1 DRAIN MODEL DESCRIPTION

1.1 Geophysical mapping and modeling

The site was intensely mapped with two surface geophysical methods: a GCM (Ground conductivity meter -- DUALEM-421S) system for the shallow subsurface (0-5 m depth) (McNeill, 1980) and a tTEM (towed-transient electromagnetic method) system for the geology (5-80 m depth) (Auken et al., 2018). The two data sets were processed and inverted independently. Both used a 1D sharp model formulation (Vignoli et al., 2015) in a spatially constrained inversion setup (Viezzoli et al., 2009) as recommended in Christiansen et al. (2016) and Frederiksen and Navarro (2021).

Parts of the site were not mapped due to fences and pavement. The gaps in the geophysical data coverage were filled in two steps. First, the resistivity models were interpolated from points (x, y, z) to a 3D grid by (S-GeMS (Remy, 2005)) using kriging, which was discretized into 10 by 10 by 1 m voxels (x, y, z, respectively) to cover the same volume as the MODFLOW model. Second, the gaps were filled using the known complex geological patterns in the area by Multi Points Statistics algorithm 'Direct Sampling' (Daly & Caers, 2010; Mariethoz et al., 2010).

1.2 Hydraulic conductivity estimation

Hydraulic conductivity of the soil was estimated using slug tests (Cheremisinoff, 1998) in areas with both low and high electrical resistivities and at different depths: 7 piezometers with a screen depth of 0.4-1.5 m (with prefix 'S' in Figure 1c and 7 piezometers with a screen depth of 2-5 m (with prefix 'D' in Figure 1c).

1.3 Geological models

As the aim was to assess drainage discharge spatial variation in different geological and topographical characteristics, four geological models were tested. Two models with homogenous hydraulic properties called 'one-zone models' and two models with heterogeneous hydraulic properties called 'Two-zone models' were developed. Four different hydraulic conductivity parameterizations were tested.

Two-zone models: Electrical resistivity values were extracted on the location of hydraulic conductivity estimates. Hydraulic conductivity estimates were then compared with electrical resistivity values. Since the aim was to create two zones, an electrical resistivity threshold value was selected between low hydraulic conductivity and high hydraulic conductivity for each two-zone model using clustering. For each model, a selected electrical resistivity threshold value was used to divide the model area into two zones. The geometric mean of hydraulic conductivity estimates was calculated for the high and low electrical resistivity zones and assigned to model cells in all model layers within the respective zone. Because the threshold value of electrical resistivity is uncertain, two threshold values were tested to assess the effect on model simulations.

One-zone models: homogenous hydraulic properties where the geometric mean of low and high hydraulic conductivity was assigned to all model cells in all model layers.

1.4 Hydrological model

A transient model was constructed using MODFLOW 6 (Langevin et al., 2017). The simulation period was from 01-Apr-2013 to 31-Mar-2015, where the first year was used as a warmup period and the second year as a calibration period. The three-dimensional model applied a uniform horizontal nodal spacing of 10 m and six layers with 120 rows (north to south) and 140 columns (east to west). The model area was kept larger than the actual study site to allow lateral flows in and out of the site. Water entered the groundwater system as a uniform recharge to the water table. The upper five layers were 1 m thick, while the bottom layer was 15 m thick. Based on our analysis, the study area has a thick clay layer below 15-20 m depth, so layers below 20 m depth were not included in the model.

1.4.1 Tile Drainage

MODFLOW 6 has a specific drain package to simulate the effects of agricultural drains, and drains are only activated in areas where the water level goes above drains (Langevin et al., 2017). The exact location of drains was unknown in the study area; therefore, drains were represented in all nodes in model layer 1 with the assumption that all regions that need to be drained have active drains. As drain depth (H_{DRN}) and drain conductance (C_{DRN}) were also unknown in study area, a H_{DRN} value (0.9 m) and a C_{DRN} value (10 m^2/day) were taken from literature (Hansen et al., 2013; Hansen, Storgaard, et al., 2019). H_{DRN} is the rate of flow of water into the drains (Langevin et al., 2017).

The drain package in MODFLOW6 is based on the principle of head-dependent flux boundary. With this boundary condition, if the groundwater head (h) in the cell falls below drain depth (H_{DRN}), the flux (Q) from the model cell to the drain drops to zero. If h in the cell raises above H_{DRN}, the Q is linearly dependent on a specified drain conductance (C_{DRN}) and the difference between the head (h) and H_{DRN} (Langevin et al., 2017) and is calculated as:

$Q = C_{DRN}(h - H_{DRN})$ (Equation 1, Langevin et al. (2017))

1.4.2 Calibration

The model was calibrated for specific yield (Sy) and specific storage (Ss) with OSTRICH tool (Matott, 2017), which is a model-independent optimization tool that includes multi-objective optimization algorithms. The Parallel Pareto archived dynamically dimensioned search (ParaPADDS) algorithm for multi-objective function optimization, and parameter estimation was used. The objective function optimized the Weighted Sum of Squared Error (WSSE) of Kling–Gupta efficiency (KGE) and percent bias (PBIAS). Sy and Ss parameters were optimized because tile drainage discharge showed sensitivity to the two parameters. The one-zone models and two-zone models were calibrated separately. After calibration, the models were run for a validation period from 01-April-2015 to 31-Mar-2016. The best solution was chosen for the analysis based on the lowest value of WSSE of KGE and PBIAS while keeping a model performance equally suitable for all four models.

1.4.3 Drainage fraction (DF)

DF is the ratio between groundwater discharge and recharge. The tile drainage discharge corresponds with the volume of water captured by the drain, and the recharge corresponds with the boundary conditions applied for each cell. It was calculated for the simulation period for each cell for layer 1 in the models, as it is the layer where recharge takes place, and the drains are located. DF is calculated as

$DF = \frac{\sum_{t=1}^{N} d}{\sum_{t=1}^{N} r}$ (Equation 2)

In eq. 2, d is tile drainage discharge volume at a specific cell, r is groundwater recharge volume at a specific cell, t is stress period (day), and N is the total number of stress periods (365 days). DF becomes zero when there is no discharge to recharge, and it becomes one when all recharge is the same as drains. DF can be above 1 due to lateral fluxes from neighboring cells or upward fluxes from deeper layers.