MINISTRY OF ENVIRONMENT AND TOURISM

WATER AUTHORITY
GOVERNMENT IMPLEMENTING AGENCY

2030 Water Resources Group

HOSTED BY WORLD BANK GROUP Water

MONGOLIA

April 2021

DIGITAL WATER PLATFORM: DEVELOPMENT OF A GROUNDWATER MONITORING PORTAL USING DISRUPTIVE TECHNOLOGY

DISCLIAMER

This work is a product of 2030 WRG with external contributions. The findings, interpretations, and conclusions expressed in this work do not necessarily reflect the views of the World Bank, its Board of Executive Directors, or the governments they represent. The World Bank does not guarantee the accuracy of the data included in this work. The boundaries, colors, denominations, and other information shown on any map in this work do not imply any judgment on the part of the World Bank concerning the legal status of any territory or the endorsement or acceptance of such boundaries.

This report has been reviewed and accepted by the 2030 WRG Mongolian Multi-Stakeholder Platform.

World Bank Country Office
14241, Seoul street 4, MCS plaza
Ulaanbaatar city, Mongolia 14241
Phone: +976 70078235
Web: www.2030wrg.org
Email: ddechinlkhundev@worldbank.org
obazargochoo@worldbank.org

Prepared and Edited by:
Munkhtsetseg Zorigt
National Consultant, 2030 Water Resources Group, World Bank

Saulyegul Avlyush, PhD
Mongolia Partnerships Coordinator, 2030 Water Resources Group

Contributed by:
Mahesh Patankar
Senior Advisor, Disruptive Technologies, 2030 Water Resources Group

Odbayar Bazargochoo
Mongolia Partnerships Coordinator, 2030 Water Resources Group

BlobCity Inc.

KPM LLC

Led and Managed by:
Dorjsuren D, PhD
Mongolia Country Coordinator, 2030 Water Resources Group

Rochi Khemka
Global Partnerships Coordinator, 2030 Water Resources Group

Contents

List of tables

List of figures

Abbreviations

AUGWB - Altai Uvur Gobi Water Basin
EICIMH - Environmental Information Center, Institute of Meteorology and Hydrology
GW – Groundwater
GUDGWB - Galba Uush Doloodiin Gobi Water Basin
GDP – Gross Domestic Production
IT – Information technology
LLC - Limited Liability Company
MET– Ministry of Environment and Tourism
ML – Machine Learning
NAMEM - National Agency of Meteorology and the Environment Monitoring
RBA – River Basin Authority
RI – Republic of India
UGGKhWB - Umard Gobiin Guveet Khalkh Water Basin
WB – World Bank
WRG – Water Resources Group

1. Introduction

Groundwater is the primary source of water supply in Mongolia and accounts for approximately 82% of water use. As such, monitoring and assessing groundwater resources is essential for policy and decision making, compliance assessments, the use of preemptive measures for groundwater exploration, and protection of the resource.

Nonetheless, water demand in Mongolia has steadily increased due to population growth and economic development in Ulaanbaatar city, and mining activity in the Southern Gobi region. To ensure groundwater is managed properly requires timely and accurate data capturing, information dissemination, and stakeholder capacity building.

The Government of Mongolia requested 2030 WRG's Mongolia team to support the development of a groundwater monitoring portal using disruptive technologies for sustainable groundwater management in the country. With the endorsement of 2030 WRG's Multi-Stakeholder Platform Steering Board in Mongolia, the team initiated this project in September 2019.

The project aimed to develop a new cutting edge digital portal and dashboard for groundwater datasets using disruptive technologies. However, Mongolia suffered from a lack of groundwater data due to data confidentiality and security issues. Thus, the project was implemented as two sub-projects.

(1) The primary project started with anonymized data sharing from selected river basins of Mongolia, with the densest and reliable data for data analysis. This resulted in the development of a new groundwater dashboard created by 2030 WRG, with the support of the Department of Groundwater Monitoring, Water Authority, and Ministry of Environment and Tourism of Mongolia. The team also created a model for predictions of groundwater levels using machine learning techniques and Artificial Intelligence (AI).

(2) The second objective of the project was upgrading the dashboard of KPM LLC in Mongolia, using the prediction models developed under this project. The revamped portal under this project has allowed accurate monitoring and predictions of the current and future groundwater resources.

All the deliverables of the above projects were handed over to the key project partners, namely Water Authority and Ministry of Environment and Tourism of Mongolia on March 15, 2021.

This report provides (a) a background assessment of the prevailing situation of groundwater monitoring and data management in Mongolia, (b) the results and outcomes of the development of a groundwater dashboard and portal, and conversion of the existing groundwater monitoring system to the portal, and (c) recommendations for further development of a digital water platform and groundwater data management in Mongolia. In summary, the project has initiated and facilitated the development of a Groundwater Management Centre for sustainable groundwater management using disruptive technology.

2. Use of disruptive technologies in groundwater management in Mongolia

2.1. Background

Mongolia is susceptible to water stress and is projected to experience significant gaps in water supply and demand in two economically significant areas: Ulaanbaatar, the capital city of Mongolia, and the southern Gobi region, Mongolia's mining hub. In view of Mongolia's growing population and prominent—but water-intensive—mining sector, Mongolia's water demand is expected to exceed supply capacity by the year 2021 (2030 WRG hydroeconomic analysis, 2014).

Total water resources of Mongolia are estimated at 564.8 million m^3/year. The majority of the water comes from surface water (98.1%), including glaciers, and only 1.9 % comes from groundwater resources (MET, 2019). However, 82 percent of total water use was contributed by groundwater resources. Moreover, water use increased 2.3 times between 2013 and 2016, with close to 95 percent of such use permitted through groundwater resources (MET, 2017). Another important issue is the use of water for the mining sector in the Southern Gobi region. According to the 2018 Annual Report of MET, water use is maximum in the agriculture and livestock sectors (305 million m^3/year or 54% respectively) and mining contributing to 16% (90 million m^3/year). Amongst all the sectors, the mining sector is likely to increase its water consumption, with a high level of gross domestic product (GDP) growth expected through exports of minerals to other countries. As such, the mining water consumption, which is dependent primarily on the non-renewable fossil water needs specific attention.

In this context, a good understanding of the water availability, particularly in the southern Gobi region, is of prime importance. The current groundwater monitoring network in Mongolia has a total of 273 monitoring boreholes and it will be extended with another 170 new boreholes in 2050, as documented in Mongolia's long-term development policy vision document (Vision 2050). Therefore, 2030WRG initiated the groundwater portal and dashboard to create a predictive analytical tool for groundwater availability using disruptive technology.

2.2. Objectives of the project

2030 WRG specified the following objectives in the application of disruptive technologies for groundwater management in Mongolia:

- Assess the current level of groundwater through an IT-enabled platform using data from initial observation wells
- Suggest a robust data capture and transfer process to store and curate the data on a central server/ repository
- Explore correlations of groundwater levels with weather patterns, rainfall and other suitable independent variables
- Use machine learning tools to identify data anomalies and provide early warnings

- Model the groundwater data with a select set of independent variables
- Disseminate learnings from Mongolia to other 2030 WRG multi-stakeholder platforms

2.3. Stakeholder engagement

This project was initiated with data collection through the officials of the Ministry of Environment and Tourism and the Water Authority of Mongolia. An initial meeting was held with Ministry officials, including the service provider, and former and current groundwater monitoring experts, who shared the data from the previous groundwater monitoring phase. Consequently, the meeting enabled the stakeholders to identify the rationale and needs for the groundwater dashboard.

The discussion highlighted the need to create a general map of the groundwater portal, such as data collection and user right of access, and to enable predictions of groundwater availability through digital technology applications.

The project regularly engaged with stakeholders through 2030 WRG multi-stakeholder platform in Mongolia, with the final deliverables endorsed through the platform's Steering Board and workstreams.

3. Background on the groundwater monitoring system in Mongolia

At the start of the project, there were three main digital groundwater databases in Mongolia:

(1) Environmental Information Center, National Agency of Meteorology and the Environment Monitoring (NAMEM) contains some of the static data on the groundwater resource, generated through the Integrated Water Resource Management project implemented by Ministry of Environment and Tourism, and annual statistical data on groundwater resource and consumption collected from provincial environmental departments and river basins – www.eic.mn;
(2) Groundwater monitoring network portal, containing dynamic data, mainly collected by dataloggers – www.groundwater.mn;
(3) Integrated water database in Kharaa River Basin, containing multiple types of data for characterizing the Kharaa River Basin, and management of water resources within the basin.

Only the database of Environmental Information Center is open to the public, with the other two databases limited by access restrictions, due to confidentiality and data security issues, as open data licensing is not practised in Mongolia. However, dynamic groundwater data is captured in the groundwater.mn portal, including data from 273 monitoring wells. Also, this monitoring network is the main data source for the government, including the Water Authority, river basin authorities and experts, and groundwater specialists. Hence, the monitoring network provided the groundwater data for the project.

The purpose of this chapter is to review the groundwater monitoring network, from the collection of data to the dissemination of results.

3.1. Groundwater monitoring network (www.groundwater.mn)

The groundwater monitoring network www.groundwater.mn compromises 273 wells (Figure 1) for observational and abstraction purposes. At the national level, monitoring potable water supply is a high priority; thus, it is assumed that priority has been given to the potable water supply wells. At the river basin scale, the territory of Mongolia is divided into 29 river basins. Out of 29 river basins in Mongolia, ten river basins do not have monitoring points.

Among RBAs, Tuul River basin, Umard Gobiin Guveet Khalkh Water Basin (UGGKhWB), Galba-Uush-Doloodiin Gobi Basin (GUDGWB), Altain Uvur Gobi Basin (AUGWB) have been covered intensively with groundwater monitoring points, due to intensifying mining activities and groundwater demand. There are 42 water aquifers that have been identified in the AUGWB and GUDGWB, respectively. Approximately, more than 50 water aquifers have been identified in the UGGKhWB. Aquifers belong to Gobi-Sumber, Dornogobi and Umnugobi provinces generally have monitoring points.

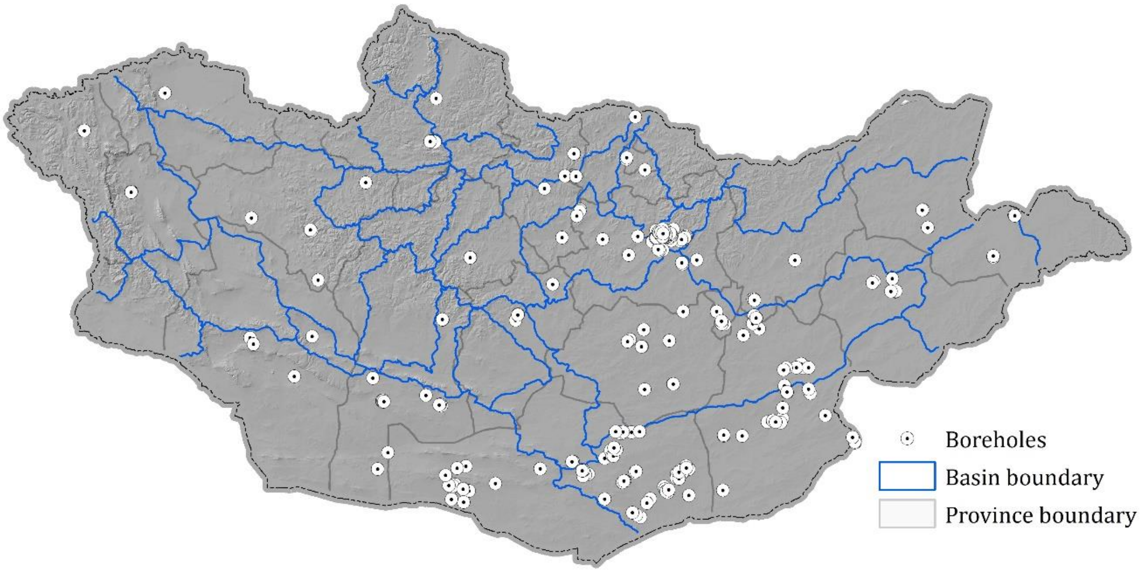


Figure 1. Groundwater monitoring network of Mongolia

At the national level, the monitoring network has covered all the province centers in Mongolia (Table 1).

Table 1. GW aquifers and monitoring boreholes of the provinces of Mongolia

№	*PROVINCE*	*GROUNDWATER AQUIFERS*	*NUMBER OF MONITORING WELLS*
1	*ARKHANGAI*	*6*	*2*
2	*BAYANKHONGOR*	*24*	*13*
3	*BAYAN-ULGII*	*3*	*2*
4	*BULGAN*	*9*	*4*
5	*DARKHAN*	*2*	*3*
6	*DORNOD*	*8*	*8*
7	*DORNOGOBI*	*14*	*31*
8	*DUNDGOBI*	*17*	*11*
9	*GOBI-ALTAI*	*13*	*4*
10	*GOBISUMBER*	*2*	*6*
11	*KHENTII*	*7*	*7*
12	*KHOVD*	*7*	*3*
13	*KHUVSGUL*	*6*	*4*
14	*ORKHON*	*2*	*2*
15	*SELENGE*	*4*	*1*
16	*SUKHBAATAR*	*10*	*8*
17	*TUV*	*6*	*13*
18	*ULAANBAATAR*	*14*	*73*
19	*UMNUGOBI*	*25*	*68*
20	*UVURKHANGAI*	*6*	*4*
21	*UVS*	*5*	*1*
22	*ZAVKHAN*	*4*	*5*

3.2. Monitoring parameters and data frequency

At present, the groundwater monitoring database is developed to contain the following types of data:

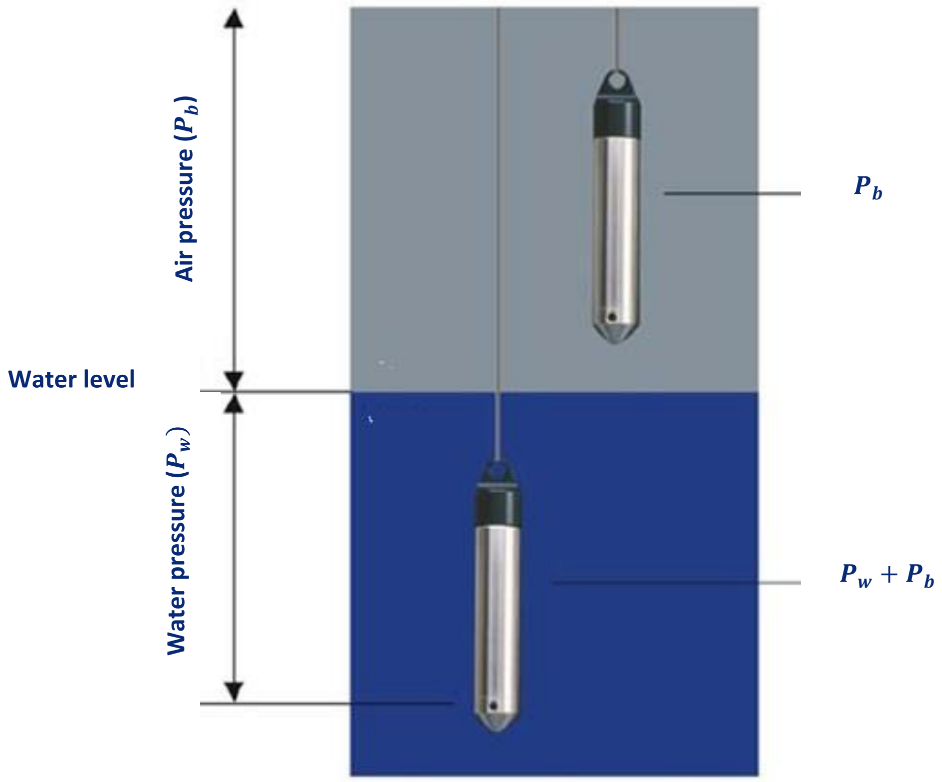
- Groundwater level, pressure, temperature, hardness, total dissolved solids and conductivity in time series;
- Geological data, such as borehole logs as static one-time entry for some wells;
- Hydrogeological data, such as depth to water surface, as static data;
- Construction data, such as casing, piezometer, and borehole development photos, as static, one-time entry;
- Type of equipment, such as pumping equipment and records, as static, one-time entry;
- Geophysical data, including surface and borehole measurements, as static, one-time entry.

During the data collection process, it became clear that many of these wells have no known construction information. Without such information, it becomes difficult to assess the spatial and vertical monitoring coverage of hydrogeological units and groundwater zones. The essential well log data and zones of completion for wells monitored in the RBAs should be incorporated into the database. This information will be used in many of the data evaluations, including zone-specific queries that have been designed and used to illustrate current groundwater conditions.

The frequency of these data is hourly (2-4 hours) and measurements have been launched since 2014 for the monitoring wells.

3.2.1. Groundwater level and its measurement

Conceptually, to measure groundwater level, the logger of the water pressure sensor is installed in the wells. There are two main types of loggers installed for the groundwater monitoring network of Mongolia, including Diver loggers of 'Van Essen" for 166 wells, and "Hobo" loggers of "Onset" installed for 66 monitoring wells. Depending on the type of equipment, the logger measures the water pressure. Variation of the air pressure may influence the measurements. It is recommended, therefore, to install barometer logger (Figure 2).



Source: Van Essen, 2009

Figure 2. General overview of GW level measurements

The collected data from the loggers are recorded as absolute pressure. This means that the pressure sensor not only measures the water pressure, but also the air pressure pushing on the water surface. Thus, air pressure should be subtracted from the absolute pressure of water to reduce the variation of the air pressure effect. Then water pressure can be converted into water level related to the vertical reference datum.

$$P_w = (P_b + P_w) - P_b$$

3.2.2. Water quality data

Water quality data in the wells are variable and insufficient in the current monitoring network. It should be noted that there is no mechanism to upload water quality data into the monitoring network.

3.2.3. Climate and water consumption data

Climate and groundwater abstraction are valuable information for groundwater monitoring and assessment. The current monitoring network has no linkage to climate data. A few of the well loggers were installed with precipitation loggers, but they are not operative. Similarly, groundwater use is not linked to monitoring wells.

3.3. Brief analysis of current situation of GW in Mongolia

We briefly analyzed data from 102 wells from the total 273 monitoring points in Mongolia (Figure 3) for the two regions of South Gobi and Ulaanbaatar, where the highest density of groundwater monitoring wells are located with the longest data records.

The monitoring points were located in the Gobi region at wells with depths ranging between 25-320 m. The depth of only one well was 24.8 m, depths of 29 wells were in the range of 33-93.5 m, depths of 28 wells were in the range of 100-188 m, depths of 5 wells were in the range of 204-243 m, and the depths of 3 wells were in the range of 290-320 m. The selected 44 wells in the Southern Gobi region were in the deep zone up to 300 m below the ground. For the Tuul river basin in the Ulaanbaatar area, a total of 84 monitoring well are registered in the network, of which 16 are non-operative. Of these, 58 wells were selected in the shallow and intermediate zone up to 88 m below the ground surface. The depths of 49 wells is in the range of 8-47 m and depths of 19 wells is in the range of 60-88 m.

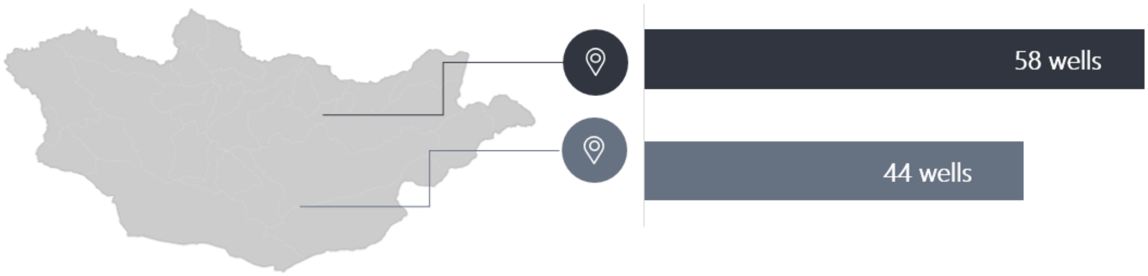


Figure 3. GW monitoring boreholes used for the analysis

The groundwater data analysis conducted for the two selected regions is explained in the following sections.

3.3.1. South Gobi region

In the data analysis phase, the Umnugobi province has been selected due to the density of the network of groundwater monitoring points. Four river basins, UGGKhWB, GUDGWB, AUGWB and Ongi river basin are located in the Umnugobi province territory. Umnugobi province monitors groundwater levels through an extensive network of 67 monitoring points that belong to GUDGWB and AUGWB. Only 44 of them were usable for this analysis (Table 2). The number of monitoring points with information for about 4 years (2015-2019) were 44, with 19 of them were newly established in 2018-2019, approximately. Dataloggers of four monitoring points seemed to have stopped functioning.

The network monitoring points have been classified to the extent possible by zone of completion and location within a subbasin. Monitoring points in the network include: no monitoring points in the shallow zone, 23 monitoring points in the intermediate zone, and 21 monitoring points in the deep zone, ranging between 24-320 m depth, of the selected 44 wells. The median depth was 101.7 m, which is selected as a delineating border between

intermediate and deep zones. The depths of the first and second quartiles were 60.3 m and 132.5 m, respectively. 2030 WRG did some analysis on existing data gathered since 2015 for South Gobi Area and the results are shown below.
In Umnugobi province, a total of 39 aquifers have been identified, consisting of 21 aquifers belong to Altai Uvur Gobi Water Basin and 18 aquifers belong to Galba Uush Doloodiin Gobi Water Basin. Based on hydrogeologic maps and mining license maps, a pattern of monitoring points emerged. In addition to monitoring points of potable water supply for soum or province centers, monitoring points have mostly utilized existing wells along with known groundwater aquifers, which have been surveyed largely for mining exploration.

Table 2. Information of the monitoring boreholes in the South Gobi region

Well location	Total number of wells	Number of wells selected	Installed new	Online wells	Active online wells
Tsogttsetsii	12	6	6	-	
Khankhongor	7	3	4	-	
Bayan-Ovoo	10	5	5	2	2
Khanbogd	19	16	1	3	2
Nomgon	2	1			
Noyon	1	1			
Khankhongor	2	-	2	2	2
Bayandalai	2	2		1	1
Gurvantes	12	10	1		
Total	*67*	*44*	*19*	*8*	7

Data for more than 70% of the selected wells are incomplete and inconsistent. For instance, Figure 4 shows the there is incomplete historical data for well 52 (namely GUD 52) with a gap for the period June 2018 to August 2019 (Figure 4).

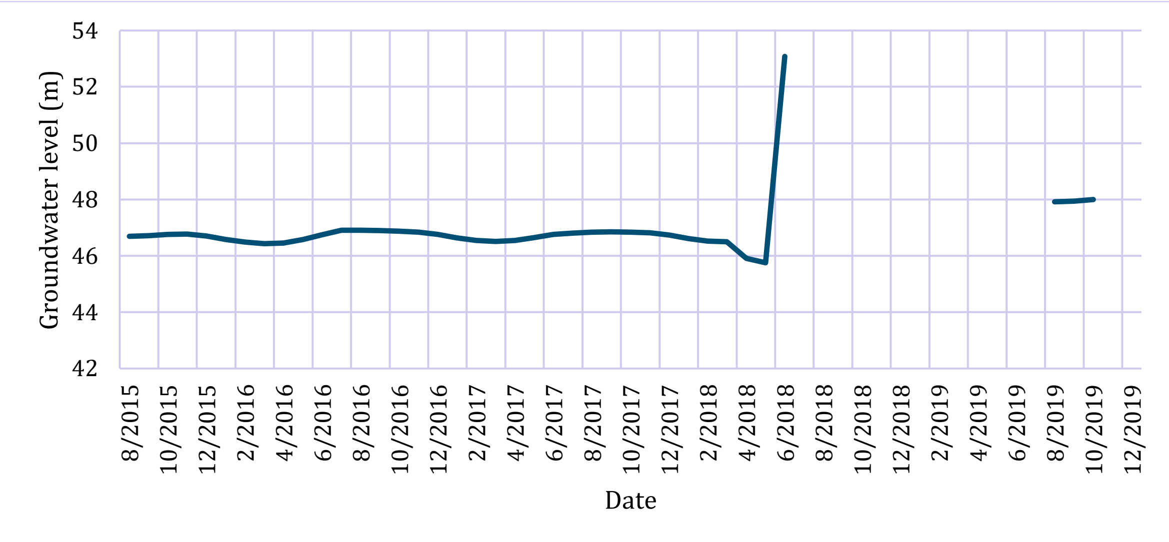


Figure 4. Inconsistency of the groundwater level of well 52 in the Galba Uush Dolood Gobi basin

The boxplots of the water level of the wells show that variations of the water level are low. However, there are still certain anomalies present (Figure 5).

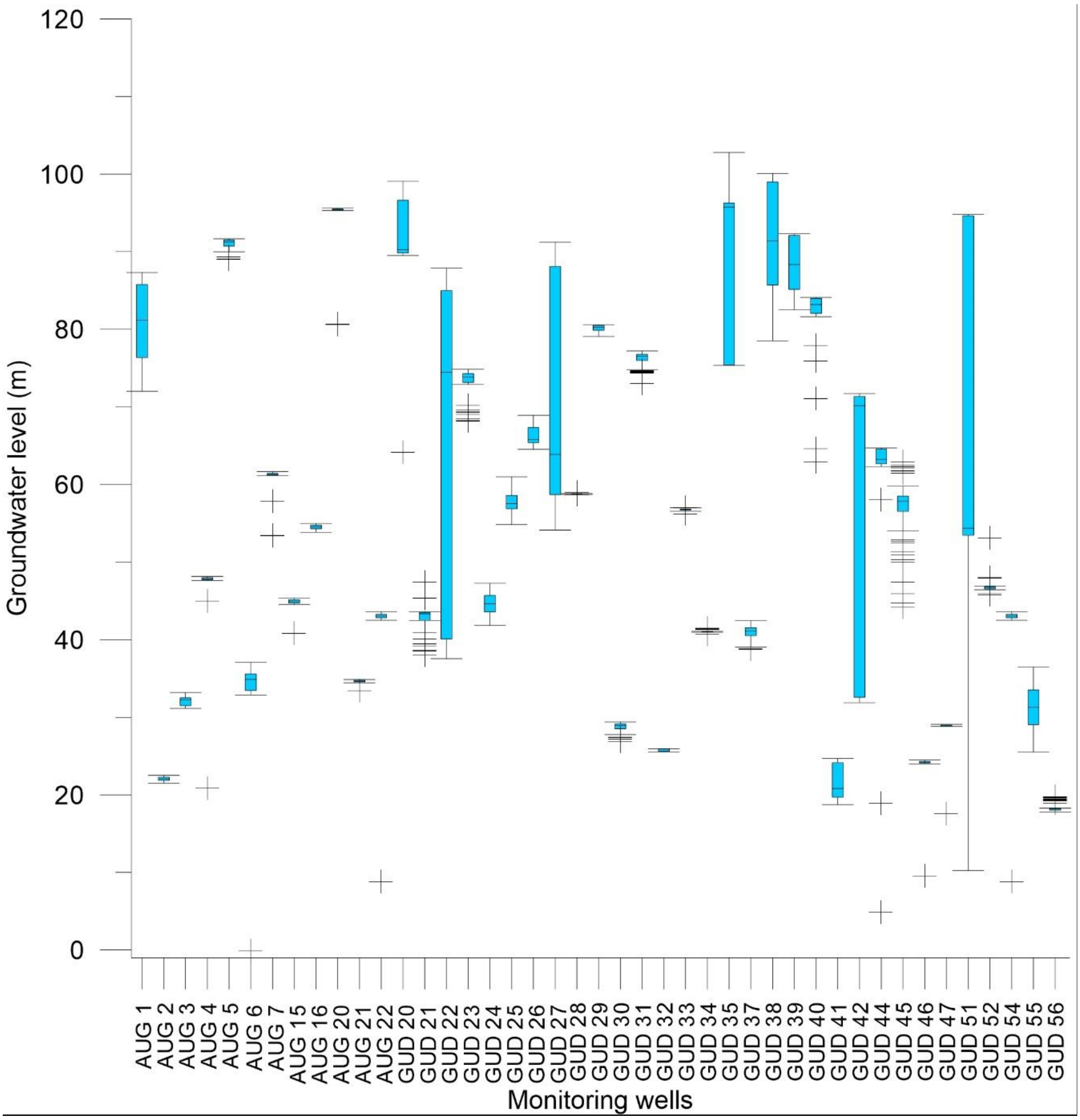


Figure 5. Groundwater level variations of the monitoring wells in the South Gobi region

Trend analysis for the groundwater levels of the monitoring wells in the Umnugobi province is illustrated in Table 3. Three monitoring wells are located near soum centers, four wells are located within the influence zone of a mining license area and the rest of the wells are in areas with little anthropogenic impact. Two monitoring wells are in the influence zone of soums and have a constant decreasing trend. The wells located in the influence zone of a mining license have an increasing trend until 2017 and a decreasing trend since 2017, plotting a downward parabola. The details have been discussed below (Table 3).

Table 3. Trends of the GW level at the monitoring boreholes in the South Gobi regions

Well location	Monitoring well description	Groundwater level trend description	Groundwater level trend	Linkage with precipitation and other factors

Noyon	There is only one monitoring well (100.1 m) at deep zone. It can be said that the monitoring well is located in natural settings with minimal anthropogenic effect.	The trend of groundwater level is a constant increase by 11.1 m between 2015-2019.	⇧	No impact - consistent with the general trend of precipitation of this period.
Sumber-Ovoot	There are two wells at intermediate level (50.3 m) and one monitoring well at deep (100.6 m) zones. It can be said that the well is located in the natural zone with minimal anthropogenic effect.	The trend of groundwater level is a constant decrease by 0.28-0.70 m between 2015-2018, and an increase by 0.02-0.42 m in 2019 compared to 2018. The deep monitoring well has a similar pattern of a decrease by 1.36 m between 2015-2018 and an increase by 0.16 m in 2019.	⇘ ⇗	Seasonal fluctuation of intermediate groundwater monitoring wells followed by seasonal fluctuation of the precipitation trend, but the general trend has been different. The precipitation did not affect the deep well based on this snapshot of data.
Nariin Sukhait	There are three monitoring wells at shallow (33.6 m) and one monitoring well at deep (70.5 m) zones. All of the wells are located in or near the mining license area.	Except for one monitoring well, two wells at intermediate shallow wells and on deep well have a trend of water level increase by 0.15-0.95 m between 2015-2017 and decrease by 0.25-1.17 m in 2017-2019. One monitoring well has a constant increase of 3.2 m in 2019.	⇗ ⇘	The trend could have been more likely to be influenced by mining activities rather than precipitation trends.
Tooriin shand	Two monitoring wells at deep (101.5 m) zone are located in natural settings.	The trend of change in the groundwater level is a decrease of 0.68-0.93 m between 2015-2019.	⇘	The precipitation might have influenced the seasonal groundwater level fluctuations, yet it needs further

				studying. The general trend in groundwater level does not match the precipitation trend pattern in Gurvantes soum.
Gurvan tes	One monitoring well at deep (101.7 m) zone is located in the influence zone of the soum center.	The trend of the groundwater level is a decrease of 1.49 m. It should be noted that the datalogger level and groundwater level do not match closely.	↓	No sign of effect of precipitation.
Bayandalai	One monitoring well at intermediate (40.7 m) and one well at deep (86.6 m) zones are located next to the soum center.	Due to the lack of sufficient data in 2017, and a total lack of data in 2018, it is impossible to say anything about the trend in the deep well monitoring point. In addition, it should be noted that the datalogger level in deep well and groundwater level in deep well is not exactly matched. The intermediate monitoring well point has a groundwater level trend of decrease by 0.31 m between 2015-2019, if one day data on Aug-18, 2019 is excluded.	↘	The groundwater level at intermediate monitoring well has shown a seasonal fluctuation with an almost half year lagged behind precipitation. Compared with only precipitation data, it can be said that there is anthropogenic impact besides of precipitation variation pattern.

3.3.2. Tuul river basin

Tuul river basin is the most crowded basin of Mongolia with the capital city, Ulaanbaatar, located in the basin, hosting over half the national population. The basin area covers 15,059 km^2 with a length of 1.2-4 km. The water supply for Ulaanbaatar is provided by groundwater from 10 wellfields in the alluvial, diluvial and proluvial deposits (Dorjsuren et al 2015). There are 84 monitoring points in the Tuul river basin, of which only 58 were accessible during the project. These monitoring points are located in the shallow (0-50 m) and deep zone (>50 m). The median depth was 30 m. The depths of the first and third quartiles were 22.9 m and 45 m, respectively. In this report, 58 wells in Ulaanbaatar in the Tuul river basin were selected and analyzed for groundwater level data between the period of 2015 and 2020 (Table 4).

Table 4. GW monitoring boreholes in the Tuul river basin

Wellfields	Total number of monitoring wells	Operative	Non operative	Wells with online logger
Deed	12	8	4	2
Tuv	15	15	-	1
Gachuurt	8	6	2	1
Makh Kombinat	3	3	-	
Uildver	10	9	1	1
Nisekh	3	3	-	1
Yarmag	4	3	1	1
Uliastai valley	4	2	2	1
Others	25	14		
Total	**84**	**74**	**10**	**8**

Data for more than 71% of the selected wells is incomplete and inconsistent. For instance, Figure 6 shows there is incomplete historical data for well 13 (namely Tuul 13) with a gap for the period September 2016 to May 2017 and since June 2019.

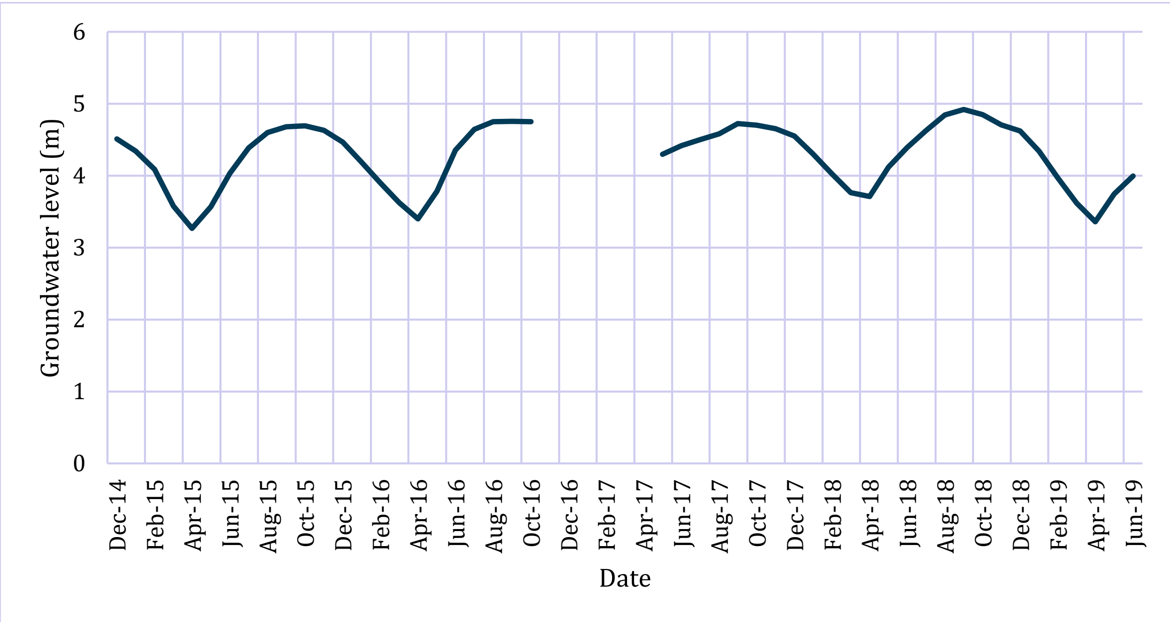


Figure 6. Inconsistency of the groundwater level of Well 13 in the Tuul river basin

Boxplots of the groundwater level show higher variations but with fewer outliers than the South Gobi region (Figure 7).

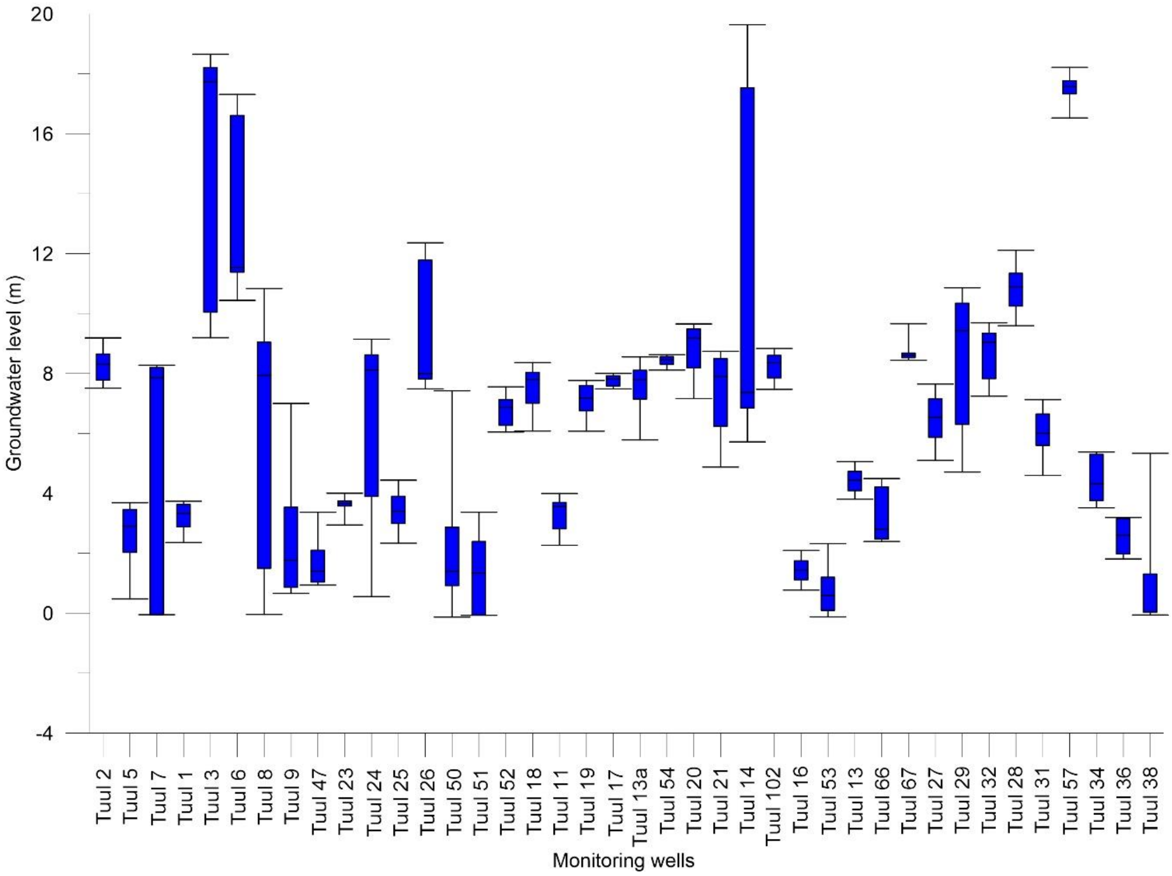
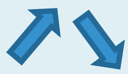


Figure 7. Groundwater level variations of the monitoring wells in the Tuul river basin

Groundwater level data is captured for 12 wells with data ranging from January 2015 to June 2020. The result of the trend analysis shows that most of the wells have seasonal fluctuations and are weakly correlated with precipitation (Table 5).

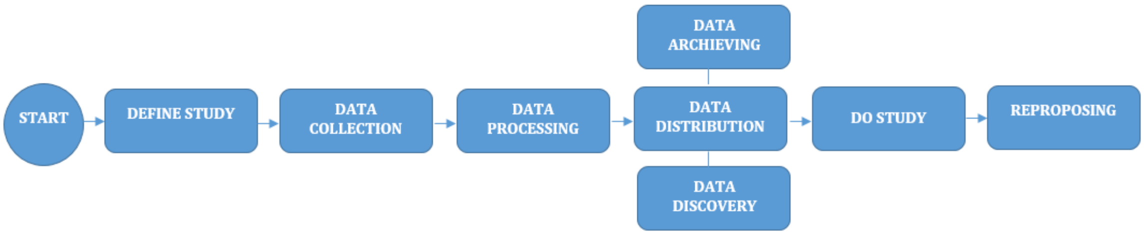
Table 5. Trend analysis for the monitoring boreholes with consistent data between 2015-2020 in the Tuul river basin

Wells location	Monitoring well description	Groundwater level trend description	Groundwater level trend	Linkage with precipitation and other factors
Deed	One monitoring well is in the shallow zone (38 m) under the anthropogenic impact.	The trend of Groundwater level has increased between 2017-2018, and decreased in 2019.		No sign of effect of precipitation.
Tuv	6 wells are located in this wellfield 14.5-29 m from	There is a constant trend for the 2 wells		There is a weak correlation between

	the ground surface. Three of them are under the natural condition and rest are under the anthropogenic impact.	(12,20), but third well (13) has decreased by 0.2 m.	↘	groundwater levels and precipitation and 6 wells have seasonal fluctuation.
Gachuurt	Total of 3 monitoring wells are located in the shallow zone (20-50 m) under the anthropogenic impact.	Base on groundwater level data between 2015-2020, one well's water level has decreased and two wells are constant.	↘	No sign of effect of precipitation.
Uildver	Two monitoring wells are located in the shallow zone under the anthropogenic impact.	One well's groundwater level has decreased between 2015-2017 and increased over 2018-2020. Another well's water level has fluctuated seasonally since 2018.	↘ ↗	No sign of effect of precipitation, but seasonal variations noted.

3.4. Analysis for groundwater data management and recommendations for further development

To achieve the objective of the groundwater monitoring portal, additional tools such as data input, initial processing, validation, reporting and visualization are needed. From a broad perspective, every process involving data can be a part of data management. Hence, the data management model adapted from Data Documentation Initiative (Thomas et al. 2009) and as defined by Fitch, P. et al. (2016) is depicted in Figure 8, showing how each step can been applied to groundwater data management.



Source: Thomas (2009)

Figure 8. Data management

- Define study: Define objectives and data requirements to initiate data management. Depending on the objective, types of data, accuracy, interval could be defined in more detail.
- Data collection: Collecting data as per objectives and requirements as mentioned in the previous step.
- Data processing and validation: Data is preprocessed into appropriate resolutions and formats such that it is suitable for further steps along with quality assurance and validations.
- Data archiving: Data is archived in either a central place or distributed places with a possibility to data exchange.
- Data distribution: Either publicly or user-controlled access can be implemented.
- Data discovery: User can answer a question with the help of data.
- Do study: Depending on whether the user achieved a goal or not in this stage, further development ideas or subsequent repurposing data can arise.
- Followed by above-mentioned concept, data management analysis and a set of recommendations have been produced in the next subsections.

3.4.1. Data collection

Data collection in Mongolia is a multi-faceted issue, characterized by institutional, legal, cultural, and technical dimensions. Technical issues are easier to identify and correctable. The common technical issues are:

- Groundwater measurements are made incorrectly in the field, such as erroneous reading of level gauges;
- Field technician fails to calibrate gauges or dataloggers;
- The data collector does not perform a data quality evaluation;
- Errors are introduced into the data as it is imported into whatever method is being used for its storage. Typical errors include transcription errors through manually inputting data and unit conversion errors or decimal mis-conversion or entering;
- Erroneous data is recorded due to datalogger glitch or malfunction;
- Testing data is included in the data recordings.

Several of the above mentioned issues have been identified in the groundwater monitoring network collected through dataloggers. Through brief analyses, the dominance of the issues has been identified (Figure 9).

- Data recordings have been stopped
- Erroneous data have been recorded
- Datalogger recorded data during downloading data process
- Testing data is included in the data recordings

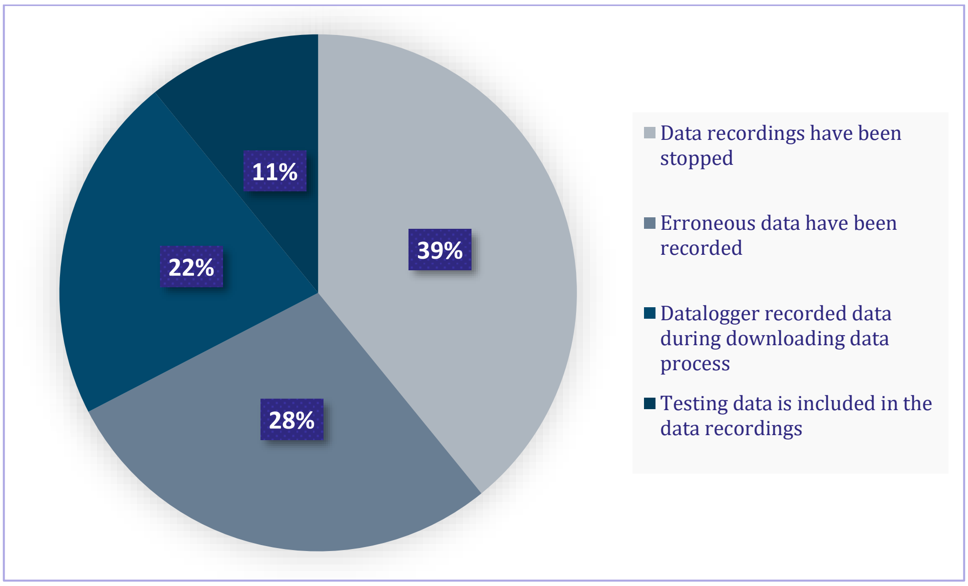


Figure 9. Erroneous data records

It should be noted that (1) the source of the issues was not always clear, and the erroneous data could have been introduced through a combination of more than one reason, and (2) certain issues cannot be identified unless there is a written document on standard operating procedure and logs.

Another important issue to be noted is the possibility of sensor depth change during data collection and the logger relaunching process. On several occasions, suspicious changes have been observed in groundwater level time series data without any apparent reason. It could be either substantial abstraction over a period, or change in logger depth. Yet, it was indiscernible due to lack of abstraction data. The most common issue was the termination of recordings of a parameter, which is mainly groundwater hardness data. The second common issue was the records of erroneous data, mostly due to glitches in the dataloggers. However, it is easily amendable either manually or automatically. Third common issue which could be also easily amendable is records of data during data downloading and logger relaunching process. Again, the erroneous data can be easily discerned from the general pattern of data and amended via manual or automated processes. Finally, the test data also can be easily discerned based on technician's logs and easily amendable via manual data processing.

For instances, a total of 46 errors have been noted in the groundwater time-series data of 44 monitoring points in the South Gobi region. The ubiquity of the errors in the database strongly indicates that appropriate data handing and data validation should be in place. In other words, the data collector does not perform a data quality evaluation.

3.4.2. Data quality evaluation

Primary validation: This type of validation may be performed by visual inspection of the data and is intended to catch errors. Some examples in the case of groundwater level data are:

- A recorded depth to water that is greater than the depth of the well;
- Errors due to a change in the height of the measuring point;
- Errors due to assigning the groundwater level recordings to the wrong well or screen due to careless labeling while entering a file.

Secondary validation: The purpose of this step is to determine whether the selected values are errors or natural extremes. Simple statistical methods are available to check whether apparently erroneous values are statically likely.

Examples of validation steps for groundwater level measurements:

- Contour maps of groundwater level data for a certain period or date can be used to identify outliers. Outliers show up as a high levels of contours around the location of the concerned well. The values causing the supposed discrepancies in the regional contour map should be checked.
- Preparing time series graphs with multiple hydrographs and the visual inspection of the graphs may show erroneous measurements. Comparing hydrographs would normally show similar trends and fluctuations, and if not, may indicate erroneous groundwater level measurements.
- Some simple statistical methods are available to check groundwater level measurements e.g. deviation from the mean or the median. Data series can be checked for values which differ more than three times the standard deviation from the mean or median.

The groundwater database has manual and automated data correction capacities. The automated data correction capacity is based on moving average technique. In general, it would be desirable to have more than one technique available for selection for automated error correction.

Tertiary validation: Tertiary validation involves advanced techniques for analysis and validation of spatial and temporal data. This includes, for instance, advanced statistics and comparison of different data types. Advanced statistical methods can be used for detection of outliers. Parametric and non-parametric statistical trend detection techniques have been developed for time series to provide most likely estimates of the water level changes over time and the corresponding confidence interval.

3.4.3. Data storage

At present, no historical data has been entered or stored into the groundwater database. Yet, it is advisable to adapt a consistent Electronic Data Format (EDF) or other standards, such as those of the Open Geospatial Consortium (OGC), the International Organization for Standardization (ISO) and World Meteorological Organization (WMO). Inconsistencies may require alignment of the data at five levels: systems, syntax, structure, semantics and pragmatics. At present, the only minor issue is the deployment time of the dataloggers, which have not been consistent. Mostly, dataloggers have an option to start at a specified date/time in the future and those options can be utilized in order to solve this type of issue.

It is advisable to have meta-data corresponding to the data anticipated in the future. There is a significant international adoption of the ISO/TC2116 standards. Most of the data are available from RBAs. However, the groundwater monitoring database lacks majority of the above-mentioned data. Ideally, all of the information listed above should be stored electronically and in a central repository. However, storing the above data in a digital database requires a major effort, which may not be feasible in the first phase of groundwater assessment. In the case of limited budget, it is recommended to give high priority to electronic storage of groundwater variables.

Unless the data is collected by a sole collector, and handled, processed, and stored by the same person, the best practice is to have standard operating procedures (SOP), which is a manual describing the protocols for collecting and handling data in the field and in the office/lab settings. It ensures uniformity, consistency, and completeness of data sets. The SOP will enable technical personnel to understand and replicate the work at hand and force them to conduct complete and thorough surveys. SOPs should be established at the outset of an investigation to establish that work has been conducted to a consistent and scientifically defensible standard.

The SOP needs to contain several elements: 1. Step-by-step guidelines, and 2. Field notes (e.g. date/time of datalogger data collection and re-launching, or battery replacement etc.). In addition to enforcing the collection of specific data elements, the data repository (database or GIS) should contain mandatory fields that match those on the field data collection sheets. By enforcing the mandatory nature of these fields, the data cannot be saved into the system, unless they are completed. Other basic steps to maintain the integrity of the data include taking succinct and clear field notes and permanently archiving them. SOP is essential for quality assurance.

3.4.4. Data processing

The processing of data should allow for interpretation of the groundwater resource. Examples of the type of data processing that would be applied to the database include the following:

- Common ways of presenting groundwater level data by means of contour maps (spatial image) or hydrographs (time series).
- Water level elevation with respect to a reference level, usually mean sea level
- Single date or average value over a period.
- Care should be taken that only measurements from the same time or period and same hydrostratigraphic unit are used; otherwise misleading results will be obtained.
- Provide a means of relating the impact of natural and human influences to the groundwater resource.
- Plot precipitation on the secondary axis to provide a visual observation of the response of the groundwater resource to this event.
- Abstraction rates can also be helpful when interpreting groundwater hydrographs due to potential anthropogenic influence on the environment.

3.4.5. Data reporting and presentation

The current groundwater monitoring network has a wide range of functionality in terms of generating data reports, including:

- Graphical representations of data, comparing several parameters of one monitoring point for the selected timeframe
- Graphical representations of data, comparing several monitoring points for the selected parameter for the selected timeframe
- One-time well construction report
- Data input report
- Log of user access etc.

4. Development of the groundwater dashboard (www.iucn.blobcity.com)

4.1. Process description of the groundwater dashboard development

Visualizing the groundwater monitoring data and managing it from a web-based portal or dashboard is valuable and timely information for stakeholders. Based on the analysis of the existing groundwater monitoring system described in the previous chapters, the 2030 WRG team identified a need to develop a groundwater dashboard for a graphical illustration of groundwater and related data. Moreover, current monitoring wells will be extended by about 170 new boreholes in 2050 towards Mongolia's long-term development policy vision 2050, with machine learning applied for predictions. In order to achieve this objective, two main issues needed to be addressed:

- The first one was sharing the location of wells. The location of the wells is currently not shareable and is kept confidential according to Article 13.1.5. of the Law on State and official secrets of Mongolia.

To solve this problem, the project collaborated with government officials and the Water Authority, and agreed to share anonymized data. The coordinates of the location of the monitoring wells were not shared with the project partners and members. Only the time-series of groundwater data was anonymized and coded for use.

- Secondly, the project could not be implemented for all 273 wells at the national level due to legal issues and data availability. As such, priority basins for the dashboard and prediction analysis needed to be selected.

Through agreement with the Water Authority, the project selected two regions, namely Tuul river basin and South Gobi region.

A new dashboard was developed by 2030 WRG with the support of BlobCity Inc. for the groundwater monitoring of Mongolia. The first step was building the data management concept to a new dashboard based on a review of the current groundwater monitoring network. The next step was to design and develop the dashboard with five sections including (1) Single well analysis, (2) Single well deep dive, (3) Multiple well analysis, (4) Anomaly detection and (5) Prediction.

The next step was to upload groundwater metrics to the dashboard. A total of 44 wells' metrics were taken from the monitoring network www.groundwater.mn. The selected metrics of data recorded on the monitoring wells included: groundwater pressure, temperature, conductivity and pH. In addition, climate metrics like air temperature, precipitation and wind speed from the closest station were gathered. All data for period from 2014 to 2019 is covered and uploaded to the dashboard. Users can see the plots of groundwater metrics for the selected time duration.

The final stage of the project developed a prediction model using machine learning techniques for predictions of the groundwater level on the dashboard.

4.2. Statistical methods and disruptive technology used for the dashboard

4.2.1. Anomaly detection

The dashboard involved new innovative functions. Due to the lack of data evaluation and erroneous data records observed in the current groundwater network, anomaly detection was critical to correct the data entries. 2030 WRG identified statistical methods to identify sudden changes in groundwater levels, which could be categorized as anomalous readings. These anomalies could be due to incorrect data capture by meters or due to actual sudden changes in water levels that deviate from the general trend. The key approaches adopted included:

- Exploratory data analysis (EDA): The first step is to develop a deep understanding of the available data using an EDA. The analysis provides an understanding of data availability of all wells for different time periods, and allows for observing trends and seasonal patterns that the data might or might not follow.
- Statistical techniques: Two statistical methods of time series models were used to detect anomalies such as 30 Day-Moving Averages (MA) and 7 Day-Weighted Moving Average. Moreover, three limits were determined using Standard deviations from those moving averages like ±1.5 SD, ±1.75 SD and ±2 SD (Figure 10).

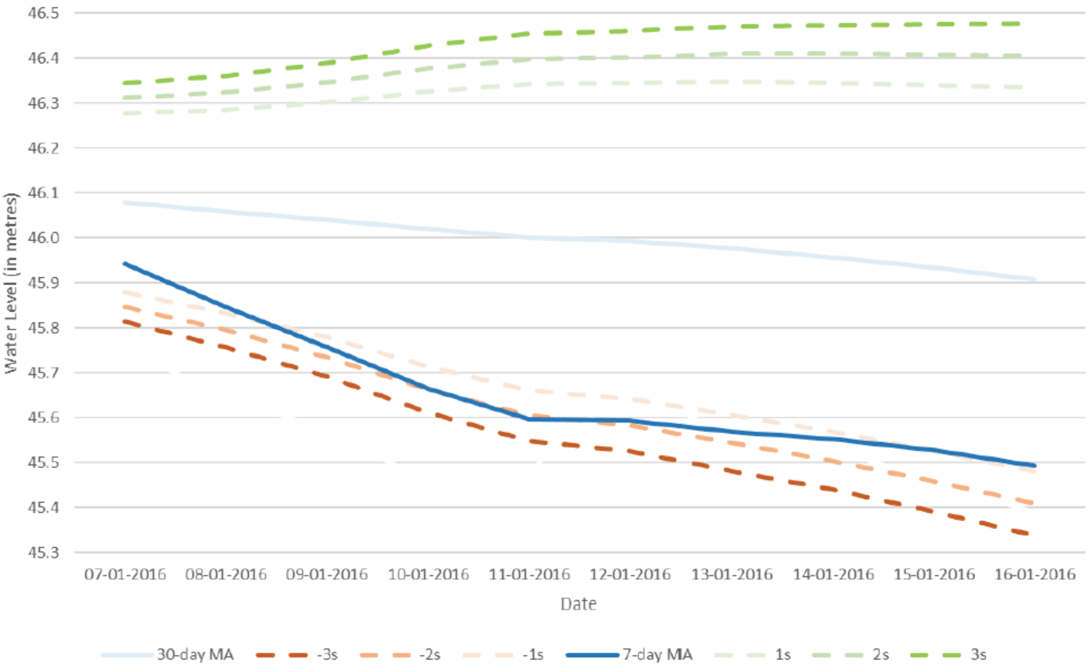


Figure 10. Anomaly detection

If there is an anomaly detected in the data, the raw data requires checking for understanding the underlying rationale for the shift.

4.2.2. Methods of prediction analysis

The groundwater level prediction was conducted in two ways. Specifically, time series models of the groundwater level and independent variables for the prediction model were widely used to train the system for groundwater level predictions. Time series models typically include AR (Auto Regressive), ARMA (Auto Regressive Moving Average), ARIMA (Auto Regressive Integrated Moving Average) and SARIMA (Seasonal Auto Regressive Integrated Moving Average). Developing groundwater data using AI for predictive analysis is essential. It is easier to train the system than to program the system manually with recent trends in ML and AI (Jordan and Mitchell, 2015). In order to develop groundwater level predictions, the following steps were implemented (Gupta 2021):

Step 1: Exploratory data analysis

To develop a deep understanding of the available data, an EDA (Exploratory Data Analysis) was done. This meant getting an understanding of data availability and checking any trend or seasonal patterns.

Step 2: Locking analysis metrics

To begin predictive analysis, two sets of water level predictions were made. The first set was an Evaluation Forecast, with the main purpose of showcasing accuracy of the predictions made. The time periods identified were:

- Training Data (Time period to learn from): 54 months (Jan 2015 to Jun 2019)
- Testing Data (Time period to forecast for and compare against): 12 months (Jul 2019 to Jun 2020)

The second set was a One-Step Ahead Forecast, with the purpose of churning predictions for a time period with no actual data available. The time periods identified for this set were:

- Training Data (Time period to learn from): 66 months (Jan 2015 to Jun 2020)
- Time Period of Forecast: 11 months (Jul 2020 to May 2021)

Statistical modeling techniques used for the prediction analysis included:

- Linear regression (to understand important variables against the one being tested)
- Seasonal ARIMAX (to predict additional variables: runoff and abstraction)
- Random Forest Regression (to predict groundwater level while using additional variables/features)

For the programming language, Python on Jupyter Notebooks (using BlobCity AI Cloud) was used to implement the forecasting exercise.

Step 3: Feature engineering

Feature engineering is a very important part of any prediction exercise. It typically represents 2 primary actions, which are: (1) Transforming existing variables to a format that the machine learning model can understand, and (2) Creating new variables/features using domain knowledge to include into the prediction exercise, to help the machine learning model better capture the changes in the main/target variable.

Step 4: Prediction of external variables (runoff and abstraction)

Before the water level could be predicted, the external variables with less data had to be predicted. This is because these external variables will become an input to the main prediction exercise conducted for water level predictions. These predictions were done using SARIMAX (Seasonal Auto Regressive Integrated Moving Averages), which is a very strong time-series modelling technique with a focus on capturing seasonal trends in data and extrapolating the same very effectively on future predictions.

Step 5: Predictions of groundwater level
Eight model iterations were executed to ensure that the best results were generated. Linear Regression was performed to understand variable importance, whereas multiple iterations of Random Forest Regression were performed using different model parameters, along with training and testing data splits, all with a focus on increased accuracy and reduced error % from the last iteration.

Step 6: Model evaluation
Model evaluation was tested using two metrics including R^2 (R squared) and MAPE (Mean Absolute Percentage Error). This measure is the average of the percentage errors. Being a percentage error, it is a relatively simple but intuitive metric to track R^2 (R squared). Put simply, this metric denotes how much change of the main/target variable (water level in our case) is explained by the additional variables provided. This is also a percentage error and, similar to MAPE, is a simple and intuitive metric to track. Two of these metrics used together are very effective in evaluating a model's performance. The final model was chosen based on low values of these metrics.

4.2.3. Performance of the predictive analysis

To predict groundwater levels, 12 wells data was selected for the model development based on consistency and the length of the dataset. Another criterion was the government's decision on the basin to be selected for the predictive analysis. These wells were located in the 20-50 m zone from the ground surface in the wellfields of the Tuul river basin.

The independent variables of the groundwater level were air temperature, precipitation, river runoff and abstraction. The records of the periods of the datasets was different for each data. For instance, groundwater level data was available between Jan 2015 to Jun 2020; air temperature and precipitation data were available for Jan 2015 to Dec 2020, runoff data was covered for Jan 2015 to Dec 2019 and abstraction data was recorded until end of 2020 (Figure 11). The challenging part was to secure data for the same time period and frequency for all metrics for each well. To overcome this, the team used lapse rates of temperature and precipitation from the closest meteorological stations around the wells based on the elevation. There were three stations located in the basin and air temperature and precipitation lapse rate was -0.6° C/100 m and 19 mm /100m, respectively.

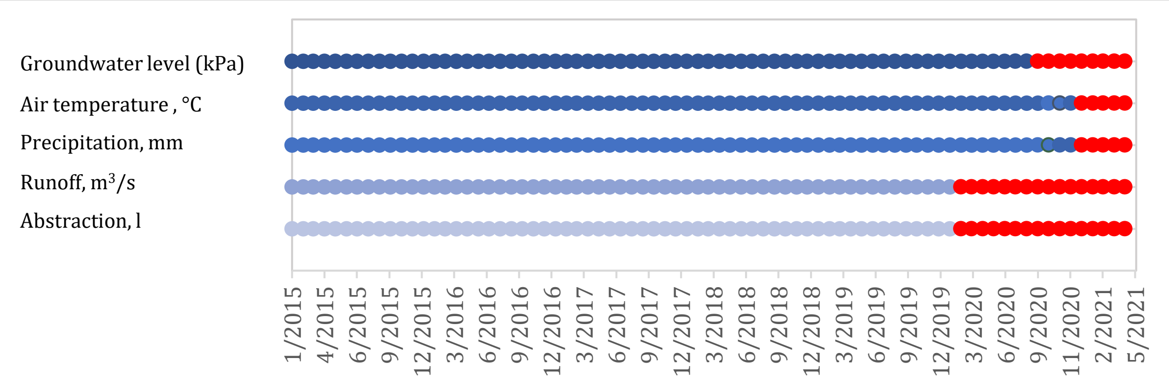


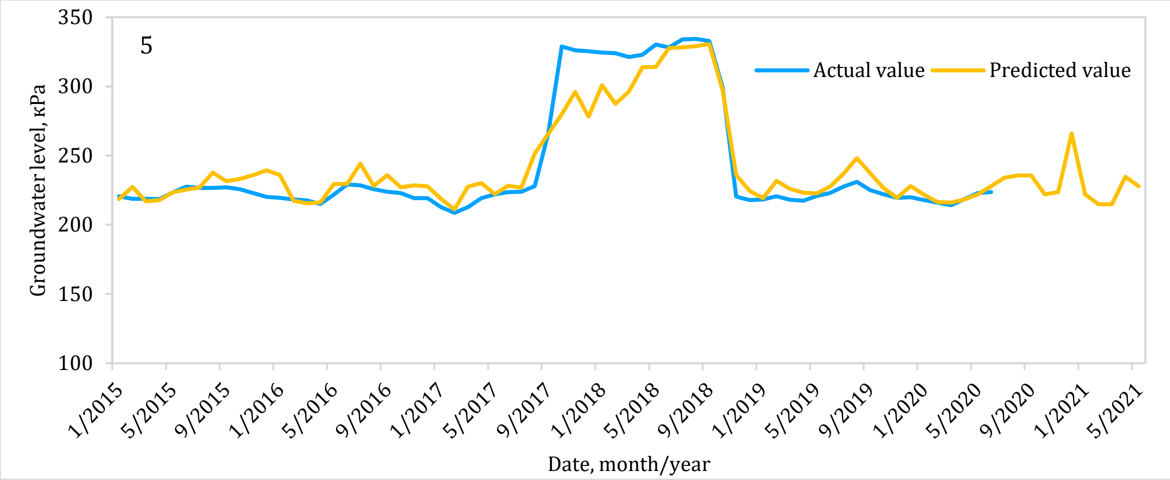
Figure 11. Data consistency of the groundwater prediction analysis

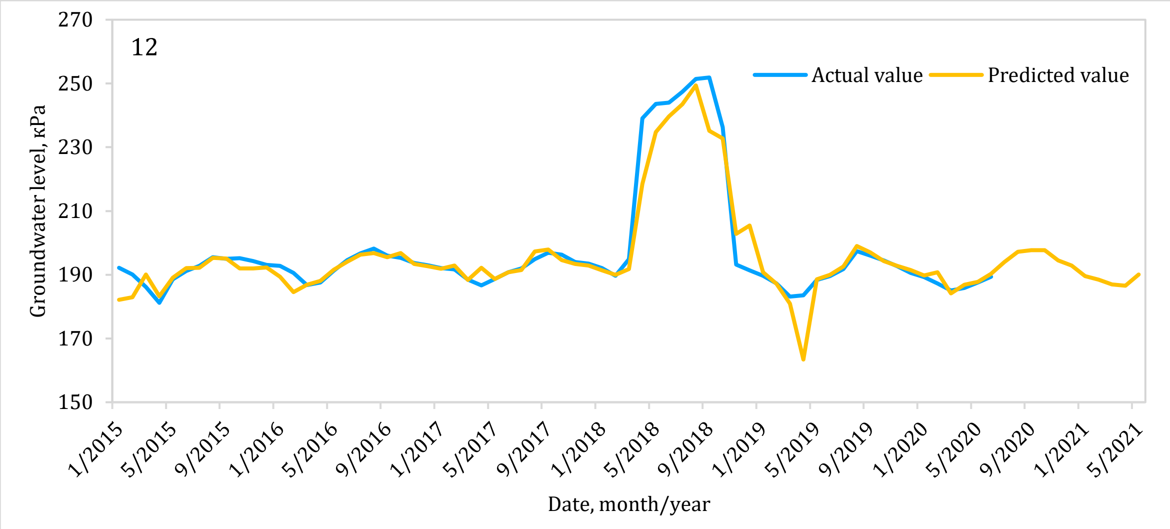
The prediction model was successfully applied to complete records of the measured groundwater level between 2015-2020 for 12 wells in the Tuul river basin. For the model, MAPE ranged between 0.52-5.71% and values of the R^2 were between 0.67-0.98 (Table 6).

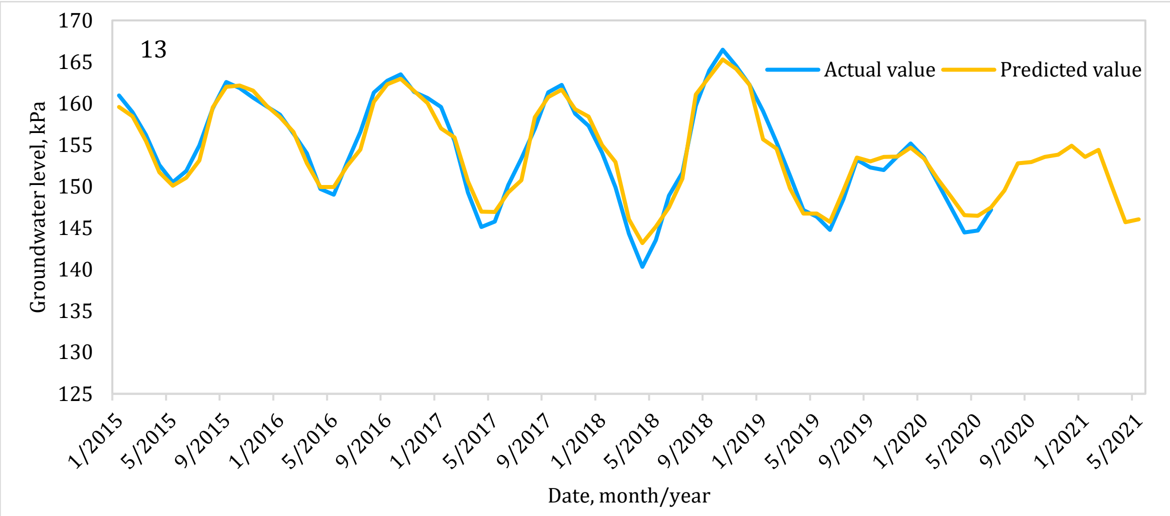
Table 6. Accuracy of the results of the groundwater level prediction model

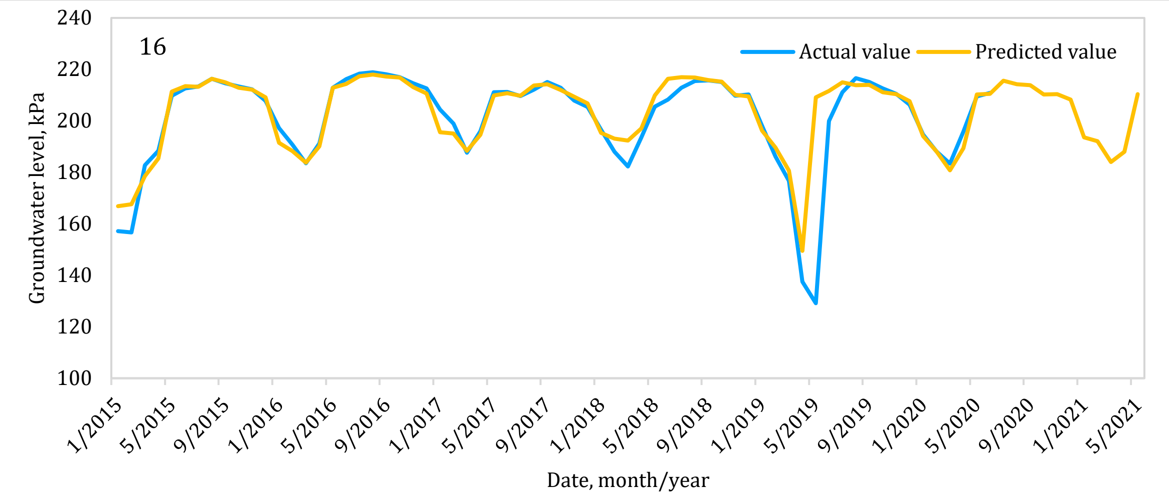
Well number	MAPE, %	R^2
5	3.45	0.92
12	1.37	0.91
13	0.65	0.96
16	2.30	0.67
19	1.61	0.69
20	0.52	0.98
22	2.38	0.92
30	3.60	0.71
33	5.38	0.89
39	5.71	0.87
41	4.55	0.70
43	4.78	0.90

The predictors' leading periods define the prediction duration. NANEM forecasts air temperature and precipitation seasonally until May 2021. Therefore, the prediction was undertaken until May 2021. The training periods of the prediction was from Jan 2015 to Jun 2020 and forecasting period between July 2020 to May 2021 (Figure 12).

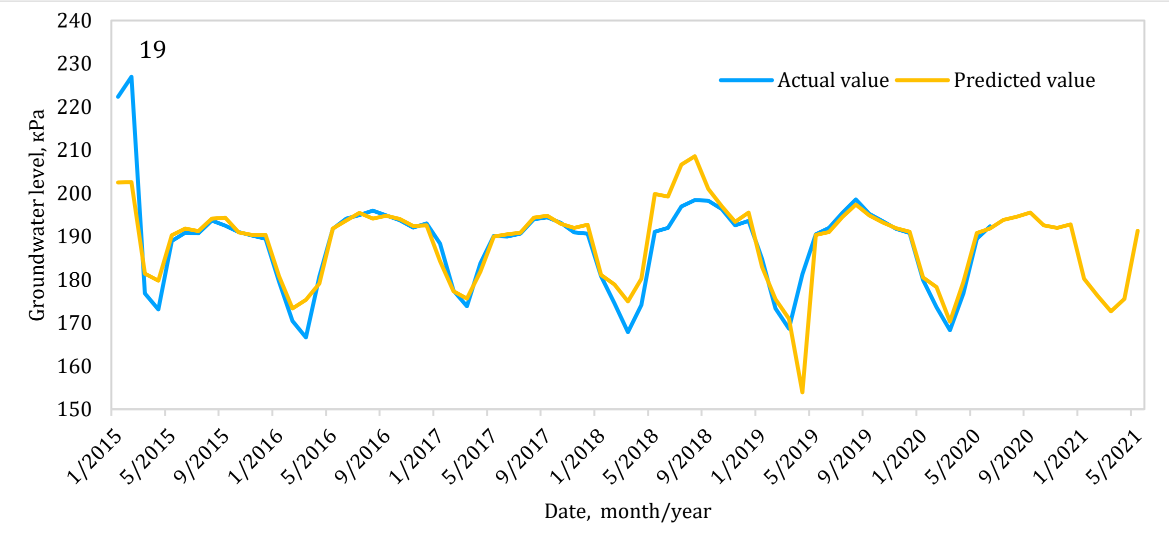

5
Actual value
Predicted value
Groundwater level, кPa
350
300
250
200
150
100
1/2015
5/2015
9/2015
1/2016
5/2016
9/2016
1/2017
5/2017
9/2017
1/2018
5/2018
9/2018
1/2019
5/2019
9/2019
1/2020
5/2020
9/2020
1/2021
5/2021
Date, month/year


12
Actual value
Predicted value
Groundwater level, кPa
270
250
230
210
190
170
150
1/2015
5/2015
9/2015
1/2016
5/2016
9/2016
1/2017
5/2017
9/2017
1/2018
5/2018
9/2018
1/2019
5/2019
9/2019
1/2020
5/2020
9/2020
1/2021
5/2021
Date, month/year

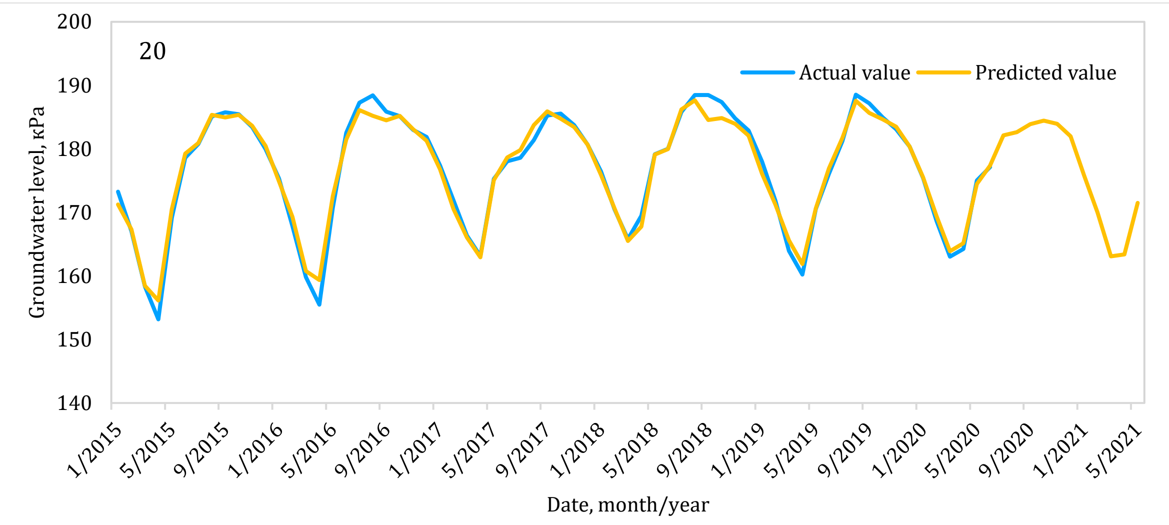

13
Actual value
Predicted value
Groundwater level, кPa
170
165
160
155
150
145
140
135
130
125
1/2015
5/2015
9/2015
1/2016
5/2016
9/2016
1/2017
5/2017
9/2017
1/2018
5/2018
9/2018
1/2019
5/2019
9/2019
1/2020
5/2020
9/2020
1/2021
5/2021
Date, month/year



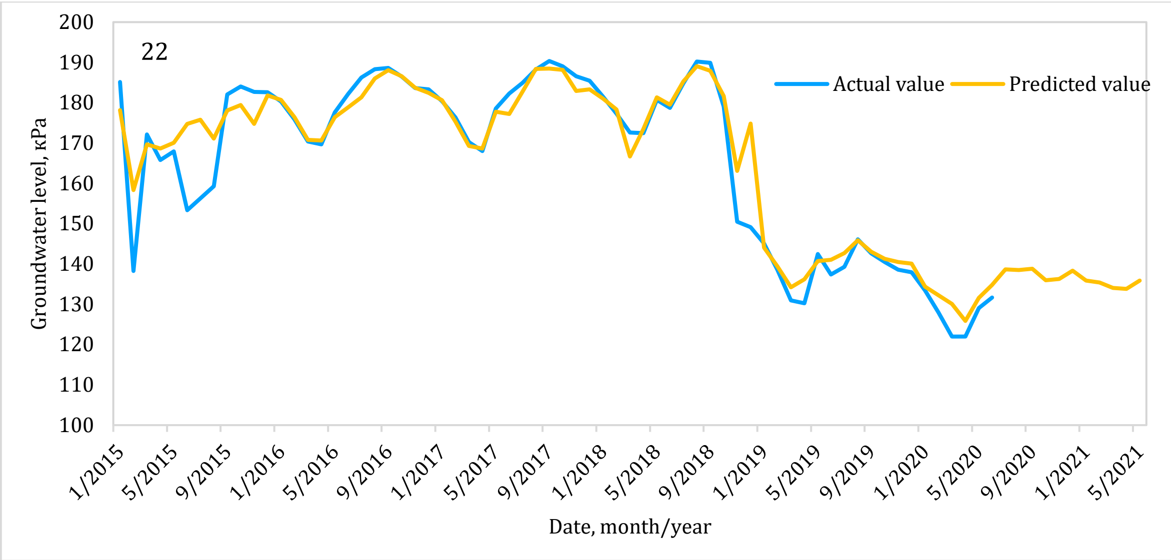
16
Actual value
Predicted value
Groundwater level, kPa
240
220
200
180
160
140
120
100
1/2015
5/2015
9/2015
1/2016
5/2016
9/2016
1/2017
5/2017
9/2017
1/2018
5/2018
9/2018
1/2019
5/2019
9/2019
1/2020
5/2020
9/2020
1/2021
5/2021
Date, month/year

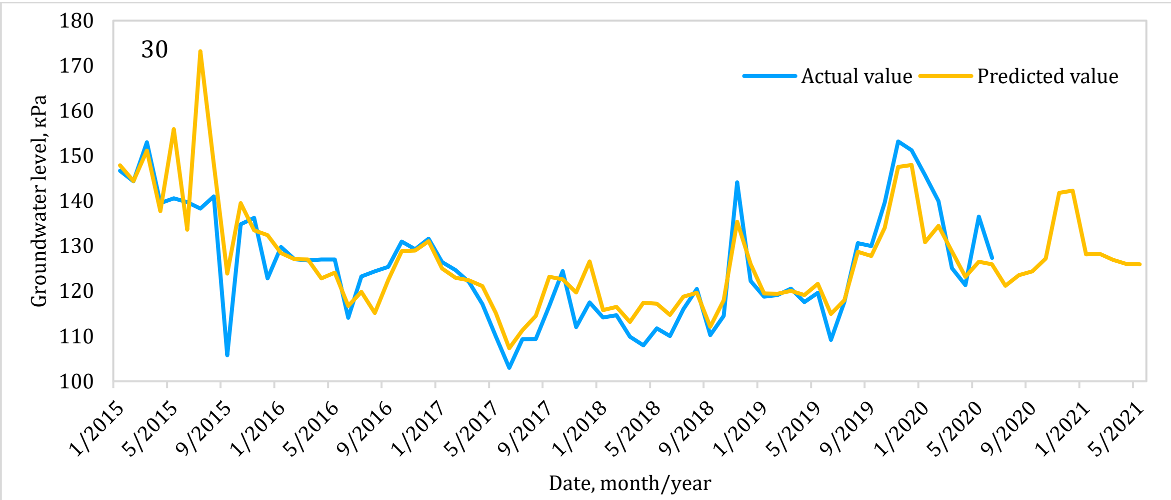


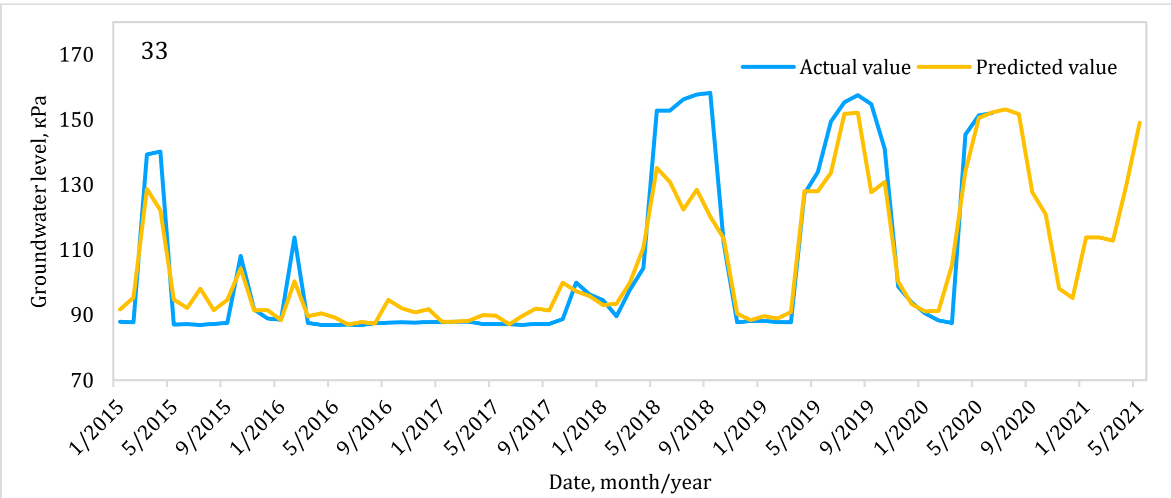
19
Actual value
Predicted value
Groundwater level, кPa
240
230
220
210
200
190
180
170
160
150
1/2015
5/2015
9/2015
1/2016
5/2016
9/2016
1/2017
5/2017
9/2017
1/2018
5/2018
9/2018
1/2019
5/2019
9/2019
1/2020
5/2020
9/2020
1/2021
5/2021
Date, month/year

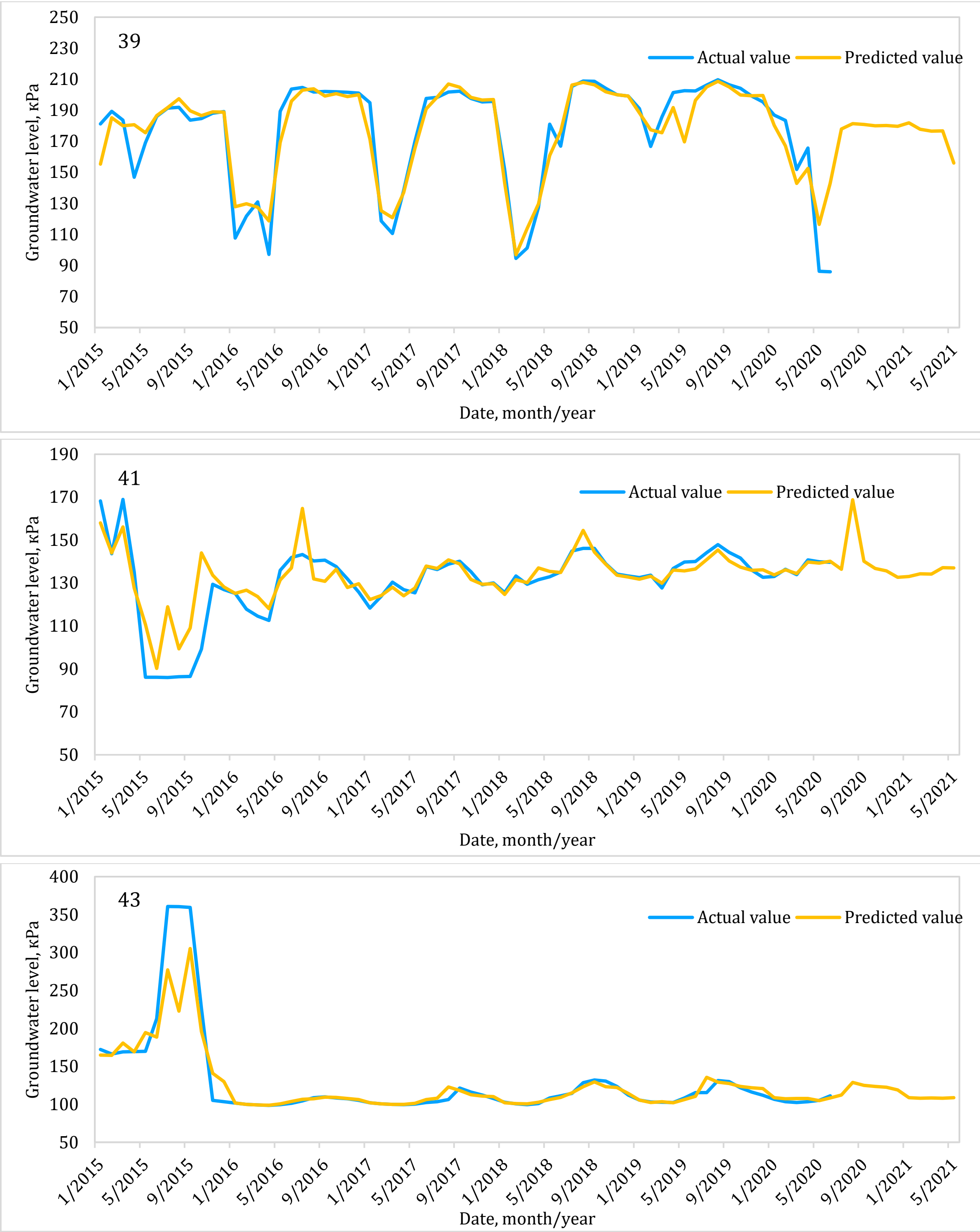


20
Actual value
Predicted value
Groundwater level, кPa
200
190
180
170
160
150
140
1/2015
5/2015
9/2015
1/2016
5/2016
9/2016
1/2017
5/2017
9/2017
1/2018
5/2018
9/2018
1/2019
5/2019
9/2019
1/2020
5/2020
9/2020
1/2021
5/2021
Date, month/year


22
Actual value
Predicted value
Groundwater level, кPa
200
190
180
170
160
150
140
130
120
110
100
1/2015
5/2015
9/2015
1/2016
5/2016
9/2016
1/2017
5/2017
9/2017
1/2018
5/2018
9/2018
1/2019
5/2019
9/2019
1/2020
5/2020
9/2020
1/2021
5/2021
Date, month/year


30
Actual value
Predicted value
Groundwater level, кPa
180
170
160
150
140
130
120
110
100
1/2015
5/2015
9/2015
1/2016
5/2016
9/2016
1/2017
5/2017
9/2017
1/2018
5/2018
9/2018
1/2019
5/2019
9/2019
1/2020
5/2020
9/2020
1/2021
5/2021
Date, month/year


33
Actual value
Predicted value
Groundwater level, кPa
170
150
130
110
90
70
1/2015
5/2015
9/2015
1/2016
5/2016
9/2016
1/2017
5/2017
9/2017
1/2018
5/2018
9/2018
1/2019
5/2019
9/2019
1/2020
5/2020
9/2020
1/2021
5/2021
Date, month/year



Source: Gupta (2021)

Figure 12. Actual and predicted ground water level until May 2021 for 12 wells

From the results of the groundwater predictions, the most successfully evaluated well was Well 20 and the highest error occurred for Well 39. Well 39 recorded higher variations of the groundwater level than others (Table 7).

Table 7. Predicted values of the groundwater level of the 12 wells in the Tuul river basin

DATE, MONTH/YEAR	WELL NUMBER											
	5	12	13	16	19	20	22	30	33	39	41	43
	PREDICTED GROUNDWATER LEVEL, KPA											
7/2020	234.0	194.0	149.5	215.6	193.8	182.2	138.6	121.2	153.2	177.9	136.4	112.1
8/2020	235.7	197.2	152.8	214.3	194.6	182.6	138.5	123.6	151.8	181.4	168.8	129.0
9/2020	235.7	197.7	152.9	213.8	195.5	183.9	138.8	124.4	127.7	180.9	140.2	125.1
10/2020	222.0	197.7	153.6	210.2	192.6	184.4	136.0	127.2	120.9	180.1	136.8	123.6
11/2020	223.7	194.5	153.8	210.5	192.0	183.9	136.2	141.8	98.1	180.1	135.7	122.6
12/2020	265.9	192.9	154.9	208.2	192.8	182.0	138.3	142.3	95.2	179.7	132.7	118.8
1/2021	222.1	189.5	153.5	193.6	180.2	175.8	135.9	128.2	113.8	181.9	133.1	108.8
2/2021	214.9	188.5	154.4	192.1	176.3	170.0	135.4	128.3	113.8	177.8	134.3	108.2
3/2021	214.8	187.0	150.0	184.1	172.7	163.1	134.1	126.9	112.9	176.5	134.2	108.5
4/2021	234.9	186.6	145.7	188.0	175.5	163.4	133.8	126.0	130.0	176.7	137.1	107.9
5/2021	227.8	190.1	146.0	210.3	191.3	171.5	135.9	126.0	149.1	156.1	137.1	108.8

4.3. Components of the dashboard

Based on the preliminary analysis of data availability and objectives of the portal, the groundwater dashboard has been organized into 5 sections (Figure 13) including (1) Single well analysis, (2) Single well deep dive, (3) Multiple well analysis, (4) Anomaly detection and (5) Prediction.

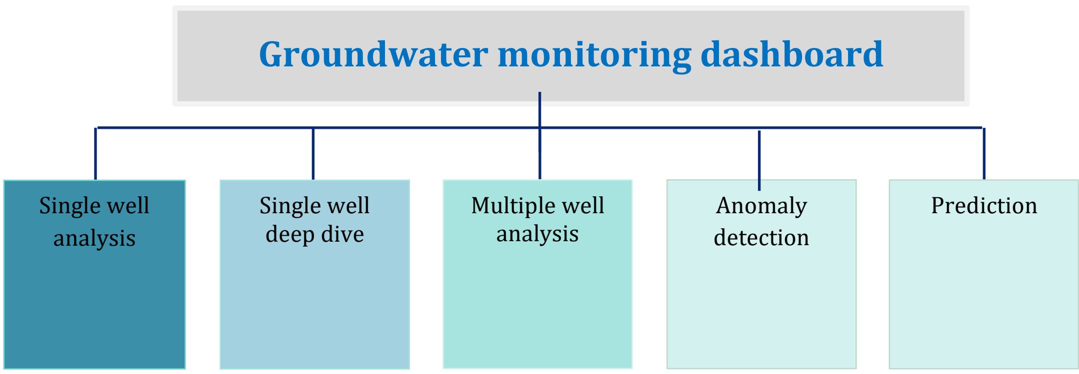


Figure 13. Dashboard structure

Each section of the dashboard is briefly explained below.

4.3.1. Single well analysis

This section allows the user to analyze the Water Level of an individual well across different time periods (Figure 14). Along with the Water Level, the user can also select any parameter (Hardness, Temperature, Pressure etc.) and view the readings for the same time periods, thus enabling the user to identify a correlation between the water level readings and the parameters. As an example, Figure 14 shows the user can view the Water Level, Pressure and Air Temperature across 2015 and 2016 for Province 1 - Watershed 1 - Well 1.

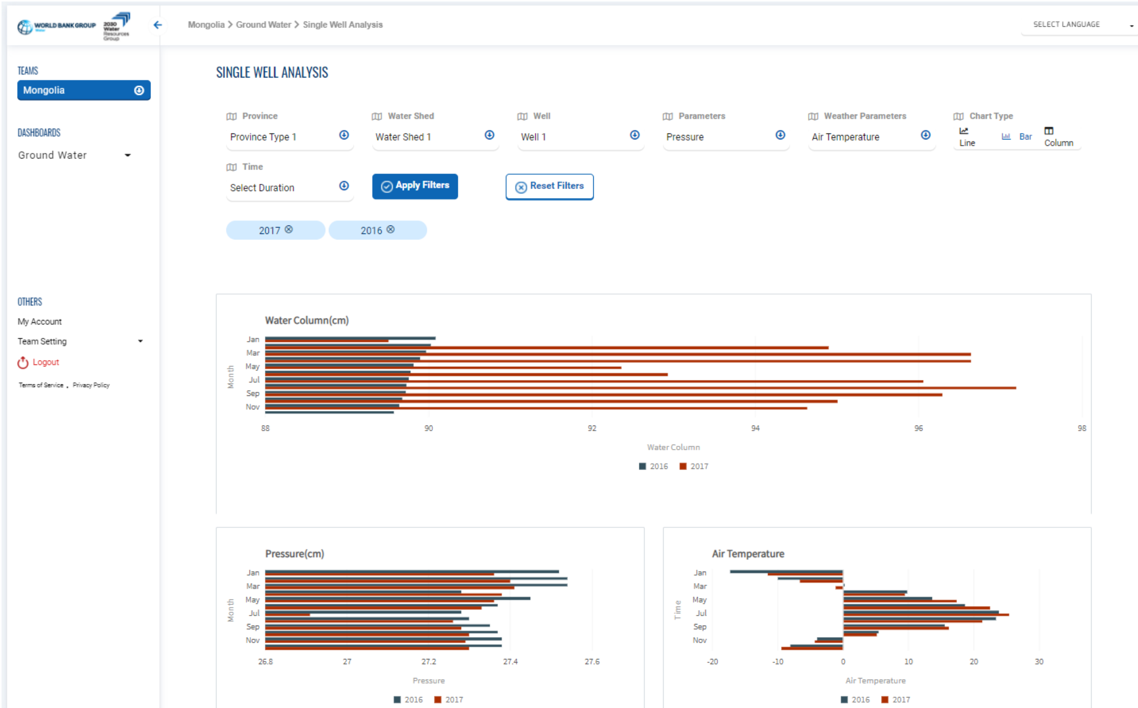


Figure 14. Appearance of the Single well analysis menu

4.3.2. Single well analysis deep dive

In this section, the user can select any fixed time period and analyze the water level readings along with the readings of multiple parameters together for the selected time period. This enables the user to deep dive into the impact of water levels on multiple parameters. For instance, the user can analyze the Pressure, Conductivity and Temperature readings across 2016 for Province 1- Watershed 1 -Well 1 as shown in Figure 15.

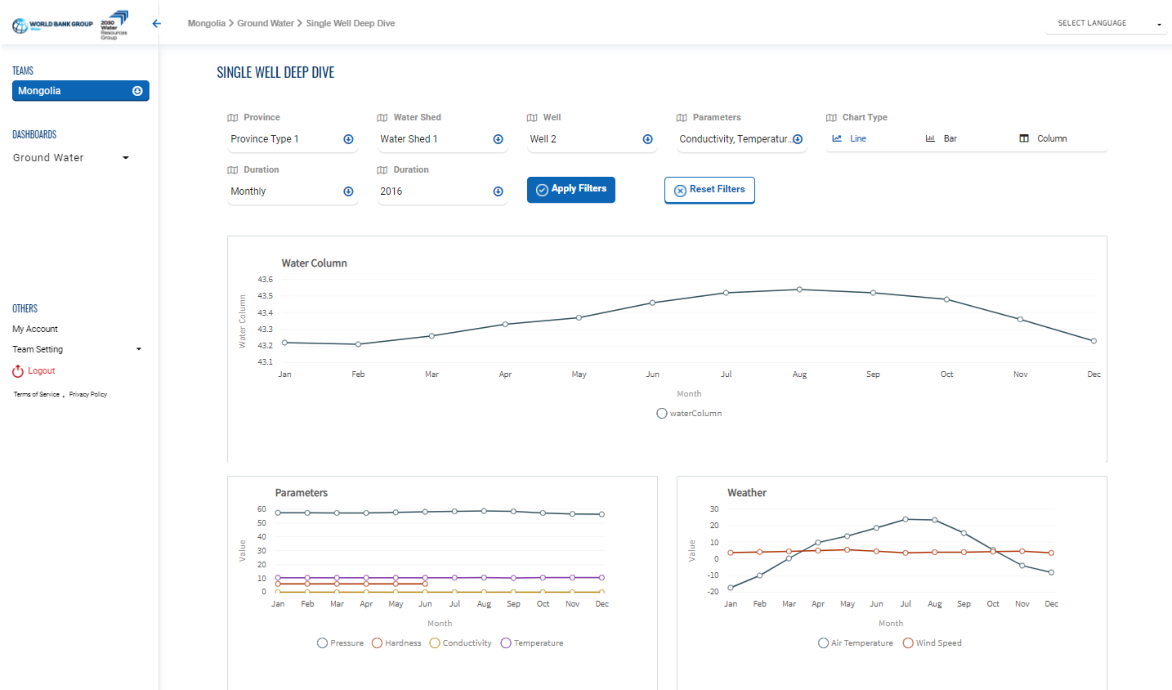


Figure 15. Appearance of the Single well deep dive menu

4.3.3. Multiple well analysis

In this section, users who have permission to view data across multiple watersheds can compare the water levels and parameter readings across wells for a selected time period. This helps users understand how water levels differ across major watersheds. For example, shown in Figure 16, a user can view how the groundwater level for two different watersheds (Province 1- WaterShed1- Well 1 and Province 1-WaterShed 1- Well 3) varies across 2016 (Figure 16).

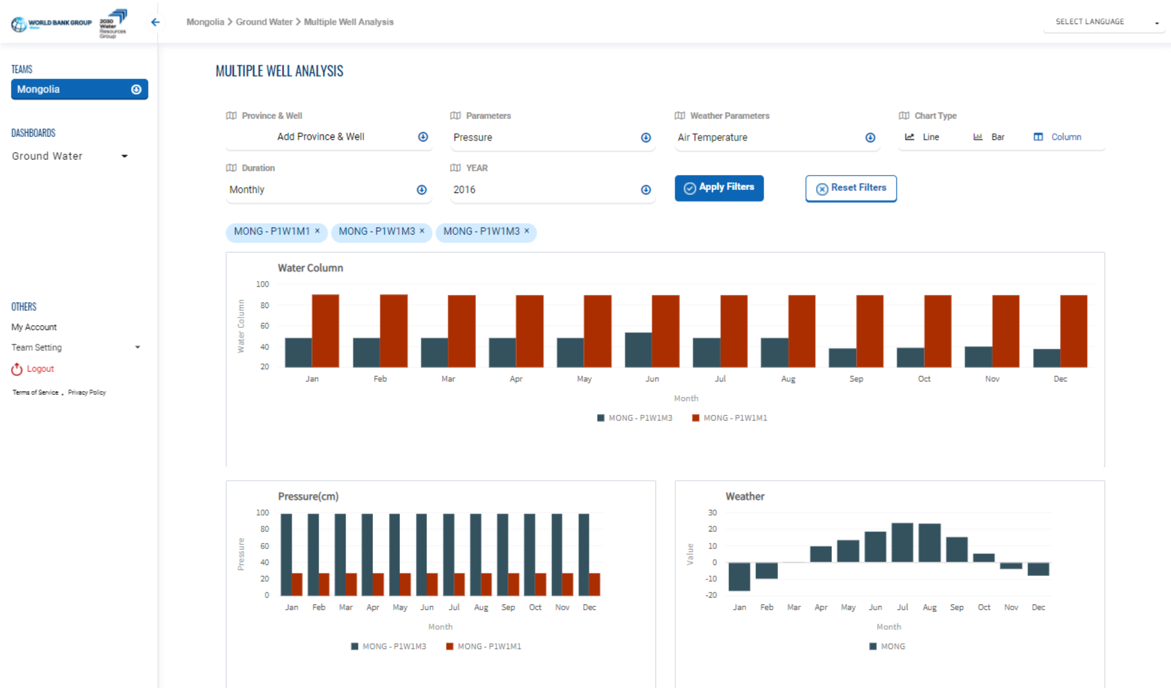


Figure 16. Appearance of the Multi well analysis menu

4.3.4. Anomaly detection

In this section, the user can view all the anomalous readings for any selected metric in that given time period. Metrics highlighted as green indicate that the particular parameter has seen an anomalous increase, whereas metrics highlighted as red indicate that the parameter has seen an anomalous decrease (Figure 17).

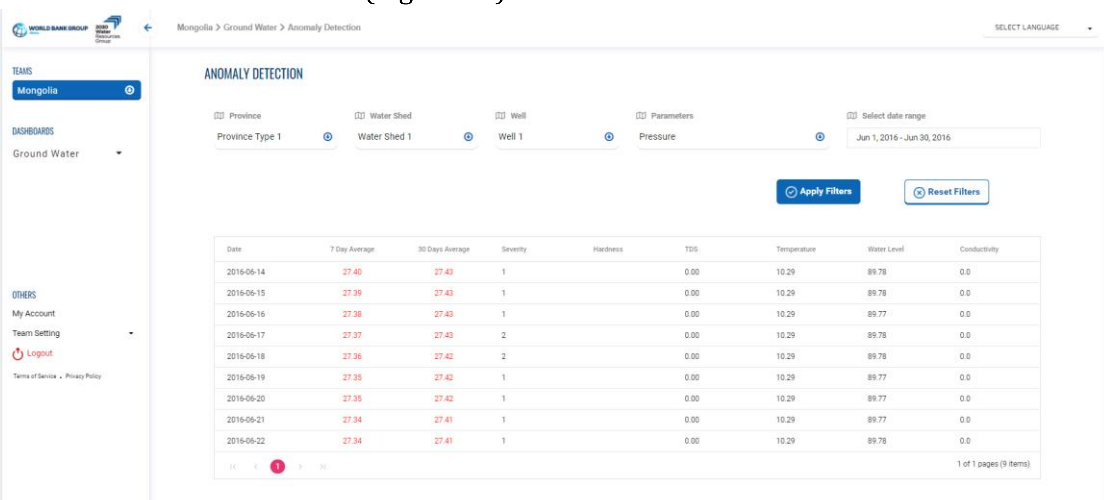


Figure 17. Appearance of the Anomaly detection menu

4.3.5. Prediction analysis

In this section, users can view the predicted water levels across multiple well. In this case, water levels have been predicted till May 2021. The user can view the historical water level trend in conjunction with the predicted trend. Also, the user can compare the '2021 year to date' average water level to the '2020 actual' average water level to be able to understand if the overall water level has increased or decreased when compared to the same time period of the previous year. As shown in Figure 18, the user can observe that for Well 20, the average water level can be expected to decrease by 1.07% in Jan-May 2021 when compared to Jan-May 2020.

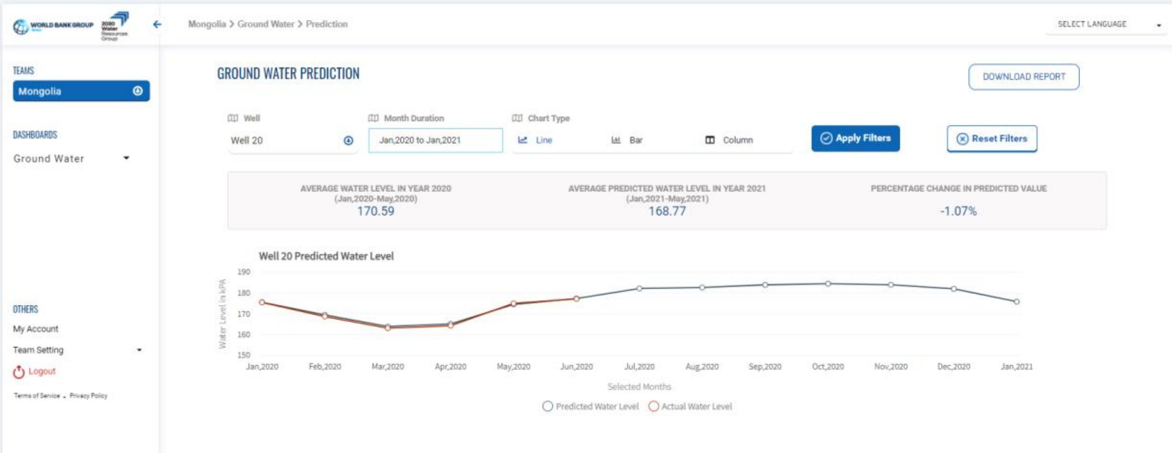


Figure 18. Appearance of the Prediction menu

4.3.6. Dashboard filters

The dashboard has several filters to select the following sections.

- *Select province:* User can select the provinces where the well of interest is located.
- *Select Watershed:* User can select the Watershed where the well of interest is located.
- *Select wells:* User can select the well of interest from the list.
- *Select parameters:* Pressure, Conductivity, Total Dissolved Solids, Temperature, Hardness
- *Select weather parameters:* Air Temperature, Precipitation and Wind Speed
- *Select duration:*
  - *Hourly:* The user can select a specific date for which the hourly readings for that date will be displayed.
  - *Daily:* The user can view the daily readings for either all days in a given month or in a given week.
    - *Daily+Weekly:* The user can view the readings for the 7 days of a given week. For example, the use can select 2016 week 36-Sep 5 which indicated 36th Week of 2016 which begins on September 5th (Monday) and ends on September 11th.
    - *Daily + Monthly*: The user can view the daily readings for all the days in the selected month. A user can select January 2019 and see the daily readings from 1st - 31st January.

- *Monthly* - The user can select any year and see the monthly readings in that year.
- *Custom Date* - The user can select a random date range (e.g. 14th Dec 2019 - 20th Jan 2020) and accordingly view data for all days in that range.
- *Custom Month* - The user can select a random month range (e.g. July 2018 – Aug 2020).
- *Past 6 Months* - This shows the readings for the last 6 months of available data.

4.3.7. User management

There are two types of user administration on the dashboard as Admin and Team settings (Figure 19).

- Users can be added to the portal through the "Admin" section. This section will only be visible to system administrators.
- Added users can be granted the following permissions:
  - *Super Admin* - Gives the user System Administrator privileges with the ability to create and manage teams and view all the dashboards across the entire system. Admins can select if a user can have access to multiple dashboards at a time
  - *Users* - Ability to access the dashboards and data for only teams that the user is present in.

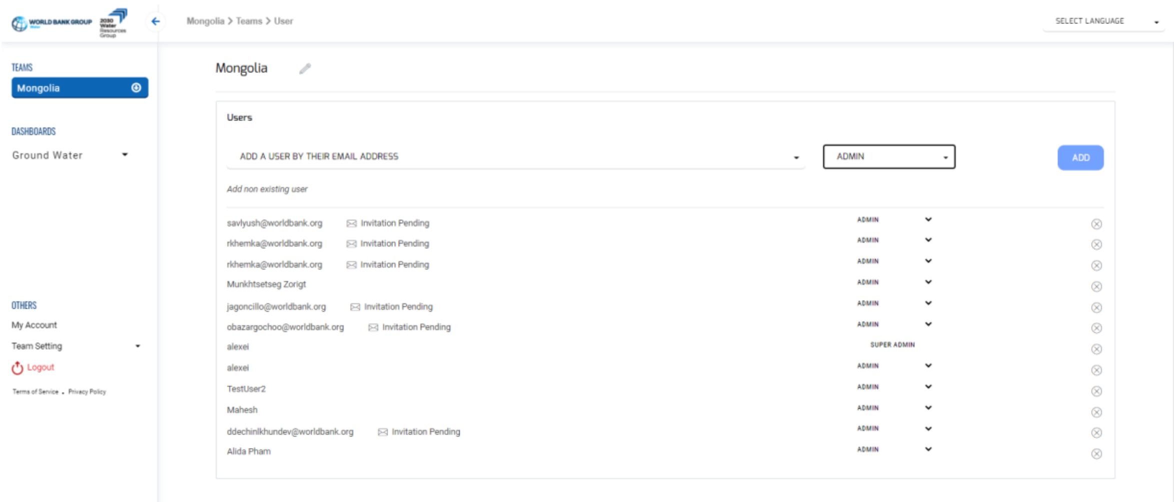


Figure 19. Dashboard menu of the User management

The users are available to provide following settings in the Team settings section (Figure 20).

- Admins can create multiple teams with unlimited number of members within a team.
- Users can be given the following permissions within a team through the Team Settings section.
  - *Admin* - These users can upload data to the portal and add new users to the dashboard. Admins of a given team do not have access to data from any other team unless he/she has been explicitly added to other teams as well.

- *Users* - These users can only view the dashboard but cannot make changes to team members and cannot upload data.

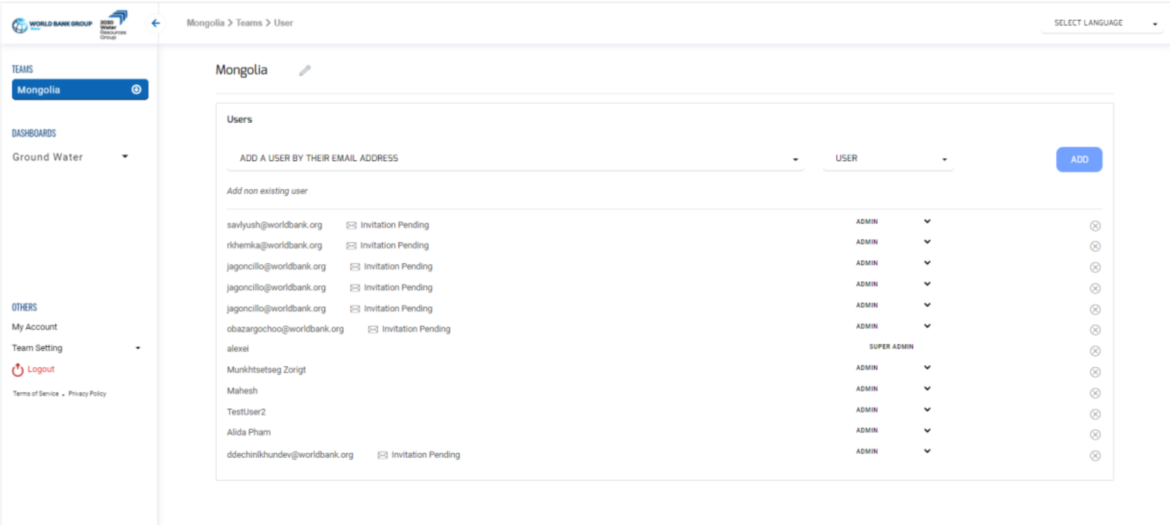


Figure 20. Dashboard menu of Team settings

4.3.8. Data upload

Users can upload new data sets on their own with the help of the data upload portal which is accessible through the **Team Settings** section. Users will have to upload their data files in a fixed .csv format.

4.3.9. Upload climate data

Weather Data Sets can be uploaded from the Weather Data Upload section. The data will have to be uploaded in the following format (Table 8):

Table 8. Meteorological data format for uploading to the dashboard

Date	Soum	Daily average air temperature [°C]	Daily average wind speed [m/s]	Daily precipitation [mm]
2015.01.01	Gurvantes	-10.5	1.1	
2015.01.02	Gurvantes	-8.3	3.9	
2015.01.03	Gurvantes	-5.1	1.5	
2015.01.04	Gurvantes	-6.8	4.4	
2015.01.05	Gurvantes	-11.4	3.3	
2015.01.06	Gurvantes	-11.1	3.9	
2015.01.07	Gurvantes	-8.1	4.1	
2015.01.08	Gurvantes	-7.6	2.8	
2015.01.09	Gurvantes	-8.6	2.5	
2015.01.10	Gurvantes	-8.7	2.3	
2015.01.11	Gurvantes	-7.3	1.9	

5. Converting current groundwater monitoring network to the portal (www.groundwater.mn)

5.1. Process description of upgrading existing monitoring network

To enable stakeholders to use the groundwater portal, existing groundwater monitoring system at the national level groundwater.mn was upgraded.

KPM LLC was selected to upgrade the former groundwater monitoring network along with BlobCity Inc. and the 2030 WRG team. The former monitoring network was only used by RBAs and government officials. In order to support multi-stakeholder participation, the upgraded portal added other users from the public and decision makers and was divided into three categories with different privileges for the portal. Moreover, the portal was upgraded to upload precipitation data from meteorological stations and water abstraction data connected with monitoring wells. In addition, the groundwater level prediction model developed by the 2030 WRG team was applied to the portal. Groundwater level predictions allow the groundwater managers, experts, decisionmakers and other stakeholders to plan and manage the resource.

5.2. Data entries to the portal

5.3.1. Uploading climate data into the dashboard

It is useful to have precipitation data at the monitoring wells for analysis and prediction. However, there are not enough precipitation loggers installed with groundwater loggers in the wells in Mongolia. Therefore, measured climate data from the meteorological station was connected with the wells using the Thiessen polygon method. The river basins of Mongolia were divided into the Thiessen polygons based on the meteorological stations. Each polygon contains only a single meteorological station. If monitoring wells are located in a certain polygon, they were connected with the precipitation data from that station (Figure 21).

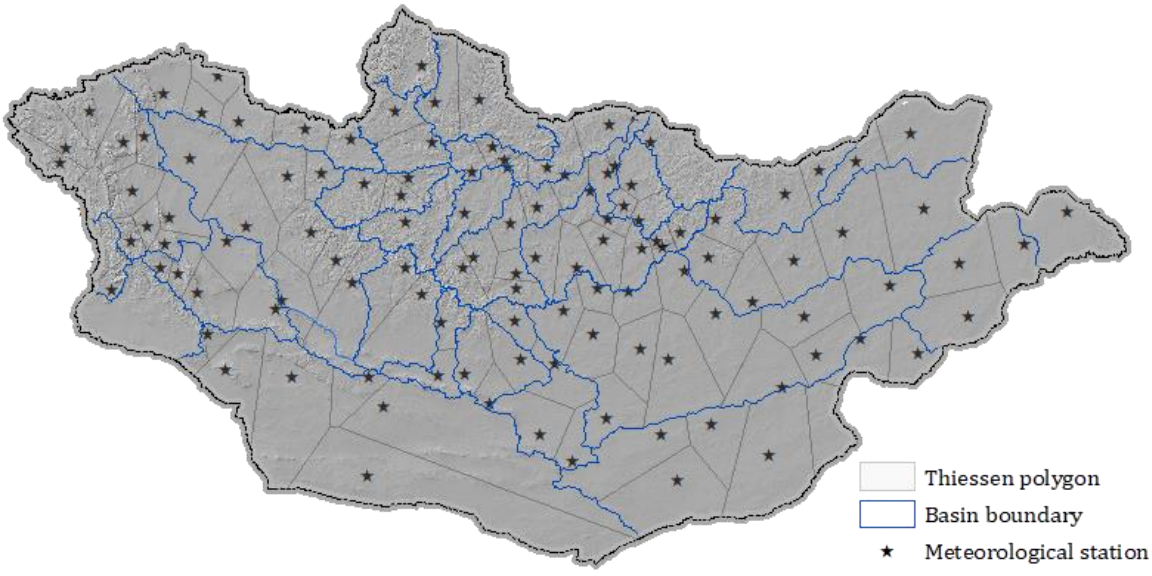


Figure 21. Thiessen polygons of Mongolia based on meteorological stations

5.3.2. Uploading water consumption data

The groundwater monitoring system covers monitoring wells, but in some cases, abstraction wells have not installed a water meter to measure how much water is extracted from that well. As such, it is challenging to link the water abstraction data with the groundwater level. In most cases, water abstraction data recorded the total volume of groundwater extracted from several wells. Therefore, the abstraction data will be uploaded manually or by files.

5.3.3. Prediction model

The prediction model code in the Jupyter Notebooks was developed by BlobCity Inc. and provided to KPM (Figure 23) for application to the portal.

```
2021-04-01  W43 2021-04-01    NaN
2021-05-01  W43 2021-05-01    NaN

[924 rows x 3 columns]

In [23]: # Define function for ADF (Augmented Dickey Fuller) test
from statsmodels.tsa.stattools import adfuller

# Define ADF function
def adf_test(timeseries):
    # Perform Dickey-Fuller test:
    print ('Results of Dickey-Fuller Test:')
    dftest = adfuller(timeseries, autolag = 'AIC')
    dfoutput = pd.Series(dftest[0:4], index = ['Test Statistic','p-value','#Lags Used','Number of Observations Used'])
    for key,value in dftest[4].items():
       dfoutput['Critical Value (%s)'%key] = value
    print(dfoutput)

# Apply ADF test on all wells
for well in df_runoff.Well.unique():
    df2 = df_runoff[df_runoff.Well == well]
    df2.dropna(axis = 0, how = 'any', inplace = True)
    df2 = df2[['Value']]

    print("Well: " + well)
    adf_test(df2)
    print("\n")


# Results:
# Stationary: W5, W12, W13, W16, W19, W20, W22, W30, W33, w39, W41, W43
# Not stationary: None

# Refer this link to interpret results:
# https://www.analyticsvidhya.com/blog/2018/09/non-stationary-time-series-python/

<ipython-input-23-174339394c0e>:17: SettingWithCopyWarning:
A value is trying to be set on a copy of a slice from a DataFrame

See the caveats in the documentation: https://pandas.pydata.org/pandas-docs/stable/user_guide/indexing.html#returning-a-v
iew-versus-a-copy
  df2.dropna(axis = 0, how = 'any', inplace = True)
<ipython-input-23-174339394c0e>:17: SettingWithCopyWarning:
A value is trying to be set on a copy of a slice from a DataFrame

See the caveats in the documentation: https://pandas.pydata.org/pandas-docs/stable/user_guide/indexing.html#returning-a-v
iew-versus-a-copy
  df2.dropna(axis = 0, how = 'any', inplace = True)
<ipython-input-23-174339394c0e>:17: SettingWithCopyWarning:
```

Figure 22. Prediction code on the Python in the Jupyter Notebook

5.3.4. User management

The analysis was performed to evaluate whether and where highly developed groundwater database is desirable, and, if that is the case, in defining the types of data and scope of the data processing, considering the given situation and the budgetary and organizational conditions. Database users and their requirements can be diverse, leading to a substantial challenge. This analyses therefore addresses three main questions:

- Who will be possibly using the database? And why?

- What types of data is needed? How extensive does the data need to be?
- In which format should the data be reported? How much processing is required?

Table 9. Brief analysis of key stakeholders

	Government officials	River basin administrative unit	Water users (individual or entities – NGOs and companies	Research institutes and research-oriented companies
Need of information	Yes – high	Yes – high	Yes/No Citizens/ Companies generally do not require access to detailed information, unless their interest/right is violated. Companies/NGOs need information depending on their activities	Yes – high
Types of information	Trend of groundwater at regional/ river basin levels; groundwater quality, water reserve amount (by aquifer, basin, renewable, non-renewable etc.), water consumption, usage, and re-usage information. Other supporting data	Trend of groundwater level at regional/ river basin level, groundwater quality, water reserve amount (by aquifer, basin, renewable, non-renewable etc.), water consumption, usage, payment information, and other supporting decision-making information.	Local/related information on water quality, resource amount, whether there is sufficient water etc. and other relevant information	Depending on research question
Levels of processing/Accuracy	Consolidated/ semi-processed	Consolidated/semi-processed information,	Either consolidated/ semi-	Mostly raw data or consolidated/

	information, suited for decision making and axillary semi-processed data for interpretation. Processing - nationally, regionally, locally, monthly, seasonally, yearly	suited for decision making and axillary semi-processed data for interpretation. Processing - regionally, locally, monthly, seasonally, yearly	processed information or highly processed information. Locality and temporal accuracies are dependent on question	semi-processed information or highly processed information. Locality and temporal accuracies are dependent on question
Data usage purpose	Decision making, policy making, intervention measures, precautionary measures – i.e. scenario evaluation	Decision making, regulation/compliance, intervention measures	Being informed, surveying & researching	Conducting research
Reporting frequency	Quarterly	Quarterly	Not relevant	Not relevant
Reporting format	Report, map, charts	Report, map, charts	Map, charts	Raw data

Based on the above questionnaires (Table 9), it was decided to create three user types who can log in and view the current and future groundwater level and upload data. These users are:

1. Admin
2. River basin
3. Guest (public and decision makers)

5.4. Upgraded structure of the monitoring portal

Upgrading the current groundwater monitoring network www.groundwater.mn was developed in accordance with the dashboard design by BlobCity. The web portal included dashboard screens that give users a graphical view of the groundwater data. The groundwater monitoring portal allowed access to three types of users with different scopes (Figure 23).

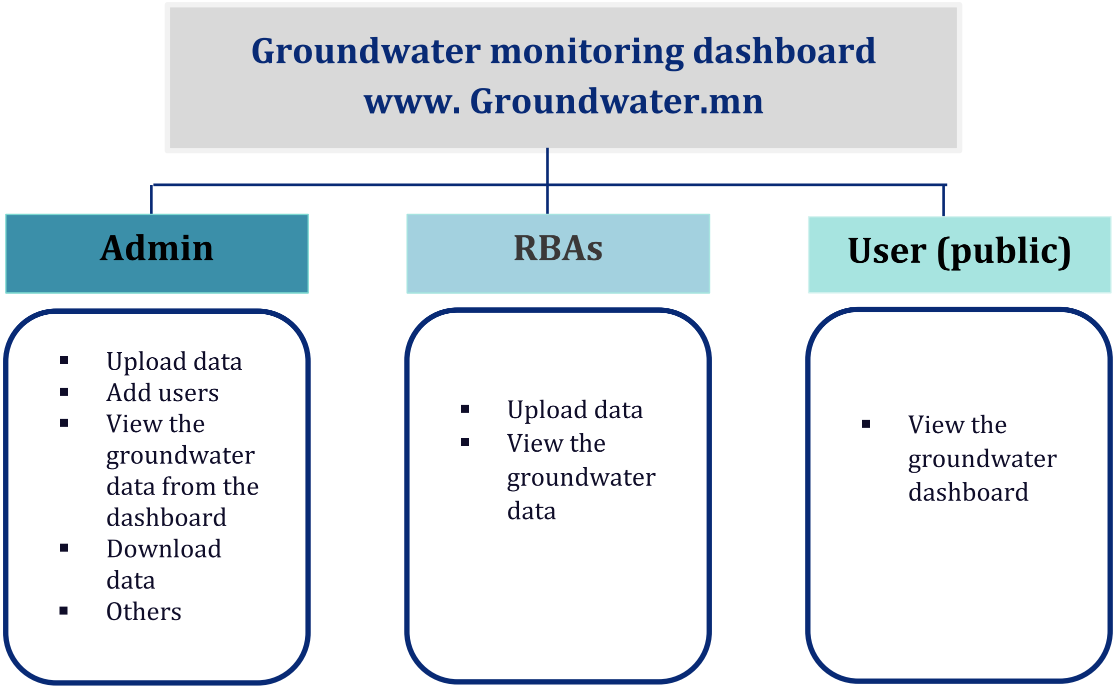


Figure 23. User management of the Mongolian groundwater dashboard

After logging in, the users were presented with the GW monitoring portal home page. The user can select the well (Figure 24) and view the groundwater and relevant datasets. Each user has different functionalities to view and edit the groundwater metrics (KPM 2021).

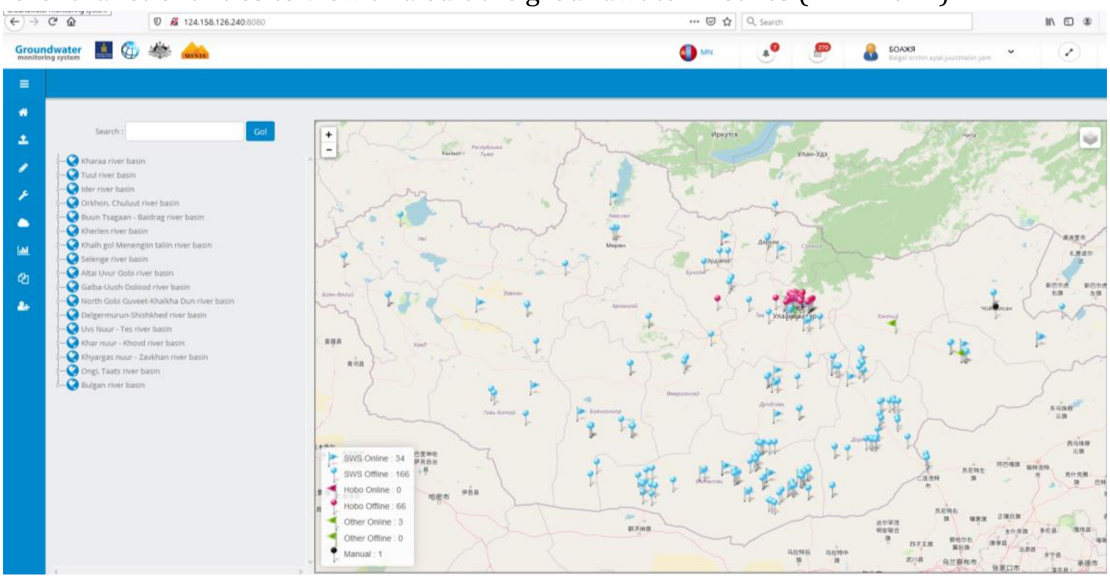


Figure 24. Main menu of the portal

5.5. User management

5.5.1. “Admin” users

Admin users are the main users for managing the portal with several privileges. Within the groundwater portal, the admin has access to the following functions:

- User management: In this section, the admin can manage the user’s rights and settings of the portal. The admin can enable actions, such as provide login id, view and upload data and grant privileges.
- River basin management: The user’s management of the River basin provides the capability to see, export data and report.
- Monitoring wells management: This function allows to admin users to create, register and maintain the logger of the monitoring wells.
- Barologger management: In this section, the users can manage different types of barologgers, which measure atmospheric pressure and temperature. The settings of the section demonstrate registration number, type of the equipment, serial number and installed date. Moreover, the users can see and edit data from the barologgers.
- Data logger management: The data logger management section provides detailed information about the groundwater loggers. The users can see the registration and serial number, installed date and other relevant information of the equipment. They also can manage the data to edit.
- Company management: This section proposes to store groundwater monitoring data generated by private sector and companies. It promotes systemizing the database and arrangements for data sharing with other parties. The companies can register their loggers for the portal and manage it.

5.5.2. “RBAs” users

The users of RBAs are typically groundwater experts and specialists that collaborate on primary groundwater data. A basin has a set of permissions that allow or restrict access to the different basins. As an RBA user, they have credentials for their river basin and the following main rights:

- Log in to the portal
- Get detailed information of the monitoring wells
- Upload data (groundwater metrics, climate and abstraction data)
- Reporting

5.5.3. “Guest” users

Guest users can access the portal using their credentials such as phone number or email address. They can see the three sections of the dashboard: Single well analysis, Single well deep dive and Multiple wells analysis (Figure 25). Guest users can see the plots of the groundwater data for the selected wells.

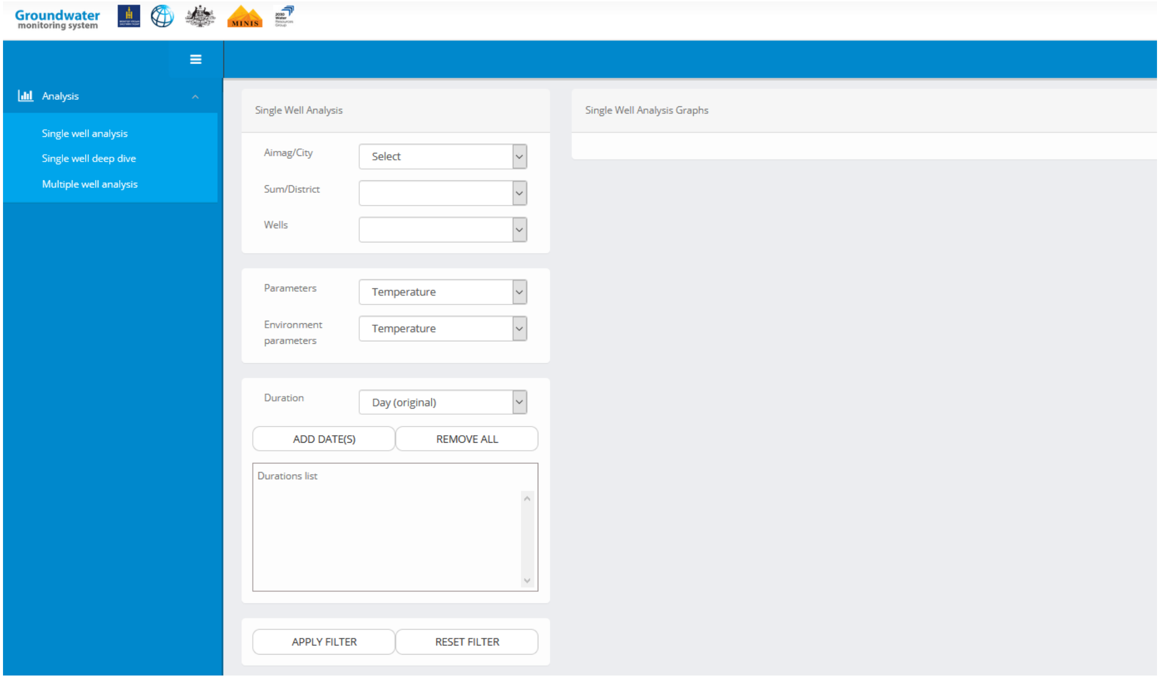


Figure 25. Appearance of the dashboard menu for the Public users

5.6. Dashboard menus

The current monitoring network has been developed to facilitate three types of dashboards including Single well analysis, Single well deep dive and Multiple well analysis. The dashboard is accessible for all users to see groundwater level and other parameters.

Single well analysis : The users can see the groundwater level at different time scales in this section (Figure 26).

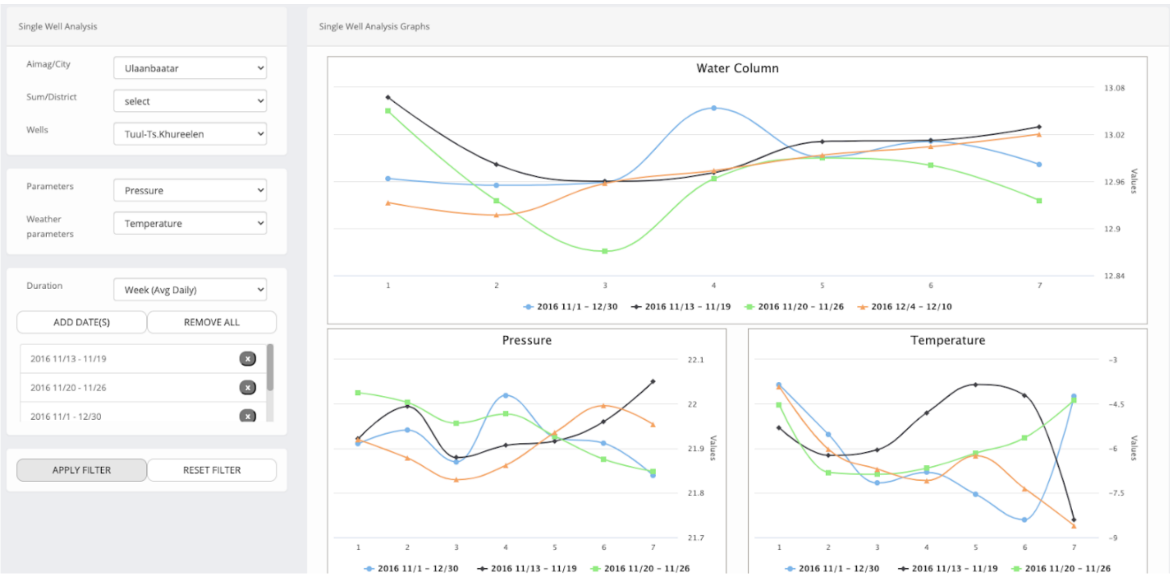


Figure 26. Single well analysis section

Single well deep dive analysis: In this section, the user can view the plots for the groundwater level with other parameters like water temperature, precipitation and air temperature etc. (Figure 27).

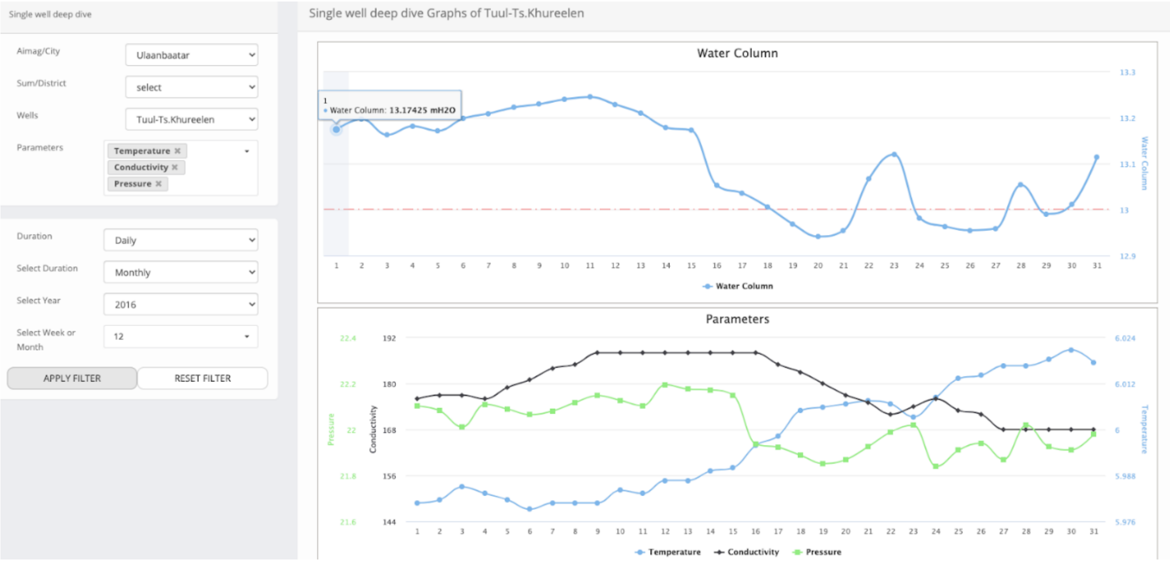


Figure 27. Single well deep dive analysis section

Multiple wells analysis: In this section, users can compare the groundwater level for different wells (Figure 28). Metrics can be displayed showing the range in the data for a selected time period.

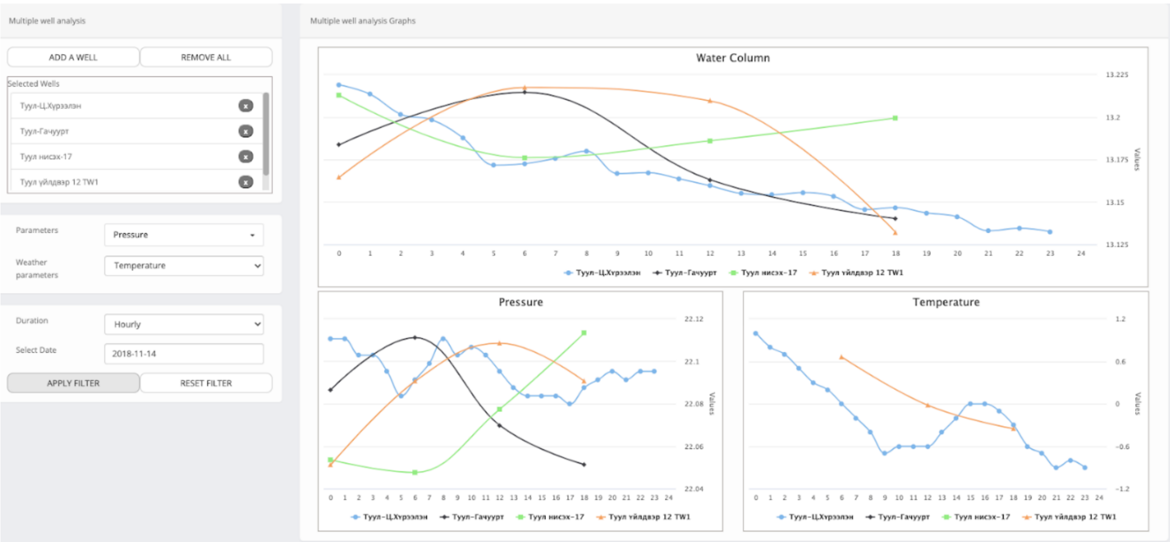


Figure 28. Multiple well analysis section

6. Conclusions

Based on the analysis of the existing groundwater monitoring system of Mongolia, concerns regarding the following issues were identified:

1. Data management: The groundwater data system in Mongolia does not currently prioritize consistent and regular data collection and uploading, and lacks a data management system. The data cannot be fully trusted in the absence of any rigorous or defined validation efforts. Without a data management system, erroneous data can be introduced:
   - Groundwater data is not consistently collected and uploaded to the network as evaluated for 102 monitoring wells.
   - The recipient of the data does not perform a data quality evaluation.
   - Errors are introduced into the data, as it is imported into whatever method is being used for its storage.
2. Limited users: There is a limitation for user management which does not support multi stakeholder participation.
   - Only RBAs and government officials can access the groundwater monitoring system, and other stakeholders are unable to use the data for decision making and resource management.

2030 WRG enabled improved groundwater management, including a dashboard and an upgraded portal for the groundwater monitoring system. The following activities were successfully implemented by 2030 WRG:

1. The existing groundwater monitoring network was upgraded and converted to a dashboard, including access to three types of users to the portal.
2. Uploading and use of climate and abstraction data as independent variables for the prediction was completed.
3. The use of disruptive technology for a predictive model for groundwater levels for 12 wells in the Tuul river basin was applied successfully. This highlighted the ability of disruptive technology to support groundwater management.

7. Recommendations

The results of this project have a number of important implications for the future development of a groundwater digital portal to enhance groundwater management and sustainable resource use:

- Continued efforts are needed to make the monitoring portal accessible to all the stakeholders.
- It is crucial to manage groundwater resources using robust and validated monitoring data.
- Such data should be accessible to all stakeholders.
- In addition, the monitoring portal provides valuable information to manage and plan groundwater, which is the primary source for water supply.

The team has outlined key recommendations for groundwater monitoring and sustainable operation of the portal (Table 10).

Table 10. Recommendations on groundwater monitoring objectives and sustainable operation of the portal

Measures	Coordination		
	MET, WA	RBAs	Specialists
DATA COLLECTION FROM THE MONITORING POINTS:			
• Uploading historical groundwater data into the portal and data validation;		√	√
• Collecting and distributing data according to user guides/manuals, SOP of the data loggers			√
• Analyzing collected data and explaining anomalies, if detected		√	√
• Uploading water quality data into the portal and providing guidance for that	√		
• Registering the monitoring wells as observational and abstraction wells	√	√	√
• Installing water meters for the direct monitoring of groundwater abstraction in the abstraction wells and uploading water meter readings to the portal		√	√
GROUNDWATER DATA PROCESSING:			

• Providing guidance for groundwater data validation	√	√	√
• Providing guidelines and standards for data collection and processing		√	√
• Establishing evidence to improve the monitoring network and portal	√		
• Providing advance support to collect groundwater data annually	√		
• Uploading the independent variables which affected groundwater levels		√	√
SUSTAINABLE USE AND DEVELOPMENT OF THE PORTAL:			
• Establishing secure servers and using a mirror server for the groundwater database and the monitoring portal		√	√
• Providing field and laboratory instruments for the groundwater monitoring	√	√	
• Providing detailed hydrogeological and aquifer behavior for the monitoring wells into the portal			√

Overall, this project has proven that disruptive technology is a useful and robust tool for managing and predicting groundwater metrics for accurate and efficient data to meet the needs of stakeholders.

8. References

Dorjsuren D., Purevkhuu D., Ganchimeg O., Altankhuyag A., and Ankhbayar A. (2015). Report on reestimation of the centralized and non-centralized water resources in Ulaanbaatar city - 1. Ulaanbaatar

MET. Ministry of Environment and Tourism (2018). Integrated Water Resources management plan of the Galba Uush Doloodiin Gobi basin, Ulaanbaatar, Mongolia.

KPM (2021). User guide on groundwater monitoring portal (Admin)

KPM (2021). User guide on groundwater monitoring portal (River basin authority)

KPM (2021). User guide on groundwater monitoring portal (Guest)

BlobCity (2020). Report on Groundwater dashboard.

Gupta S (2021). Mongolia Groundwater prediction. BlobCity

Fitch P., Brodaric B., Stenson M., Booth N. (2016). Integrated Groundwater Data Management. In: Jakeman A.J., Barreteau O., Hunt R.J., Rinaudo JD., Ross A. (eds) Integrated Groundwater Management. Springer, Cham. https://doi.org/10.1007/978-3-319-23576-9_26

Jordan MI and Mitchell TM (2015). Machine-Learning: Trends, perspectives, and prospects. Science 17. Vol 349:6245. Pp 255-260. DOI: 10.1126/science.aaa8415

Thomas W, Gregory A, Gager J, Kuo I-L, Wackerow A, Nelson C (2009). Data Documentation Initiative (DDI) technical specification part I: overview, pp 1–103. Retrieved from http://www.ddialliance.org/

Van Essen Instruments (2009). Diver product manual. Delft, The Netherlands HL341v3

Vision 2050 (2020). Action plan for 2021-2030 of Mongolia 's Long-term development policy "vision 2050", Ulaanbaatar