permitivity is then converted to VWC using the Topp equation [151, 152], which is specified as:

$$VWC = 4.3 \times 10^{-6}\epsilon_a^3 - 5.5 \times 10^{-4}\epsilon_a^2 + 2.92 \times 10^{-2}\epsilon_a - 5.3 \times 10^{-2}$$

where ϵ_a is the dielectric permittivity of the soil and is between 2–7 [151]

2. Wireless communication (WSN) Layer: LoRaWAN communication protocol is used to establish connection between nodes. Nodes are configured for efficient power use but with limited down link rate. Such implementation might not be suitable for applications with interactive and continuous data flow requirements. However, since the aim of this work is to acquire soil data from ground over time, no significant communication is needed top-down. Moreover, nodes will be configured in sleep-more wake-up-less mode where node-gateway communication is initiated once a day for efficient power utilization. This is also due to the fact that soil properties do not significantly vary over a short period of time and continual frequent monitoring of a soil results in better understanding of the soil over time than any single measurement. Thus, the adopted approach fits well to the task requirements of this work.
3. Back-haul communication: The back-haul layer manages the Internet communication between the LoRa gateway and the network server for sensed data transmission. Since the area of interest for this work is a remote rural area with no or minimal communication infrastructure, we will explore possible transmission options to successfully transmit the sensed data to a network server. Cellular or satellite communication link will be considered emphasizing cost-effective and reliable connectivity over the TCP/IP network protocol. The network server also handles network functionality including routing of data between nodes and application, schedules communication sessions and manages gateway status [153]. The Message Queue Telemetry Transfer (MQTT) communication protocol is used over the TCP/IP network protocol to achieve such functionality. The MQTT facilitates gateway-application communication through a publish/subscribe method. The things network (TTN) network server will be considered for the MQTT back-end host.
4. Application Layer: manages application definitions and data usability functionality. The business model for the soil monitoring system including data visualization, pre-processing and analysis are all defined in this layer. A temporal data of over two years will be collected, requiring a time-series persistent data management and monitoring capability. Accordingly, the PostgreSQL data management system is identified appropriate to use.

The general architecture of the proposed system is presented in Figure 3.10. The overall communication of the network is broken into four distinct levels:

- Sensor-to-node: wired communication between sensors and LoRa nodes to acquire ground data.
- Node-to-gateway: LoRaWAN wireless communication to exchange information once a day over the entire monitoring duration. The gateway uses external omnidirectional antenna and establish a reliable communication over a wide area.
- Gateway-to-network server: defines the interaction between the gateway and the cloud in IoT architecture. This communication enables integration of the WSN with the Internet and makes the data acquired through the WSN accessible to the world. Wireless TCP/IP protocol is used to establish this communication.
- network server-to-application server: this communication also uses the TCP/IP communication protocol over the wireless connection. The network server transfers gateway uploaded data to respective applications through appropriate application technology. MQTT is used to provide this interaction. MQTT is a topic-based publish/subscribe open protocol that facilitates communication between applications and nodes [154]. Soil moisture and temperature readings

are defined as topics in which communication is established on. The gateway then publishes data in the topics, to be retrieved by the application interface, through subscription. The MQTT broker runs on the network server and manages smooth communication [155].

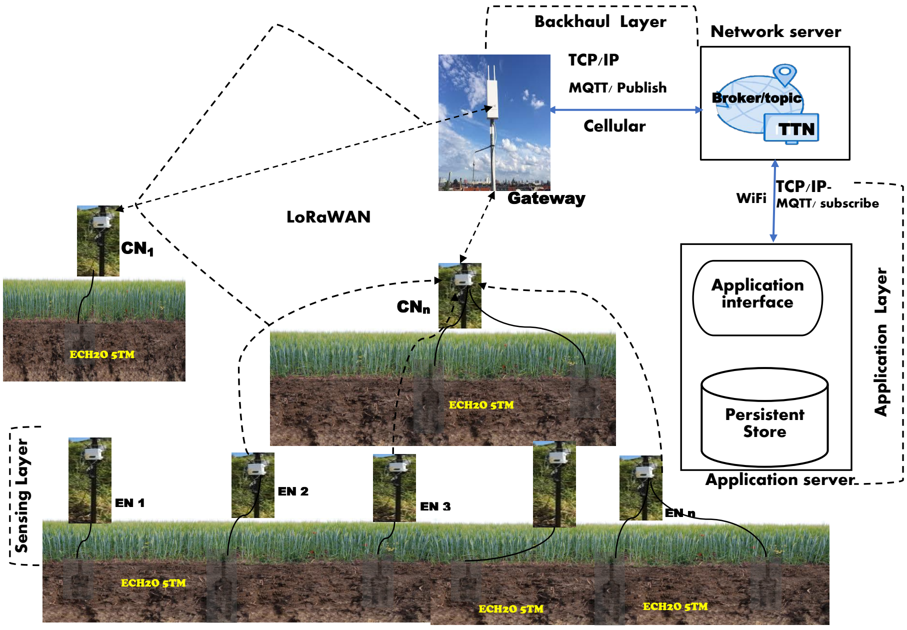


Figure 3.10: Conceptual diagram: WSN-based IoT soil monitoring system

3.5.5 Summary

WSN has an immense opportunity to acquire accurate and real-time *in-situ* data. Most developing countries have significant financial and communication limitations to deploy technological interventions and monitor crop productions at farm-field level. This work is an attempt to fill this gap by providing a long-range, power-efficient, outdoor and (near) real-time soil monitoring system with reasonable cost and minimal maintenance requirements. WSN based on the LoRaWAN protocol is chosen as suitable to obtain information that answers our research hypothesis. The system aims to cover smallholder farms in rural Ethiopia. In-door and experimental applications and simulations of a WSN have shown to be successful in most aspects. IoT has also proved to be robust for applications in urban areas where uninterrupted power and communication supply exists. However, deployment of the network in a harsh environment like agricultural fields might be a challenge, unless careful design and implementation strategy is applied. Environmental and technical parameters that affects efficient communication will be assessed to determine optimal node and gateway distributions with minimal data loss rate or noise interruptions. Detailed design considerations and nodes deployment scheme is further discussed in section 3.9. In addition to soil data, the work also aims to assess power consumption, transmission delay, signal strength, link quality ,fault identification and tolerance of the deployed network. As such, empirical analysis on the performance of the network will be conducted. Experimental assessment of the spatial, temporal and energy characteristics of the network will help to evaluate outdoor wireless sensor network implementations and improved designs.

3.6 Participatory soil macro-nutrient analysis

3.6.1 Concepts

Soil is one of the most significant attribute for better farm productivity and, thus, food security. Soil tests are commonly carried out in laboratories far from farm fields, where farmers are unable to frequently travel. The cost of these tests are also unaffordable by most farmers, holding them back from checking the health and nutrient level of their field. Local agricultural extension workers also do not have this data unless a country-wide study is conducted, and this is infrequent. As a result, up-to-date soil data is uncommon, especially in developing countries. On the other hand, most of Africa's soil is reported to be old and infertile, hence, less productive [156]. Sustainable soil management is thus crucial to keep pace for better yield and reduced land degradation. Sustainable soil management, however, shall be based on various soil properties which significantly vary in space and time; regular maintenance and timely communication of such variations to stakeholders of the agri-chain is required for change to take place. Doing so is resource-intensive and requires substantial and recurrent investments which can often not be afforded.

Alternative and cost-effective methods are required to fill this gap. Remote sensing and satellite imagery system is one option in this context [156]. Alternatively, mobile soil testing kits can be more appropriate for smallholder farmers at remote farm locations. In particular, self-operated and easy-to-use soil test kits can be of more importance. This can help farmers to observe and better understand their soil, regularly monitor their farms and improve farm management practice or take appropriate improvement actions. Such knowledge also helps farmers to regulate the input amounts they use on the farm, which in turn assists in improved crop yield and resource efficiency. It is also evident that crops will be more nutrient-balanced.

The usability of such soil testing kits, however, is limited and mostly used by farmers in developed countries. This is due to low awareness and limited distribution of such tools. The low literacy rate of smallholder farmers in most developing countries also contributes to the limited use of these kits, as some training and careful interpretation of results is required. If soil test kits are combined with easy to use, localized and efficient analysis platforms, farmers and agriculture extension workers can easily conduct the analysis at their pace and fulfill the information requirements regularly. Building the capacity and knowledge of the local communities on soil and environment management practices by creating an engaging environment is one way to fill the information gap and increase the usability of mobile test kits. Citizen science is recommended as one option for a timely collection of soil data [156].

Citizen science or public participation in scientific investigations is not a new concept. It has been increasingly advocated as a means for scientists to address large-scale data limitations [157–159]. Citizens voluntarily serve as sensors and provide information about surroundings, trends, past events and more, which can be further investigated by researchers to produce usable information. Particularly, environmental monitoring and observation can benefit highly from citizen science as it can offer real-time, *in-situ* data about a situation. Thus, it is vital to establish collaboration, participation and communication between scientists and the community. In recent years, the participation and involvement of citizens has shown tremendous growth due to technological proliferation: social media, interactive web interfaces, smart sensors and scientific measurement tools and smart phones. Citizens have become more environmentally aware and big data is generated that can be used to add scientific value. A number of research projects have also showcased the impact of citizen science on such environmental monitoring [160]. When the data collected by citizens involves spatial data such as location, it is termed as geographical citizen science [161]. This voluntary participation has been given various classifications based on the service and involvement of volunteers [162]. Active or passive classification describe citizen science as intentional or an unconscious participation of citizens, respectively. Another classification is that of implicit versus explicit, which defines the direct and indirect geo-spatial data provision by participants, respectively.

No common framework or tools exist for citizen science research approach as every work has its own specification and needs for mode of participation [159]. However, voluntary and sustained engagement of citizens is what every citizen science aspires in the process of the scientific investigation. It is natural for participants to be subjective or less accurate during measurements, unless concise procedures are defined and proper guidance and follow-up is used. Thus, it is also required to apply a reliable methodology that is easy to use for participants, yet, at the same time, generates comparable information with standard measurements.

According to [163], three common approaches exist to citizen science projects: Contributory, Collaborative and Co-created.

Contributory approach is the most commonly practiced approach in which citizens collect data using precisely defined protocols by the researchers and with limited participation in the data analysis.

Collaborative approach is when citizens are also involved in re-structuring the research design, data analysis and result dissemination to others. Participants are involved actively in some research activity such as analyzing samples, interpreting data and inferring conclusions. Communicating with the community can also be a task that citizens may be involved in. They may also assist in design and refinement of data collection protocols.

Co-creation approach is a bottom-up approach in which participants actively engage in the research design and are involved in most parts of the scientific investigation life cycle. This approach is also known as partnership approach.

The difference between these approaches is summarized in Table 3.4, where Y stands for full public participation and y for partial public participation in the step. It is often the case that a hybrid of these approaches is implemented. However, which

Table 3.4: Models for public participation in scientific research: (Source: [163])

Scientific process step	Contributory project	Collaborative project	Co-creation project
Problem definition			Y
Requirement analysis			Y
Hypothesis formulation			Y
Data collection methodology design			Y
Data collection	Y	Y	Y
Data/sample analysis	y	Y	Y
Data interpretation and synthesis		y	Y
Dissemination of findings	y	y	Y
Execution of actions based on findings			Y
Discussion of results			Y

approach to adopt depends on the requirements and the spatial temporal extent of the work. For instance, a contributory citizen science approach is more suitable when a large volume of data covering wide geographical locations is collected. On the other hand, if repeated measurements of a specific target of smaller domain is priority, a co-creation approach is more appropriate [164].

3.6.2 Operationalization: Approach and workflow

Prior research has presented the micro- and macro-nutrient deficiencies of soils in the Amhara National Regional State (ANRS) and recommended re-vitalizing actions such as application of fertilizers [165, 166]. This, however, needs to be supported by frequent diagnostic analysis as the presence of nutrients varies temporally. Research works involving the community for such continual and spatially distributed data collection process is thus required. Local and regional-level agriculture and natural resource management offices, farm extension workers, national and international aid organizations, research institutes and the farmer community, as a whole, needs to come together and work hand-in-hand to improve the status of the soil. Easy-to-use, affordable and robust tools can be used to bring such collaborations for soil data collection and analysis.

This work proposes a collaborative and co-created digital citizen participatory approach to collect and analyze soil macro-nutrient data at farm field. The work also follows an active and implicit geographical citizen science approach in which participants capture and transmit geographical data of samples taken. Strong collaboration needs to be surfaced for successful realization of this objective sustainably. Accordingly, possible local, regional and national collaborations are explored and identified before the actual work commences. Figure 3.11 indicates the possible collaborations lines of stakeholders.

Figure 3.11: Potential stakeholders and role identified for participatory data collection

The workflow of the system is categorized in two phases: the preliminary phase and the operational phase.

In the preliminary phase, campaigns and outreach activities are conducted to create awareness on the research and of the need for active involvement of the aforementioned stakeholders. The aim is also to explore and refine the stakeholder list and their possible roles. Volunteer facilitators from agricultural extension offices, aid organizations and higher academic institutes are recruited and first round training and discussion will be conducted. Knowledge exchange and methodological refinement may also happen at this stage. A second round training is offered, together with trained facilitators, to recruit volunteer citizens for participation. High school students are targeted as volunteer participants for various reasons including knowledge build and environment awareness. The trainings are designed to be interactive and explanatory, covering issues on safety procedures, sample selection and soil macro-nutrient analysis using off-the-shelf test kits. Schools and extension offices are then equipped with lab kits to be distributed to participants during field visits. A lab kit consists of NPK test kits, safety procedures and equipment, data collection and analysis guidelines and required field equipment. Volunteer participants then work closely with the researchers and facilitators and design convenient field visit time to conduct the investigations. The spatio-temporal execution of the investigation needs to be designed carefully considering the pre-cropping, growing and harvest seasons. The frequency of field visit will also be agreed upon and decided through discussion during this phase.

In the operational phase, participants visit pre-identified sample areas, record the soil type and texture, collect soil samples, prepare soil composites and conduct the NPK analysis, following the developed procedures and guidelines. Each sample taken will be location and time-stamped to ensure the reliability and also record the temporal variability. Once the soil analysis findings have been obtained, participants translate the results to a spatial representation using a digital camera on a digital device used, and upload the data. The system acknowledges every recipient and allows participants to submit a maximum of three such data elements per sample point. An offline standard of data transmission platform will be designed and integrated in the client machine, onto which participants can upload content frequently. The uploaded data will then be synced with a central repository when an Internet connection is available. Doing so avoids the connection barrier commonly found in rural areas while creating an opportunity to acquire data regularly. The system rewards participants based on their participation, which enhances participants' motivation while also creating chances to recruit more volunteers, hence ensuring sustainability of participants. The workflow of the system is shown in figure 3.12.

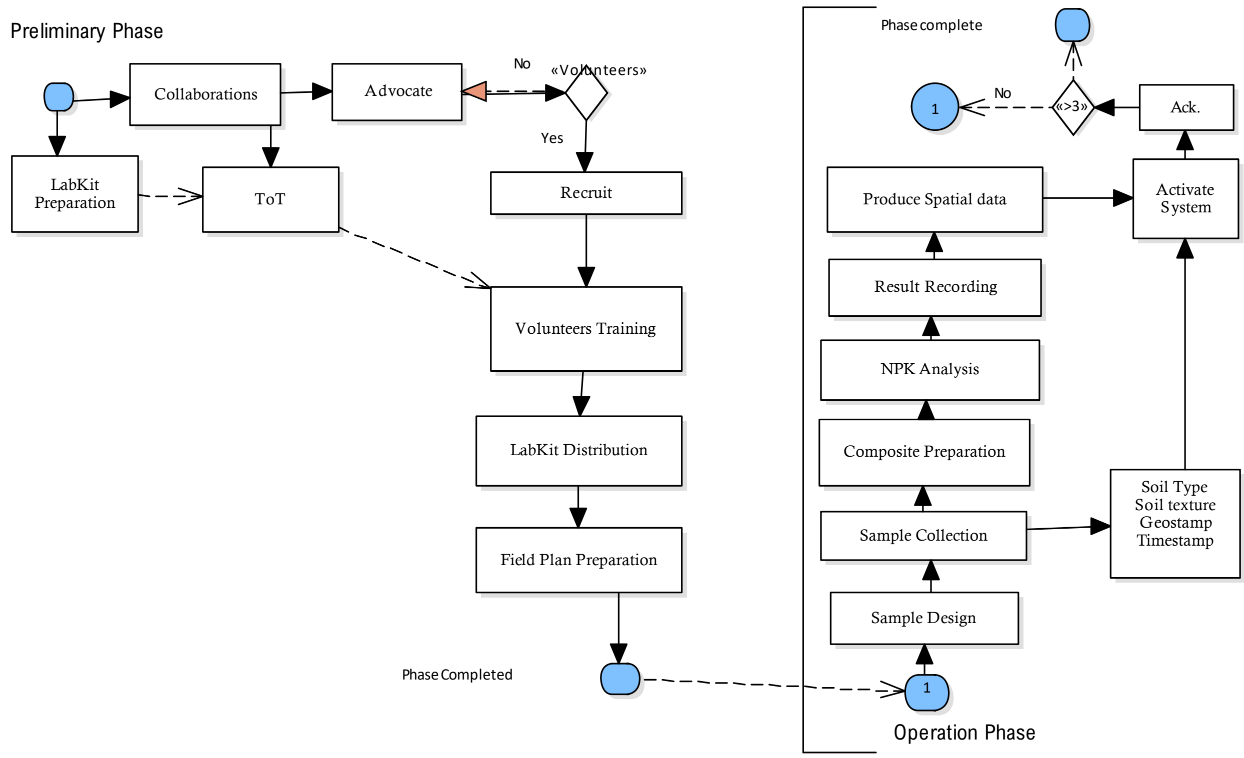


Figure 3.12: Workflow diagram for participatory soil data collection

Several work exist that utilize a web platform for participatory data collection and transmission purposes. The open data kit *ODK* platform [167], Sapelli Platform [168], and FPIC [169] are worth mentioning in this context.

From a software engineering perspective, a system of four tiers is designed with a layered architecture of model view controller *MVC* and persistent data management approach as shown in figure 3.13. The deployment follows a client-server architecture of thin clients, with views deployed on mobile devices, while other layers are maintained at a central server. The system implementation is done using appropriate and freely available open source tools.

Model Layer holds all business definitions and data management component of the system. This layer controls user management, data classification and data preprocessing tasks, in addition to responding to user needs.

Controller Layer is the service layer of the system and acts as an interface between the model and end users. In addition, this layer holds a temporary buffering component that responds to user requests at times when a connection cannot be established with the model. The buffer is a mere replicate of the repository of the model layer which is synced once communication is established. The model pushes data to the buffer frequently to ensure reliability.

View Layer provides the user interface as interaction platform for users. It presents the form in which data is entered by users, once proper authentication is passed.

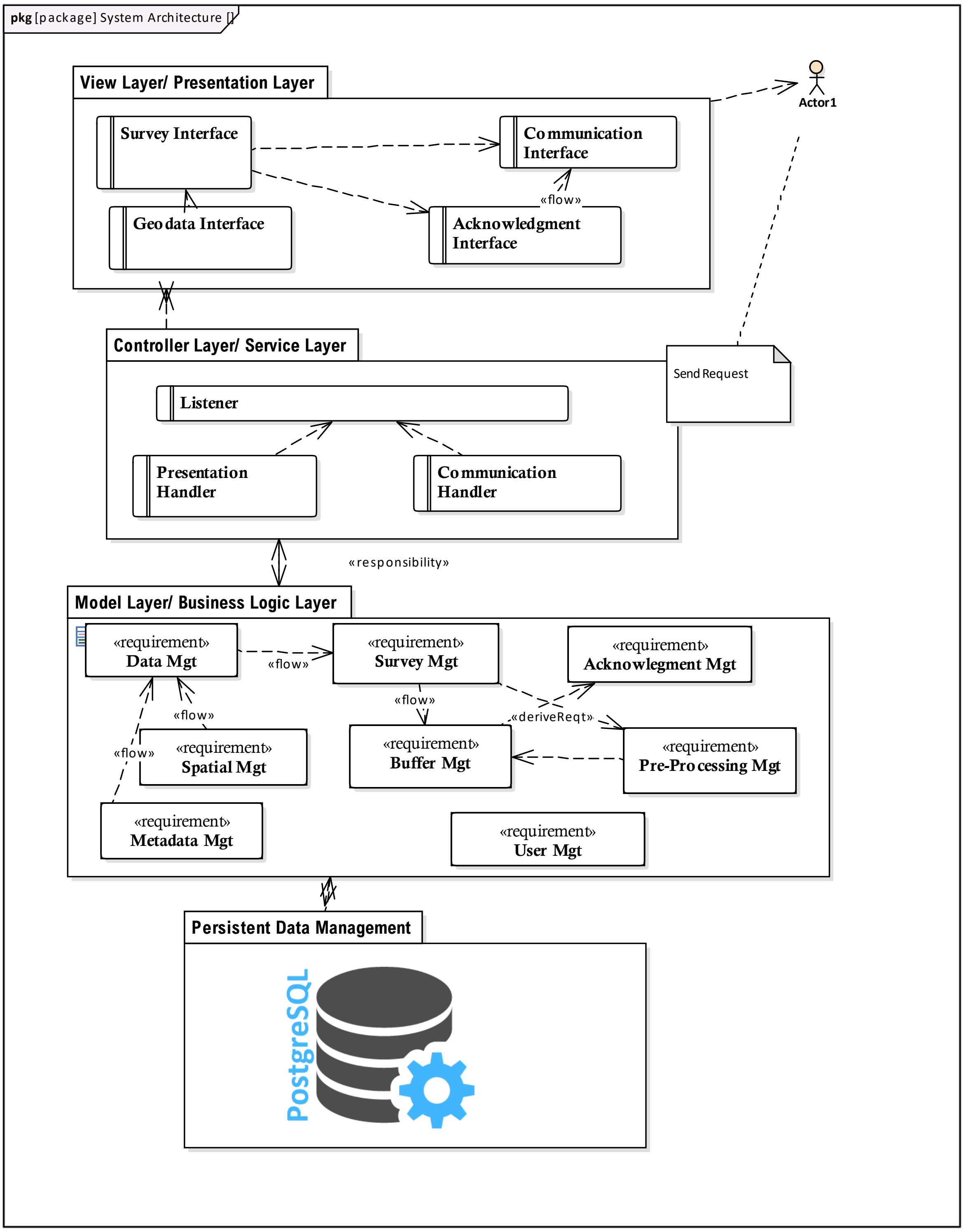


Figure 3.13: System architecture diagram for participatory soil data collection

It also displays results returned by the model layer. The user interface is designed so that minimal knowledge of interaction is needed and it supports the local language of participants. Most part of the survey displayed is populated automatically by the system, reducing the data entry requirements from participants.

3.6.3 Tools and methods

An iterative object-oriented (IOO) system development approach is used throughout the life cycle of the system. The system construction is broken down into phases: requirement analysis and specification, system design, system implementation, system deployment and testing, and system maintenance. Tasks are executed iteratively going back and forth between phases, as needed. As part of the requirement analysis focal group discussions, interviews and closed question surveys are conducted and expert knowledge is acquired. On the design phase, tools needed for soil NPK testing, safety procedures, samples, database and other sub-systems are designed. For the implementation, Python programming language is used with the Django framework to support the MVC architecture. The persistent data management is implemented using PostgreSQL database system for its flexibility and strong support of spatial data and queries. In each phase, deliverable are produced using appropriate IOO tools.

3.6.4 Summary

The involvement of citizens in continuously monitoring their environment has significant impact to create better understanding and knowledge among members. Soil macro-nutrient analysis often requires expensive lab works and takes considerable time. A system that integrates soil test kits with an appropriate participatory platform can be an alternative for feasible and regular soil monitoring and analysis. In addition to the continual and real-time data flow, this approach also assists in technology transfer to the community. It also is a tool for 'democratization of science and technology' [170]. With such a system, citizen knowledge and understanding of the farm soils can be improved with possible cost-effective diagnostic actions to take effect. The direct involvement of the community also has impact on acceptance and usability of further research in the area. However, active and sustained volunteer participation of citizens is crucial for such system to be successful. Strong collaboration among stakeholders is also required to run such system consistently.

3.7 Agricultural advisory system

3.7.1 Concepts

Soil is the basis for crop growth: a crop's nutrient absorption ability is determined by the soil's water retention capacity. The soil texture also affects air and water movement for crops to use and much more. It is thus important to understand which soil is suitable for which crop type. The inherent properties of a land also needs to be aligned with its intended purpose, which helps to achieve better productivity while keeping the land healthy. One of the main goals of agriculture production is to get maximum yield with minimal operational cost or efficient use and applications of such resources. Smallholders invest on their farm and wait until harvest time to know how much it pays off. This hinders the farmers from exploring market and income possibilities for high or low productions. There is also significant yield gap most smallholders are faced with. Land suitability analysis can help to match crops with appropriate soil types. It can be used to understand the inter-dependency that exist along the crop production chain: spatial references, crops, weather, soil, farm practices and other inputs. Estimation of crop yields made before harvest is also essential to inform farmers on required mitigation actions and empower stakeholders with information for crop management and other logical business decisions. A properly designed knowledge base that integrates all these information is required for better advisory and decision support in the catena. Such systems are particularly significant in smallholder farmlands which are inaccessible or where advisory professionals are scarce, both in quality and quantity. However, the efficiency of such a system heavily relies on the availability and quality of data sources. In developing countries, however, data scarcity is typically severe while timely yield estimations and crop-land match is significantly required. Accordingly, there is a need where all possible data sources are explored exhaustively. A holistic approach that integrates primary and secondary data sources, experts and local knowledge is thus needed.

Data mining and machine learning algorithms have been making a significant contribution in this context. Machine learning (ML) systems use the concept of learning from experience and train a machine to get an optimal estimate of a final outcome. Artificial Neural Network (ANN) is a machine learning algorithm that represents and processes concepts based on the human brain's neural representation [171]. It models input-output interactions and correlations as interconnected neurons of the human brain and mimics their functionality. Such relationships are defined through the analysis of a large number of input and output scenarios, which are then used to predict future instances. ANN supports one or more layers with activation function, connected in such a way that every connection signals new message and thus easily identify correlations between inputs to generate unseen outputs. The connected objects are referred to as nodes while the connections are known as synapses [97]. Known values are presented as input nodes, which are linked to unknown and inferred hidden nodes. Each link has a weight coefficient, which is generated to minimize the error value between desired output and known value of the model using back-propagation algorithms. There can be one or more such connections depending on the data hierarchies and level of precision required. An ANN with two or more layers between the input and output layers is referred as Deep Neural Netowrk (DNN) [] Data sets from different and heterogeneous sources can be fed to an ANN model at different levels and new information can be inferred precisely. ANNs are capable of processing, classifying and estimating a situation, if adequate and accurate data is provided. Their nonlinearity, inferring unknown knowledge, adaptivity, fault tolerance, generalization and input-output mapping also renders them preferable to such multi-variable dependent operations [171].

3.7.2 Operationalization: Approach and workflow

In this work, land suitability analysis and crop yield estimation for smallholder farmlands are studied as use cases using ANNs algorithm by integrating outputs of previous objectives. Interactions between the spatial and temporal data layers produced are modeled to determine the most likely factors that could limit crop growth as well as the level of suitability of the soil to selected crops. Two crops growing in the study area are targeted for study based on their socio-economic significance. The obtained information can then be used to make comprehensible recommendations regarding farm management and other input specifications. ANNs are recently getting more attention in research works of classification and accurate predictions [172, 173]. We propose to develop a model for yield forecasting and land suitability based on temporal agricultural data, soil type, texture, moisture, temperature and NPK levels, farm input and management practices and weather parameters and using appropriate ANN.

Land suitability is a method of assessing possible uses of land for various purposes [174]. Land suitability analysis, in our context, is a measurement and evaluation of the conduciveness of a farmland, in terms of the soil's moisture, temperature, texture, type and NPK level, for a specific crop to grow. Knowledge of these parameters can assist in management of water resources, identify and select appropriate crop type and adjust fertilization inputs, and subsequently maximize crop yield. In particular, it is the aim of this work to analyze the interplay between soil moisture and macronutrient dynamics with plant physiology. Soil is classified based on the aforementioned chemical and physical properties and associated with crop requirements to predict its suitability. The crop type, its root depth, water requirement and growing duration are data required to know a crop. The soil data is acquired through 3.5 and 3.6 while remote sensing, experts and literature are consulted to identify the specific-crop requirements. The Sen2Agri system is used to retrieve and analyze satellite images to identify crop types and vegetation index products. This will give an insight in the most dominantly cultivated crops and based on the findings, crops to be considered for the study are selected. The *in-situ* data produced from this work is then integrated with all the mentioned data for knowledge base construction and to monitor the impacts of soil properties on the growing progress of crops over the cropping season. These attributes are used as inputs to the ANN model, which will be trained with 85% of acquired data while the remaining 15% is used for validation. The final outcome of this model, which is a nominal classification of suitability level, will be used as input to the yield estimation model. Other parameters to be considered for the yield estimation model include weather parameters, elevation, farmland size, input use and other farm management practices. The weather data will be obtained from local, national and international weather stations. Experts and farmers are surveyed to further refine the target crop types of the study. The economic significance and productivity level of the crops is considered in this respect. Based on interactions observed and the crop status data obtained, in-season optimal crop-yield estimation will be formulated through the model. Formulating the yield estimation and land suitability context into classification problems, ANN technique can be used to produce accurate findings [175, 176]. The outcome of such system can be used to assist on possible informed and logical decision makings. Figure 3.14 presents the workflow of this objective.

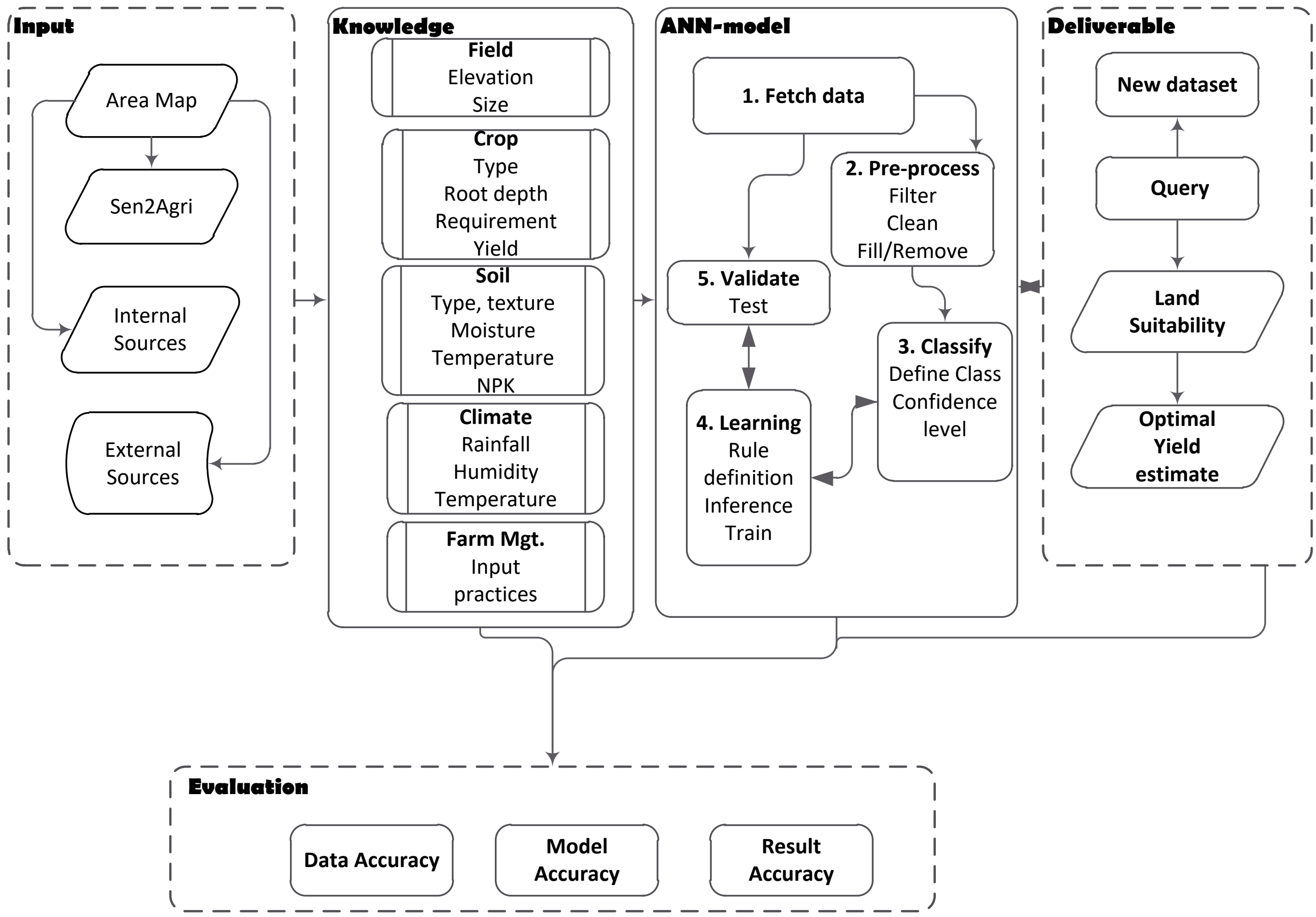


Figure 3.14: Workflow diagram: Agircultural expert system

3.8 Study area

The Beshilo basin is found in Amhara National Regional State (ANRS), located between 38.2 deg E–39.6 deg E longitudes and 10.8 deg N–11.9 deg N latitudes. The basin is defined around Beshilo river, one of the largest tributaries to the Blue Nile River. It runs over 16 woredas from three zones in the region: North Wollo, South Wollo and South Gondar [177]. The altitude of the basin ranges from 1170 masl to 4260 masl and covers an area of around 13,000 km^2 [178]. Figure 3.15 shows the topography of basin. With an annual average rainfall range of 825–1470 mm, an annual temperature range from 13 °C to 30 °C (maximum) and -10 °C to -15 °C (minimum) and annual potential evapo-transpiration (PET) range of 1060–1920 mm, the basin holds four common agro-climatic zones (AEZs), as seen in Figure 3.15. The Dega AEZ encompasses the highlands with elevation range of 2000–2800 masl, cool and humid climate, average annual rainfall from 1200–2200 mm and average annual temperature of 12–16 °C. The Weina-Dega AEZ describes the midlands at around 1500–2000 masl elevation, cool and

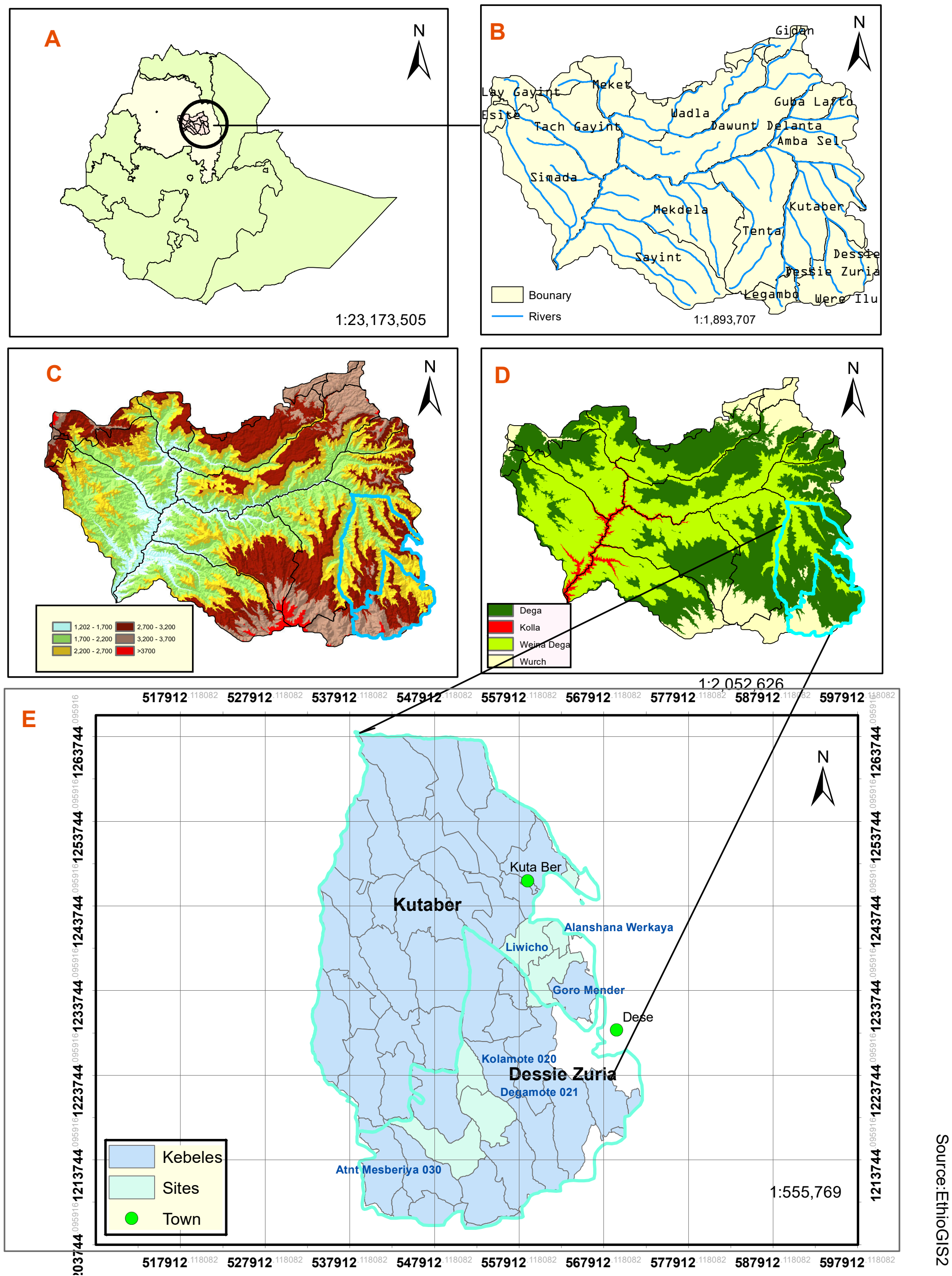


Figure 3.15: Beshilo basin and selected woredas

sub-humid climate with average annual rainfall from 800 to 1200 mm and average annual temperature of 16–20 °C. Kolla AEZ is a lowland area below 1500 masl elevation, warm and semi-arid climate, average annual rainfall of 200 to 800 mm and average annual temperature of 20–27.5 °C. Wurch AEZ covers highlands beyond 2800 m elevation with cold and moist climate, 2200 mm average annual rainfall and ≤ 11.5 °C average annual temperature [179] Most of the woredas in the basin are densely populated, with more rural settlements and agriculture-dominated livelihood. Woredas, especially of the South Wollo zone, are known to be least food-secure and many are short of producers to even support themselves, all year long. The slightly small farmland owned by farmers and the unreliability of the long-rainy season (Meher) attributes to these poor farm yield. Most farmers of the region own on average 0.82 ha of land and practice farming during the short (Belg) season [180]. This zone is also known for sometimes with rigorous drought and often referred to as the "heart of the country's famine belt" [181]. Irregular and erratic rain fall, soil degradation and nutrient depletion, small farmlands and crop pest manifestation have all hindered the productivity of farmlands of the zone [182]. As a result, most households of the area have been receiving food aids and other assistance from government and non-governmental aid organizations [182].

As case study area for this work, purposeful sample selection is used and two woredas of South Wollo zone have been selected from the basin. Soil moisture varies in spatially and temporally. Soil texture, topography, weather factors, vegetation covers and farm management practices all affect this variability. These parameters also vary among different AEZs while they tend to be relatively similar within. Thus, the AEZ label can be used as deterministic of these variant properties and represent all well [183]. Kutaber and Dessie Zuria Woredas have been chosen mainly because of the representative AEZ distribution they have, the severity of the problems mentioned above and the relevance to address them and accessibility of the locations. The woredas are known with bimodal rainfall pattern and two cropping seasons. The AEZ distribution of the woredas is 63% Dega, 22% Woina Dega and 14% Wurch. A summarized description of the two woredas is presented in Table 3.5. The cereals produced

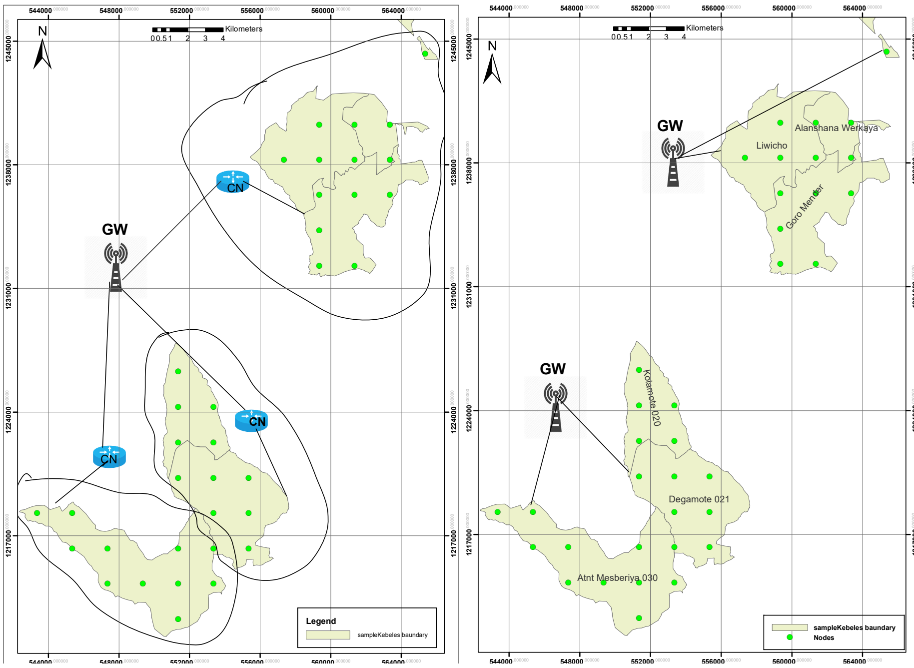
Table 3.5: Kutaber and Dessie Zuria woredas summary

Summary	*Dessie Zuria*	*Kutaber*
Area	990 km^2	1216 km^2
Elevation (m)	2300–3500	1970–3100
Total population	156,679	201,433
Rural population density	125per km^2	164per km^2
Major crop types	Teff, Pulses, Barley, Sorghum, Wheat	source(FEWSNet)
Major soil types	Leptosol, Cambisol, Vertisol,	source(EthioSIS)
Soil texture	Sandy-loam	source(EthioSIS)
Average annual rainfall	1150 mm	1083 mm
Mean annual temperature	10.5 °C	14 °C

are mostly for own consumption while Khat cultivation is practiced to some extent as cash crop [184]. Of all the area, about 37% (595 km^2) is used for farming agriculture, which is permanently cultivated and mostly rain-fed.

3.9 Sampling design and network layout

Both technical and environmental dimensions need to be assessed before establishing an out-door WSN. Devices are resource-constrained, topography and other human-induced factors can hinder better performance, various interferences can compromise quality of data. These and other elements need to be considered. In this project, the aim is to cover representative agricultural fields in the watershed and obtain soil data all along. A single-hop star topology is the simplest arrangement of nodes in WSN deployment where each end node directly connects to a central gateway. Such design is suitable when few nodes exist and are distributed in a small-scale coverage, less than 1 km. A multi-hop topology is much preferred if a wider area coverage is required and several nodes are available, with considerable maintenance cost. In this work, a two-tier star topology is proposed and nodes are arranged in a cluster format. Nodes which cannot directly connect the gateway (GW) are referred as end nodes (ENs) while nodes in close proximity to the GW are referred as controller nodes (CN). ENs are clustered and connect to CN which in turn connects to the GW as shown in Figure 3.16. Such an arrangement not only reduces signal loss over long-distance propagation, but also enables modular node management and maintenance, efficient power use and minimal communication between gateway and end nodes.



(a) Setup 1 (b) Setup 2

Figure 3.16: Sampling design alternatives

The design and sampling layout procedure is classified into tasks that are described below:

Zonation Two clusters are initially defined, one per woreda of the study area: Dessie cluster and Kutaber cluster. Three smaller clusters are then defined per outer cluster, representing the three AEZs of the woredas. Accordingly, Zone 1 (Dega), Zone 2 (Weina Dega), and Zone 3 (Wurch) are set in both Dessie cluster and Kutaber cluster. We refer the reader to Figure **??** for the AEZ classes of the study area. With in clustered strata, 36 sample fields are identified using 2 km by 2 km grid sampling approach. These fields will be used as sample populations for all activities of this project and nodes are placed randomly within each field.

Role Definition Nodes are designated different roles and responsibilities as EN, CN and GW. EN collects and transfers data either to GW or CN. A CN receive other nodes' data and transmit it to GW, in addition to the normal sensing task. Zones have at least one EN and zero or more CNs, based on the distance from the GW. The GW node relays the received data to the cloud.

Distribution design With average 0.82 ha plot per household of the total of 595 km^2 crop land of the strata, an estimate of 72,561 farmlands are found in the two woredas. And for the selected kebeles, there are about 6,585 farmlands. And from the above percentage ratio of each AEZ, a rough estimate distribution of these locations can be made as follows: Zone 1 (63%): 4149, Zone 2 (22%): 1,449 and Zone 3 (14%): 922. This implies more nodes go to Zone 1, specifically, 16 nodes while 6 and 3 nodes are deployed in Zone 2 and Zone 3, respectively. A geometric random graph model proposed in [185] will be used to determine the maximum range for successful connection establishment between ENs and a GW. According to this paper, given a finite area A, average nodal degree E > 0 and nodes density n the connectivity radius is defined as:

$$r(n) = \sqrt{A \cdot (\lg n + \epsilon \cdot \lg n)/(\pi \cdot n)}$$

For the GW, two modes of implementation are designed as shown in Figure 3.16: (1) a GW is placed at a position equidistant to all zones and nodes directly connected to it. (2) a GW is placed at the center of each zone and connect to nodes only in that region. In both cases, avoidance of obstruction of direct line of sight to and from ENs or CNs and availability of direct power source is also considered for the GW. It is thus required to place the gateway at higher elevation, roughly 5 m above the ground. Since the interest of this work is in soil data of farmlands, sensor distribution will be done only on farmlands. Nodes are placed 1 m above the ground, avoiding unsuitable locations such as fallow land and field boundaries. CNs are set in a zone only if the distance between EN and GW exceeds r(n). Depth of sensor deployment is determined once the crop requirements are identified.

Network Deployment Nodes and GWs are placed according to the design and checked for their proper integration. The network is assumed to be static with all nodes placed in a specific location and remain fixed in that position throughout the application. However, with the empirical investigation, the location and elevation of nodes can be reviewed and improved subsequently. The deployment follows an incremental approach where one zone is setup first and its performance is monitored. Considering the empirical investigation, the next zone will be set, and so on.

Evaluation and Testing Both black box and white box testing is conducted on the network to evaluate the proper functioning of hardware installation and the system's robustness and accuracy of the application designed. Empirical evaluation is conducted on the overall performance of the design.

From a technical point of view, computational parameters such as range, throughput and link-budget are considered for optimal implementation. Adjustment to the network setup is carried out based on the received signal strength level (RSSI), and power consumption. The adjustment will be both in azimuth and elevation of the gateway position. While doing so, due consideration is given to the below.

Scalability Currently, we have available 30 nodes and 3 gateways for deployment. The network is designed following an incremental approach and starts with a single setup, but nodes and gateways can easily be added to or removed from the network with no significant impact. The number of CNs is minimized to the extent possible which reduced coupling. Ideally, the LoRa gateway is said to support thousands of nodes. This functionality is evaluated in the empirical investigation of this work as well.

Transmission Rate Continuous data reads from sensor nodes puts a high power demand and significantly affects the lifetime of the network. On the other hand, soil moisture and temperature do not significantly vary over short time intervals, and this can be used as an opportunity to limit the data transmission rate. Thus, minimal communication is aimed for but without compromising the amount of data obtained. Hence, ENs collect data every hour but transmit to the GW only twice a day and they hibernate for the rest of the time. This specification, however, will be reviewed during the experimental evaluation phase so as to achieve optimal transmission rate. An external memory with better access times may be added to provide secondary storage and to alleviate the application size constraints imposed by on-chip memory capacity of the nodes.

Autonomy : As the WSN is deployed in remote farm fields and far from the researchers base, there is little expertise to frequently visit and maintain the network. As such, the network shall operate autonomously and be non-intrusive to farmers. To this end, and also for safety of the network, cables are cut as short as possible and proper casing and covering is performed during installation. Moreover, embedded software carries an intrinsic risk of failure away from 'home' and with proper identification of hardware, is used to remotely monitor the network and also to easily identify faults.

The participatory soil analysis work is also conducted in the sampling area, taking random points with-in. Kit hubs are set at selected high schools found in the study area. Five level-one and two level-two high schools are found here. The level-two high schools are used as kit hubs as they encompass our target audiences: grades 9–12. Each kit hub is equipped with a soil test kit, which is a package of guidelines, bottles, caplets, color charts, tubes, chemicals, shovel and a tablet for the result uploading. The intended data collection and analysis is conducted once a month on a practical session of a course identified as close to this project activity. Dessie University is a nearby higher education institute found in the study area, which runs agriculture study, among other programs. Students and faculty of this department are identified as possible collaborators to the mentioned field activity through training and facilitating the kit use. In addition, agriculture development agents (DA) of local offices also accompany the team in every field and data collection visit. The observation is planned to be conducted over two years for a total of 16 field visits per school. The kit hubs need to be re-filled with required equipment, as needed.

3.10 Data sources

The main purpose of this research is to establish a data infrastructure for soil data collection and study the correlation between soil properties, the farm management and farm productivity. Use cases of agricultural decision support systems are also produced to showcase the significance of such analysis and the data acquisition system. In this regard, additional datasets, both primary and secondary, are required. Moreover, the performance and quality of the data acquisition system needs to be evaluated, requiring base datasets. The table below presents a summary of the required datasets and the possible sources.

Table 3.6: External data and possible sources

Dataset	Description	Data type	Sources
Crop data	Dominant crop types, crops characteristics, socio-economic significance	Primary/secondary	Central Statistics Agency (CSA), FAO, Literature and Experts, Sentinel-2
Weather data	Climate data of the area: Temperature, rainfall, Precipitation, humidity, evapo-transpiration	Primary/secondary	Ethiopian Meteorological Agency (EMA)
Potential production	The possible production of the land under observation	Secondary	Sentinel-2, Literature and experts
Actual production	Temporal yield statistics of farm fields	Primary	CSA, survey and literature
Topography map	Geo-spatial representation of the area	Primary	Satellite Imagery
Soil moisture	Large-scale remotely captured data	Secondary	Soil Moisture Active Passive (SMAP)
Others	Farm management and inputs	Primary/secondary	Survey, observation, literature , experts, report, repositories

3.11 Anticipated challenges and possible coping mechanisms

In large-scale field work and in remote areas like ours, much can go wrong and we anticipate various challenges, both technical and logistical. Deployment of the WSN, for one, is susceptible to logistic, communication and human error issues. These include:

Lack of farmer awareness Since the network is deployed inside farmlands, establishing trust with farmers is a concern. With no experience of this kind of technological interventions in the area, sincere discussion and information exchange to reach consensus is needed. An early deliverable in this context is required as farmers will be more accommodating if this is the case. Moreover, working closely with agriculture extension offices, local and regional government bodies, community associations and aid organizations found in the area, we hope to minimize the challenge.

Lack of higher management awareness Crucial to address as the WSN uses the free RF in the unlicensed spectrum and no well-documented regulation exists on its usage. Through sharing neighboring countries' experiences and with proper information exchange, the information gap can be reduced and the matter can be settled.

Incorrect sensor installation Since the sensors use dielectric conductivity of the soil to measure soil moisture, interfering objects shall be avoided. In this context, sample areas with minimal electric conductivity need to be determined prior to implementation. Moreover, too tight and too wide holes affect the accuracy of the sensor readings. Removal of air gaps is needed and extremely compacted land surfaces shall be avoided.

Backhaul communication We use a cellular network for gateway to network server communication. The gateway uses a direct power supply. Nevertheless, in remote farming areas, disruptions of either happens frequently, and fall-back options are required. Thus, the gateway has a backup battery that operates for at-least three hours a day when direct power is interrupted. Data store-and-sync capability is also added to the gateway for temporary buffering during periods of disruption. During a long outage, this allows to collect the information from the buffer directly.

Network performance and data accuracy Sensor readings may significantly vary when disturbance such as movement on the surface happens, especially with unstable soil. A human or animal can step on the sample point and affect the readings. Proper installation points are identified and fallow or slope points are excluded. A data validation and filtering scheme is used to clean outlier readings. The network performance also depends on the line of sight (Los) and absence of blockages between nodes. The gateway is thus, placed in a position where minimal blockings exists and in a higher place to ensure better Los. Literature indicate the potential of large distance coverage (up to 50 km) and optimal power utilization (up to five years) achieved by the LoRaWAN WSN protocol [186]. However, implementation of the protocol is in an early stage and its performance evaluation is mostly indoors, controlled or simulation-based. This creates a concern and requires careful design for large-scale outdoor deployments. Continual monitoring and network maintenance is carried out based on experimental findings and network performance improvement measures are taken when needed.

From the participatory data collection and analysis aspect, some of the anticipated challenges are the following.

Volunteerism Considering the literacy and livelihood status of the area's population, the recruitment of volunteers can be a bottleneck for the work. Most farmers are illiterate and spend much of their time in farms, working the whole day. Misconception on the objectives of the project may arise, and thus volunteer participation may be low, which will jeopardize the work. To minimize such risks, a collaborative research approach is envisioned. Possible collaborating governmental and non-governmental bodies are identified so that existing public participation schemes are used. Local-level administrative entities and agriculture extension agents are also of high interest in this respect. We also target youths of the community, preferably school students aged 15–18 and research institutes found in the area. The accessibility and local knowledge of the selected population is crucial on top of the basic communication skills needed for the work. The volunteerism can also be aligned to some school subjects taken by participants, and secure participation. Care shall be taken, however, with the quality of the data obtained (see next).

Data quality With soil analysis conducted by non-professionals and through test kits, reliability and data quality are a concern. Faults in the kits used, sample preparation error and analysis readings inaccuracy, the quality of the camera used for spatial data generation and user's operational accuracy can be some causes. Training participants, detailed guideline and procedure preparation, and proper calibration and evaluation of kits are foreseen countermeasures. With spatial and temporal stamp on every data acquired, the reliability of data is evaluated with other submissions in close space/time proximity. Increased sample density is another mitigation scheme planned. The system allows multiple entries of a single reading. System level data validation and cleaning is then implemented to avoid any inconsistent readings. A pre-processing module is designed which is trained with valid data and filters users' readings accordingly. Extension agents, volunteer school teachers and university students also play a significant role to ensure the quality and reliability of the whole process. They are considered as first level quality assurance agents.

Time investment A significant amount of time is expected to be spent by participants. Going to the field, collecting samples and conducting the analysis requires considerable time, which might make participants less interested. We plan to use parallelization and a team approach to address this requirement. Volunteers work more as a team than as individual, and perform tasks in parallel. Sample collection, composite preparation, soil testing and uploading results can all be distributed within a team. A trade-off has to be made between time spent and knowledge acquired, as the project's indirect objective is also to install knowledge and technology transfer to the community. The concept of collaborative competition is another strategy planned to keep participants motivated in conducting the research. A team competition in which teams are evaluated based on the frequency of their participation and quality of process, with a reward mechanism in place can increase participation rates.

Sustainability Financial and logistic support is needed to mobilize volunteers and get the work done. Adequate budget allocation is thus required to keep the project going. Possible sources of such input will be explored and secured. Local and international aid organizations working with youths, capacity development, agriculture transformations and natural resource management will be brought on board. In the long run, it is also expected for such entities to takeover and sustain the work.

3.12 Summary

Successful realization of the objectives of this work depends on the spatial and temporal availability of data from the ground. Accurate and real-time readings over the study area are needed to regularly monitor the soil. Setting up a proper data acquisition infrastructure can help to achieve this. This work proposes integrating a WSN and citizens' participatory methods for such data flow. Existing data sources shall also be considered as additional source and need to be acquired. With all the data obtained, various decision support and advisory outputs can be produced. This chapter presents approaches to be used for implementation of *in-situ* data collection schemes and development of agriculture expert system based on ANN algorithm.

Chapter 4

Research Plan

The research work is broken down to tasks and activities which are planned as follows:

ID	Task Name	Activity	Start	Finish	Duration
1	Problem Investigation	Proposal work	7/10/2018	2/20/2019	162d
2		Literature Review	7/10/2018	12/28/2018	124d
3		Progress Report	2/26/2019	2/26/2019	1d
4	Phase I execution	Inception, site selection	4/1/2019	4/19/2019	15d
5		Existing source review, Survey	4/10/2019	9/30/2019	124d
6		Network design & Calibration	5/1/2019	5/30/2019	22d
7		Network deployment & configuration	6/3/2019	6/28/2019	20d
8		Data Acquisition	7/1/2019	9/1/2021	568d
9		Network evaluation/ empirical	7/1/2019	12/30/2019	131d
10		Network maintenance & expansion	9/16/2019	9/30/2019	11d
11		Observation analysis	7/1/2019	9/30/2019	66d
12		Paper write-up/ empirical findings	10/1/2019	1/31/2020	89d
13		Stakeholder meet-up	9/16/2019	9/30/2019	11d
14	Phase II execution	System design and development	1/24/2020	10/15/2020	190d
15	Phase III execution	Facilitators' recruit and training	9/1/2020	9/15/2020	11d
16		School visit and kit-hub setup	12/16/2020	1/1/2021	13d
17		Volunteers recruit and training	1/1/2021	2/26/2021	41d
18		Soil analysis and data acquisition	3/10/2021	9/30/2021	147d
19		Performance evaluation	3/22/2021	9/30/2021	139d
20		Paper write-up/ preliminary findings	6/1/2021	9/30/2022	349d
21	Phase V completion	Data analysis and knowledge build	3/30/2021	2/1/2022	221d
22		Deliverables and Thesis write-up	1/11/2021	2/1/2021	16d
23		Result communication	7/1/2022	7/1/2022	1d

Figure 4.1: Execution plan of the proposed work

Chapter 5

Conclusion

Soil is an important resource for Ethiopian farmers, for so many reasons, but significantly for agriculture production. Agriculture is dominantly rain-fed, generated by subsistence farmers with an average of less than one ha of cultivated land which has poor soil fertility. The soil nutrient depletion and degradation has a negative impact on food production potential. Crop yield is low mainly due to unhealthy soil which is caused by various biophysical, environmental and management challenges. Continual real-time observation and monitoring of the soil status, is thus crucial to overcome this low yield. Doing so has been a challenge and time consuming because of limited resources and in-situ data acquisition infrastructures. An alternative is to look for feasible and optimal methods and tools that can fill this gap. Technology advancements such as IoT are enabling tools for such on-the-spot monitoring and data acquisition. Moreover, citizens also play significant role in monitoring and sharing such local information. Integrating collaborative participatory approach with IoT for soil monitoring and analysis is the way forward envisioned by this work. Through simple, efficient, feasible and accessible soil kits, citizens can be involved in observing and understanding the nutrient content of farmlands. Sensing objects are embedded in fields in three dimensions: time, space, and depth to obtain additional soil parameters such as moisture and temperature. The acquired information can further be processed to see spatial and temporal variations of soil and study the interplay between soil properties and crop growth. The outcome of such analysis further broadens the knowledge of stakeholders and assist in agricultural advisory and decision support. It can be used to depict the investment farmers need to put, estimate yields as exists and what to do to improve the yield, determine the favorable soil conditions to specific crop and hence better productivity. This work is thus, an attempt to significantly contribute to:

- Real-time soil moisture and temperature acquisition system using WSN
- Real-time soil macro-nutrient analysis through affordable, accessible and easily utilized soil test kits
- Participatory farm-field data collection by engaging citizens
- Integration of heterogeneous data sources and development of knowledge base
- Development of an inference engine for agriculture advisory system
- Facilitate efficient information acquisition and flow, both in quantity and quality

Accurate measurements of soil nutrient level, moisture and temperature can help stakeholders take timely and logical corrective actions, which are key factors for better crop productivity and regulation of nutrient uptake as well as determining the type, amount and timing for fertilizer applications. The outcome of this project can be an important input to economic models that assess food security and land use at different spatial scales. The research creates a win-win-win state; enhance crop yield of farmers, increased adaptation and resilience to climate change and reduce the land degradation of the area. It is our belief that stakeholders of the agrichain will takeover this project and sustain continual monitoring of farmlands, in larger-scale. It can be extended to a larger scale with minimal investment either with sensors' network rotation and make use of existing resources at different sites of interest or with new set up on existing architecture.

Bibliography

[1] Committee on World Food Security. Coming to terms with terminology: Food security, nutrition security, food security and nutrition, food and nutrition security. *Proceedings of the 39th session of the Committee on World Food Security [internet]*, 2012.

[2] UNDP. Sustainable Development Goal 2. online, 2018. URL http://misc.undp.org/content/undp/en/home/susstainable-development-goals/goal-2-zero-hunger.html. last accessed: May, 2018.

[3] Mercedes de Onis, David Brown, Monika Blossner, and Elaine Borghi. Levels and trends in child malnutrition. joint child malnutrition estimates. Technical report, UNICEF and WHO and World Bank, 2015.

[4] FAO, IFAD, UNICEF, WFP, and WHO. The State of Food Security and Nutrition in the World 2017. Technical report, Food and Agriculture Organization of the United Nations, 2017. URL http://misc.fao.org/3/a-I7695e.pdf. last accessed: May, 2018.

[5] United Nations, General Assembly. Agriculture development, food security and nutrition. Technical report, United Nations, General Assembly, 2017. URL https://reliefweb.int/sites/reliefweb.int/files/resources/N1724950.pdf. last accessed: May 2018.

[6] ECESA, FAO, IFAD, and WFP. HLPF thematic review of SDG2: End hunger, achieve food security and improved nutrition, and promote sustainable agriculture. Technical report, United Nations, 2017. URL https://sustainabledevelopment.un.org/content/documents/14371SDG2_format.revised_FINAL_28_04.pdf. last accessed: March 2018.

[7] International Monetary Fund. The Federal Democratic Republic of Ethiopia: Poverty Reduction Strategy Paper Growth and Transformation Plan 2010/11–2014/15 – Volume II. Country report no. 11/305, International Monetary Fund, 2011. URL https://misc.imf.org/en/Publications/CR/Issues/2016/12/31/The-Federal-Democratic-Republic-of-Ethiopia-Poverty-Reduction-Strategy-Paper-Growth-and-25284. last accessed: March 2018.

[8] World Food Program. Country Programme Ethiopia (2012-2017). Standard project report, World Food Program, 2016. URL https://docs.wfp.org/api/documents/e6f36b2703664a2ca81a8573ed6872b0/download/. last accessed: March 2018.

[9] Ethiopia. YaÍkonomi lemātenā tebeber ministér. *Survey of the Ethiopian Economy: Review of Post-reform Developments (1992/93 - 1997/98)*. Ministry of Economic Development and Cooperation (MEDaC), 1999. URL https://books.google.nl/books?id=oouBtgAACAAJ.

[10] USAID. As Ethiopia Restores Watersheds, Landless Youth Find Sweet Rewards. online, 2017. URL https://misc.usaid.gov/results-data/success-stories/landless-youth-find-sweet-rewards-watershed-rehabilitation. last accessed: March 2018.

[11] Joint Government and Humanitarian Partners. Ethiopia–The 2017 Humanitarian RequirementsD́ocument. Technical report, Government of Ethiopia, 2017.

URL https://misc.unicef.org/ethiopia/ECO_2017_HRD.pdf. last accessed: March 2018.

[12] Aregay Waktola. *Exploratory study of two regions in Ethiopia to identify target areas and partners for intervention*. DCG, 1999.

[13] Amare Haileslassie, Fitsum Hagos, Everisto Mapedza, Claudia W Sadoff, Seleshi Bekele Awulachew, Solomon Gebreselassie, and Don Peden. *Institutional settings and livelihood strategies in the Blue Nile Basin: implications for upstreamdownstream linkages*, volume 132. IWMI, 2009.

[14] FEWS NET. ETHIOPIA Food Security Outlook: June 2018 to january 2019. Report, World Food Programme, 2018. URL http://fews.net/sites/default/files/documents/reports/ETHIOPIA_Food_Security_Outlook_June2018_final_0.pdf. last accessed: July, 2018.

[15] Government of Ethiopia and UN Office for the Coordination of Human Affairs. Ethiopia: 2018 Humanitarian and Disaster Resilience Plan. Plan report, HUMANITARIAN COORDINATOR FOR ETHIOPIA, 2018. URL https://misc.humanitarianresponse.info/sites/misc.humanitarianresponse.info/files/documents/files/ethiopia_2018_humanitarian_and_disaster_resilience_plan_2.pdf. last accessed: July, 2018.

[16] Sarah K Lowder, Jakob Skoet, and Terri Raney. The number, size, and distribution of farms, smallholder farms, and family farms worldwide. *World Development*, 87:16–29, 2016.

[17] Geoffrey Livingston, Steven Schonberger, and Sara Delaney. Sub-Saharan Africa : The state of smallholders in agriculture. online, 2011. URL https://pdfs.semanticscholar.org/f2cb/d3f72cb333c1cc6fd3eba6d5bc8bb8c89469.pdf. Paper presented at the IFAD Conference on New Directions for Smallholder Agriculture.

[18] ALEMAYEHU SEYOUM TAFFES Se, Paul Dorosh, and Sinafikeh Asrat Gemessa. Crop production in Ethiopia: regional patterns and trends. *Food and agriculture in Ethiopia: Progress and policy challenges*, 74:53, 2012. URL http://ebrary.ifpri.org/cdm/ref/collection/p15738coll2/id/127350.

[19] N Berhanu, A Berhanu, and G Samuel. Land tenure and agricultural development in Ethiopia. *Ethiopian Economic Association/Ethiopian Economic Policy Research Institute*, 2002. URL https://misc.eeaecon.org/sites/default/files/publications/Land%20Tenure%20and%20agricultural%20Development%20in%20Ethiopia_research%20report_0.pdf.

[20] Porter JR, Xie L, Challinor AJ, Cochrane K, Howden SM, Iqbal MM, Lobell DB, and Travasso MI. Food security and food production systems. In *Climate Change 2014: Impacts, Adaptation, and Vulnerability. Part A: Global and Sectoral Aspects*, pages 485–533. Cambridge University Press, 2014. doi: https://hdl.handle.net/10568/68162.

[21] Daniele Giovannucci, Sara Scherr, Danielle Nierenberg, Charlotte Hebebrand, Julie Shapiro, Jeffrey Milder, and Keith Wheeler. Food and Agriculture: The future of sustainability. A strategic input to the Sustainable Development in the 21st Century (SD21) project. Technical report, United Nations Department of Economic and Social Affairs, Division for Sustainable Development, 2012. URL https://sustainabledevelopment.un.org/content/documents/agriculture_and_food_the_future_of_sustainability_web.pdf.

[22] Dacian Ciolos and Andris Piebalgs. Sustainable agriculture for the future we want. *European Commission. Agric. and Rural Development.*, 2012. URL https://ec.europa.eu/agriculture/sites/agriculture/files/events/2012/rio-side-event/brochure_en.pdf.

[23] FAO. The State of Food and Agriculture: Innovation in family farming. report 0081-4539, FAO, 2014. URL http://misc.fao.org/3/a-i4040e.pdf. last accessed: February 2018.

[24] B Kelsey Jack. Constraints on the adoption of agricultural technologies in developing countries. *White paper, Agricultural Technology Adoption Initiative, Boston: J–PAL (MIT) and Berkeley: CEGA (UC Berkeley)*, 2011.

[25] Organization for Economic Co-Operation and Development. Adoption of technologies for sustainable farming systems. In Centre français d'exploitation du droit de copie (CFC), editor, *Wageningen Workshop Proceedings*, 2001.

[26] Ibrahim Kasirye. Constraints to agricultural technology adoption in Uganda: evidence from the 2005/06–2009/10 Uganda national panel survey. *Research series, Economic Policy Research Centre (EPRC)*, 2013. URL https://misc.africaportal.org/publications/constraints-to-agricultural-technology-adoption-in-uganda-evidence-from-the-200506-200910-uganda-national

[27] World Bank. World Development Report 2008 : Agriculture for Development. online, 2007. URL http://hdl.handle.net/10986/5990. last accessed: February 2018.

[28] Zebib Kassahun Weldegiorges. Benefits, constraints and adoption of technologies introduced through the eco-farm project in ethiopia. Master's thesis, Norwegian University of Life Sciences, Ås, 2015.

[29] Manzoor Hussain Memon, Khalid Khan, Muhammad Yasir Abbass, Gulawar Khan, and Muhammad Abdul Kamal. Impediments to Technology Adoption: A Case Study of Peach Production in District Swat, Pakistan. *Journal of Managerial Sciences*, 9(2):226–242, 2015.

[30] Jock R Anderson and Gershon Feder. Agricultural extension. *Handbook of agricultural economics*, 3:2343–2378, 2007.

[31] C. Ragasa and J. Mazunda. The impact of agricultural extension services in the context of a heavily subsidized input system: The case of Malawi. *World Development*, 105:25–47, 2018. doi: https://doi.org/10.1016/j.worlddev.2017.12.004. Cited By :1.

[32] Ethiopian Agricultural Transformation Agency and Ministry of Agriculture. National Strategy for Ethiopia's Agricultural Extension System; Vision, Systemic Bottlenecks and Priority Interventions, 2017.

[33] World Food Program. The State of Food Insecurity in the World: Economic Crisis–Impacts and Lessons Learned, 2009.

[34] USAID Collaborative Research Support Programs Team. Amhara National Regional State Food Security Research Assessment Report. Assessment report, USAID, 2000. URL https://misc.ctahr.hawaii.edu/sm-crsp/phase1/pdf/amhara.pdf. last accessed: February 2018.

[35] International Crops Research Institute for the Semi-Arid Tropics ICRISAT. Feeding degraded soils in Ethiopia to feed the people and the environment. Technical report, International Crops Research Institute for the Semi-Arid Tropics ICRISAT, March 2018.

[36] The Federal Democratic Republic of Ethiopia Central Statistics Agency. Agricultural sample survey, area and production of major crops. Survey, CSA, May 2016. URL file:///C:/Users/amtej/Downloads/Report%20on%20Area%20and%20production%20of%20major%20crops-%202010%20Meher%20season.pdf. last accessed: March 2018.

[37] Hendrik Wulf, Michael E. Schaepman Titia Mulder, Armin Keller, and Philip C. Jorg. Remote Sensing of Soils. Document, Zurich Open Repoitory and Archive, University of Zurich, ZORA, 1 2015. URL http://misc.zora.uzh.ch/id/eprint/109992/1/2015_Wulf_Remote_sensing_of_soils_BAFU_report_dpi300_v%20.pdf.

[38] Zerger, Andre and Rossel, RA Viscarra and Swain, DL and Wark, Tim and Handcock, Rebecca N and Doerr, VAJ and Bishop-Hurley, GJ and Doerr, ED and Gibbons, PG and Lobsey, Craig. Environmental sensor networks for vegetation, animal and soil sciences. *International Journal of Applied Earth Observation and Geo-information*, 12(5):303–316, 2010.

[39] E Khamala et al. Review of the available remote sensing tools, products, methodologies and data to improve crop production forecasts. *Review of the available remote sensing tools, products, methodologies and data to improve crop production forecasts.*, 2017.

[40] Katarzyna Dabrowska-Zielinska, Maria Budzynska, Radoslaw Gurdak, Jan Musial, Alicja Malinska, Martyna Gatkowska, and Maciej Bartold. Application of Sentinel-1 VH and VV and Sentinel-2 for soil moisture studies. In *Active and Passive Microwave Remote Sensing for Environmental Monitoring*, volume 10426, page 104260C. International Society for Optics and Photonics, 2017.

[41] Burzykowska, Anna, Velasco, Almudena, Klaase, Annemarie, Huber, Silvia, Geerders, Paul, Dost, Remco, Vrielink, Arjen, Haas, Eva, de By, Rolf A, and Aparicio, Evelyn. How can Earth Observation support agriculture development in rural areas?: EO4SD-Agriculture and Rural Development cluster. *A Better World*, 4:18–23, 2018.

[42] Frederic Achard, Javier Gallego, Guido Lemoine, Pierre Defourny, Andrew Davidson, Thierry Fisette, Heather McNairn, Bahram Daneshfar, Jacques Delincé, Danilo Mollicone, John Latham, Shibendu Ray, Neetu, and Oscar Rojas. *Global Strategy to improve Agricultural and Rural Statistics (GSARS);Handbook on Remote Sensing for Agricultural Statistics*. GSARS, 2017. URL http://gsars.org/wp-content/uploads/2017/09/GS-REMOTE-SENSING-HANDBOOK-FINAL-04.pdf.

[43] Bartholomeus, Harm, Kooistra, Lammert, Stevens, Antoine, van Leeuwen, Martin, van Wesemael, Bas, Ben-Dor, Eyal, and Tychon, Bernard. Soil organic carbon mapping of partially vegetated agricultural fields with imaging spectroscopy. *International Journal of Applied Earth Observation and Geoinformation*, 13(1): 81–88, 2011.

[44] Sirpa Thessler, Lammert Kooistra, Frederick Teye, Hanna Huitu, and Arnold K Bregt. Geosensors to support crop production: Current applications and user requirements. *Sensors*, 11(7):6656–6684, 2011.

[45] Shengyao Jia, Hongyang Li, Yanjie Wang, Renyuan Tong, and Qing Li. Hyperspectral Imaging Analysis for the Classification of Soil Types and the Determination of Soil Total Nitrogen. *Sensors*, 17(10):2252, 2017.

[46] Joseph V Sinfield, Daniel Fagerman, and Oliver Colic. Evaluation of sensing technologies for on-the-go detection of macro-nutrients in cultivated soils. *Computers and Electronics in Agriculture*, 70(1):1–18, 2010.

[47] Yufeng Ge, J Alex Thomasson, and Ruixiu Sui. Remote sensing of soil properties in precision agriculture: A review. *Frontiers of Earth Science*, 5(3):229–238, 2011. doi: 10.1007/s11707-011-0175-0. URL https://doi.org/10.1007/s11707-011-0175-0.

[48] Telmo Adão, Jonáš Hruška, Luís Pádua, José Bessa, Emanuel Peres, Raul Morais, and Joaquim João Sousa. Hyperspectral imaging: A review on UAV-based sensors, data processing and applications for agriculture and forestry. *Remote Sensing*, 9(11):1110, 2017.

[49] Mahmoud Omid. Is Internet of Things (IoT) the future of agriculture? online, 2017. URL https://misc.researchgate.net/project/Real-world-applications-of-Internet-of-Things-IoT-in-agricultural-settings. last accessed: September, 2018.

[50] Jorge Ibarra-Esquer, Félix González-Navarro, Brenda Flores-Rios, Larysa Burtseva, and María Astorga-Vargas. Tracking the evolution of the internet of things concept across different application domains. *Sensors*, 17(6):1379, 2017.

[51] Peter Middleton, Peter Kjeldsen, and Jim Tully. Forecast: The Internet of Things, worldwide, 2013. online, December 2013. URL https://misc.gartner.com/newsroom/id/2636073. last accessed: September, 2018.

[52] Louis Columbus. 2017 Roundup Of Internet Of Things Forecasts. online, December 2017. URL https://misc.forbes.com/sites/louiscolumbus/2017/12/10/2017-roundup-of-internet-of-things-forecasts/#71211b621480. last accessed: September, 2018.

[53] anonymous. Why IoT, big data and smart farming are the future of agriculture. online, February 2017. URL http://misc.fao.org/e-agriculture/news/why-iot-big-data-smart-farming-are-future-agriculture. last accessed: September, 2018.

[54] Coen J Ritsema, Henk Kuipers, Leon Kleiboer, Erik Van Den Elsen, Klaas Oostindie, Jan G Wesseling, Jan-Willem Wolthuis, and Paul Havinga. A new wireless underground network system for continuous monitoring of soil water contents. *Water resources research*, 45(4), 2009. doi: doi:10.1029/2008WR007071.

[55] N. Dlodlo and J. Kalezhi. The internet of things in agriculture for sustainable rural development. In *2015 International Conference on Emerging Trends in Networks and Computer Communications (ETNCC)*, pages 13–18. IEEE, May 2015. doi: 10.1109/ETNCC.2015.7184801.

[56] R.N. Rao and B. Sridhar. IoT based smart crop-field monitoring and automation irrigation system. In *2nd International Conference on Inventive Systems and Control (ICISC)*, pages 478–483. IEEE, 2018.

[57] Prashaunsa J Nachankar, Mayur G Somani, Deeksha M Singh, and Sunil N Katkar. IoT in agriculture. *Decision making*, 1(2):3, 2018.

[58] Danco Davcev, Kosta Mitreski, Stefan Trajkovic, Viktor Nikolovski, and Nikola Koteli. IoT agriculture system based on lorawan. In *2018 14th IEEE International Workshop on Factory Communication Systems (WFCS)*, pages 1–4. IEEE, 2018.

[59] Aalaa Abdullah, Shahad Al Enazi, and Issam Damaj. Agrisys: A smart and ubiquitous controlled-environment agriculture system. *Big Data and Smart City, IEEE*, 2016.

[60] Renfei Wu, Ying Xu, Long Li, Junli Zha, and Renfa Li. Design and Implementation of an Intelligent Monitoring System Based on ZigBee for the Agricultural Greenhouse. In *Advanced Technologies in Ad Hoc and Sensor Networks*, pages 195–203. Springer, 2014.

[61] Jesús Martín Talavera, Luis Eduardo Tobón, Jairo Alejandro Gómez, María Alejandra Culman, Juan Manuel Aranda, Diana Teresa Parra, Luis Alfredo Quiroz, Adolfo Hoyos, and Luis Ernesto Garreta. Review of IoT applications in agro-industrial and environmental fields. *Computers and Electronics in Agriculture*, 142:283–297, 2017.

[62] Renee Cho. Revitalizing Africa's Soils. online, September 2014. URL https://blogs.ei.columbia.edu/2014/09/24/revitalizing-africas-soils/. last accessed: September, 2018.

[63] G Bakker, JP Okx, M Assen, and T Solomon. Ethiopian soil laboratory infrastructure. Technical report, Alterra, Wageningen-UR, 2016.

[64] E Ben-Dor, Sabine Chabrillat, JAM Demattê, GR Taylor, J Hill, ML Whiting, and S Sommer. Using imaging spectroscopy to study soil properties. *Remote sensing of environment*, 113:S38–S55, 2009.

[65] S Mirzaee, S Ghorbani-Dashtaki, J Mohammadi, H Asadi, and F Asadzadeh. Spatial variability of soil organic matter using remote sensing data. *Catena*, 145:118–127, 2016.

[66] Andreas Steinberg, Sabine Chabrillat, Antoine Stevens, Karl Segl, and Saskia Foerster. Prediction of common surface soil properties based on vis-NIR airborne and simulated EnMAP imaging spectroscopy data: Prediction accuracy and influence of spatial resolution. *Remote Sensing*, 8(7):613, 2016.

[67] Y Li, D Chen, CN Walker, and JF Angus. Estimating the nitrogen status of crops using a digital camera. *Field Crops Research*, 118(3):221–227, 2010.

[68] Mark A Liebig, John W Doran, and John C Gardner. Evaluation of a field test kit for measuring selected soil quality indicators. *Agronomy journal*, 88(4):683–686, 1996. doi: https://doi.org/10.2134/agronj1996.00021962008800040030x.

[69] Rigor G Regalado and Jennifer C Dela Cruz. Soil pH and nutrient (Nitrogen, Phosphorus and Potassium) analyzer using colorimetry. In *Region 10 Conference (TENCON), 2016 IEEE*, pages 2387–2391. IEEE, 2016.

[70] Aree Choodum, Khanitta Malathong, Niamh NicDaeid, Wadcharawadee Limsakul, and Worawit Wongniramaikul. A cost effective hydrogel test kit for pre and post blast trinitrotoluene. *Forensic science international*, 266:202–208, 2016.

[71] LP Brandenberger, TJ Bowser, H Zhang, LK Carrier, and ME Payton. Evaluation of Testing Kits for Routine Soil Analyses. *Discourse Journal of Agriculture and Food Sciences*, 4(2346–7002), 2016. URL misc.resjournals.org/JAFS.

[72] Ben A Faber, A James Downer, Dirk Holstege, and Maren J Mochizuki. Accuracy Varies for Commercially Available Soil Test Kits Analyzing Nitrate–Nitrogen, Phosphorus, Potassium, and pH. *HortTechnology*, 17(3):358–362, 2007.

[73] James Sumberg, Christine Okali, and David Reece. Agricultural research in the face of diversity, local knowledge and the participation imperative: theoretical considerations. *Agricultural systems*, 76(2):739–753, 2003.

[74] Jonathan Hellin, MR Bellon, Lone Badstue, John Dixon, and Roberto La Rovere. Increasing the impacts of participatory research. *Experimental Agriculture*, 44 (1):81–95, 2008.

[75] Salvatore Ceccarelli and Stefania Grando. Decentralized-participatory plant breeding: an example of demand driven research. *Euphytica*, 155(3):349–360, 2007.

[76] Melissa Eitzel, Jessica Cappadonna, Chris Santos-Lang, Ruth Duerr, Sarah Elizabeth West, Arika Virapongse, Christopher Kyba, Anne Bowser, Caren Cooper, Andrea Sforzi, et al. Citizen science terminology matters: Exploring key terms. *Citizen Science: Theory and Practice*, pages 1–20, 2017.

[77] R. J Murphy. Facilitating citizen science through gamification. *Department of Design Science and Information Systems, Florida International University*, 2015.

[78] Oded Nov, Ofer Arazy, and David Anderson. Technology-mediated citizen science participation: A Motivational Model. In *ICWSM*, 2011.

[79] Oded Nov, Ofer Arazy, and David Anderson. Dusting for science: motivation and participation of digital citizen science volunteers. In *Proceedings of the 2011 iConference*, pages 68–74. ACM, 2011. doi: 10.1145/1940761.1940771.

[80] Dana Rotman, Jen Hammock, Jenny Preece, Derek Hansen, Carol Boston, Anne Bowser, and Yurong He. Motivations affecting initial and long-term participation in citizen science projects in three countries. *iConference 2014 Proceedings*, 2014.

[81] Eskender Beza, Jonathan Steinke, Jacob Van Etten, Pytrik Reidsma, Carlo Fadda, Sarika Mittra, Prem Mathur, and Lammert Kooistra. What are the prospects for citizen science in agriculture? evidence from three continents on motivation and mobile telephone use of resource-poor farmers. *PloS one*, 12(5): e0175700, 2017.

[82] Sebastian Deterding. Situated motivational affordances of game elements: A conceptual model. In *Gamification: Using game design elements in non-gaming contexts, a workshop at CHI*, 2011.

[83] Richard M Ryan and Edward L Deci. Intrinsic and extrinsic motivations: Classic definitions and new directions. *Contemporary educational psychology*, 25(1):54–67, 2000. doi: 10.1006/ceps.1999.1020[PubMed].

[84] Jonathan Martin, Diego Torres, Alejandro Fernandez, Santiago Pravisani, and Guillaume Briend. Using citizen science gamification in agriculture collaborative knowledge production. In *Proceedings of the XIX International Conference on Human Computer Interaction*, page 37. ACM, 2018. doi: 10.1145/3233824.3233859.

[85] Eskender Andualem Beza. *Citizen science and remote sensing for crop yield gap analysis*. PhD thesis, Wageningen University, 2017.

[86] Wayne Shelley, Russell Lawley, and David A Robinson. Crowd-sourced soil data for Europe. *Nature*, 496(7445):300–300, 2013.

[87] Linda Davies, JNB Bell, James Bone, Martin Head, Lawrence Hill, C Howard, Sarah J Hobbs, David T Jones, Sally A Power, N Rose, et al. Open Air Laboratories (OPAL): A community-driven research programme. *Environmental Pollution*, 159(8-9):2203–2210, 2011.

[88] GLOBE. The GLOBE Program A Worldwide Science and Education Program: Pedosphere (Soil) Community. online, 2018. URL https://misc.globe.gov/web/soil/overview. last accessed: October, 2018.

[89] David G Rossiter, Jing Liu, Steve Carlisle, and A-Xing Zhu. Can citizen science assist digital soil mapping? *Geoderma*, 259:71–80, 2015.

[90] Jayavardhana Gubbi, Rajkumar Buyya, Slaven Marusic, and Marimuthu Palaniswami. Internet of Things (IoT): A vision, architectural elements, and future directions. *Future generation computer systems*, 29(7):1645–1660, 2013.

[91] Guest Writer. IoT Applications in Agriculture;The demand for growing population can be successfully met with IoT. online, January 2018. URL https://misc.iotforall.com/iot-applications-in-agriculture. last accessed: October, 2018.

[92] S Banu. Precision Agriculture: Tomorrow's Technology for Today's Farmer. *Journal of Food Processing & Technology*, 6(8):1, 2015.

[93] Athanasios Balafoutis, Bert Beck, Spyros Fountas, Jurgen Vangeyte, Tamme Wal, Iria Soto, Manuel Gómez-Barbero, Andrew Barnes, and Vera Eory. Precision agriculture technologies positively contributing to ghg emissions mitigation, farm productivity and economics. *Sustainability*, 9(8):1339, 2017.

[94] Sjaak Wolfert, Lan Ge, Cor Verdouw, and Marc-Jeroen Bogaardt. Big data in smart farming–a review. *Agricultural Systems*, 153:69–80, 2017.

[95] Rob Lokers, Rob Knapen, Sander Janssen, Yke van Randen, and Jacques Jansen. Analysis of Big Data technologies for use in agro-environmental science. *Environmental Modelling & Software*, 84:494–504, August 2016.

[96] Saiyyad Mohmmad Ali Muzffar Ali. A Survey on Agriculture Crop Recommender System. *International Journal of Advanced Research in Computer Engineering and Technology (IJARCET)*, 7, February 2018. ISSN 2278 – 1323.

[97] Rohit Kumar Rajak, Ankit Pawar, Mitalee Pendke, Pooja Shinde, Suresh Rathod, and Avinash Devare. Crop recommendation system to maximize crop yield using machine learning technique. *Int Res J Eng Technol*, 4(12):950–953, 2017.

[98] Raheela Shahzadi, Javed Ferzund, Muhammad Tausif, and Muhammad Asif Suryani. Internet of Things based Expert System for Smart Agriculture. *(IJACSA) International Journal of Advanced Computer Science and Applications*, 7(9):341–350, 2016.

[99] Jeetendra Shenoy and Yogesh Pingle. IoT in agriculture. In *Computing for Sustainable Global Development (INDIACom), 2016 3rd International Conference on*, pages 1456–1458. IEEE, 2016.

[100] Hoefsloot, Peter, Ines, Amor VM, van Dam, Jos, Duveiller, Gregory, Kayitakire, Francois, and Hansen, James. Combining crop models and remote sensing for yield prediction: Concepts, applications and challenges for heterogeneous smallholder environments. Technical report, CCFAS-JRC Workshop, June 2012.

[101] Kazeem O Rauff and Rasaq Bello. A review of crop growth simulation models as tools for agricultural meteorology. *Agricultural Sciences*, 6(09):1098, 2015. doi: http://dx.doi.org/10.4236/as.2015.69105.

[102] René Gommes. An introduction to the art of agrometeorological crop yield forecasting using multiple regression. *Dhaka FAO. 38p*, 2001.

[103] Bouman B.A.M., van Diepen, C.A. Vossen P., and van der Wal T. Simulation and systems analysis tools for crop yield forecasting. In *Applications of Systems Approaches at the Farm and Regional Levels Volume 1*, pages 325–340. Springer, 1997. doi: https://doi.org/10.1007/978-94-011-5416-1_24.

[104] Bruno Basso, Davide Cammarano, and Elisabetta Carfagna. Review of crop yield forecasting methods and early warning systems. In *Proceedings of the First Meeting of the Scientific Advisory Committee of the Global Strategy to Improve Agricultural and Rural Statistics*, pages 18–19, 2013. doi: 10.1017/CBO9781107415324.004.

[105] Arianna Di Paola, Riccardo Valentini, and Monia Santini. An overview of available crop growth and yield models for studies and assessments in agriculture. *Journal of the Science of Food and Agriculture*, 96(3):709–714, 2016. doi: 10.1002/jsfa.7359.

[106] Tapio J Salo, Taru Palosuo, Kurt Christian Kersebaum, Claas Nendel, Carlos Angulo, Frank Ewert, Marco Bindi, Pierluigi Calanca, Tommy Klein, Marco Moriondo, et al. Comparing the performance of 11 crop simulation models in predicting yield response to nitrogen fertilization. *The Journal of Agricultural Science*, 154(7):1218–1240, 2016. doi: 10.1017/s0021859615001124.hal-01413572.

[107] P Thimme Gowada, Sunil A Satyareddi, and SB Manjunath. Crop Growth Modeling: A Review. *Journal of Agriculture and Allied Sciences*, 2, 2013.

[108] C.A. van Diepen, J. Wolf, H. van Keulen, and C. Rappoldt. WOFOST: a simulation model of crop production. *Soil use and management*, 5(1):16–24, 1989. doi: https://doi.org/10.1111/j.1475-2743.1989.tb00755.x.

[109] J.W Jones, G Hoogenboom, C.H Porter, K.J Boote, W.D Batchelor, L.A Hunt, P.W Wilkens, U Singh, A.J Gijsman, and J.T Ritchie. The DSSAT cropping system model. *"European Journal of Agronomy"*, 18(3-4):235–265, 2003. ISSN 1161-0301. doi: https://doi.org/10.1016/S1161-0301(02)00107-7. URL http://misc.sciencedirect.com/science/article/pii/S1161030102001077. Modelling Cropping Systems: Science, Software and Applications.

[110] S. Moulin, A. Bondeau, and R. Delecolle. Combining agricultural crop models and satellite observations: from field to regional scales. *International Journal of Remote Sensing*, 19(6):1021–1036, 1998. doi: 10.1080/014311698215586.

[111] Ashish Singh, A.K. Singh, A.N. Mishra, and C.B. Singh. Yield Forecast of Rice Crop in Eastern UP Using Simulation Model. *International Journal of Current Microbiology and Applied Sciences*, 6(8):1005–1009, 2017. doi: https://doi.org/10.20546/ijcmas.2017.608.124.

[112] Tesfaye Fantaye. Application of the GYGA approach to Ethiopia. online, 2016. URL http://misc.yieldgap.org/gygamaps/pdf/Details%20on%20data%20and%20methods%20for%20GYGA%20protocols%20applied%20to%20Ethiopia.pdf. last accessed: September, 2018.

[113] Anonymous. Food, climate, disasters, analyses and more...; The three rules of crop yield forecasting. online, 2014. URL http://wergosum.com/the-three-rules-of-crop-yield-forecasting/. last accessed: October, 2018.

[114] R. Gommes. Non-parametric crop yield forecasting, a didactic case study for Zimbabwe. *Remote sensing support to crop yield forecast and area estimates. ISPRS Archives*, 36(8):W48, 2006. URL http://misc.fao.org/nr/climpag/pub/non_parametric_yield_forecasting_FAO.pdf.

[115] R. Gommes, F. L. Snijders, J. Q. Rijks, D. Rijks, J. M. Terres, and P. Vossen. FAO crop yield forecasting philosophy in national food early warning-systems. *David Lubin Memorial Library, Food and Agriculture Organization of the U. N.*, 1998.

[116] Paul C Doraiswamy, Sophie Moulin, Paul W Cook, and Alan Stern. Crop yield assessment from remote sensing. *Photogrammetric engineering & remote sensing*, 69(6):665–674, 2003. doi: 10.14358/PERS.69.6.665.

[117] Jianmao Guo, Qi Wang, Tengfei Zheng, Junyi Shi, and Jinhui Zhu. Simulation of regional winter wheat growth by using remote sensing data and crop model. In *Remote Sensing and Modeling of Ecosystems for Sustainability IX*, volume 8513, page 85130E. International Society for Optics and Photonics, 2012. doi: 10.1117/12.930387.

[118] Jianqiang Ren, Zhongxin Chen, Huajun Tang, Fushui Yu, and Qing Huang. Simulation of regional winter wheat yield by combining epic model and remotely sensed LAI based on global optimization algorithm. In *Geoscience and Remote Sensing Symposium (IGARSS), 2011 IEEE International*, pages 4058–4061. IEEE, 2011. doi: 10.1109/IGARSS.2011.6050123.

[119] VK Dadhwal. Crop growth and productivity monitoring and simulation using remote sensing and GIS. *Satellite Remote Sensing and GIS Applications in Agricultural Meteorology*, pages 263–289, 2003. doi: 10.1007/978-3-642-12261-3_27.

[120] Jiaxuan You, Xiaocheng Li, Melvin Low, David Lobell, and Stefano Ermon. Deep Gaussian Process for Crop Yield Prediction Based on Remote Sensing Data. In *AAAI*, pages 4559–4566, 2017. doi: 10.1109/MWSCAS.2006.381794.

[121] Kentaro Kuwata and Ryosuke Shibasaki. Estimating crop yields with deep learning and remotely sensed data. In *Geoscience and Remote Sensing Symposium (IGARSS), 2015 IEEE International*, pages 858–861. IEEE, 2015. doi: 10.1109/IGARSS.2015.7325900.

[122] Benjamin W Heumann, Stephen J Walsh, and Phillip M McDaniel. Assessing the application of a geographic presence-only model for land suitability mapping. *Ecological informatics*, 6(5):257–269, 2011. doi: 10.1016/j.ecoinf.2011.04.004.

[123] Bert Rijk. Integration of sensor data in crop models for precision agriculture: Towards a farmer's decision support system. Thesis report GIRS, Centre for Geo-Information; Wageningen University, 2013.

[124] Roel J Wieringa. *Design science methodology for information systems and software engineering*. Springer, 2014. ISBN 978-3-662-43839-8. doi: https://doi.org/10.1007/978-3-662-43839-8_1.

[125] Herbert Alexander Simon. *The sciences of the artificial*, volume 1968. MIT press, 1969. ISBN 9780585360102. URL https://monoskop.org/images/9/9c/Simon_Herbert_A_The_Sciences_of_the_Artificial_3rd_ed.pdf.

[126] Jordi Inglada, Marcela Arias, Benjamin Tardy, Olivier Hagolle, Silvia Valero, David Morin, Gérard Dedieu, Guadalupe Sepulcre, Sophie Bontemps, Pierre Defourny, et al. Assessment of an operational system for crop type map production using high temporal and spatial resolution satellite optical imagery. *Remote Sensing*, 7(9):12356–12379, 2015.

[127] Sen2-Agri Consortium. Detailed description of the Sen2-Agri system. online, 2018. URL http://misc.esa-sen2agri.org/operational-system/system-description/. Accessed: 2018-11-4.

[128] Sen2-Agri Consortium. Detailed description of the Sen2-Agri system. online, 2018. URL http://misc.esa-sen2agri.org/operational-system/system-description/.

[129] Cosmin UDROIU, Laurentiu NICOLA, Mickael SAVINAUD, Nicolas BELLEMANS, Philippe MALCORPS, Sophie BONTEMPS, and Pierre DEFOURNY. Sentinel-2 Agriculture. online, 2018. URL http://misc.esa-sen2agri.org/wp-content/uploads/Sen2Agri_SUM_v2.8.pdf. Accessed: 2018-11-4.

[130] Ethiopian Agricultural Transformation Agency (ATA). Soil Fertility Status and Fertilizer Recommendation Atlas of Amhara National Regional State, Ethiopia. unpublished, 2016. obtained: personal contact.

[131] Agricultural Transformation Agency (ATA). Ensemble model machine learning approach digital soil fertility map of soil-test results coupled with different satellite imageries has changed fertilizer advisory service in ethiopia. online, 2019. URL http://misc.ata.gov.et/programs/highlighted-deliverables/ethiosis/. last accessed: January, 2019.

[132] Rashid Alahmedi, Manjari Asawa, and Ariel Bream. Accelerating the Impact of IoT Technologies. online, 2018. URL https://misc.weforum.org/projects/accelerating-the-impact-of-iot-technologies. Accessed: 2018-20-11.

[133] Knud Lasse Lueth. The effect of the Internet of Things on sustainability. online, January 2018. URL https://iot-analytics.com/effect-iot-sustainability/.

[134] James Manyika, Michael Chui, Jacques Bughin, Richard Dobbs, Peter Bisson, and Alex Marrs. *Disruptive technologies: Advances that will transform life, business, and the global economy*, volume 180. McKinsey Global Institute San Francisco, CA, 2013.

[135] ITU. Series Y: Global information infrastructure, Internet protocol aspects and next-generation networks next generation networks–frameworks and functional architecture models. *International Telecommunication Union, Recommendation ITU-T Y*, 2060, 2012.

[136] Andrea Zanella, Nicola Bui, Angelo Castellani, Lorenzo Vangelista, and Michele Zorzi. Internet of things for smart cities. *IEEE Internet of Things journal*, 1(1): 22–32, 2014. doi: 10.1109/JIOT.2014.2306328.

[137] Andrea Biral, Marco Centenaro, Andrea Zanella, Lorenzo Vangelista, and Michele Zorzi. The challenges of M2M massive access in wireless cellular networks. *Digital Communications and Networks*, 1(1):1–19, February 2015. doi: https://doi.org/10.1016/j.dcan.2015.02.001.

[138] Partha Pratim Ray. A survey on Internet of Things architectures. *Journal of King Saud University–Computer and Information Sciences*, 30(3):291–319, July 2018. doi: https://doi.org/10.1016/j.jksuci.2016.10.003.

[139] WiSUN ALLIANCE. IEEE 802.15.4g Standards. Online, 2015. URL https://misc.keysight.com/upload/cmc_upload/All/Slide_IOT_Part_2.pdf. last accessed: September, 2018.

[140] Keith E Nolan, Wael Guibene, and Mark Y Kelly. An evaluation of low power wide area network technologies for the Internet of Things. In *Wireless Communications and Mobile Computing Conference (IWCMC), 2016 International*, pages 439–444. IEEE, September 2016. doi: 10.1109/IWCMC.2016.7577098.

[141] Calum McClelland. IoT Connectivity -– Comparing NB–IoT, LTE–M, LoRa, SigFox, and other LPWAN Technologies. Online, August 2018. last accessed: September, 2018.

[142] SIG Weightless. LPWAN technology decisions: 17 critical features. Online, 2016. URL http://misc.weightless.org/about/weightlessp. Accessed: 2018-9-12.

[143] Mahmoud Elkhodr, Seyed Shahrestani, and Hon Cheung. Emerging wireless technologies in the internet of things: a comparative study. *arXiv preprint arXiv:1611.00861*, 8(5), October 2016. doi: 10.5121/ijwmn.2016.8505. URL http://arxiv.org/abs/1611.00861.

[144] Bhupjit Singh and Bipjeet Kaur. Comparative study of internet of things infrastructures & security. *Global Wireless Submit*, 2016, 2016.

[145] P Biggs, J Garrity, C LaSalle, A Polomska, and R Pepper. Harnessing the Internet of Things for global development. *International Telecommunication Union*, 2016.

[146] Tarun Agarwal. Wireless sensor network architecture and its applications. *Sensor & sensor node-Scott Broscious*, 2016. doi: https://misc.elprocus.com/architecture-of-wireless-sensor-network-and-applications/.

[147] Kamaldeep Kaur, Parneet Kaur, and Er Sharanjit Singh. Wireless sensor network: Architecture, design issues and applications. *International Journal of Scientific Engineering and Research (IJSER)*, 2(11):1–10, November 2014. doi: ISSN(Online):2347-3878. Paper ID: 13111401.

[148] Ian F Akyildiz, Weilian Su, Yogesh Sankarasubramaniam, and Erdal Cayirci. Wireless sensor networks: a survey. *Computer networks*, 38(4):393–422, March 2002. doi: https://doi.org/10.1016/S1389-1286(01)00302-4.

[149] Antonio Liñań Colina, Alvaro Vives, Marco Zennaro, Antoine Bagula, and Ermanno Pietrosemoli. Internet of Things (IoT) in 5 days. Online, June 2016. URL http://misc.iet.unipi.it/c.vallati/files/IoTinfivedays-v1.1.pdf. Accessed: 2019-1-3.

[150] H Jacob and G Clarke. Methods of soil analysis, part 4, physical method. *Soil Science Society of America, Inc*, 2002.

[151] Decagon Devices. *5TM Water Content and Temperature Sensors: Operator's Manual*. Decagon Devices, 2010. URL http://misc.agrolan.co.il/UploadProductFiles/Manual-5TM.pdf.

[152] G Clarke Topp, JL Davis, and Aa P Annan. Electromagnetic determination of soil water content: Measurements in coaxial transmission lines. *Water resources research*, 16(3):574–582, 1980.

[153] The Things Industries. Building a global open LoRaWAN network. Online, 2019. URL https://misc.thethingsnetwork.org/docs/network/. Accessed: 2019-1-7.

[154] Urs Hunkeler, Hong Linh Truong, and Andy Stanford-Clark. MQTT-S—A publish/subscribe protocol for Wireless Sensor Networks. In *Communication systems software and middleware and workshops, 2008. comsware 2008. 3rd international conference on*, pages 791–798. IEEE, 2008.

[155] The HiveMQ Team. MQTT Essentials Part 3: Client, Broker and Connection Establishment. online, 2015. URL https://misc.hivemq.com/blog/mqtt-essentials-part-3-client-broker-connection-establishment/. last accessed: January, 2018.

[156] Montpellier Panel. No ordinary matter: Conserving, restoring and enhancing Africa's soils. *London, Agriculture for Impact*, 2014. URL misc.ag4impact.org.

[157] Devictor, Vincent and Whittaker, Robert J and Beltrame, Coralie. Beyond scarcity: citizen science programmes as useful tools for conservation biogeography. *Diversity and distributions*, 16(3):354–362, 2010.

[158] Hervé Goëau, Hervé Glotin, Willem-Pier Vellinga, Robert Planqué, Andreas Rauber, and Alexis Joly. LifeCLEF bird identification task 2014. In *CLEF: Conference and Labs of the Evaluation forum*, 2014.

[159] Poonam Yadav and John Darlington. Conceptual frameworks for building online citizen science projects. *arXiv preprint arXiv:1704.05084*, 2017. ISSN 2330-8001. doi: 10.15346/hc.v3i1.12.

[160] Cornwell, Myriah L and Campbell, Lisa M. Co-producing conservation and knowledge: Citizen-based sea turtle monitoring in North Carolina, USA. *Social Studies of Science*, 42(1):101–120, 2012. doi: https://doi.org/10.1016/j.biocon.2016.05.015.

[161] Muki Haklay. Citizen science and volunteered geographic information: Overview and typology of participation. In *Crowdsourcing geographic knowledge*, pages 105–122. Springer, 2013.

[162] Maria Arias de Reyna and Joana Simoes. Empowering citizen science through free and open source GIS. *Open Geospatial Data, Software and Standards*, 1(1): 7, 2016.

[163] Rick Bonney, Heidi Ballard, Rebecca Jordan, Ellen McCallie, Tina Phillips, Jennifer Shirk, and Candie C Wilderman. Public participation in scientific research: Defining the field and assessing its potential for informal science education. A CAISE Inquiry Group Report. *Online Submission*, 2009.

[164] Tweddle, John C and Robinson, Lucy D and Pocock, MJO and Roy, Helen E. *Guide to citizen science: developing, implementing and evaluating citizen science to study biodiversity and the environment in the UK*. NERC/Centre for Ecology & Hydrology, 2012. ISBN 978-1-906698-37-9. URL http://misc.ukeof.org.uk.

[165] Kebrom Tekle and Lars Hedlund. Land cover changes between 1958 and 1986 in kalu district, southern Wello, Ethiopia. *Mountain research and development*, 20(1):42–51, 2000.

[166] Tekalign Mamo and Gezahegn Ayele. Poverty, land resource management and gender participation in libokemkem district of northern Ethiopia. *Journal of Agriculture and Rural Development in the Tropics and Subtropics (JARTS)*, 104 (1):51–64, 2003.

[167] Carl Hartung, Adam Lerer, Yaw Anokwa, Clint Tseng, Waylon Brunette, and Gaetano Borriello. Open data kit: tools to build information services for developing regions. In *Proceedings of the 4th ACMIEEE international conference on information and communication technologies and development*, page 18. ACM, 2010.

[168] Matthias Stevens, Michalis Vitos, Julia Altenbuchner, Gillian Conquest, Jerome Lewis, and Muki Haklay. Introducing Sapelli: a mobile data collection platform for non–literate users. In *Proceedings of the 4th Annual Symposium on Computing for Development*, page 17. ACM, 2013.

[169] Jerome Lewis and Teodyl Nkuintchua. Accessible technologies and FPIC: independent monitoring with forest communities in Cameroon. *Participatory Learning and Action*, 65(13):151–165, 2012.

[170] Karin Backstrand. Civic science for sustainability: Reframing the role of experts, policy–makers and citizens in environmental governance. *Global Environmental Politics*, 3(4):24–41, 2003. doi: 10.1162/152638003322757916. URL https://doi.org/10.1162/152638003322757916.

[171] Simon Haykin. *Neural networks: a comprehensive foundation.* Prentice Hall PTR, 1994.

[172] M. Paul, S.K.Vishwakarma, and A.Verma. Analysis of soil behaviour and prediction of crop yield using data mining approach. In *2015 International Conference on Computational Intelligence and Communication Networks (CICN)*, pages 766–771, Dec 2015. doi: 10.1109/CICN.2015.156.

[173] Georg Ruß, Rudolf Kruse, Martin Schneider, and Peter Wagner. Data mining with neural networks for wheat yield prediction. In *Industrial Conference on Data Mining*, pages 47–56. Springer, 2008.

[174] FAO. *A framework for land evaluation: FAO Soils bulletin 32*. FOOD AND AGRICULTURE ORGANIZATION OF THE UNITED NATIONS, 1976. ISBN 92-5-100111-1. URL http://misc.fao.org/docrep/x5310e/x5310e04.htm#chapter%203:%20land%20suitability%20classifications. last accessed: January, 2018.

[175] F Wang. The use of Artificial Neural Networks in a Geographical Information System for Agricultural Land-Suitability Assessment. *Environment and Planning A: Economy and Space*, 26(2):265–284, 1994. doi: 10.1068/a260265. URL https://doi.org/10.1068/a260265.

[176] Igor Oliveira, Renato L.F. Cunha, Bruno Silva, and Marco A.S. Netto. A scalable machine learning system for pre-season agriculture yield forecast. *arXiv preprint arXiv:1806.09244*, 2018.

[177] Sileshi Tadesse Wodaje. Land degradation vulnerability assessment using GIS and remote sensing in Beshilo River Basin, Ethiopia. Master's thesis, NTNU, 2016.

[178] Aster Denekew Yilma and Seleshi Bekele Awulachew. Characterization and Atlas of the Blue Nile Basin and its Sub basins. *International Water Management Institute (IWMI)*, 2009. doi: https://hdl.handle.net/10568/38169. URL http://publications.iwmi.org/pdf/H042502.pdf.

[179] Mekonnen Getahun. Characterisation of agricultural soils in CASCAPE intervention woredas of Amhara region. Technical report, Wageningen University and Research (WUR):BDU-CASCAPE, 2015.

[180] Peter D Little, M Priscilla Stone, Tewodaj Mogues, A Peter Castro, and Workneh Negatu. "churning" on the margins: How the poor respond to drought in South Wollo, Ethiopia. *BASIS Brief*, 21, 2004.

[181] Peter D Little, M Priscilla Stone, Tewodaj Mogues, A Peter Castro, and Workneh Negatu. 'moving in place': Drought and poverty dynamics in South Wollo, Ethiopia. *The Journal of Development Studies*, 42(2):200–225, 2006.

[182] HEA. HEA LZ profile - Abay Beshilo Basin livelihood zone (abb) Amhara Region, Ethiopia 2007. Online, 2012. URL http://misc.heacod.net/countries/ethiopia/reports/hea-lz-profile-abay-beshilo-basin-livelihood-zone-abb-amhara-region. Accessed: 2019-1-6.

[183] Agriculture Organization of the United Nations. Soil Resources and Conservation Service. *Agro-ecological zoning: Guidelines*. Number 73. Food & Agriculture Org., 1996.

[184] Anne M Cafer. Water poverty? a political ecology approach to the study of irrigation in Ethiopia. *Agrarian Frontiers*, page 88, 2013.

[185] T. K. Philips, S. S. Panwar, and A. N. Tantawi. Connectivity properties of a packet radio network model. *IEEE Transactions on Information Theory*, 35(5): 1044–1047, September 1989. ISSN 0018-9448. doi: 10.1109/18.42219.

[186] The Things Network. Ground breaking world record LoRaWAN packet received at 702 km (436 miles) distance. online, September 2017. URL https://www.thethingsnetwork.org/article/ground-breaking-world-record-lorawan-packet-received-at-702-km-436-miles-distance. last accessed: January, 2019.

List of Figures

List of Tables