

# Groundwater dynamics in the Indus revealed by integrated flow modeling and satellite data

Li Huang<sup>1</sup>, Susanna Werth<sup>2</sup>, Dimitrios Stampoulis<sup>3</sup>, Glen Low<sup>4</sup>, and John Sabo<sup>5</sup>

<sup>1</sup>School of Sustainable Engineering and the Built Environment, Arizona State University

<sup>2</sup>Virginia Polytechnic Institute and State University

<sup>3</sup>Arizona State University

<sup>4</sup>The Earth Genome

<sup>5</sup>Tulane University

November 23, 2022

## Abstract

In the Indus River Basin, groundwater plays a key role in mitigating the water storage fluctuations due to climate variation and meeting the rapidly increasing water demand in agriculture dominated basins. A comprehensive understanding of groundwater dynamics is essential for a transition to more efficient and sustainable water resources management. To gain detailed insight response of water flows and storage in the Indus aquifers to agricultural activities, we build a high resolution 3D regional groundwater flow model for the entire basin. However, in practice, regional flow models, as they are most widely used, suffer from calibration challenges. To address the sparsity of in-situ groundwater data in the region and to acquire a realistic reproduction of flow dynamics, we calibrate the model using both in-situ and satellite-based estimates of ground states. We test the advantage of such a multi-objective approach by comparing its results with a single-objective approach in which we constraint the model parameter only against in-situ data. We examine and discuss the model results for flow and storage conditions, which reveal: 1) depth to water table has decreased (1998-2007) almost exclusively in urban areas (1 m), and 2) groundwater storage depletion averaged ~5cm in equivalent water thickness basin-wide over 20 years of simulations. Groundwater storage depletion results primarily from intensive groundwater withdrawal to meet extensive irrigation demands. Optimizing crop patterns and associated groundwater extraction in space and time could improve groundwater conditions.

## **Groundwater dynamics in the Indus revealed by integrated flow modeling and satellite data**

**Li Huang<sup>1,2</sup>, Susanna Werth<sup>3</sup>, and Dimitrios Stampoulis<sup>2</sup>, Glen Low<sup>4</sup>, John Sabo<sup>2</sup>**

<sup>1</sup> School of Sustainable Engineering and the Built Environment, Arizona State University, Tempe, AZ 85287, USA.

<sup>2</sup> Future H2O, Knowledge Enterprise Development, Arizona State University, Tempe, AZ 85287, USA.

<sup>3</sup> School of Geographical Sciences and Urban Planning, Arizona State University, Tempe, AZ 85287, USA.

<sup>4</sup> The Earth Genome. PO Box 1337 Mill Valley, California 94942

Corresponding author: John Sabo ([john.l.sabo@asu.edu](mailto:john.l.sabo@asu.edu))

### **Key Points:**

1. Construction and calibration of a high-resolution groundwater flow model for the entire Indus River Basin by using GRACE and in-situ data
2. GRACE calibrated flow model outperforms the model calibrated by in-situ water level measurements alone
3. Models calibrated by the multi-objective approach produce more accurate spatial and temporal details about the groundwater system

## Abstract

In the Indus River Basin, groundwater plays a key role in mitigating the water storage fluctuations due to climate variation and meeting the rapidly increasing water demand in agriculture dominated basins. A comprehensive understanding of groundwater dynamics is essential for a transition to more efficient and sustainable water resources management. To gain detailed insight response of water flows and storage in the Indus aquifers to agricultural activities, we build a high resolution 3D regional groundwater flow model for the entire basin. However, in practice, regional flow models, as they are most widely used, suffer from calibration challenges. To address the sparsity of in-situ groundwater data in the region and to acquire a realistic reproduction of flow dynamics, we calibrate the model using both in-situ and satellite-based estimates of ground states. We test the advantage of such a multi-objective approach by comparing its results with a single-objective approach in which we constraint the model parameter only against in-situ data. We examine and discuss the model results for flow and storage conditions, which reveal: 1) depth to water table has decreased (1998-2007) almost exclusively in urban areas (<-25 m) but has slightly increased in agricultural land adjacent to rivers (>1 m), and 2) groundwater storage depletion averaged ~5cm in equivalent water thickness basin-wide over 20 years of simulations. Groundwater storage depletion results primarily from intensive groundwater withdrawal to meet extensive irrigation demands. Optimizing crop patterns and associated groundwater extraction in space and time could improve groundwater conditions.

Keywords: groundwater dynamics; regional flow model; MODFLOW; GRACE; Indus

## 1. Introduction

The Indus River Basin (IRB) has the world's largest irrigation system. Groundwater provides close 35% of the basin's overall water supply, of which 94% is for irrigated agriculture [[Pakistan Bureau of Statistics, 2017](#)]. Agriculture is the largest single sector of Pakistan's economy; hence, groundwater plays a vital role in Pakistan's economic development [[Dalin et al., 2017](#); [Gleeson et al., 2012](#)]. Groundwater depletion in the IRB is among the highest in the world [[MacDonald et al., 2016](#)]. As such, quantitative estimates of temporal and spatial variability of groundwater level and storage are paramount for water resources management and planning [[Y. Wada et al., 2014](#); [Yoshihide Wada et al., 2010](#)].

Regional groundwater flow modeling is a valuable quantitative tool to support water resources decision making, in a variety of relevant contexts including, flux dynamics [[Aliyari et al., 2019](#); [Furlong et al., 2011](#); [N. R. Rossman et al., 2018](#); [Sahoo and Jha, 2017](#); [Sridhar et al., 2018](#); [Zhou and Li, 2011](#)], water resources management and planning [[Cao et al., 2013](#); [Nathan R. Rossman and Zlotnik, 2013](#); [Tian et al., 2015](#)], and assessment of groundwater system responses to stresses [[Furlong et al., 2011](#); [Hassen et al., 2019](#); [Imaz-Lamadrid et al., 2019](#)]. However, in practice, large scale distributed models are computationally expensive and time consuming [[B Wu et al., 2015](#)]. Moreover, calibration and validation of large-scale, regional models are cumbersome [[Alexander Y Sun et al., 2012](#)]. One significant obstacle to calibration is the lack of consistent, regional-scale, hydrogeological information to represent the real physical structure of groundwater system and to capture the groundwater flow head fluctuations [[de Graaf et al., 2017](#)]. A second reason is that the model process is iterative, involving model

conceptualization, construction and calibration, which brings in structural uncertainty [Xu et al., 2017a]. Finally, the model output has significant uncertainty associated with the imperfection of input parameter values and structures [Xu et al., 2017b].

In situ groundwater level measurements alone do not suffice for calibrating regional models, especially when these point measurements are unevenly distributed. Other data, such as evapotranspiration, groundwater surface water exchange fluxes and environmental tracers can be used as supplemental data to improve model calibration [Castro and Goblet, 2003; Schilling et al., 2019]. These supplemental datasets have similar limitations (few or uneven point estimates). By contrast, satellite data eliminate many of the aforementioned limitations. Launched in 2002, GRACE monitors Earth's gravity anomaly [Tapley et al., 2004]. With the continuous refinement in GRACE satellite data post processing techniques, GRACE products have significantly increased the spatial localization and amplitude of recovered terrestrial Total Water Storage anomalies (TWSA) [Scanlon et al., 2016], providing much needed information for inferring changes in the groundwater system. By the end of 2019, there will be more than 16 years of continuous TWSA data available. In this vein, GRACE observations have been used to calibrate surface hydrology models [Long et al., 2017; Scanlon et al., 2018; S. Werth and Güntner, 2010], to detect groundwater variations [Richey et al., 2015; A. Y. Sun, 2013; S Werth et al., 2017], to infer evapotranspiration and groundwater recharge [Long et al., 2014; Pan et al., 2017; Q Wu et al., 2019], and to estimate change in groundwater storage and water table [Seyoum et al., 2019; Stampoulis et al., 2019].

GRACE has been used in combination with hydrologic models to estimate trends in groundwater depletion in parts of the Indus River Basin. For example, GRACE based gravity anomalies from 2003 and 2010 in the Punjab were converted into GW storage changes by subtracting simulated surface water storage from the Variable Infiltration Capacity (VIC) hydrological model [Iqbal et al., 2017]. The authors found that the most depleted area in Punjab is Bari Doab. However, a comparison of groundwater depletions calculated from GRACE and piezometric data in the same region revealed a gap of 0.16 billion cubic meters per year. In this same study, a Visual MODFLOW model was used to validate the GRACE calculated groundwater storage change and the authors found that the satellite-estimated anomalies generally agree with the groundwater trends simulated by groundwater flow model even though there are gaps in simulated average depletion rates and some disagreement in depletion trends [Iqbal et al., 2017]. Methods for fusing GRACE into hydrologic models for calibration and estimation of groundwater trends need further improvement.

In this study, we built a regional groundwater flow model for the entire IRB in order to explore future interventions that can be used to predict groundwater depletion in the Indus river basin. In doing this, we address the aforementioned data challenges by developing new way for integrating satellite-based (GRACE) data products in to the groundwater model as an additional calibration constraint as well as an independent method for validation. Our premise is that a groundwater model calibrated and validated with satellite data products will be more robust than the traditional approach constrained by in situ data only. In the following sections, we first introduce the study area, applied data sets and methods, including the modeling scheme and the calibration approach (Section 2-4). Next, we present basin scale model results and discussion (Section 4) and conclude by comparing the reliability of model estimates of change in groundwater storage using in situ and satellite data products (Section 5).

## 2. Study Domain and Data

### 2.1. Site Description

The Indus River Basin is located in south Asia, in between the coordinates 66°E-82.5°E and 23.5°N-37.5°N, covering a total area of ~1 million km<sup>2</sup> (Figure 1). The basin is one of the most intensively developed, agriculture dominated basin in the world [[Watto et al., 2018](#)]. It contains the world's largest irrigation system-Indus Basin Irrigation System (IBIS).

The basin has an arid to semi-arid climate with a mean annual precipitation of 365 mm. The Indus River originates in the Tibetan Plateau and the Himalayas, it is fed by at Panjnad by five major tributaries that merge with the ~2900 km long mainstream, and subsequently it flows into the Arabian Sea [[Inam et al., 2007](#)]. Annual peak flow occurs between June and late September, during the southwest monsoon. The waters of the Indus River and its tributaries are heavily appropriated for irrigation in this relatively arid area and the river is a lifeline for the economy and culture of the region [[Fahlbusch et al., 2004](#)]. The arid climate, high population density (145 people/km<sup>2</sup>) and intensive agriculture results in major impacts on the water cycle in the IRB. 90% of the original forests are lost, many channels and rivers have shrunk and have been modified [[Inam et al., 2007](#)].

The Indus River Basin is shaped by the collision of the Indian Plate and Eurasian Plates around 45 million years ago and the erosion after. The Indus River follows the Indus-Tsanpo suture zone and flows along the active strike-slip plate boundary within its foreland. The IRB has various landscape types including glacier, high mountains, alluvial plains and desert and ranges from 6677 to 0 m above mean sea level from north to south. The Indus River Plain Aquifer, including the Upper Indus, the Middle Indus and the Lower Indus, is part of the Indo-Gangetic Basin (IGB) alluvial aquifer system [[Bonsor et al., 2017](#)]. The aquifer system, formed from sediments eroded from Himalayas and redistributed by the Indus River, is unconsolidated alluvium deposition in most parts. Finally, it is often considered as a single highly permeable homogenous aquifer, but with significant spatial variability in permeability, hydraulic conductivity and storage characteristics [[Richts et al., 2011](#)] (Figure 1).

### 2.2. Data sources

The datasets used to establish the ArcGIS database and construct the flow model for this study are detailed in Table 1. All data, unless explained elsewhere, were obtained from the Lahore Water and Sanitation Agency (WASA) and the Water and Power Development Authority (WAPDA). A Digital Elevation Model (DEM), lithological map, and soil and deposits thickness were used to define the aquifer extent. Land use, soil type, and aquifer information were used to define aquifer parameters. River networks and irrigation canals were used to define drainage system. Precipitation, evapotranspiration, irrigation and groundwater pumping data served as model inputs to drive the flow system. Streamflow and groundwater level monitoring data were used to calibrate the model. We acquired 82887 groundwater values from 1998 to 2017 (20 years) for 4096 wells spread unevenly across Indus basin within both Pakistan and India (Figure 1) from WASA, the Programme Monitoring and Implementation Unite (PMIU) and the Water Resources Information System of India (India\_WRIS). Most wells from India have continuous pre-monsoon and post-monsoon measurements from 1998-2017, most of the wells from WASA have continuous yearly Measurements from 2000-2014 and most of the wells from PMIU have continuous pre-monsoon and post-monsoon measurements from 2003-2016. GRACE results are independent information used to crosscheck the model performance. Figure 2 presents some of the key GIS data layer.

### 3. Methods

Figure 3 shows a flowchart of the applied modeling process. As demonstrated, based on data collection and database development, a conceptual model has been constructed and revised before developing numerical model. Once the numerical model is running, we used GRACE data in addition to in situ data for model calibration and validation.

#### 3.1. Model development

We used the U.S. Geological Survey Modular Ground Water Flow Model- or MODFLOW to estimate change in groundwater levels at the basin scale in the IRB. The MODFLOW model is a distributed process-based numerical model, using finite different approach, which simulates runoff and infiltration from precipitation, as well as the interaction of surface water with groundwater in watersheds that range from a few square kilometers to several thousand square kilometers and for time periods that ranges from months to several decades [Harbaugh, 2005; Langevin et al., 2017; Markstrom et al., 2008]. In a groundwater flow system, natural discharge and recharge are estimated within the EVT, DRN, RCH packages. Human interventions (e.g., pumping) can be simulated by the WEL package, and canals and irrigation return are simulated by DRN and RCH packages. Modifying cropping patterns can be simulated by RCH and EVT package. Specifically, groundwater surface-water interactions are simulated by Streamflow Routing package (SFR). By adjusting the input parameters within relevant packages, such as well volumetric recharge rate, infiltration rate, maximum ET rate, extinction depth, etc., the model is used to predict the response of the flow system to external stresses (natural stresses or human interventions).

The governing equation of groundwater flow model can be described by:

$$\frac{\partial}{\partial x} \left( K_x \frac{dh}{dx} \right) + \frac{\partial}{\partial y} \left( K_y \frac{dh}{dy} \right) + \frac{\partial}{\partial z} \left( K_z \frac{dh}{dz} \right) + q_{ss} = S_s \frac{dh}{dt}$$

where  $K_x$ ,  $K_y$  and  $K_z$  are the hydraulic conductivities aligned with the x, y, and z coordinate directions;  $h$  is the potentiometric head;  $q_{ss}$  represents the sources and(or) sinks of water;  $S_s$  is the specific storage of the aquifer. Based on the 3-D differential equation, all parameters in the flow model including hydraulic conductivity, specific yield, effective porosity, total porosity are spatially distributed, and the recharge and discharge rates are both spatially and temporally variable. The simulation results can be used to calculate the groundwater storage change based on water balance formula:

$$\Delta S_{GW} = P - ET + \Delta S_B + \Delta S_S + \Delta S_o$$

where,  $\Delta S_{GW}$  is the groundwater storage change;  $P$  is the precipitation,  $ET$  is the evapotranspiration over the same period,  $\Delta S_B$  is the net volume of flow exchanged in the boundary,  $\Delta S_S$  is the net volume of flow exchanged between groundwater and stream, and  $\Delta S_o$  is the net volume of other sinks or sources, such as pumping.

In this study, we constructed a 3D multi-layer groundwater flow model for the Indus River Basin (IRB model) using MODFLOW-2005. The surface domain was delineated into 28 sub-basins. The subsurface was divided into three layers, each with 35752 active cells (5km×5km), represent the unconfined alluvial aquifer. Aquifer extent was estimated based on cross-section details for the mainstream (along A-A' cross-section line in Figure 1). The thickness of layer 1 ranges from ~10m to 100m, with a mean of ~60m. Layer 2 is ~50-200m in thickness, with a mean of ~160m. The thickness of layer 3 ranges from ~100 to 360m with a mean of ~200m. The zonation of parameters, values of hydraulic conductivity and specific yield, ET rate and extinction depth were estimated based on the lithology, soil and land cover

information, and published references (Table 2, [[Ahmad et al., 2010](#); [Alam and Olsthoorn, 2014](#); [Bonsor et al., 2017](#); [Chandio et al., 2012](#); [Hussain et al., 2017](#); [Khan et al., 2008](#); [Shakoor et al., 2018](#); [Usman et al., 2015](#)]). The initial condition of groundwater flow system of IRB is projected based on observed groundwater level and USGS report [[Greenman et al., 1967](#)] (Figure 2).

### 3.2. Model calibration

#### 3.2.1. In-situ data calibration process

We apply the trial-and-error approach and the Parameter ESTimation (PEST) program to calibrate the IRB models by adjusting individual model input parameter structures and values and assessing the difference between observed and model simulated values of hydraulic head and hydrographs [[Doherty, 1994](#)]. The goal of the calibration is to minimize the root-mean-square error of simulated and measured values (RMSE), the calibrated model was evaluated by using the Pearson correlation coefficient (R):

$$O_1 = \text{RMSE} = \left( \frac{1}{n} \sum_{i=1}^n (S_i - M_i)^2 \right)^{\frac{1}{2}}$$

$$R = \left( \sum_{i=1}^n (S_i - \bar{S})^2 \cdot \sum_{i=1}^n (M_i - \bar{M})^2 \right)^{-\frac{1}{2}} \cdot \sum_{i=1}^n (S_i - \bar{S})(M_i - \bar{M}),$$

where  $S_i$  and  $M_i$  are the  $i$ th simulated and measured values;  $\bar{S}$  and  $\bar{M}$  are the average simulated and measured values;  $n$  is the total number of simulated and observed values.  $R$  lies between -1 and 1. A value of 1 indicates a perfect positive linear correlation.

The in-situ data calibration process (Cal-I) was divided into two steps: steady-state model calibration and transient model calibration. A steady-state flow model was first constructed and calibrated to reproduce the initial flow condition. Values and zonation of hydraulic conductivity and storage parameters are adjusted in this step. In the next step, we performed a 20-stress period transient simulation and then calibrated the model using the observed data for the period 1998-2017 as listed in Table 1, anisotropy ration  $K_v/K_z$  (vertical-to-horizontal hydraulic conductivity) was also adjusted in this calibration. In Figure 4 we compared simulated and observed groundwater levels of representative wells for the southeastern portion of the IRB (Punjab). This focused calibration was aimed to capture the groundwater dynamics of the highly stressed area of Punjab which has strong data support. In the next calibration step, we are applying satellite data to improve model simulations in regions of the IRB with sparse ground data.

#### 3.2.2. GRACE calibration and validation

The Gravity Recovery and Climate Experiment (GRACE) mission measures large-scale mass changes of earth system with unprecedented accuracy ([[Scanlon et al., 2016](#)]) and hence, provides an alternative way to estimate groundwater storage change

$$\Delta S_{GW} = \Delta S_{TWS} - (\Delta S_{SW} + \Delta S_{SM} + \Delta S_{SI}),$$

where  $\Delta S_{GW}$  is the groundwater storage change;  $\Delta S_{TWS}$  the total water storage change;  $\Delta S_{SW}$  is surface water storage change;  $\Delta S_{SM}$  is soil moisture storage change;  $\Delta S_{SI}$  is snow and ice water storage change. To isolate GRACE estimated GW storage change ( $\Delta S_{GW}$ ), we derived and removed storage changes in surface water, soil moisture, snow and ice using estimates of these changes from the Global Land Data Assimilation System (GLDAS). GLDAS estimates of soil moisture are considered robust ([[Rodell et al., 2004](#)]), while the others are less reliable; however, these three sources of water balance change are negligible in the IRB. In Pakistan—and in most

agricultural settings in the world—groundwater pumping is unregulated and undocumented. Understanding this missing but critical component of the water budget can be achieved more efficiently with satellite data. Hence we further calibrated and validated our MODFLOW model using GRACE data [[Castellazzi et al., 2016](#); [Long et al., 2017](#); [Alexander Y Sun et al., 2012](#)]; that is, we adjusted the input pumping rates in certain heavily depleted areas and areas without in-situ data (Cal-II). The Nash-Sutcliffe efficiency coefficient was used to indicate the agreement between GRACE-based and flow model simulated groundwater storage change averaged within the entire Indus river basin and on annual basis [[Nash and Sutcliffe, 1970](#)].

$$O_2 = NSE = 1 - \frac{\sum_{i=1}^n (\Delta S_{GW_i}^G - \Delta S_{GW_i}^S)^2}{\sum_{i=1}^n (\Delta S_{GW_i}^G - \overline{\Delta S_{GW}^G})^2}$$

Where  $\Delta S_{GW_i}^G$  and  $\Delta S_{GW_i}^S$  are the  $i$  th groundwater storage changes obtained from GRACE observation and model simulation;  $\overline{\Delta S_{GW}^G}$  is the mean of GRACE-based groundwater storage changes;  $n$  is the total number of simulated and observed periods. NSE can range from  $-\infty$  to 1. A value of 1 indicates a perfect match, and a value smaller than 0 indicates a worse agreement than the temporal mean of the observation would provide.

### 3.3. Multi-objective parameter estimation and optimization

In this study, we first use trial-and-error, by running the forward simulation to calibrate the parameter *structures* (e.g, zonation, ratio of horizontal to vertical hydraulic conductivity, etc.). Because of the large degrees of freedom in a large-scale regional model, model calibration by manual adjustment of input parameters is only the first recourse. To improve the parameter space examined, trial-and-error approaches are followed by application of the inverse method to estimate the numerical solutions of parameters by optimizing the objective functions. Model calibration is a nonunique process [[Konikow and Bredehoeft, 1992](#)], especially in large scale model with many degrees of freedom and sparse data. A single objective function may not adequately emphasize important model characteristics, and multi-objective calibration has been found to provide more consistent and efficient results [[Yapo et al., 1998](#)]. In this study, we used an algorithm to recalibrate the model by simultaneously optimizing the objectives  $O_1$  (minimize  $O_1$  and maximize  $O_2$ ). Most groundwater flow modes are nonlinear, a general nonlinear model can be expressed as:

$$y_i = f_i(x, y, z, t, p_1, \dots, p_m) + \varepsilon_i$$

where  $y_i$  and  $f_i$  are the measured and calculated dependent variables at the  $i$  th observation point;  $x, y, z$  and  $t$  are the independent variables;  $p_1$  through  $p_m$  are the input parameters;  $\varepsilon_i$  is the residual term. Thus, the parameter estimation process can be formulated as:

$$\text{Min } O_1 = \text{Min} \left[ \left( \frac{1}{n} \sum_{i=1}^n (y_i - f_i(x, y, z, t, p_1, \dots, p_m))^2 \right)^{\frac{1}{2}} \right]$$

$$\text{Max } O_2 = \text{Max} \left[ 1 - \frac{\sum_{i=1}^n (y_i^* - f_i^*(x, y, z, t, p_1, \dots, p_m))^2}{\sum_{i=1}^n (y_i - \bar{y}_i)^2} \right]$$

$$\text{Subject to } p_j^a \ll p_j \ll p_j^b$$

where  $y_i$ ,  $f_i$ ,  $y_i^*$ ,  $f_i^*$  represent the measured and calculated values for two dependent variables;  $p_j^a$  and  $p_j^b$  are the lower and upper bounds on the  $j$  th parameter.

## 4. Results and Discussion

### 4.1. Multimethod to enhance model performance

Regional groundwater flow models are valuable part of the toolkit supporting water resources planning and adaptive management, especially in regions with limited surface water resources and highly stressed aquifers. Unfortunately, a distributed large-scale process-based model is computationally expensive and more importantly, hard to calibrate because of temporally sparse or unevenly distributed ground data (ie., head measurements). Interestingly, remote sensing products have become increasingly relevant and fusing these multisensory data into groundwater flow models can improve estimation of groundwater storage change.

In this study we use a combination of rich (but unevenly distributed) in situ water level measurements and satellite-based (GRACE) groundwater storage changes for calibrating regional groundwater models. We first run the IRB model calibrated with in situ water level data (Cal-I) resulting in RMSE of 1.34m, NRMSE (normalized RMSE) of ca. 4% and NSE of 0.25. (Figure 5). These values suggest that the IRB Cal-I model is valid; however, predicted (Cal-I) and measured (GRACE) change in groundwater storage are incoherent in time suggesting that that the model captures only local groundwater dynamics for the spatial domain with adequate in situ data.

Satellite-based data improve model performance. The IRB Cal-II using GRACE data, led to better indicators of model fit. RMSE improved to 0.91m, NRMSE increased to 0.97, and NSE was to 0.71 (Table, Figure X?). The most significant changes made during the calibration is to the pumping rates and their distributions. The initial pumping inputs are estimated based on data obtained from WASA, in a single domain of relatively small area compared to the IRB (Table 3). After calibration, we add more wells to heavily depleted area (Table 3).  $\Delta S_{GW}^S$  generated from IRB Cal-II clearly better captured groundwater dynamics—including coherence of modeled (Cal-II) and observed (GRACE) groundwater storage changes (Figure 5). Clearly, the IRB model could be improved further to better simulate the flow in extremely wet and dry years.

### 4.2. Groundwater flow and storage changes

The calibrated IRB model robustly reproduces long term groundwater flow and change in groundwater levels across the Indus River Basin (Figure 6). Simulated groundwater levels vary gradually from ~200m to ~300m in the piedmont area to 0-10m southwest near the Arabian Sea. Groundwater in the basin generally flows from the piedmont area towards the center the basin and discharges at the edge of the alluvial fan in the middle and lower parts of this basin, then enters the Arabian Sea.

In Figure 4c, we present the temporal variation of observed and simulated depth to water table (DTW) at 10 selected observation wells. DTWs mostly have a decreasing trend (water table is rising) in heavily irrigated areas (e.g., wells 157, 220, 460, 500, and 1176). By contrast, DTWs exhibit an increasing trend in urban areas (e.g., wells 287, 559 and 2721). Similarly, DTW trends also increased in areas where surface water has been diverted to the IBIS irrigation system (e.g., wells 106 and 595).

Depth to water table (DTW) has increased in most parts of the basin in the past 20 years (Figure 6). DTW largely increased in parts of Ravi, lower and central Punjab, with average rate larger than 1m/year. More than 75% of the Ravi basin has experienced increases in DTW of between 1 to 5m in the last two decades (1998 to 2017) and in some cases more than 25m. Groundwater storage depletion in those areas results mainly from intensive groundwater withdrawal due to the uneven distribution of irrigation demands.

The simulated annual water budget over 1998-2017 is listed in Table 4. Although the annual recharge from precipitation varies largely between 19.22 to 39.34 billion cubic meter (BCM) with a general increase over the last twenty years, groundwater water levels have not increased concomitantly because groundwater withdrawals have increased by nearly 7 BCM. Even after considering recharge returned to groundwater through irrigation infiltration, groundwater storage has generally decreased basin wide. For the period of 1998 to 2017, we estimate the cumulative groundwater depletion to ~32.5 BCM, which is equivalent to a reduction of 5cm in water thickness across the entire IRB.

#### 4.3. Pumping Estimates.

The agricultural area in the IRB mostly lies in the arid and semi-arid portions of the basin, thus groundwater serves as a primary water source for this land use activity. In this study, both irrigation and pumping are inferred from previous published references and they are distributed to cells based on known abstraction locations and land use land cover data (Table 3). As described in our calibration results, there is a big gap between actual groundwater depletion and our IRB Cal-I simulated depletion even though the simulated groundwater levels have good match with in site measurement. By contrast, the IRB Cal-II model generates calibrated pumping inputs by virtue of fusion with GRACE. According to the calibrated IRB Cal-II model, total annual groundwater withdrawals in the Lahore area have increased from 1.68 BCM to 1.94 BCM, with a yearly rate increase of 13 Million cubic meters (MCM). The total withdrawal in the Punjab area has increased from 58.06 BCM to 65 BCM, while the agriculture withdrawal has increased from 55.1BCM to 61.3BCM. The overall pumping across IRB has a yearly increasing rate of 0.35BCM. As present, the IBIS manages water resources based on per/post -independence historical allocations without differentiating between the waterlogged area and groundwater depletion area. Thus, optimizing crop patterns and associated groundwater extraction in space and time could improve groundwater system conditions compared to current situation.

## 5. Conclusions

A regional groundwater flow model calibrated using in situ data alone yielded incomplete solutions and satellite observations from GRACE helped us further constrain model parameters over the study period and, thus, enhanced the model performance. This study presents an innovative method for calibrating regional groundwater flow models more effectively and efficiently using both in situ measurements and remote sensing products by adjusting both input parameter structures and values following the iterative modeling process.

Major findings from this study include the following. First, although the current GRACE observations are limited by coarse resolution, GRACE proved very useful when integrated with physical process-based regional flow modeling. Second, the GRACE calibrated flow model, outperforms the model calibrated by in situ water level measurements alone, and better captures the overall groundwater dynamics. Last, the calibrated model produces more accurate spatial and temporal details about the groundwater system with less uncertainties.

The calibrated model provides valuable insights about groundwater flow dynamics in the Indus River Basin, and is useful for evaluating potential management strategies for more sustainable groundwater development in this region.

## 6. Acknowledgements

This research was carried out with support from the Future H2O of Arizona State University and Levi Strauss and Co (GR36657). We thank N. Alam and S. Ali for their invaluable assistance. The authors are grateful to all the organizations, institutes, governments, and universities for providing the valuable data sets. The data sets produced by this study are available at <http://www.hydroshare.org/resource/b7edfd376ae0439eb892cc4152872219>.

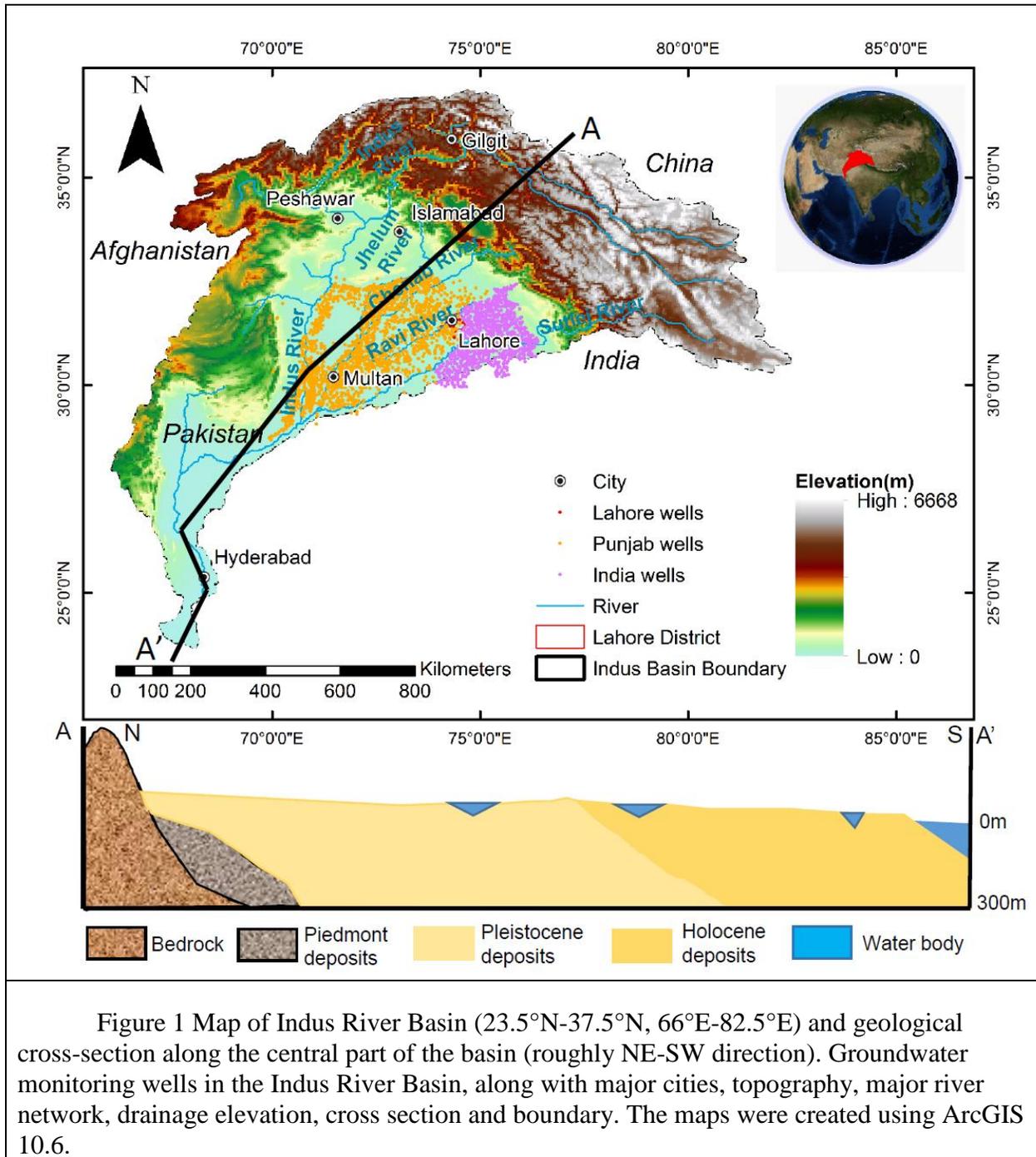
## References

- Ahmad, Z., R. Kausar, and I. Ahmad (2010), Implications of depletion of groundwater levels in three layered aquifers and its management to optimize the supply demand in the urban settlement near Kahota Industrial Triangle area, Islamabad, Pakistan, *Environ Monit Assess*, 166(1-4), 41-55.
- Alam, N., and T. N. Olsthoorn (2014), Punjab scavenger wells for sustainable additional groundwater irrigation, *Agr Water Manage*, 138, 55-67.
- Aliyari, F., R. T. Bailey, A. Tasdighi, A. Dozier, M. Arabi, and K. Zeiler (2019), Coupled SWAT-MODFLOW model for large-scale mixed agro-urban river basins, *Environ. Model. Softw.*, 200-210.
- Bonsor, H. C., et al. (2017), Hydrogeological typologies of the Indo-Gangetic basin alluvial aquifer, South Asia, *Hydrogeol J*, 25(5), 1377-1406.
- Cao, G. L., C. M. Zheng, B. R. Scanlon, J. Liu, and W. P. Li (2013), Use of flow modeling to assess sustainability of groundwater resources in the North China Plain, *Water Resour Res*, 49(1), 159-175.
- Castellazzi, P., R. Martel, D. L. Galloway, L. Longuevergne, and A. Rivera (2016), Assessing groundwater depletion and dynamics using GRACE and InSAR: Potential and limitations, *Groundwater*, 54(6), 768-780.
- Castro, M. C., and P. Goblet (2003), Calibration of regional groundwater flow models: Working toward a better understanding of site-specific systems, *Water Resour Res*, 39(6), SBH131-SBH1325.
- Chandio, A. S., T. S. Lee, and M. S. Mirjat (2012), The extent of waterlogging in the lower Indus Basin (Pakistan) - A modeling study of groundwater levels, *J Hydrol*, 426-427, 103-111.
- Dalin, C., Y. Wada, T. Kastner, and M. J. Puma (2017), Groundwater depletion embedded in international food trade, *Nature*, 543(7647), 700-+.
- de Graaf, I. E. M., R. L. P. H. van Beek, T. Gleeson, N. Moosdorf, O. Schmitz, E. H. Sutanudjaja, and M. F. P. Bierkens (2017), A global-scale two-layer transient groundwater model: Development and application to groundwater depletion, *Adv Water Resour*, 102, 53-67.
- Doherty, J. (1994), PEST: a unique computer program for model-independent parameter optimisation, *Water Down Under 94: Groundwater/Surface Hydrology Common Interest Papers; Preprints of Papers*, 551.
- Furlong, B. V., M. S. Riley, A. W. Herbert, J. A. Ingram, R. Mackay, and J. H. Tellam (2011), Using regional groundwater flow models for prediction of regional wellwater quality distributions, *J Hydrol*, 398(1-2), 1-16.
- Gleeson, T., Y. Wada, M. F. P. Bierkens, and L. P. H. van Beek (2012), Water balance of global aquifers revealed by groundwater footprint, *Nature*, 488(7410), 197-200.
- Greenman, D. W., W. V. Swarzenski, and G. D. Bennett (1967), *Ground-water hydrology of the Punjab, West Pakistan, with emphasis on problems caused by canal irrigation*, US Government Printing Office.
- Harbaugh, A. W. (2005), *MODFLOW-2005: The US Geological Survey Modular Ground-water Model--the Ground-water Flow Process*, US Geological Survey.
- Hassen, I., E. Milnes, H. Gibson, and R. Bouhlila (2019), Impact of groundwater flow across tectonic aquifer compartments in a Miocene sandstone aquifer: three-dimensional hydrogeological modeling of the Kasserine aquifer system in central Tunisia and northeastern Algeria, *Hydrogeol J*.

- Hussain, Y., S. F. Ullah, M. B. Hussain, A. Q. Aslam, G. Akhter, H. Martinez-Carvajal, and M. Cárdenas-Soto (2017), Modelling the vulnerability of groundwater to contamination in an unconfined alluvial aquifer in Pakistan, *Environ Earth Sci*, 76(2).
- Imaz-Lamadrid, M. A., J. Wurl, and E. Ramos-Velázquez (2019), Future of coastal lagoons in arid zones under climate change and anthropogenic pressure. A case study from San Jose Lagoon, Mexico, *Resources*, 8(1).
- Inam, A., P. D. Clift, L. Giosan, A. R. Tabrez, M. Tahir, M. M. Rabbani, and M. Danish (2007), The geographic, geological and oceanographic setting of the Indus River, *Large rivers: geomorphology and management*, 333-345.
- Khan, S., T. Rana, H. Gabriel, and M. K. Ullah (2008), Hydrogeologic assessment of escalating groundwater exploitation in the Indus Basin, Pakistan, *Hydrogeol J*, 16(8), 1635-1654.
- Konikow, L. F., and J. D. Bredehoeft (1992), Ground-water models cannot be validated, *Adv Water Resour*, 15(1), 75-83.
- Langevin, C. D., J. D. Hughes, E. R. Banta, R. G. Niswonger, S. Panday, and A. M. Provost (2017), Documentation for the MODFLOW 6 Groundwater Flow Model Rep. 2328-7055, US Geological Survey.
- Long, D., L. Longuevergne, and B. R. Scanlon (2014), Uncertainty in evapotranspiration from land surface modeling, remote sensing, and GRACE satellites, *Water Resour Res*, 50(2), 1131-1151.
- Long, D., Y. Pan, J. Zhou, Y. Chen, X. Y. Hou, Y. Hong, B. R. Scanlon, and L. Longuevergne (2017), Global analysis of spatiotemporal variability in merged total water storage changes using multiple GRACE products and global hydrological models, *Remote Sens Environ*, 192, 198-216.
- MacDonald, A. M., et al. (2016), Groundwater quality and depletion in the Indo-Gangetic Basin mapped from in situ observations, *Nat Geosci*, 9(10), 762-+.
- Markstrom, S. L., R. G. Niswonger, R. S. Regan, D. E. Prudic, and P. M. Barlow (2008), *GSFLOW, Coupled Ground-Water and Surface-Water Flow Model Based on the Integration of the Precipitation-Runoff Modeling System (PRMS) and the Modular Ground-Water Flow Model (MODFLOW-2005)*, US Department of the Interior, US Geological Survey.
- Nash, J., and J. Sutcliffe (1970), River flow forecasting through conceptual models part I—A discussion of principles, *J Hydrol*, 10(3), 282-290.
- Pakistan Bureau of Statistics, G. o. P. (2017), Pakistan Statistical Yearbook 2016, *Government of Pakistan*.
- Pan, Y., C. Zhang, H. L. Gong, P. J. F. Yeh, Y. J. Shen, Y. Guo, Z. Y. Huang, and X. J. Li (2017), Detection of human-induced evapotranspiration using GRACE satellite observations in the Haihe River basin of China, *Geophys Res Lett*, 44(1), 190-199.
- Richey, A. S., B. F. Thomas, M. H. Lo, J. T. Reager, J. S. Famiglietti, K. Voss, S. Swenson, and M. Rodell (2015), Quantifying renewable groundwater stress with GRACE, *Water Resour Res*, 51(7), 5217-5238.
- Richts, A., W. F. Struckmeier, and M. Zaepke (2011), WHYMAP and the groundwater resources map of the world 1: 25,000,000, in *Sustaining groundwater resources*, edited, pp. 159-173, Springer.
- Rodell, M., P. Houser, U. Jambor, J. Gottschalck, K. Mitchell, C.-J. Meng, K. Arsenault, B. Cosgrove, J. Radakovich, and M. Bosilovich (2004), The global land data assimilation system, *Bulletin of the American Meteorological Society*, 85(3), 381-394.

- Rossman, N. R., and V. A. Zlotnik (2013), Review: Regional groundwater flow modeling in heavily irrigated basins of selected states in the western United States, *Hydrogeol J*, 21(6), 1173-1192.
- Rossman, N. R., V. A. Zlotnik, and C. M. Rowe (2018), Using cumulative potential recharge for selection of GCM projections to force regional groundwater models: A Nebraska Sand Hills example, *J Hydrol*, 561, 1105-1114.
- Sahoo, S., and M. K. Jha (2017), Numerical groundwater-flow modeling to evaluate potential effects of pumping and recharge: implications for sustainable groundwater management in the Mahanadi delta region, India, *Hydrogeol J*, 25(8), 2489-2511.
- Scanlon, B. R., Z. Z. Zhang, H. Save, D. N. Wiese, F. W. Landerer, D. Long, L. Longuevergne, and J. I. Chen (2016), Global evaluation of new GRACE mascon products for hydrologic applications, *Water Resour Res*, 52(12), 9412-9429.
- Scanlon, B. R., et al. (2018), Global models underestimate large decadal declining and rising water storage trends relative to GRACE satellite data, *P Natl Acad Sci USA*, 115(6), E1080-E1089.
- Schilling, O. S., P. G. Cook, and P. Brunner (2019), Beyond Classical Observations in Hydrogeology: The Advantages of Including Exchange Flux, Temperature, Tracer Concentration, Residence Time, and Soil Moisture Observations in Groundwater Model Calibration, *Rev Geophys*, 57(1), 146-182.
- Seyoum, W. M., D. Kwon, and A. M. Milewski (2019), Downscaling GRACE TWSA data into high-resolution groundwater level anomaly using machine learning-based models in a glacial aquifer system, *Remote Sens-Basel*, 11(7).
- Shakoor, A., M. Arshad, R. Ahmad, Z. M. Khan, U. Qamar, H. U. Farid, M. Sultan, and F. Ahmad (2018), Development of groundwater flow model (MODFLOW) to simulate the escalating groundwater pumping in the Punjab, Pakistan, *Pak. J. Agric. Sci.*, 55(3).
- Sridhar, V., M. M. Billah, and J. W. Hildreth (2018), Coupled Surface and Groundwater Hydrological Modeling in a Changing Climate, *Groundwater*, 56(4), 618-635.
- Stampoulis, D., et al. (2019), Model-data fusion of hydrologic simulations and GRACE terrestrial water storage observations to estimate changes in water table depth, *Adv Water Resour*, 128, 13-27.
- Sun, A. Y. (2013), Predicting groundwater level changes using GRACE data, *Water Resour Res*, 49(9), 5900-5912.
- Sun, A. Y., R. Green, S. Swenson, and M. Rodell (2012), Toward calibration of regional groundwater models using GRACE data, *J Hydrol*, 422, 1-9.
- Tapley, B. D., S. Bettadpur, M. Watkins, and C. Reigber (2004), The gravity recovery and climate experiment: Mission overview and early results, *Geophys Res Lett*, 31(9).
- Tian, Y., Y. Zheng, B. Wu, X. Wu, J. Liu, and C. M. Zheng (2015), Modeling surface water-groundwater interaction in arid and semi-arid regions with intensive agriculture, *Environ Modell Softw*, 63, 170-184.
- Usman, M., R. Liedl, and A. Kavousi (2015), Estimation of distributed seasonal net recharge by modern satellite data in irrigated agricultural regions of Pakistan, *Environ Earth Sci*, 74(2), 1463-1486.
- Wada, Y., D. Wisser, and M. F. P. Bierkens (2014), Global modeling of withdrawal, allocation and consumptive use of surface water and groundwater resources, *Earth Syst Dynam*, 5(1), 15-40.

- Wada, Y., L. P. van Beek, C. M. van Kempen, J. W. Reckman, S. Vasak, and M. F. Bierkens (2010), Global depletion of groundwater resources, *Geophys Res Lett*, 37(20).
- Watto, M. A., S. Bashir, and N. K. Niazi (2018), Pakistan heading for groundwater crisis, *Nature*, 554(7692), 300-300.
- Werth, S., and A. Güntner (2010), Calibration analysis for water storage variability of the global hydrological model WGHM, *Hydrol Earth Syst Sc*, 14(1), 59-78.
- Werth, S., D. White, and D. Bliss (2017), GRACE Detected Rise of Groundwater in the Sahelian Niger River Basin, *Journal of Geophysical Research: Solid Earth*.
- Wu, B., Y. Zheng, X. Wu, Y. Tian, F. Han, J. Liu, and C. M. Zheng (2015), Optimizing water resources management in large river basins with integrated surface water-groundwater modeling: A surrogate-based approach, *Water Resour Res*, 51(4), 2153-2173.
- Wu, Q., B. Si, H. He, and P. Wu (2019), Determining regional-scale groundwater recharge with GRACE and GLDAS, *Remote Sens-Basel*, 11(2).
- Xu, T., A. J. Valocchi, M. Ye, and F. Liang (2017a), Quantifying model structural error: Efficient Bayesian calibration of a regional groundwater flow model using surrogates and a data-driven error model, *Water Resour Res*.
- Xu, T., A. J. Valocchi, M. Ye, F. Liang, and Y. F. Lin (2017b), Bayesian calibration of groundwater models with input data uncertainty, *Water Resour Res*, 53(4), 3224-3245.
- Yapo, P. O., H. V. Gupta, and S. Sorooshian (1998), Multi-objective global optimization for hydrologic models, *J Hydrol*, 204(1-4), 83-97.
- Zhou, Y., and W. Li (2011), A review of regional groundwater flow modeling, *Geoscience Frontiers*, 2(2), 205-214.



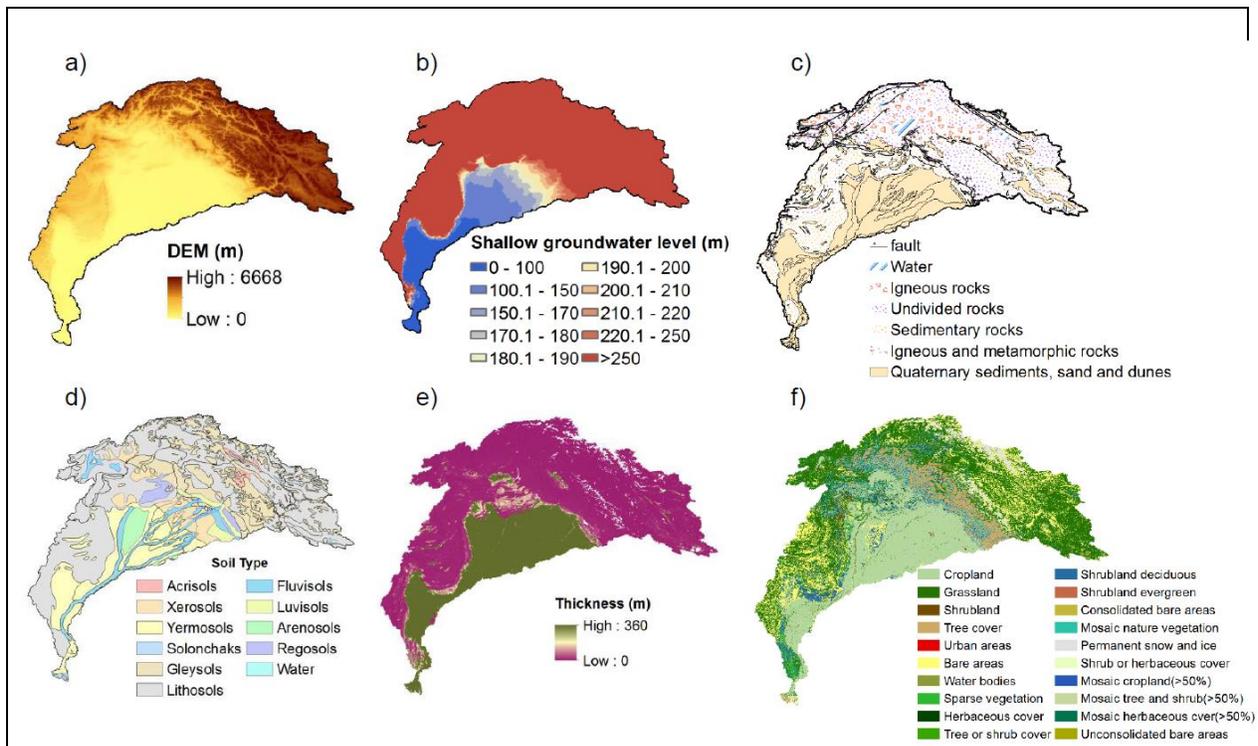


Figure 2 a) DEM of the Indus River Basin (IRB); b) Initial groundwater level of the IRB; c) Lithological map of the IRB; d) Soil types of the IRB; e) Average soil and deposits thickness of the IRB; f) Land cover and land use map of the IRB for year 1998.

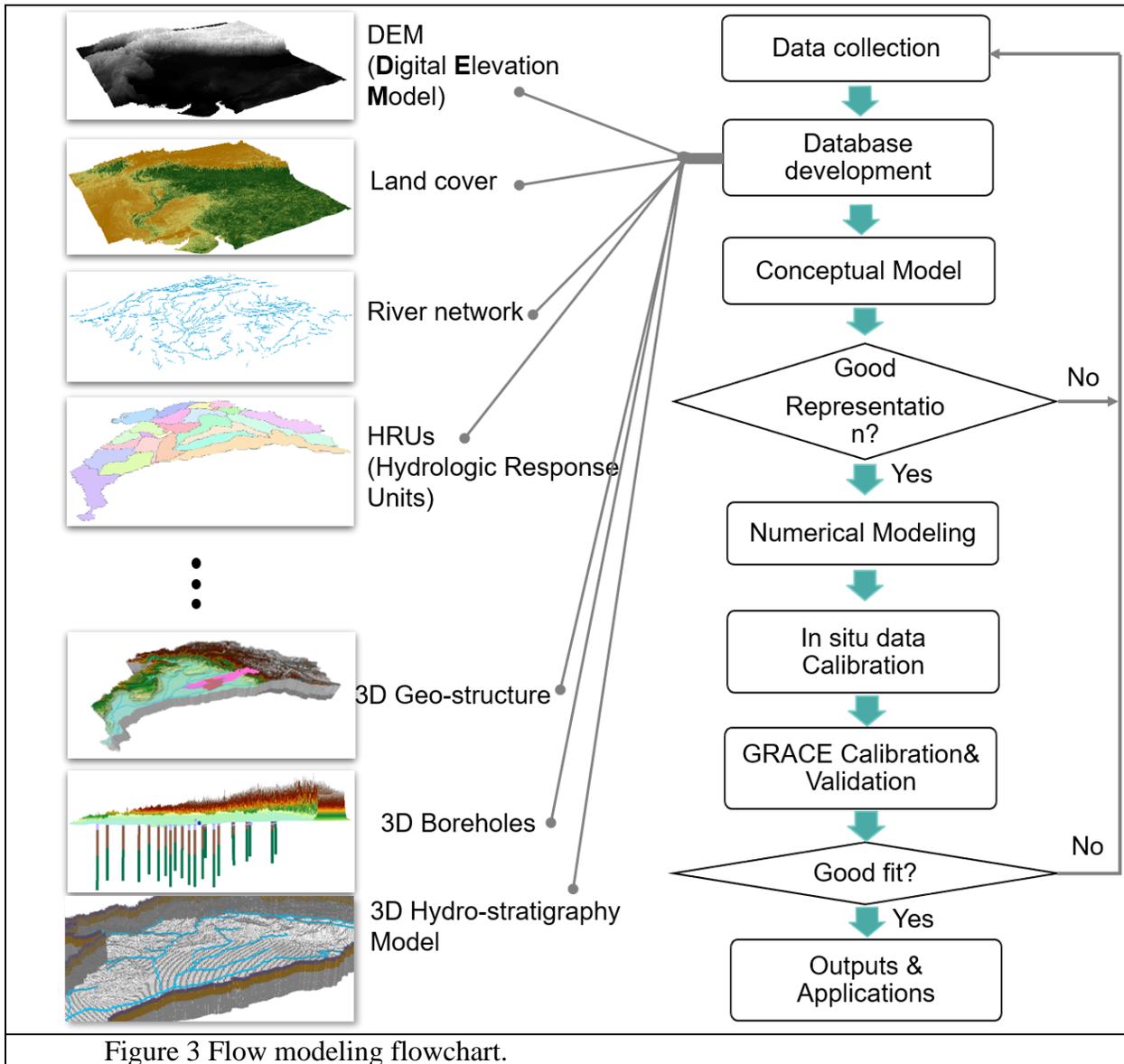


Figure 3 Flow modeling flowchart.

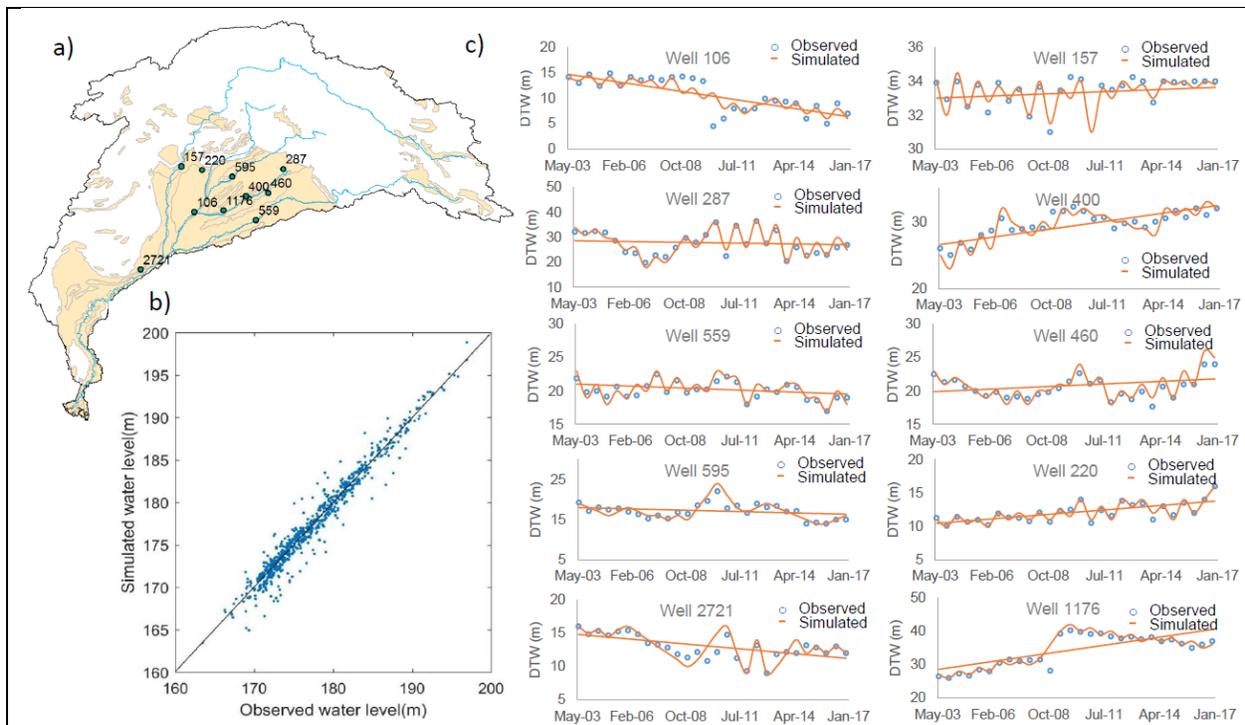


Figure 4 a) Locations of the selected 10 wells; b) Comparison of observed and simulated groundwater levels in Indus River Basin; C) Temporal variation of observed and simulated depth to water table (DTW) at 10 selected observation wells.

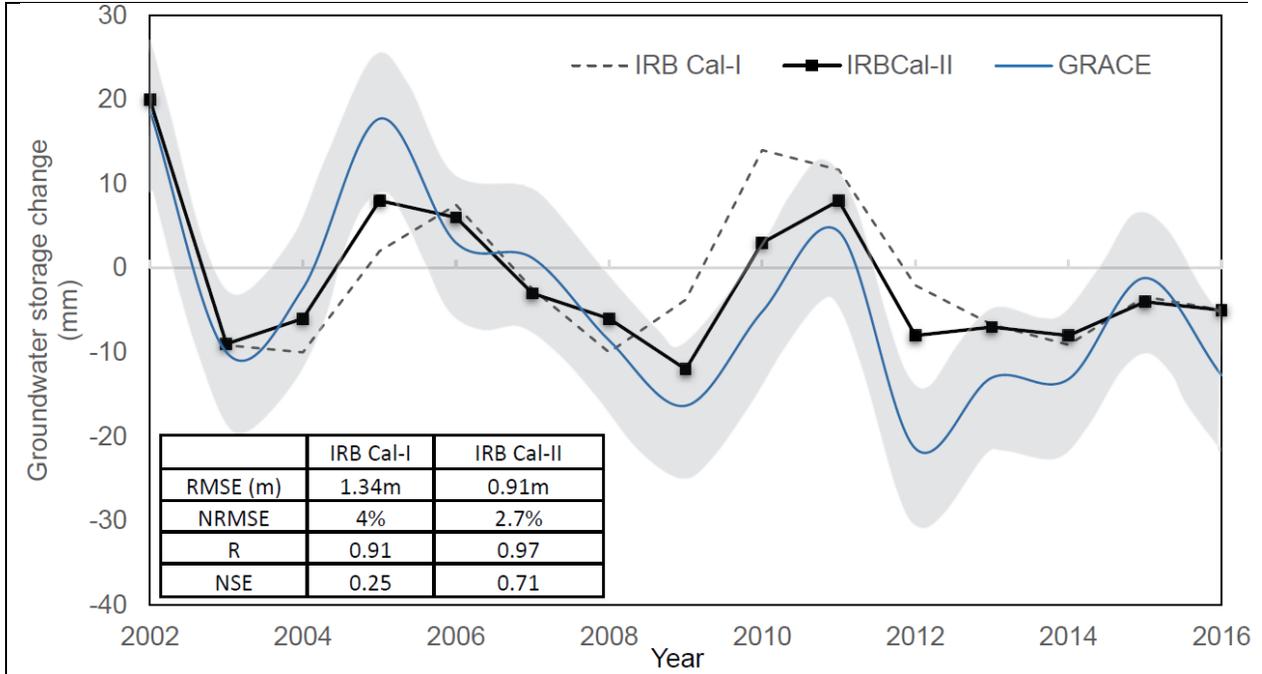


Figure 5 Comparison between  $\Delta S_{GW}^G$  and  $\Delta S_{GW}^S$  (equivalent water thickness averaged across the IRB) of Cal-I (just land-based well data) and Cal-II (well data plus GRACE), where the gray area corresponding to estimated confidence zone of  $\Delta S_{GW}^G$ .

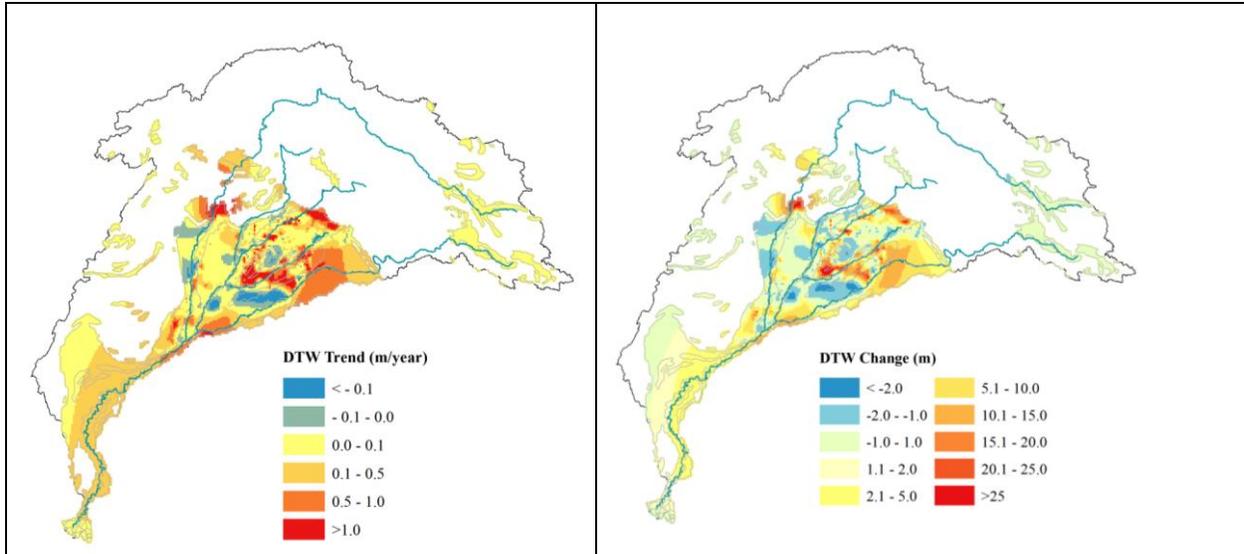


Figure 6 a) Depth to water table trend over 1998-2017; b) Depth to water table change over 1998-2017.

Table 1 Datasets and sources

Category	Data	Temporal resolution	Spatial resolution	Data format	Sources
Model setup and parameterization	DEM	2008, one time	90m	Raster	HydroSHEDS
	River network	2013, one time	15 second	Shapefile	HydroSHEDS
	Aquifer characteristics	Various, one time	Various	Documents	USGS, IWaSP, WASA, WAPDA, IGRAC
	Soil type	2007, one time	1:5 000 000	Raster	FAO-UN DSMW,
	lithological map	one time	1:1 000 000	Shapefile	GLiM
	Soil and deposits thickness	one time	90m	Raster	NASA
Initial condition	Land cover	1998-2015, yearly	300m	Raster	ESA CCI LC L4 LCCS
	Groundwater contour map	one time	1:1 000 000	Shapefile	USGS
Inputs	Precipitation	1998.01-2018.12, yearly	0.25 degree	NetCDF	NASA TRMM (TMPA/3B43) Rainfall Estimate L3
	Evapotranspiration	2000-2010, yearly	500m	Raster	NASA MOD16A3 L4
	Irrigation data	yearly	Various	Documents	PMIU
Calibration	Groundwater pumping	1998-2017, Daily, yearly	Various	Excel	WASA
	Stream flow	Daily 1998-2016,	Various	Excel	PMIU
	Groundwater level	various, yearly	4096 wells	Excel	PMIU, WASA, India_WRIS
	GRACE	Jan 2004-Dec 2009	0.5 degree	Raster	JPL-RL04M
	Soil Moisture	1998-2017, monthly	0.25 degree	Raster	GLDAS

Table 2 Values of key hydrogeological parameters of Indus Basin from previous publications

Location	Horizontal hydraulic conductivity (Kh, m/day)	Vertical hydraulic conductivity (Kv, m/day)	Specific yield (Sy)	Sources
Rechna Doab	0.05-200	0.005-20	0.05-0.25	[ <a href="#">Khan et al., 2008</a> ]
Soan River	2-170	-	-	[ <a href="#">Ahmad et al., 2010</a> ]
Thal Doab	155-362	-	-	[ <a href="#">Hussain et al., 2017</a> ]
Indo-Gangetic basin	25-500	-	0.1-0.25	[ <a href="#">Bonsor et al., 2017</a> ]
Lower Chenab Canal	24-264	0.1-0.33	0.1-0.15	[ <a href="#">Usman et al., 2015</a> ]
Punjab	35	8.75	0.35	[ <a href="#">Alam and Olsthoorn, 2014</a> ]
Lower Chenab Canal	1-265	1-15	0.05-0.25	[ <a href="#">Shakoor et al., 2018</a> ]
Lower Indus Basin	0.5-31	0.06-10	-	[ <a href="#">Chandio et al., 2012</a> ]

Table 3 Pumping inputs and calibrated values

Type	Location	Pumpage (m <sup>3</sup> /day)	Annual withdrawals (billion cubic meters)	Total number of wells	Calibrated pumpage (m <sup>3</sup> /day)	Calibrated total well numbers
Agricultural well <sup>1</sup>	Lahore	1143	-	2744	1680-1940	2744
WASA tubewell <sup>1</sup>	Punjab	733-9133	-	573	9000-12500	573
Non-WASA tubewell <sup>1</sup>	Punjab	1223-9786	-	153	6000-9800	3000
Total groundwater <sup>2</sup>	Indus	-	53.7-65.2	N/A	9000-12000	12000
GW agricultural <sup>2</sup>	Indus	-	50.5-61.3	N/A	9000-12000	10000

<sup>1</sup> Data obtained from WASA.

<sup>2</sup> Data obtained from World Bank. Date range from 1977-2012.

Table 4 Water budgets of IRB for year 1998-2017. Units: billion cubic meters.

Year	Recharge from Precipitation	ET	River leakage In-Out	Pumping	Return flow from irrigation	Groundwater Storage Change
1998	29.63	-44.00	22.80	-58.06	49.63	0.00
1999	25.22	-44.85	19.00	-58.66	56.79	-2.50
2000	19.22	-44.48	18.62	-59.26	64.90	-1.00
2001	23.36	-44.78	18.43	-59.86	59.85	-3.00
2002	20.50	-44.80	18.24	-60.41	66.47	5.00
2003	38.10	-43.69	18.05	-60.81	39.35	-9.00
2004	25.70	-45.28	17.48	-61.21	57.31	-6.00
2005	32.72	-43.31	17.10	-61.61	63.10	8.00
2006	35.43	-44.93	17.20	-62.01	60.31	6.00
2007	32.19	-44.42	17.29	-62.41	54.35	-3.00
2008	33.70	-44.02	14.44	-62.81	52.69	-6.00
2009	27.71	-44.89	11.59	-63.21	56.80	-12.00
2010	38.21	-45.19	15.96	-63.61	57.63	3.00
2011	32.51	-44.58	19.19	-64.01	64.89	8.00
2012	27.43	-43.17	18.24	-64.23	53.73	-8.00
2013	35.55	-44.27	17.10	-64.43	49.05	-7.00
2014	26.75	-43.98	18.62	-64.63	55.24	-8.00
2015	39.34	-43.93	16.34	-64.83	49.08	-4.00
2016	32.04	-44.18	19.51	-65.00	52.63	-5.00
2017	30.28	-44.36	17.64	-62.16	55.59	-3.00