Internship report on
Productization of Soil Moisture

Submitted By

N.Reddy Thejaswini

R170139

Under the Guidance of

S.Rajeswari

Department of Computer Science and Engineering

Rajiv Gandhi University of Knowledge Technologies (RGUKT),
R.K.Valley , Kadapa , Andhra pradesh.

as a part of

Partial fulfillment of the degree of Bachelor of
Technology in Computer Science and Engineering

Date : 03-05-2023

CERTIFICATE OF PROJECT COMPLETION

This is to certify that the report entitled **"Productization of SoilMoisture"** submitted by **N. Reddy Theaswin**i bearing ID.No. **R170139** and in partial fulfillment of the requirements for the award of Bachelor of Technology in Computer Science and Engineering is a bonafide work carried out by them under my supervision and guidence.

The report has not been submitted previously in part or in full to this or any other University or Institution for the award of any degree or diploma.

Project Guide,
S.Rajeswari,
Guest Faculty,
Dept of CSE,
RGUKT R. K. Valley

Head of Department,
N. Satyanandaram,
Assisstant Professor,
Dept of CSE,
RGUKT R. K. Valley

Declaration

I **N. Reddy Thejaswini**, **R170139** hereby declare that this report entitled **" Productization of Soil Moisture"** submitted by me under the guidance and supervision of **S. Rajeswari** is a bonified work. I also here declare that it has not been submitted previously in part of or in full to this university or any other university or institution for the award of any degree or diploma.

We will be solely responsible if any kind of plagiarism is found.

Date : 03-05-2023,
Place: R. K. Valley

N.Reddy Thejaswini,
R170139

Acknowledgement

I would like to express my sincere gratitude to S.Rajeswari, my project internal guide for valuable suggestions and keen interest throughout the period of project work.

I am grateful to N.Satyanandaram, HOD of CSE, for providing excellent computing facilities and a congenial atmosphere for progressing with my project.

At the outset, I would like to thank the Rajiv Gandhi University of Knowledge Technologies, R. K. Valley for providing all the necessary and fundamental resources for successful completion of my project work.

I express my sincere thanks to all those who contributed for the sucessfull completion of my project work.

With gratitude,

N. Reddy Thejaswini,

R170139

ABSTRACT

In this work, the proposal is made to productize the module named “ Soil Moisture”. In this work, I use the latest technologies to provide unified visibility in productizing the soil moisture module. This work the involves taking the existing technology and optimizing it for the commercial use. This includes the designing, modelling and implementing the strategies for scaling up production, ensuring the product quality. By collecting the information of “ Soil Moisture ” for a particular state from various resources we develop a sustainable module which can integrate with any other projects that wants to use the soil moisture module. we present the productization of a Soil Moisture module which provides an accurate and reliable information on Soil Moisture levels. The module has been integrated with the various state projects to provide customized Soil Moisture information f or each specific state. By analyzing the data collected by the module, farmers can make informed decisions regarding irrigation schedules, crop selection, and fertilization practices.

Table of Contents

1. INTRODUCTION

1.1 Problem Statement

The efficient utilization of water resources is the crucial for sustainable development. However, the lack of accurate and reliable information on soil moisture levels poses a significant challenge for water-based projects. Soil moisture data is essential for optimizing the irrigation schedules, managing reservoirs, and improving water allocation. However, the current methods of obtaining soil moisture data are complex, time- consuming, and require a significant technical expertise. This limits the adoption of sustainable water management practices and results in the inefficient use of water resources.

Therefore, there is a need for a solution that can provide accurate and reliable soil moisture data in a timely, and user-friendly manner to support efficient water management practices.

1.2 Introduction to module

The soil moisture module is designed to provide accurate and reliable information on soil moisture levels. The module utilizes various state scrapers to collect data from different states and perform aggregation operations on the values. The aggregated values arethen stored in t he database for future use. The front-end team can access this data and showcase it in a user-friendly manner to provide near real time visibility of soil moisture values to a particular area.

1.3 Importance of the module

The soil moisture module is a valuable tool for farmers, enabling them to make informed decisions regarding irrigation schedules, crop selection, and fertilization practices.These are collected in real time from diverseinformation sources including on ground measurements, satellite based assessments and also sent via mobile apps from diverse sources. And the information is put out in real time, along with number of analysis. The data are collated and analysed using tools from Vassar Labs, like Machine learning, Artificial Intelligence and big data. It fills the gap for a system that uses dynamic data to run the scientific, and machine learning models towards over all planning and the management of soil moisture levels.

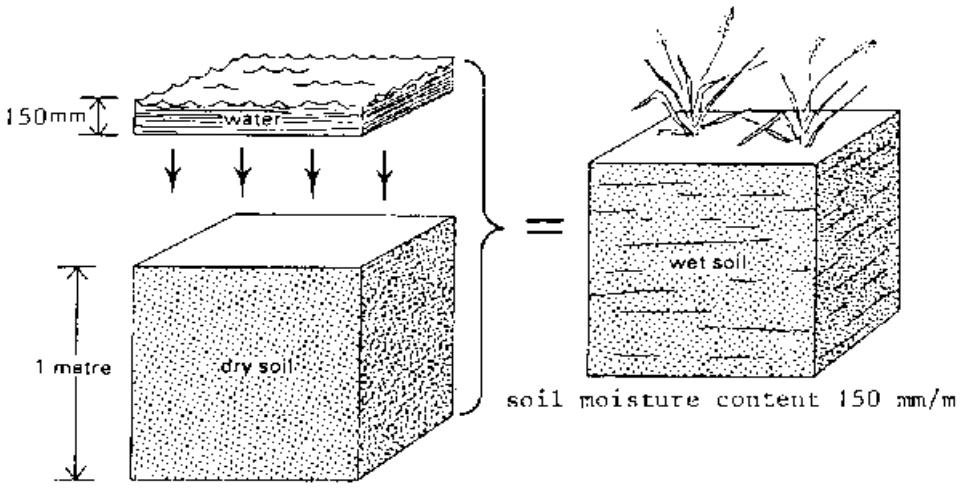


Figure 1.3: Soil Moisture Content

1.4 Importance of productizing a module

Productizing a module can lead to improved efficiency, and greater scalability, and also the potential for new revenue streams. It can also help promote standardization and collaboration across the teams, as well as to drive innovation and advanced development. By productizing a module, developers can focus on adding new features and functionality, rather than spending time building the underlying infrastructure. This can help drive innovation and allow the developers to focus on more and more complex and advanced tasks.

1.5 Scope

The scope of developing a soil moisture module is the significant, particularly in the field of agriculture and water management. Soil moisture plays a critical role in the growth and development of crops, and accurate and timely information on soil moisture levels can help farmers make informed decisions regarding irrigation schedules, crop selection, and many fertilization practices. Similarly, the water-based projects can benefit from the use of soil moisture data to optimize irrigation schedules, manage reservoirs, and improve water allocation. The development of a soil moisture module can also help promote sustainable practices by reducing water usage, improving crop yields, and promoting the conservation of natural resources. Additionally, a Soil moisture module can have economic benefits, particularly for organizations that commercialize the technology.

2. PRELIMINARIES

2.1 Collecting Soil Moisture Data

2.1.1 NRSC

Mainly the soil moisture data is collected from NRSC (National Remote Sensing Centre). The NRSC provides the soil moisture data in the form of a tif file. The tif file contains information on soil moisture levels for a specific area. Then the module utilizes this data to provide accurate and reliable information on soil moisture levels. The soil moisture levels based on their respective areas are then used to generate a heat map as shown below. We differentiate this levels of soil moisture data by adding different variating colors to each level of soil moisture .

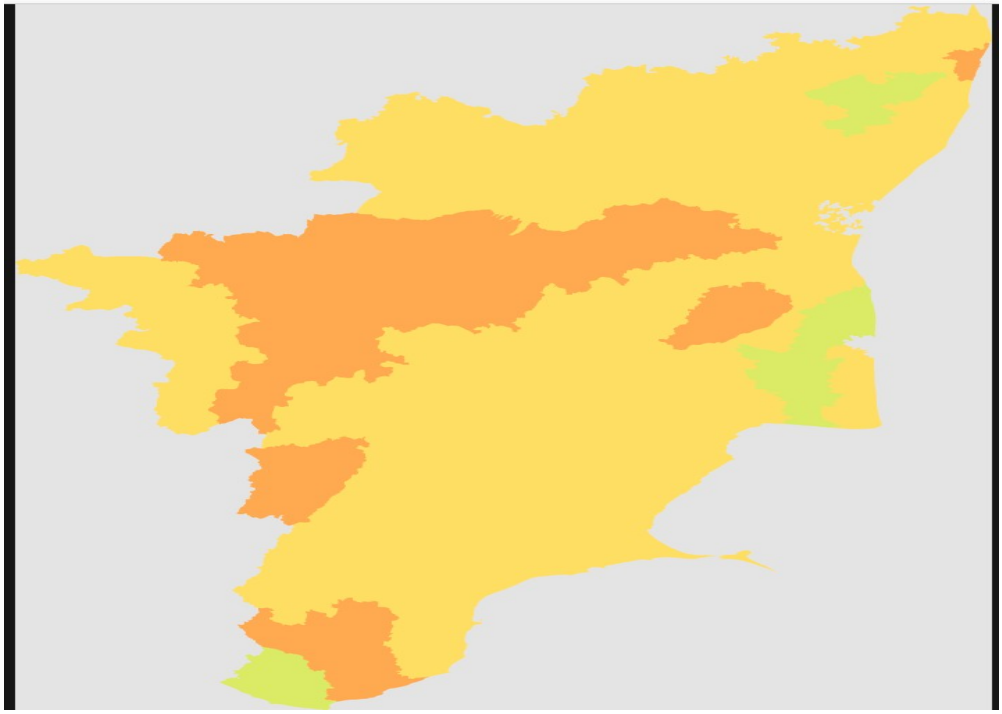


Figure 2.1.1: Heat map

2.2 Tools used

Mainly the projects is based on 3 tools:-

-> Spring Boot

-> Flink

-> Docker

2.2.1 Spring Boot

Spring Boot is an open-source Java-based framework used for creating and deploying web applications. It provides a pre-configured environment for developing and deploying web applications, making it easy for developers to create production-ready applications quickly.

Spring Boot builds on top of the Spring Framework, which is the popular Java-based framework for building enterprise- level applications. Spring Boot provides an additional features and functionality, such as auto-configuration, which simplifies the configuration of Spring-based applications, and embedded servers, which allows the applications to be run as a standalone applications without requiring a separate web server. One of the key benefits of the Spring Boot is that it reduces the amount of boilerplate code that developers need to write, allowing them to focus on writing the business logic. Spring Boot also includes a large number of pre-built components and libraries, making it easy for developers to integrate various technologies and third-party services into their applications. Overall, Spring Boot is a popular choice for building the web applications in Java, as it provides a streamlined development mode experience and a large ecosystem of tools and libraries.

2.2.2 Flink

Flink (Apache Flink) is an open-source distributed tool for stream processing framework that is designed to process large- scale data in real-time and the batch processing modes. It is a powerful and efficient platform for processing, analyzing, and querying large volumes of data, by making it a popular choice for real-time data processing in industries such as finance, e-commerce, and telecommunications.

Flink provides a unified programming model for both stream and the batch processing, allowing developers to write complex data processing pipelines that can handle both types of the data.It also includes a number of built-in libraries and many connectors for integrating with various data sources and data sinks.One of the key features of Flink is its ability to process data in a fault-tolerant and scalable manner. Flink is designed to handle failures and recover from them automatically, making it a reliable platform for processing large volumes of data. It also supports distributed processing, which allows it to scale horizontally across multiple nodes in a cluster, enabling it to handle even larger workloads.

Overall, Flink is a powerful and versatile framework for processing a large volumes of data in real- time and batch processing modes, and it has become a very popular choice for big data processing and analysis.

2.2.3 Docker

Docker is an open- source java based platform for building, shipping, and running applications in containers. Containers are lightweight, portable, and self-contained environments that encapsulate an application along with its dependencies and configurations, allowing it to run consistently across the different environments, from development to production. Docker provides a containerization solution that allows developers to package their applications into containers, which can then be easily deployed and run on any machine that supports Docker. Containers provide a consistent and isolated environment for running applications, which makes it easy to develop and test applications in a controlled environment before deploying them to production. Docker provides a number of tools and features for managing containers, such as Docker Engine, which is the core container runtime that allows you to create and manage containers, Docker Compose, which allows you to define and run multi-container applications,and Docker Swarm, which is a container orchestration tool for managing large-scale container deployments.

Overall, Docker is a powerful platform for building and deploying the applications in containers, and it has become a popular choice for developers and organizations looking to streamline their application development and deployment processes.

2.2.4 Cassandra

Cassandra is an open-source distributed NoSQL database management system that provides high scalability and fault- tolerance to handle large amounts of structured and unstructured data.Cassandra is designed to handle big data workloads across multiple nodes, offering high availability and fault tolerance. It uses a peer-to-peer architecture with no single point of failure, provides automatic data distribution, replication, and load balancing. Cassandra is schema- agnostic, meaning that it can handle a wide variety of data types and data models, including column-family, key-value, and also document-oriented data. It is optimized for read-heavy workloads and can perform well on write-heavy workloads as well.

2.2.5 Postgres

PostgreSQL, often referred to as simply Postgres, is a powerful open source relational database management system (RDBMS) known for its robustness, extensibility, and compliance with SQL standards. PostgreSQL supports a wide range of features such as support for complex queries, transactions, and concurrency control. It also provides built-in support for JSON and other semi-structured data formats. PostgreSQL is often chosen over other relational databases because of its excellent reliability, scalability and data integrity. Additionally, PostgreSQL allows for the creation of user-defined functions and custom data types, which can be written in several programming languages, including C, Python, and Java.

3. Flow Model

As the topic is soil moisture, firstly a topic named soilmoisture is created in the kafka messaging queue whenever the server is started then whenever the storm is free the topic will be received by the storm for processing. After the processing of data, the results are stored in a distributed database called " cassandra " . Zookeeper keeps tracks the topics and also helps to handle the incoming and ougoing flow of kafka queue.

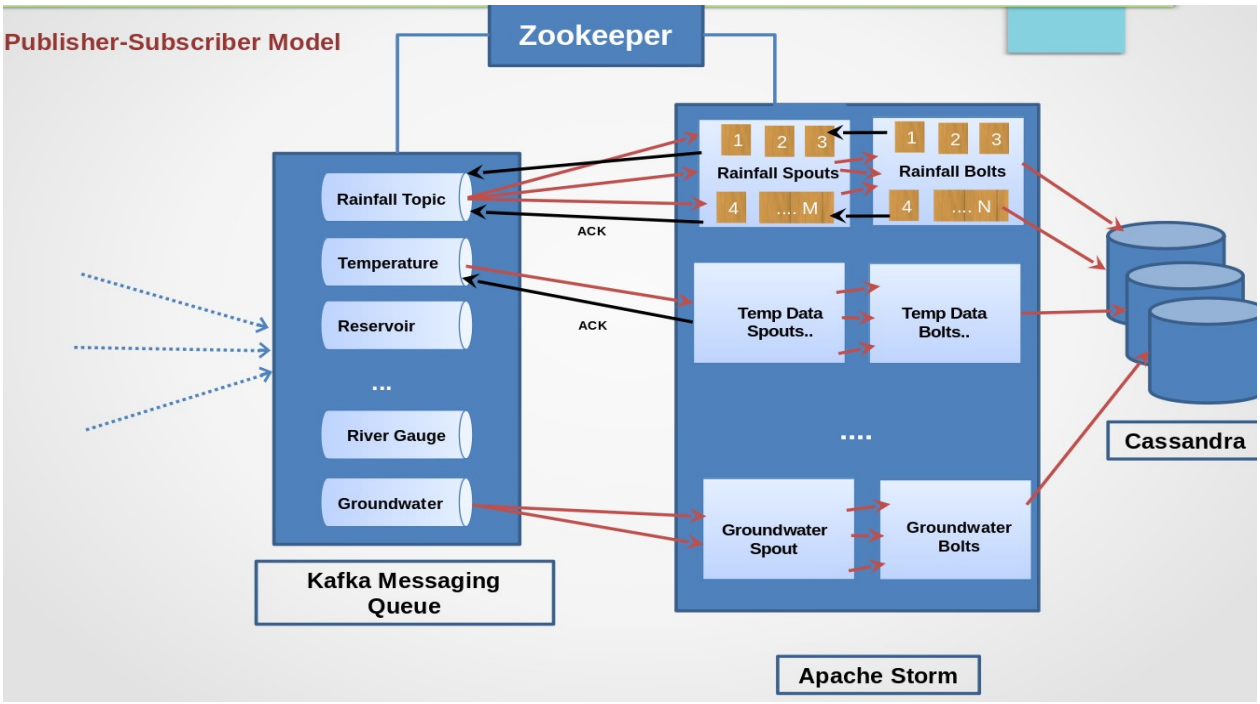


Fig 3: Publisher-Subscriber Model

4. Software Architecture

4.1 Event Driven Architecture

The event-driven architecture pattern is a popular distributed and asynchronous architecture pattern used t oproduce highly scalable applications It is also highly adaptable and can be used for small applications as well as large, complex ones. The event-driven architecture is also made up of highly decoupled, single-purpose event processing components that asynchronously receive and process events. Event channels are conduits in which events are transmitted from event emitters to event consumers.

Figure 4.1: Generic Design of an event-driven architecture

4.2 Web Application Overview/Architecture

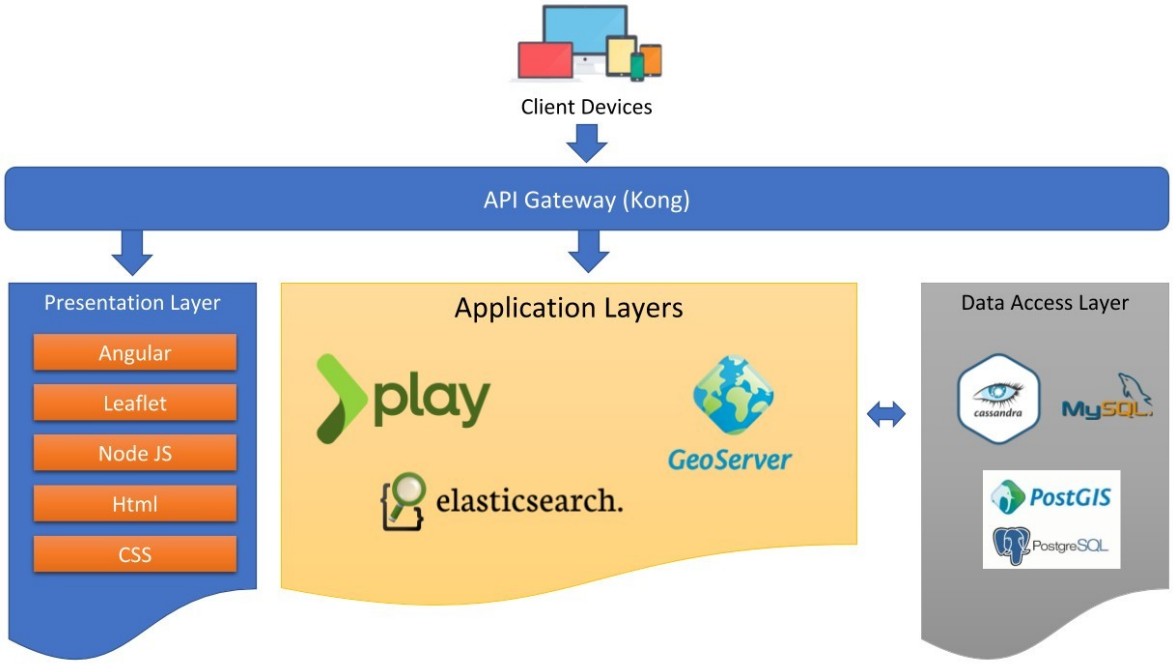


Figure 4.2: Event driven backend architecture

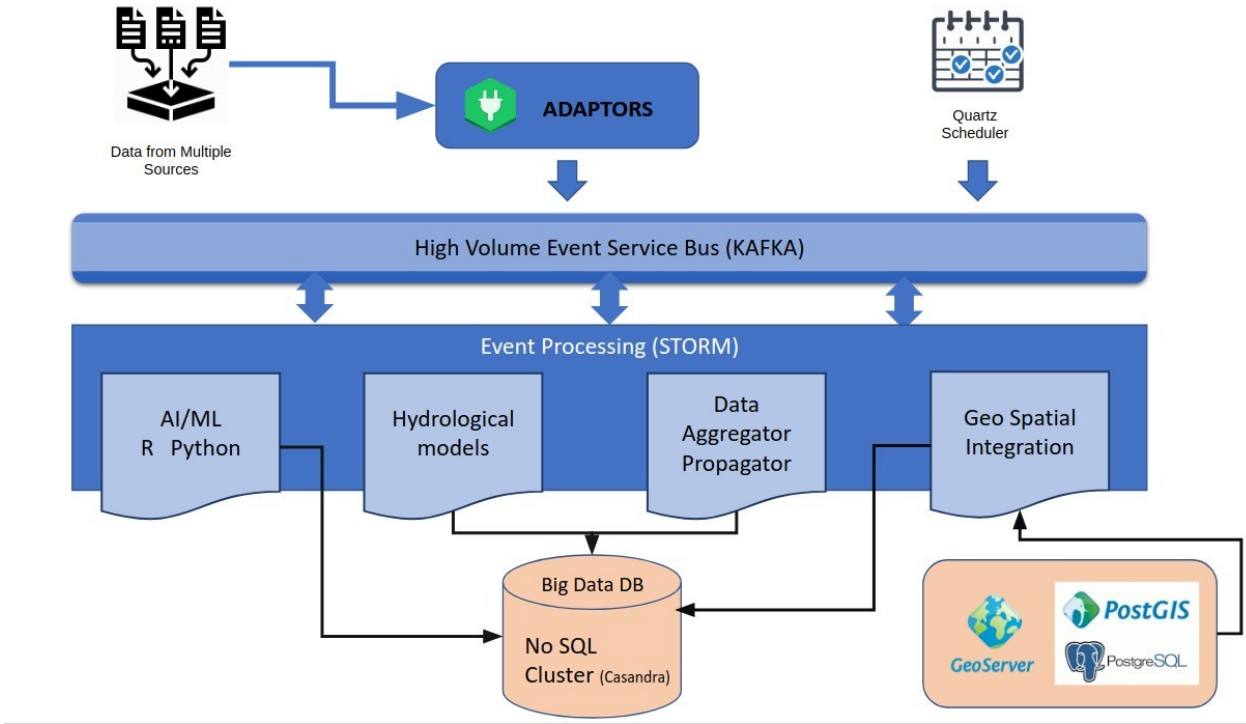


Figure 4.2: Backend architecture

5. Implementation

5.1 Scraping

Web scraping, is the process of extracting information from websites using automated tools. It involves writing code to visit a website, inspect its HTML structure, and extract the relevant data. The extracted data can be used for various purposes, such as research, analysis, designing or building applications. Web scraping can be done manually, but it is typically done automated tools such as web crawlers, which systematically browse the web and collect data using from multiple websites. When scraping data, it is important to adhere to ethical and legal guidelines, such as respecting the website's terms of use and avoiding excessive or disruptive requests, and obtaining consent if necessary. Additionally, it is important to ensure the quality scraped data, as well as to protect the privacy and security of any explicit and accuracy of the personal information that may be involved.

5.2 Aggregation

Aggregation is the process of combining multiple data points into a single summary value, either over time or over space, to gain a better understanding of the data. The types of data aggregations we are using in this project are spatial and temporal data aggregation.

5.2.1 Spatial Aggregation

Spatial aggregation refers to the process of collecting and combining data from lower-level geographic units into higher-level units. For example, in a village, data may be collected from various soilmoisture stations, sensors. This data can then be aggregated at the mandal level, which is a higher-level administrative division. The Mandal data can then be aggregated at the District level,Which is a higher-level administrative division from Mandal. The District data can then be aggregated at the State level,Which is a higher-level administrative division from District.

5.2.2 Temporal Aggregation

Temporal aggregation can also refer to the process of summarizing data over time periods, such as hourly to daily, monthly, and yearly intervals. This type of aggregation is often used in time-series analysis to identify trends and patterns over time. Hourly data can be aggregated into daily data by summing or averaging the hourly values for each day. Similarly, daily data can be aggregated into monthly data by summing or averaging the daily values for each month. Finally, monthly data can be aggregated into yearly data by summing or averaging the monthly values for each year.

5.2.3 API's

APIs (Application Programming Interfaces) are tools that allow users to access and retrieve information about data resources, such as soil moisture station data, time-series data, aggregated data etc. They can also to used to update meta data and upload several meta data of soilmoisture stations, and historical data. Overall, APIs are powerful tools that can help users access and utilize software, systems, and data in a more efficient and effective way, while also providing flexibility and customization options.

6. Testing Module

Unit testing: This involves testing of each individual function of or component of the module in isolation to ensure that they are working as expected. Unit tests can be automated and run frequently to catch any bugs or issues early on.

Integration testing: This involves testing how different components of the module work together, as well as how the module integrates with other systems or modules. This can help identify any issues that may arise from the interactions between different components.

Load testing: This involves testing the module under heavy loads or heavy traffic to ensure that it can handle the expected volume of data and user requests. This can help ito identify any performance issues or bottlenecks that may need to be addressed.

For Unit Testing,initially a new environment suitable for the working of the module is created in docker, technically we can say that we will create a suitable docker container and copy the productized soil moisture module into that container and test each and every functionality in that module as they are properly working or not. We also raise automatic events for the individual functionalities using Flink to test whether the module is able to handle or not.

7.Output

This module can provide accurate measurements of soil moisture content. The module can be integrated with other systems, such as weather monitoring stations or irrigation controllers, to provide a more complete picture of soil moisture conditions and automate irrigation schedules

A single, unified and an authoritative standalone module which can integrate into any other projects present in our company, which wants to use soilmoisture data. This can result in significant time savings for other projects and can also help to reduce the cost of integrating soil moisture measurements into their systems. Additionally, by offering a stand- alone and module, the organization can ensure consistency, accuracy in soil moisture measurements across different projects, further reducing the risk of errors or any inaccuracies. The use of a standardized, productized module can help improve the accuracy of soil moisture measurements, reducing errors and inconsistencies.

Reusability:A productized module can be reused across different projects and applications, saving time and resources in the long run.

Scalability: A productized module can be easily deployed across the different environments and integrated into different systems, making more scalable and versatile.

By productizing the soil moisture module, developers can focus on adding new features and functionality, rather than spending time building the underlying infrastructure. This can help drive innovation and allow developers to focus on more complex and advanced tasks.

8. Impact

Improved crop yields: The Soilmoisture data can help farmers optimize irrigation schedules and better manage water resources, leading to improved crop yields and increased productivity.

Water conservation: By providing accurate soil moisture data, the module can help reduce water usage and promote sustainable water management practices

Environmental sustainability: By optimizing water usage and reducing the need for fertilizers and other chemicals, the soil moisture module can help to promote environmental sustainability and reduce the negative impact of agricultural practices on the environment.

Cost savings: By reducing the water usage and improving crop yields, the soil moisture module can help farmers save money on irrigation and other inputs, leading to increased profitability.

Enhanced decision-making: The Soil moisture module can provide valuable insights into soil conditions, which can help farmers and water managers make more informed decisions regarding irrigation schedules, crop selection, and water allocation.

Overall, the impact of a soil moisture module can be significant, providing benefits to farmers, water managers, and the environment. By improving crop yields, promoting water conservation and environmental sustainability, and enhancing decision-making, the soil moisture module can help support sustainable agricultural practices and promote economic growth.

Overall, productizing the soilmoisture module can offer significant value to the other projects, enabling them to quickly and easily integrate soil moisture data and the measurements into their systems without the need for significant time and resources. This can help to accelerate the adoption of sustainable agricultural and water management practices, leading to improved crop yields, reduced water usage, and a more sustainable future.

9. Sample Screenshots

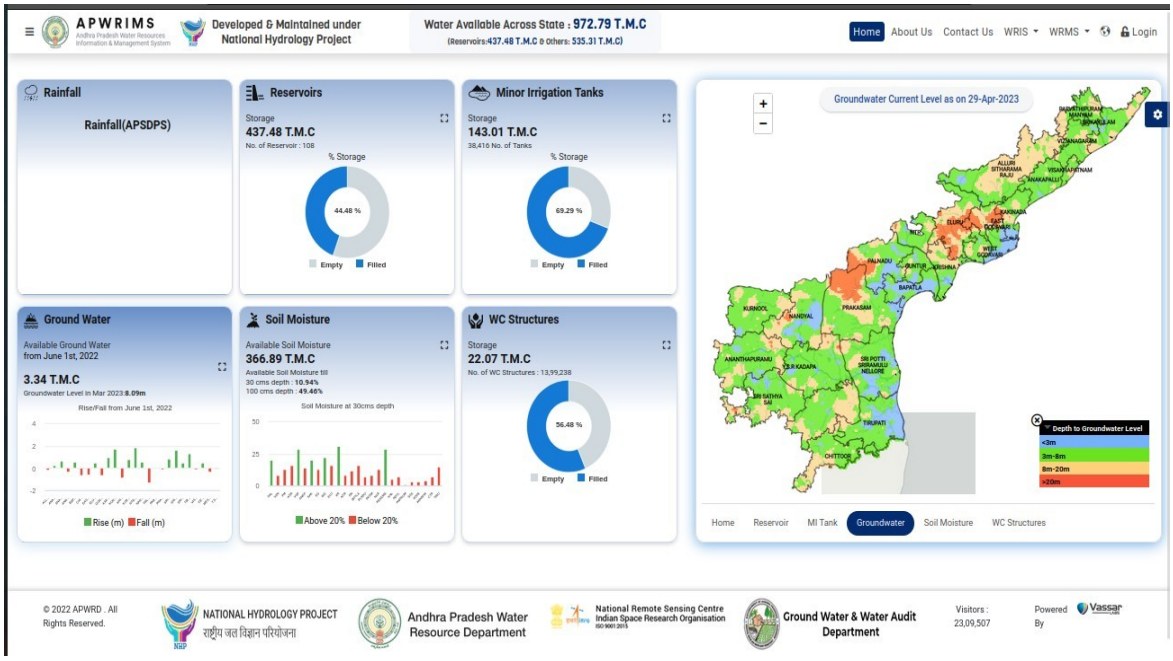


Figure 9.1 Home Page

Figure 9.2 Water Balance Page

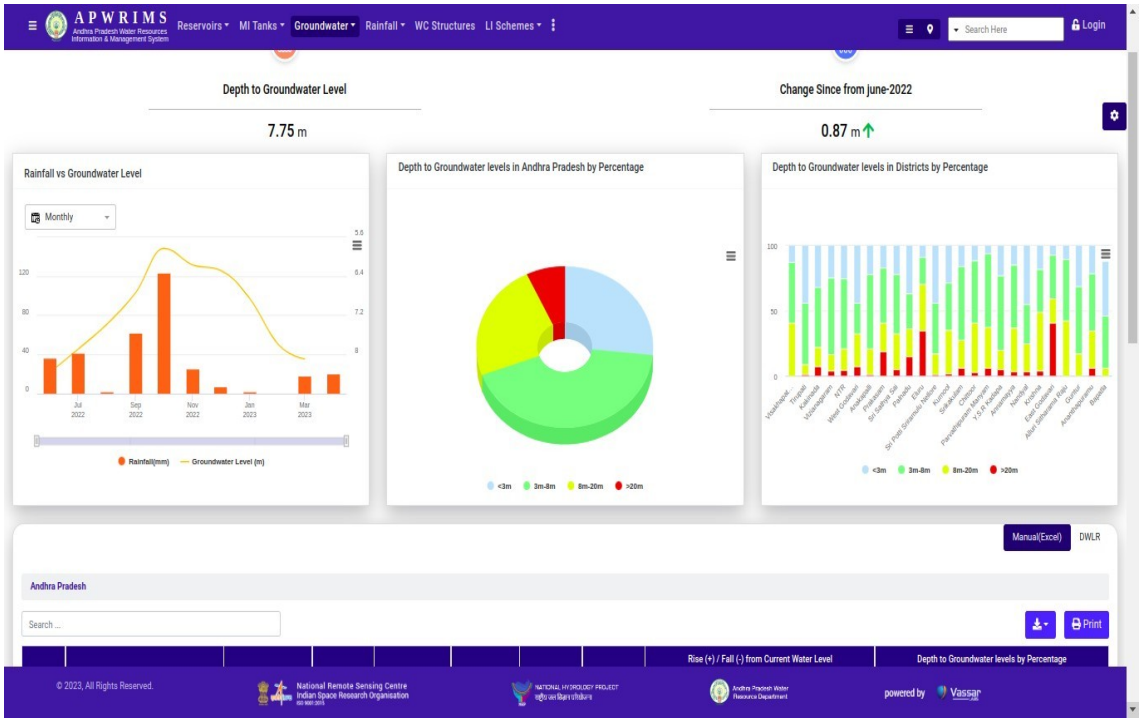
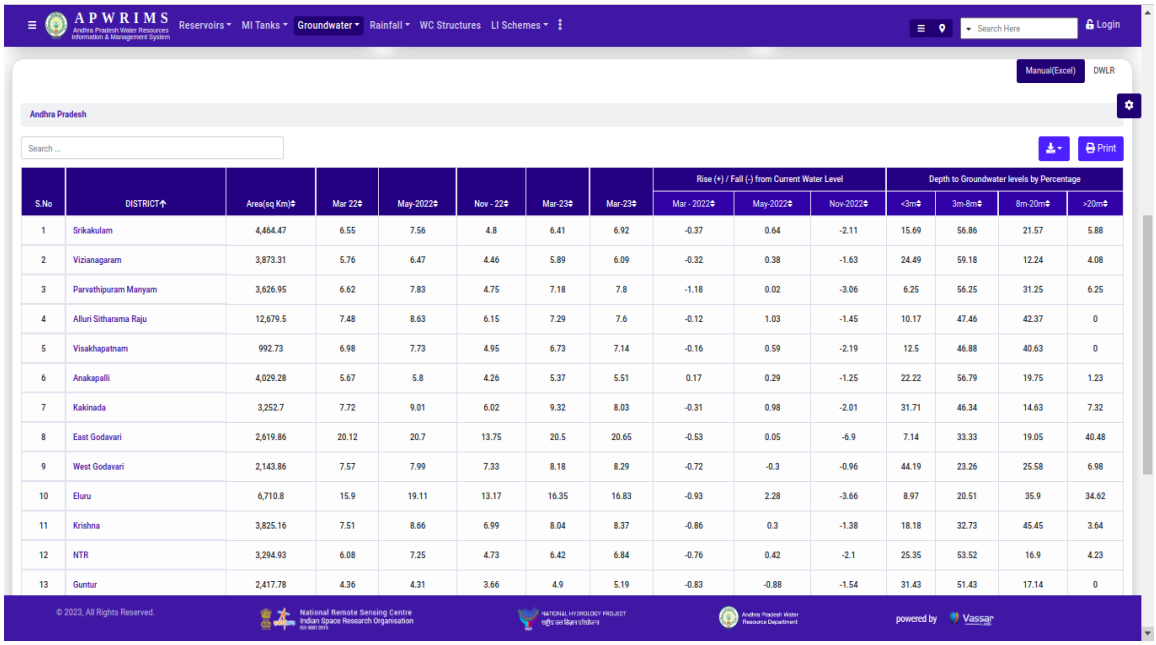


Figure 9.3 Dashboards Page



S.No	DISTRICT	Area(sq Km)	Mar 22	May-2022	Nov - 22	Mar-23	Mar-23	Mar - 2022	May-2022	Nov-2022	<3m	3m-8m	8m-20m	>20m
1	Srikakulam	4,464.47	6.55	7.56	4.8	6.41	6.92	-0.37	0.64	-2.11	15.69	56.86	21.57	5.88
2	Vizianagaram	3,873.31	5.76	6.47	4.46	5.89	6.09	-0.32	0.38	-1.63	24.49	59.18	12.24	4.08
3	Parvathipuram Manyam	3,626.95	6.62	7.83	4.75	7.18	7.8	-1.18	0.02	-3.06	6.25	56.25	31.25	6.25
4	Alluri Sitharama Raju	12,679.5	7.48	8.63	6.15	7.29	7.6	-0.12	1.03	-1.45	10.17	47.46	42.37	0
5	Visakhapatnam	992.73	6.98	7.73	4.95	6.73	7.14	-0.16	0.59	-2.19	12.5	46.88	40.63	0
6	Anakapalli	4,029.28	5.67	5.8	4.26	5.37	5.51	0.17	0.29	-1.25	22.22	56.79	19.75	1.23
7	Kakinada	3,252.7	7.72	9.01	6.02	9.32	8.03	-0.31	0.98	-2.01	31.71	46.34	14.63	7.32
8	East Godavari	2,619.86	20.12	20.7	13.75	20.5	20.65	-0.53	0.05	-6.9	7.14	33.33	19.05	40.48
9	West Godavari	2,143.86	7.57	7.99	7.33	8.18	8.29	-0.72	-0.3	-0.96	44.19	23.26	25.58	6.98
10	Eluru	6,710.8	15.9	19.11	13.17	16.35	16.83	-0.93	2.28	-3.66	8.97	20.51	35.9	34.62
11	Krishna	3,825.16	7.51	8.66	6.99	8.04	8.37	-0.86	0.3	-1.38	18.18	32.73	45.45	3.64
12	NTR	3,294.93	6.08	7.25	4.73	6.42	6.84	-0.76	0.42	-2.1	25.35	53.52	16.9	4.23
13	Guntur	2,417.78	4.36	4.31	3.66	4.9	5.19	-0.83	-0.88	-1.54	31.43	51.43	17.14	0

Figure 9.4 Tabular Data Page

10. Conclusion

In conclusion, the project aims to provide real-time insights of soil moisture. The system collect data from various sensors and websites, process it using advanced technologies such as Kafka Queue, Flink, Cassandra, and PostgreSQL, and generates dashboards, heat maps, graphs, and tables for visualization. The project use a wide range of technologies and also software tools. These technologies and tools have been chosen carefully to ensure that the system is scalable, reliable, and efficient. The project addresses several challenges in water resource management such as flood forecasting, canal management system, and water audit and budgeting. The system can help governments and organizations make data-driven decisions to manage water resources effectively.

Overall, the project has the potential to bring significant benefits to the water resource management sector by providing real-time insights, optimizing resource utilization, and improving water security. The project also has future scope for further development and expansion, such as integrating with other systems and adding new features to address emerging challenges in the field of water resource management.