STØTTET AF
Promilleafgiftsfonden for landbrug

Assessing the influence of geology and topography on fine-scale simulated tile drain flow patterns

Hafsa Mahmood[1], Raphael Schneider[2], Rasmus Rumph Frederiksen[3], Anders Vest Christiansen[1] and Simon Stisen[2]

1. Aarhus University, Department of Geoscience, Aarhus, Denmark
2. Geological Survey of Denmark and Greenland (GEUS), Department of Hydrology, Copenhagen, Denmark
3. Aarhus University, Department of Ecoscience, Silkeborg, Denmark

Corresponding Author: Hafsa Mahmood (hm@geo.au.dk)

Abstract

Study region:

Denmark

Study focus:

Tile drainage, widespread across agricultural lands in Denmark, significantly impacts the hydrological cycle. The spatial patterns of generated drain flow are challenging to quantify. We used 26 tile drain sites across Denmark to study drain response in varied topographical and hydrogeological settings on the field scale. We developed 10m resolution groundwater flow models for drain sites in MIKE SHE using National hydrological model. Joint calibration of all drain sites was conducted by evaluating PBIAS and KGE of simulated and observed drain flow data. Further, we performed a correlation analysis between physical parameters and spatial patterns of simulated drain fraction (ratio of discharge to recharge per grid cell, DF) at different spatial levels: national, regional, catchment, and field scales.

New hydrological insights:

The study achieved good predictions of drain flow dynamics in the calibrated groundwater flow models for 26 drain sites. On a national scale, the correlation of DF with topographic index variables was high. On the regional scale, Lolland-Sjaelland, and Jylland showed high correlation to topographical index variables, while Fyn showed a high correlation with clay fraction. The research provided a broad understanding of parameters controlling the spatial distribution of drain flows across Denmark. In future, the calibrated groundwater flow models can produce training dataset of DF for data-driven approaches to predict the spatial distribution of DF across Denmark

Keywords

Subsurface drains, shallow groundwater modeling, hydrogeological variables, topographical variables

Highlights

1. Jointly calibrated 10m resolution groundwater models for 26 catchments accurately predict the spatial distribution of drain flow and the temporal dynamics.
2. In areas with low hydraulic conductivity (Fyn), the average clay fraction (%) (related to the hydraulic conductivity) is the main controlling variable for drain fraction
3. In other areas, the topographical position index and topographical wetness index are more important in controlling drain flow spatial patterns.

1 Introduction

Tile drainage is a beneficial agricultural technique for regulating the subsurface water level in waterlogged regions to make areas suitable for agriculture and more productive (Jeantet et al., 2021; Prinds et al., 2019). However, there are also potential environmental consequences associated with tile-drained agricultural areas, i.e., high nutrient loads to surface waters (Hansen, Jakobsen, et al., 2019; Jacobsen & Hansen, 2016; Prinds et al., 2019; Stenberg et al., 2012). These potential environmental consequences are because tile drains significantly influence the hydrological flows – groundwater recharge, surface water fluxes, and indirectly the nutrient transport associated with the hydrological flows (Hojberg et al., 2017). The partitioning of these hydrological fluxes between groundwater and surface water is partly controlled by drain flows and their spatial variability due to differences in geology and topography.

In Denmark, around 50% of the agricultural land is tile-drained (Moller et al., 2018). Drain flow patterns are not only influenced by drain infrastructure but also geology and topography can influence the spatial drain flow patterns (Amado et al., 2017; Hansen, Jakobsen, et al., 2019; Hansen, Storgaard, et al., 2019; Motarjemi et al., 2021; Williams et al., 2015). Few studies have been conducted to improve the understanding of spatial drain flow patterns on a large scale and to identify its most influential topographical and geological variables, for example, soil properties, topographical indexes, etc. Boico et al. (2022) studied the sensitivity of hydraulic conductivities and specific yields of geological layers on spatial drain flows. However, Boico et al. (2022) did not include the influence of topographical variables; nonetheless, they recommended it for future investigations. Hansen, Storgaard, et al. (2019) studied the correlation between the spatial distribution of drain flow and topographical wetness index (TWI) in 100m resolution but was unsuccessful in finding any significant correlation. Hansen, Storgaard, et al. (2019) did study impact of deeper geological layers below the tile drain level and found geological layers below 2-5m significant for the spatial distribution of drain flow. However, both Hansen, Storgaard, et al. (2019) and Boico et al. (2022) studied these correlations based on a single drain field to understand the physical control variables that drive the spatial distribution of drain flow. To study the spatial distribution of drain flow, a variety of topographical and geological settings should be considered. Only Motarjemi et al. (2021) investigated the influence of geological and topographical indexes on yearly drain flow amounts on multiple sites. However, the spatial distribution of drain flow within the catchments was not part of the study. Motarjemi et al. (2021) found no clear correlation with topographical indexes (TI). They highlighted that TI is less important in predicting drain flow because there can be considerable variation of TI within the drain sites, which was not in the scope of their study.

This study investigates how physical variables (geology – soil, and topography) regulate the spatial distribution of drain flow in tile-drained agricultural areas. Denmark's existing national groundwater flow model has a resolution of 500m or 100m (Henriksen et al., 2020; Stisen et al., 2019) and cannot produce spatial drain flow patterns on the field scale (1-120 ha) for multiple reasons. Firstly, the model is not validated against direct drain flow observations. Secondly, the coarse resolution makes it challenging to decipher field scale controls of drain flow patterns. For example, the most relevant driver of drain flow patterns is water table depth, and there are variations in water table depth below 100m resolution, as indicated by (Koch et al., 2021). Motarjemi et al. (2021) also reaffirm that complex groundwater-drain flow patterns of field scale tile-drained catchments cannot be represented by a national scale hydrological model with a coarse resolution. To achieve our objective, first, we established, calibrated, and validated a physically distributed groundwater flow model in 10m resolution that can simulate drain flow dynamics for several field scale drain sites in Denmark. Then, we investigated the physical control variables on the model-generated spatial distribution of drain fraction (DF) (i.e., the ratio of drainage volume to recharge volume per grid).

2 Materials and methods

2.1 Data collection – drain stations

Daily drain flow data and corresponding drain sites' boundaries are available across Denmark for 26 drain sites. All drain sites were tile drained, with the size of the drain sites varying between 1 to 120 ha. In Sjælland, Lolland, and Lillebæk, data are available for four sites each. For Jylland, data is available for a total of 14 sites. Out of 14 sites, 11 sites are in situated in the the Norsminde catchment (Mid Jylland), 1 site in Vadum (Upper Jylland), 1 in Fillerup (Mid Jylland) and 1 in Ulvsborg (Mid Jylland). *Figure 1* shows all sites and their grouping on a different scale.

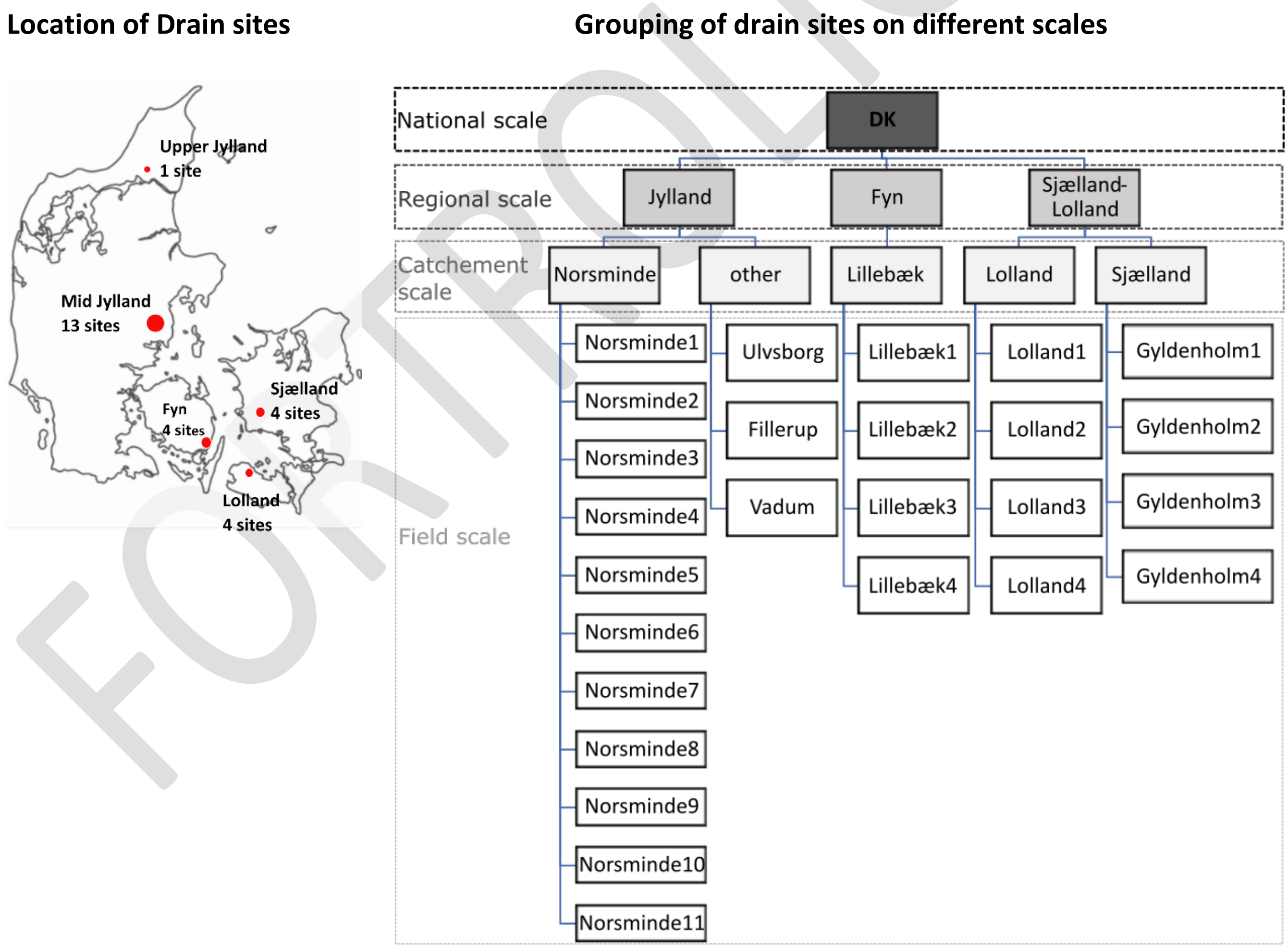

Figure 1 Drain sites selected in Denmark and their grouping on national, regional, catchment, and field scale

2.2 National hydrological model and sub-models

As a baseline model, the most recent version of the national groundwater model for Denmark in 100m resolution, referred to as the '100m DK-HIP model', is used in this study (Henriksen et al., 2020). The model describes surface and subsurface hydrological processes and structures in the MIKE SHE model code (Abbott et al., 1986; DHI, 2020). The model is forced by a national, gridded daily dataset of precipitation (10km resolution), temperature, and reference evapotranspiration (20km resolution) (Scharling, 1999a, 1999b). The model is calibrated against approximately 660,000 individual groundwater head observations and 308 stream flow observations throughout the country (Henriksen et al., 2020). The model is calibrated from 2000 to 2010 and validated for 1990 to 2019. Therefore, the national model can be used for the interval 1990 to 2019.

As the objective is to investigate the drain flow on a field scale, we set up a MIKE SHE model for each of the 26 drain sites in 10*10m grid resolution with boundary conditions from the 100m DK-HIP model. The 100m DK-HIP model is the base of all the 10m resolution groundwater models. As the 10m resolution groundwater models focus on reproducing drain flow dynamics and spatial patterns, we refined relevant topographic and drain parameters to 10m. All other subsurface descriptions are the same as in the 100m DK-HIP model. The most crucial parameters for field scale simulation included drain time constant, drain depth, topographical data, paved area fraction, and the hydraulic conductivity of the uppermost geological layer (top 2m). Our drain sites' models had no interference from streams and lakes, no pumping wells, and no irrigation. As knowledge about the exact location of drainpipes does not exist and MIKE-SHE only allows implicit representation of drains, we assume drains are in all model cells for all drain sites. We consider this assumption reasonable for the areas known to be tile drained. Therefore, all cells with groundwater levels above drain depth at a particular time will generate drain flow. The model area is extended with a buffer of 200m from all sides to alleviate the effect of the applied boundary conditions.

One of the most important parameters for our study is the hydraulic conductivity of the uppermost layer of the subsurface (2m thickness in this setup). The 100m DK-HIP model has a soil classes-based hydraulic conductivity map (Jakobsen et al., 2015), while we wanted a distributed one with higher resolution. Therefore, we developed a pedo-transfer function based on the existing 30 m clay fraction map produced by Adhikari et al. (2013). To parameterize the pedo-transfer function, we evaluated the correlation between the clay fraction map and the DK-HIP model 100m resolution hydraulic conductivity map. The pedo-transfer function converted the 30 m resolution clay fraction map to a 30m resolution hydraulic conductivity map. The pedo-transfer function-derived hydraulic conductivity map produced similar patterns to the class-based conductivity map of the 100m model.

2.3 Joint calibration of 26 sub-models

Before calibration, we ran all the 10m resolution groundwater models with parameters directly from the 100m HIP-DK model. We observed an overall underestimation of drain flows for all 26 drain sites. The calibration of the 100m HIP-DK model was based on streamflow time series and groundwater heads, but not drain flows, which might have led to this bias. Another reason for underestimation is the resolution dependency of model parameters when going from a 100m to 10m resolution. We calibrated all 26 drain sites with a refined 10m resolution keeping in view that the primary purpose of models is to estimate the drain flow dynamics.

A joint calibration is performed across all the drain sites to get one parameter set that fits all drain sites across Denmark. This is in line with the calibration scheme for the 100m HIP-DK model, which is calibrated with one parametrization to secure spatial consistency in model results. We used the OSTRICH calibration software. The Pareto archived dynamically dimensioned search (PADDS) algorithm of OSTRICH is suitable for multi-objective optimization (Asadzadeh & Tolson, 2013;

Matott, 2017). In the output of PADDS, all solutions are stored, including those along the Pareto front (i.e., solutions where no individual objective function can be improved at the cost of at least one other objective function). The objective function (Q) is comprised of the Kling-Gupta efficiency (KGE) and percentage bias (PBIAS) between observed and simulated drain flows at each of the drain sites' outlets. KGE and PBIAS are used in the objective function because the KGE represents the cumulative misfit of prediction of the drain flow dynamics and PBIAS represents the overall misfit of the simulated drain flow. Hence, the objective function was to minimize the sum of squared errors SSE for KGE and PBIAS across the 26 catchments.

$$\text{Equation 1: } Q = min[SSE(KGE) + SSE(PBIAS)]$$

In OSTRICH, two types of parameters are used for calibration; 1. main parameters, 2. tied parameters. The main parameters are directly altered and independent, while tied parameters depend on the values of the main parameters. During calibration, changes in a parameter alter the entire spatial distribution of that parameter for all regions of Denmark while keeping the same spatial relative differences among all regions of Denmark. The main parameters influencing shallow drain flow dynamics used as calibration parameters are shown in Figure 2. The selection of calibration parameters was based on past experience with the DK-Model (Henriksen et al., 2019; Højberg et al., 2015). The deeper geological layers were not included in calibration as all drain sites have relatively low influence from lateral flows. Also, we used boundary conditions of the 100m DK-HIP model and using different hydrogeological parameter values of the deeper layers would make the application of the 100m DK-HIP model's boundary conditions invalid.

The main parameters (P) used for calibration include the slope (P1) and intercept (P2) of the pedo-transfer function converting the clay fraction map into hydraulic conductivity for the upper 2m of soil; the hydraulic conductivity of the first clay layer below 2m depth for three parts of Denmark respectively (Sjælland-Lolland, Jylland, Fyn; P3, P4, P5); a factor between the hydraulic conductivity of the first sand and first clay layer for Sjælland, Jylland, and Fyn, respectively (P6, P7, P8), to ensure that the clay layers' hydraulic conductivity is lower than the sand layers'. Furthermore, the rooting depth is included (P9), and the drain time constant (P10) controls the conductance of subsurface drains, as well as the drain depth (P11). The tied parameter includes geological layers under 2m depth and rooting depth across different crop types.

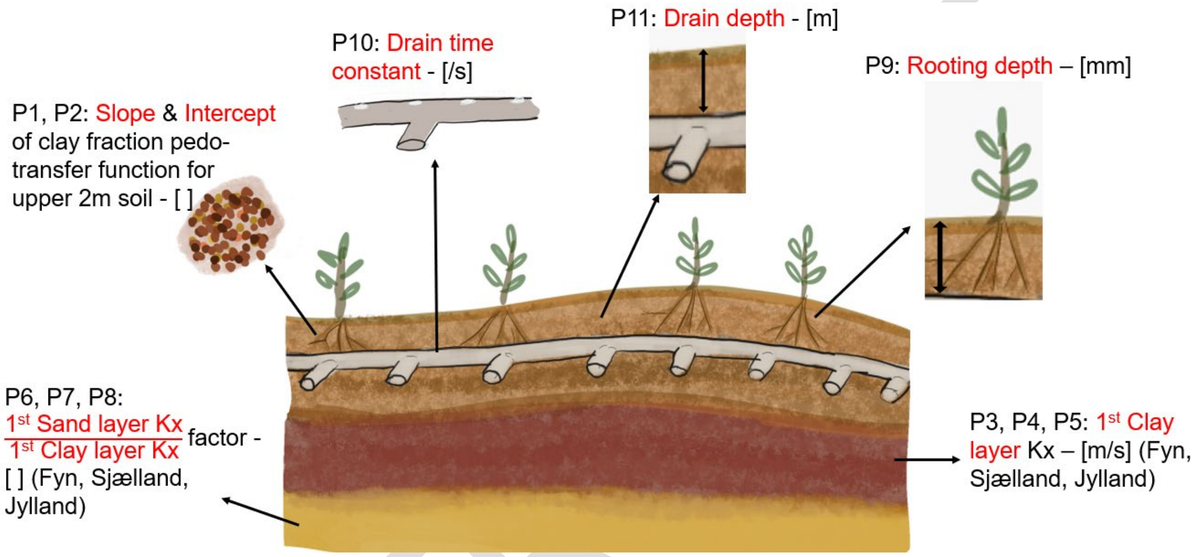

Figure 2 Parameters selected for hydrological model calibration

2.4 Drain Fraction (DF)

DF is a measure of the average drain flow to recharge ratio for the simulation period for each node in the model and is calculated as

$$\text{Equation 2: } DF = \frac{\sum_{t=1}^{N} d_t}{\sum_{t=1}^{N} r_t}$$

In eq. 2, d is the volume of drain flow at a specific cell, r is the recharge volume at the specific cell, t is the stress period (in days), and N is the total number of stress periods. A zero value of DF indicates no drain flow, while a DF value between 0 to 1 indicates recharge is higher than drain flow. DF value above 1 indicates that drain discharge is higher than recharge, and additional sources of water to the grid, e.g., lateral and upward fluxes from deeper groundwater, contribute to drain flow

2.5 Uncertainty in hydrological model and simulated DF

Five separate calibration runs were conducted with an identical setup except for different random seeds of the PADDS algorithm to get a variety of parameter combinations that give the best model performance. Different random seed values can generate slightly different outcomes for the PADDS algorithm. We selected five solutions (parameter sets) from these runs instead of one specific parameter set for the groundwater flow model. In the selection, we focused on choosing a solution on the Pareto front with equally good model performance but showing some variations in the parameter set.

For all five solutions, the spatial distribution of DF was calculated for all 26 drain sites using the spatial distribution of recharge and drain flow. Spatial mean DF was estimated by spatially averaging the DF obtained from 5 selected solutions.

2.6 Evaluation of depth to water table with spatial DF distribution

For most drain sites, relevant groundwater level observations are lacking. Hence, groundwater levels are not included in the model calibration. Nevertheless, groundwater level observations existed for two drain site sites, Norsminde3 and Gedved, which have been used to evaluate the validity of the simulated groundwater levels. The Gedved drain site is not part of the 26 drain site because it has no drain flow data between 1990 to 2019. The two catchments, Norsminde3 and Gedved, are covered by a dense network of shallow piezometers of 1.5 m depth with a screen from 0.5 m to 1.5 m, with 31 piezometers in Norsminde3 and 28 piezometers in Gedved. We gathered monthly readings of the depth of the water table from December 2019 to May 2022 (the Winter season). An observation-based estimate of drain flow probability is calculated using the monthly depth to water table readings for the winter months using drain depth as a threshold. This estimate of drain flow probability is expected to be correlated to DF. The Spearman correlation is derived between the drain flow probability estimates and model simulated spatial DF to validate the groundwater model spatial DF accuracy.

2.7 Physical control variables and Correlation analysis

The aim is to assess the physical control variables of the simulated spatial DF patterns to understand what drives the simulated drain flow spatial patterns. Topographical and geological variables are assumed as the primary physical variables and we investigate which derived indexes from those data explain the spatial variability in DF generation. All topographic variables are derived from the digital elevation model in 10m resolution. The geological variables used for correlation analysis are Clay fraction (%) in Horizon a (0-5 cm depth), Horizon b (5-15 cm), Horizon c (15-30 cm), Horizon d (30-60 cm) as developed by Adhikari et al. (2013) interpolated from their native 30m resolution to 10m; and the thickness of first clay layer and sand layer from the nationwide hydrogeological interpretation (EPA, 2020) interpolated from their native 100m resolution to 10m.

Topographical variables are TWI, Topographical Position Index (TPI), Terrain Ruggedness Index (TRI), Roughness, slope, Curvature, and Aspect (Los Huertos & Smith, 2013).

TWI represents water accumulation from its upstream area at a specific point in space. It is calculated as:

$$\text{Equation 3: } TWI = ln\,(\frac{x}{tan\,(y)})$$

In eq. 3, x is the upslope contributing area, and y is the slope angle in radian (Beven & Kirkby, 1979; Mattivi et al., 2019). TPI refers to the difference between the central pixel's elevation and its neighboring pixel's average elevation in a specific radius (Gallant & Wilson, 2000). We calculated TPI for 10m, 20m, 100m, and 200m radii. TPI represents relative differences in topography. On the other hand, TRI compares a central pixel with its neighbors by taking the absolute values of the differences between the central pixel and surrounding pixels and averaging the result (Riley et al., 1999; Wilson et al., 2007). We calculated TRI in a 30 m radius Roughness is the degree of irregularity of the catchment surface. It is estimated by the largest inter-cell difference between a central pixel and its adjacent cell (Wilson et al., 2007).

A slope is a change in elevation over a distance (Horn, 1981). Curvature represents the shape of a slope, whether it is convex or concave. The vertical shape of the slope (parallel to the slope) is called profile curvature. The horizontal shape of the slope (perpendicular to the slope) is called plan curvature. Curvature is calculated by fitting a surface to the central cell and its neighbors. It combines profile and planform curvature. Plan curvature affects the flow convergence or divergence, while profile curvature affects the flow acceleration (Zevenbergen & Thorne, 1987).

Aspect is the direction the slope facing at a specific location (Horn, 1981).

Covariance among the identified physical variables was used to exclude redundant variables. Pearson correlation between model simulated DF and the unique physical variables was determined for the catchment scale. Model simulated DF was calculated from average DF estimated from 5 selected solutions. Pixels of all drain sites in the same catchment were aggregated for catchment scale correlation. Two top correlated variables for both geology and topography were selected using the catchment scale correlations. Correlation between four shortlisted variables and model simulated DF was studied by grouping drain sites on different scales: national, regional, catchment, and field. For different scale analyses, pixels of each drain site were placed in different groups based on the grouping described in *Figure 1*.

3 Results

3.1 Groundwater models and calibration

Figure 3 shows mean model performance across 26 drain sites for the two objective functions, KGE and PBIAS, for five calibrations run with different seed numbers. Each point indicates a unique set of calibration parameters used to run the groundwater models for all 26 drain sites. The calibration results depict a significant improvement in model performance from a mean PBIAS and mean KGE of -75.6 % and -0.28 to -6.7% and 0.53 for the initial run with parameters from the 100m HIP-DK model and the calibrated 10m resolution groundwater models, respectively. The improvement of PBIAS from -75.6% to -6.7% indicates a significant decrease in the underestimation of drain flow found before the re-calibration. The mean KGE increased to 0.53, indicating a good ability to follow drain flow temporal dynamics, especially considering the precipitation uncertainty at the event scale and field scale, and the extreme peaks of drain flow. Along the Pareto front, five solutions are selected. The primary objective function (Equation 1) narrows down to the top 1% of the solutions. Five solutions are selected from the top 1% based on the lowest mean|PBIAS| and variation in the parameter sets. The calibration parameter values are described in Table 1.

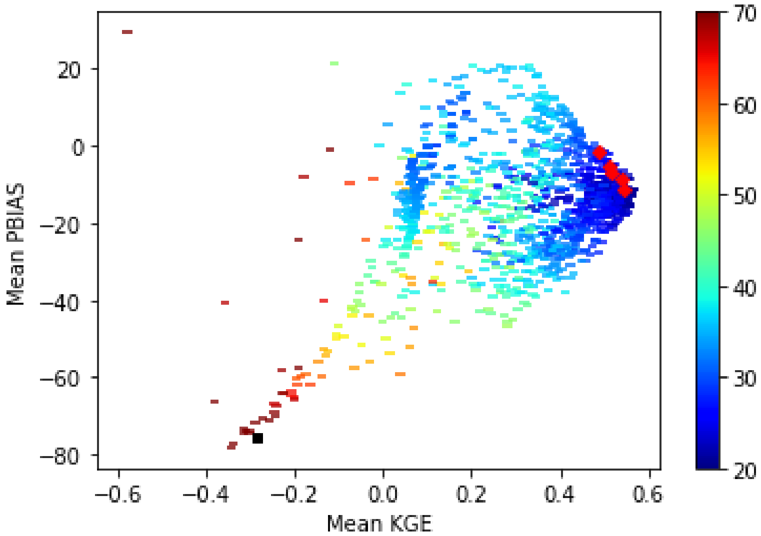

Figure 3 Mean performance of 26 drain sites over two objective functions: PBIAS and KGE. Each point indicates a unique set of parameters for running the 10m resolution groundwater flow models of 26 drain sites. The color scale represents mean|PBIAS| across 26 drain sites (the black square represents before calibration model performance, and red squares indicate the five solutions selected for analysis)

Example drain hydrographs from two drain sites, Gyldenholm4 and Norsminde1, are shown in Figure 4. The figures depict two hydrographs for each drain site, one with the highest KGE and the other with the lowest KGE value among the five selected solutions. It is clear from the hydrographs that during peak flow periods, there are differences between simulated and observed drain flows. Overall simulated drain flow is lower than observed in Gyldenholm4, which is indicated by negative PBIAS values of -20.4% and -21.7%. In Norsminde1, simulated drain flow is higher than the observed drain flow, and the positive PBIAS values of 16.5% and 15.21% indicate it.

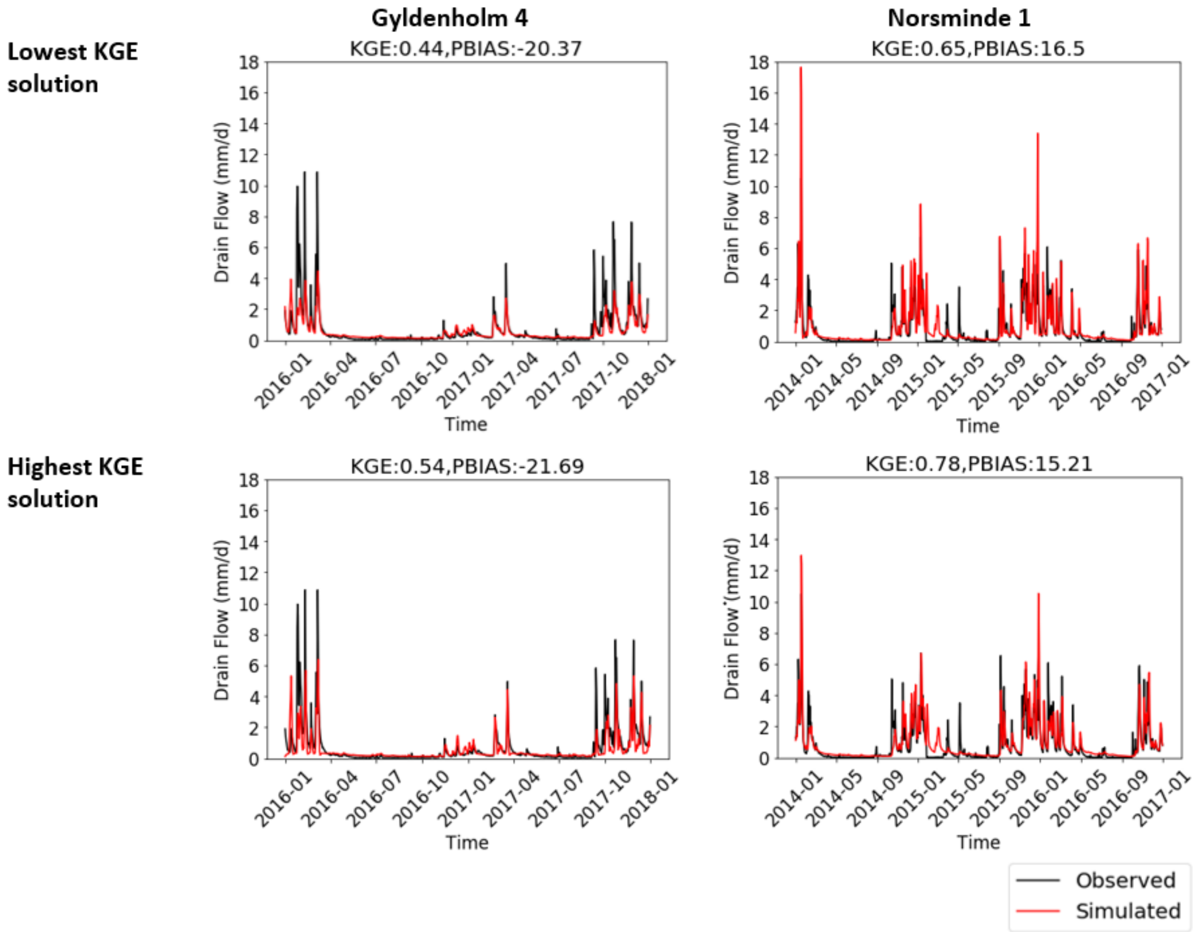

Figure 4 Hydrograph of observed (black) and simulated drain flow (red). The top row shows the lowest KGE solution of the five selected solutions, and the bottom row shows the highest KGE solution.

Mean KGE along all drain sites varies between 0.1 to 0.8 (*Figure 5*). The results show a mean KGE value above 0.4 in all drain sites except Vadum, Norsminde 8-10, and Lolland 3. This indicates that the 10m resolution groundwater models can capture the drain dynamics and seasonality on field scale drain sites after joint calibration. PBIAS values along all drain sites vary from 52% to -40% (*Figure 5*). |PBIAS| value is below 25% in all drain sites except Norsminde 4, Norsminde5, Norsminde 7, Norsminde8, Norsminde 9, Norsminde10, Norsminde11, Lillebæk3, Lolland3. The variation in KGE value across the five selected solutions is highest among the four Lillebæk catchments and Lolland3. As for the KGE, the Lillebæk and Lolland drain sites also depict higher variation in PBIAS values than other drain sites. The variation in KGE and PBIAS highlights the variation in drain flow prediction among different solutions (*Figure 5*). Variations in Lillebæk catchments are high because the variation in the parameters of the five solutions is also highest in Lillebæk catchments.

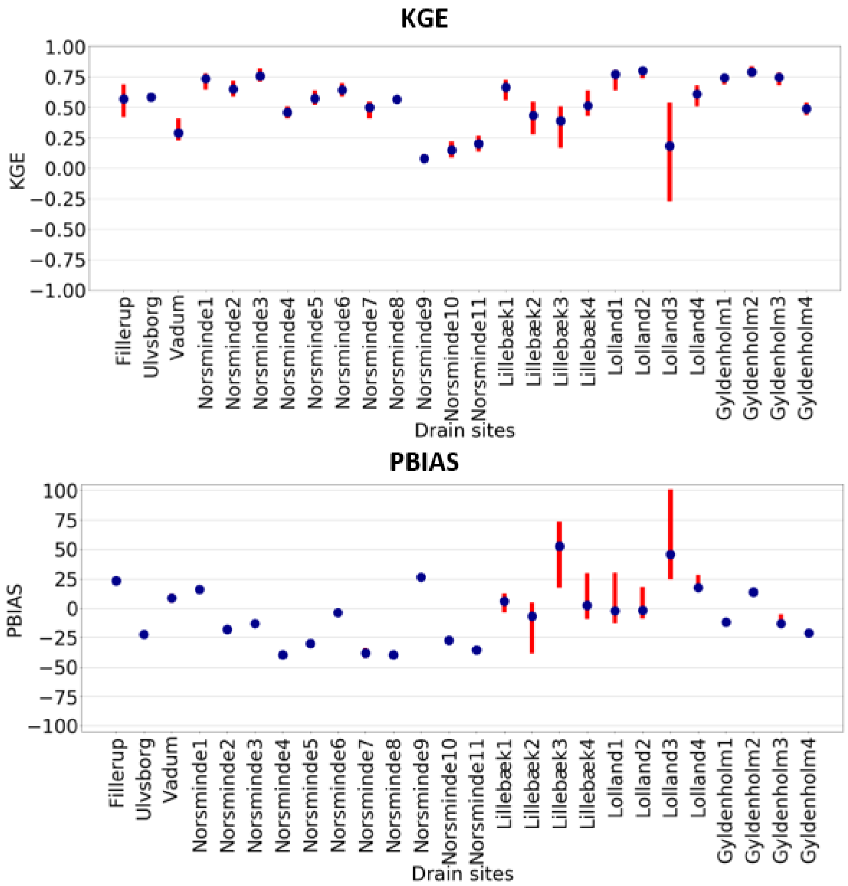

Figure 5 Mean KGE and mean PBIAS of selected solutions across 26 drain sites. The blue point shows the mean KGE/PBIAS value, while the red lines show the variation in KGE/PBIAS within the five solutions for each catchment

Table 1 Set of calibration parameters of the 5 selected solutions and parameter bounds used in the calibration

Solution no.	**P1: Slope of pedo-transfer function [-]**	**P2: Intercept of pedo-transfer function [-]**	**P3: 1st clay Kx Sjælland [m/s]**	**P6: $\frac{\text{1st sand Kx}}{\text{1st clay Kx}}$ in Sjælland [-]**	**P4: 1st clay Kx Fyn [m/s]**	**P7: $\frac{\text{1st sand Kx}}{\text{1st clay Kx}}$ in Fyn [-]**	**P5: 1st clay Kx Jylland [m/s]**	**P8: $\frac{\text{1st sand Kx}}{\text{1st clay Kx}}$ in Jylland [-]**	**P9: Rooting depth [mm]**	**P10: Drain time constant [1/s]**	**P11: Drain depth [m]**
Bounds	-0.5 to -0.15	-5.0 to 0.0	1.0E-08 to 1.0E-05	1.0E+04 to 1.0E+02	1.0E-08 to 1.0E-05	1.0E+04 to 1.0E+02	1.0E-08 to 1.0E-05	1.0E+04 to 1.0E+02	600 to 200	1.00E-9 to 1.00E-05	1.2 to 0.8
1	-0.375	0.0	3.980E-08	112	1.210E-06	111	1.000E-08	100	613	3.220E-07	1.20
2	-0.317	-0.804	1.410E-07	163	2.240E-07	10000	1.070E-08	105	600	2.360E-07	1.15
3	-0.317	-0.853	1.410E-07	163	5.290E-07	263	1.000E-08	103	600	5.680E-07	1.15
4	-0.317	-0.804	1.410E-07	163	3.510E-07	10000	1.070E-08	100	600	2.950E-07	1.20
5	-0.317	-0.8.04	1.470E-08	200	2.220E-07	100	1.070E-08	105	600	2.360E-07	1.10

3.2 Spatial distribution of DF

After reasonably calibrating the 10m resolution groundwater models, we obtained the mean drain fraction across the five selected solutions. This is because all five solutions are reasonable and including all five solutions DF will help to incorporate differences in DF due to parameters' uncertainty. *Figure 6* shows the spatial distribution of DF in two example drain sites: Gyldenholm4 and Norsminde1. In the upper row, the blue to cyan color indicates low DF, and the yellow to red color indicates high DF. Recharge from cells with a low DF (or even a DF of zero) can also contribute to drain flow as it can travel laterally to downstream cells, upwell there and potentially drain from there. That means that DF values above 1 represent areas where subsurface flow is accumulated from neighboring upstream regions and deeper layers.

In the lower row of *Figure 6*, the standard deviation of DF across the five selected solutions is shown. In Gyldenholm 4, the northwestern part shows more standard deviation than the southeastern part, while in Norsminde1, the standard deviation is higher only in high DF areas (*Figure 6*). The difference in the standard deviation in Gyldenholm4 might be because of the more rugged terrain in the northwestern part than the southeastern part (*Figure 6*).

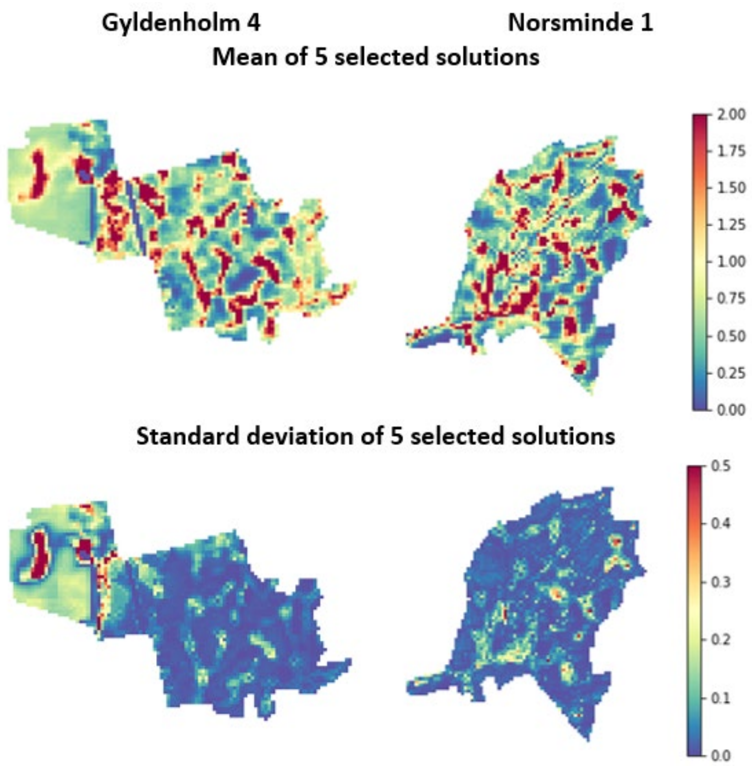

Figure 6 Mean and standard deviation of DF across the five solutions in Gyldenholm1 and FensholtD1

Figure 7 illustrates simulated and observed DF as average per drain site. It shows that our 10m resolution groundwater models underestimate DF in 10 out of 26 drain sites while the remaining 16 are overestimated. The variation in DF between the catchments is reasonably well captured with Pearson R of 0.6, indicating that the models can differentiate between high and low DFs at the field scale despite the joint calibration.

For the spatial variability in simulated DF within each drain site, Vadum has the highest standard deviation in spatial DF, followed by Gyldenholm, Norsminde, Fillerup, and Ulvsborg drain sites. Lolland and Lillebæk drain sites show the lowest standard deviation among spatial DFs.

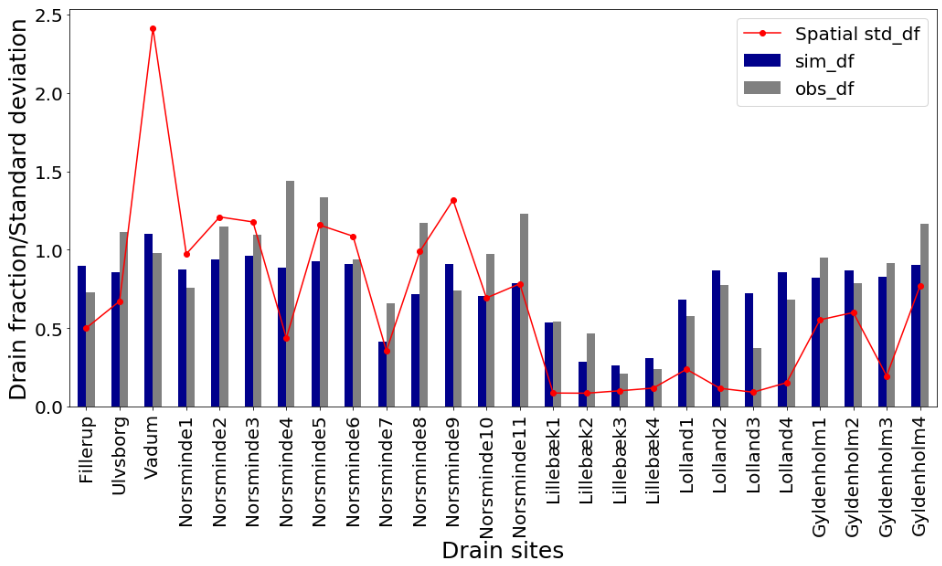

Figure 7 Means of observed and simulated DF for each catchment (grey and blue bars), together with the standard deviation of the spatial distribution of DF (red line). The y-axis represents mean DF or standard deviation of DF, respectively

Spatial comparison of simulated DF and drain flow probability

The spatial comparison of simulated DF at the piezometer locations with drain flow probability estimates based on observed groundwater levels is shown in *Figure 8*. Satisfactory Spearman correlations are observed for both Gedved (0.68) and Norsminde3 (0.50). The Spearman correlation is lower in Norsminde because of one outlier (piezometer 1). The removal of piezometer 1 increases the Spearman correlation to 0.63 (*Figure 8*).

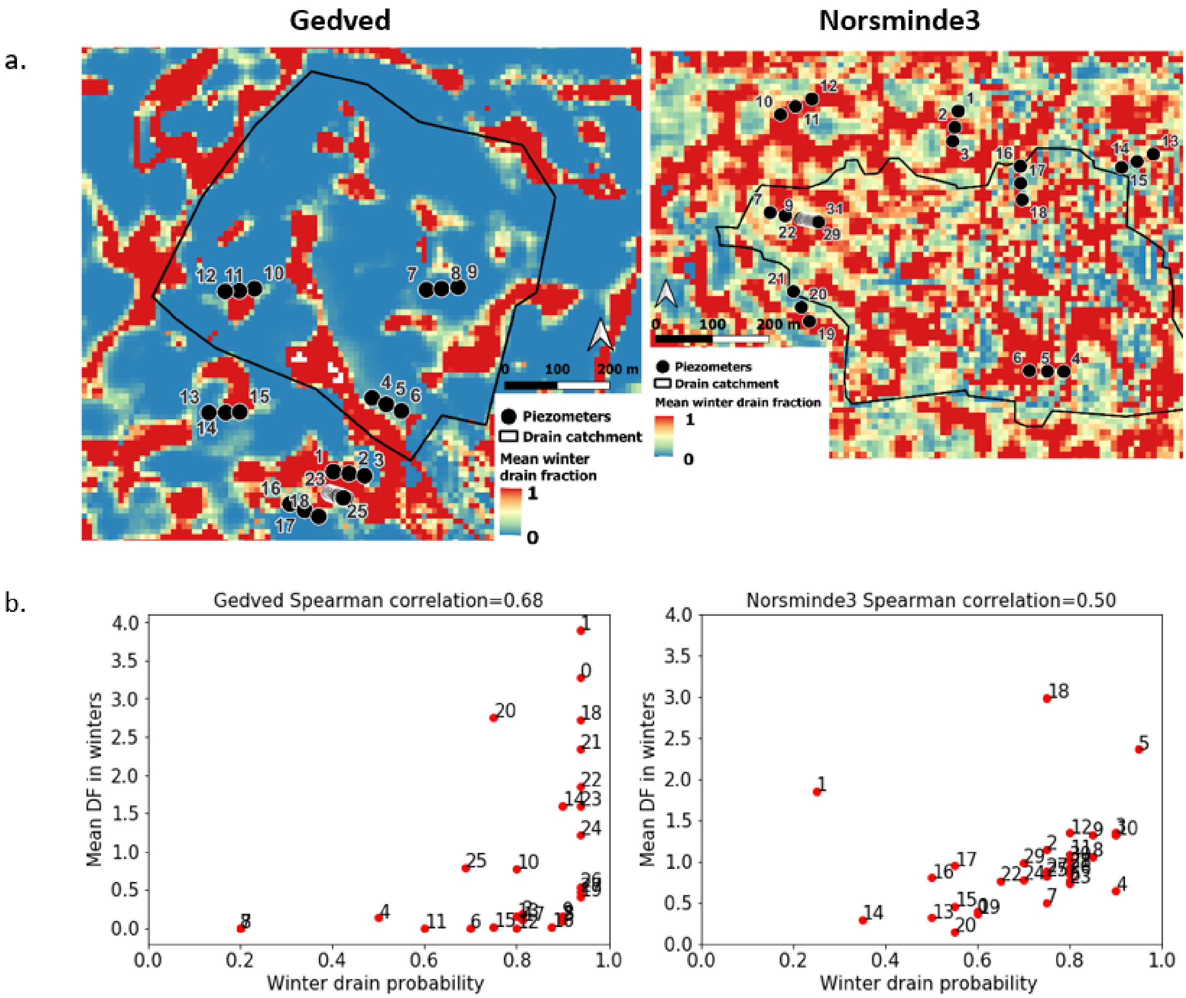

Figure 8 Point comparison of drain probability based on observed groundwater levels and simulated DF. A. Distribution of piezometers; B. Spearman correlation plot between winter drain probability and mean winter DF

3.3 Geological and topographical correlations

Before the correlation analysis, a covariance matrix is used to exclude the redundant variables that are highly correlated with a coefficient of determination above 0.9. After the covariance analysis, the redundant variables such as TPI in a radius of 10m, Curvature, Clay fraction b, and c horizon, and Roughness were excluded (*Figure 9*).

3.3.1 Covariance and initial correlations

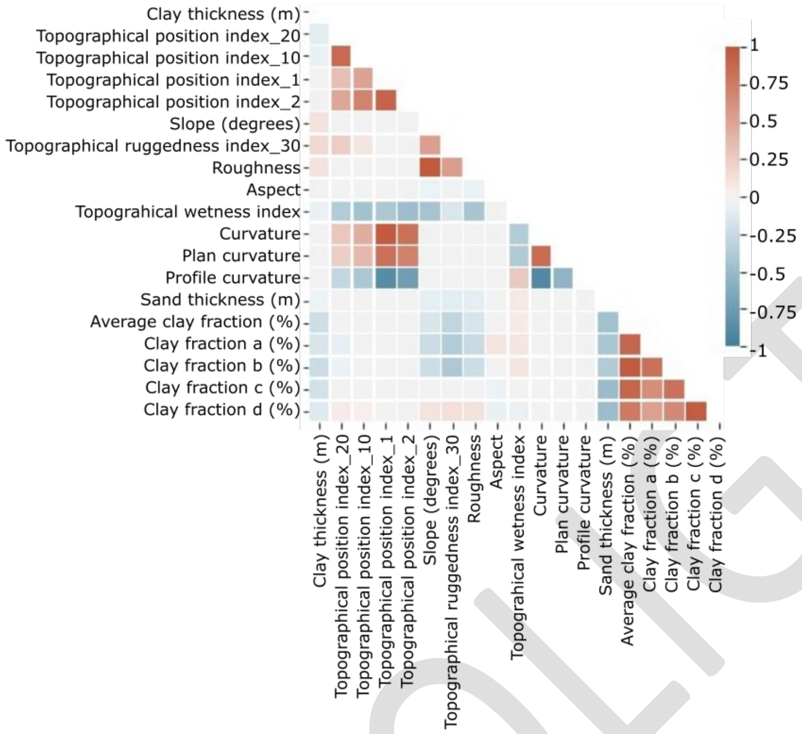

Figure 9 Covariance matrix of Physical variables

Figure 10 shows the correlation matrix between model simulated DF and non-redundant physical variables. The results showed that the Lolland region has no significant correlation with the physical variables. Lillebæk, Norsminde, and Gyldenholm show an intermediate correlation with identified physical variables. Lillebæk shows a medium correlation of 0.67, 0.5, and -0.62 with average clay fraction (%), clay thickness (m), and TRI_30, respectively. However, Gyldenholm and Norsminde display correlation only with topographical variables. Gyldenholm shows a correlation of 0.57 for TPI_20. Norsminde shows a correlation of -0.69 and 0.6 with TPI_2 and TWI.

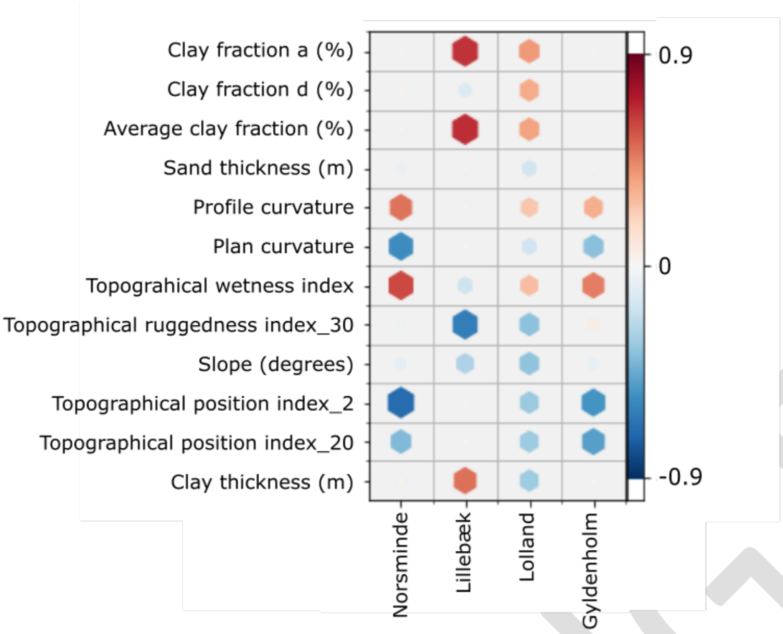

Figure 10 Correlation matrix between simulated DF and identified unique physical variables

3.4 Scale analysis across different spatial aggregations.

The correlation of four main geological and topographical variables with model simulated DF is shown in *Figure 11*. On the field scale, topographical variables display clear dominance in the correlation analysis except for Lillebæk1,2,4 and Lolland1 drain sites, where geological influence was more prominent. Lillebæk3 and Lolland2,4 show a weak correlation with TWI and average clay fraction %. On the catchment scale, Lillebæk illustrates DF correlation with clay fraction and clay thickness while Norsminde and Gyldenholm showed sensitivity to TPI and TWI; however, Lolland showed no clear trend. On the regional scale, Fyn's DF correlates with clay fraction and clay thickness, while Sjælland and Jylland showed a high correlation with TPI and TWI. A decrease in correlation was observed from the field scale to the national scale.

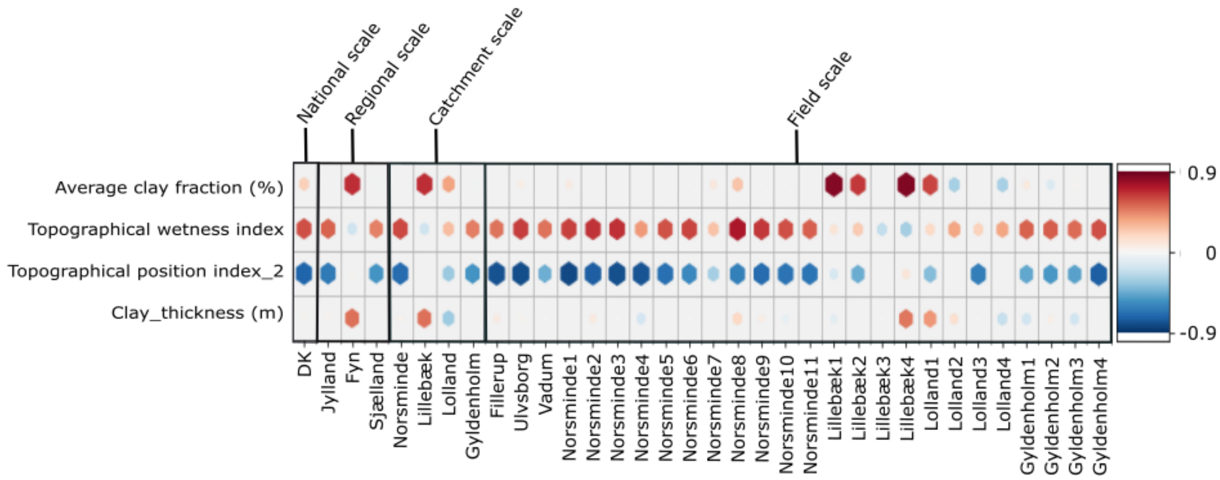

Figure 11 Scale-based correlation between model simulated DF and physical variables

4 Discussion

4.1 10m resolution groundwater models performance, equifinality, and transferability

The 10m resolution groundwater models simulate the drain flow dynamics and overall drain flow volume reasonably well, as indicated by KGE and PBIAS, considering the combined calibration of 26 drain sites. Many previous studies used models to simulate drain flow dynamics but mainly on a single catchment(Boico et al., 2022; De Schepper et al., 2017; Frederiksen & Navarro, 2021; Hansen, Jakobsen, et al., 2019; Hansen, Storgaard, et al., 2019; Salo et al., 2017; Wang et al., 2017). Only Jeantet et al. (2021) and (Motarjemi et al., 2021) modeled multiple catchments together but using lumped and machine learning models, respectively. The model performance from these previously studied lumped models, machine learning models or single catchment models is not significantly higher than our 10m resolution groundwater models.

Besides matching the overarching seasonal patterns of drain flow, with high amounts in winter related to high groundwater levels and low amounts in summer related to lower groundwater levels, our models can also capture much of the short-term dynamics. The simulated drain flows do not always capture the magnitude of the peaks, but the overall underestimation was less than 10%. In most cases, the first peak of the winter season is not simulated accurately. This could be linked to the groundwater model's inability to mimic the desiccation of clayey soils (Tang et al., 2011). In desiccation, clayey soils develop cracks in the topsoil surface that allow water to pass rapidly through the cracks creating a fast response in the groundwater table rise (and, consequently, the drain flow) until the clay saturates and swells. Another major source of uncertainty in simulated drain flow arises from climate forcing since precipitation is obtained from the national rain gauge network, potentially struggling with field scale variations in precipitation. However, PBIAS was high >±25% for many individual catchments. This high magnitude of PBIAS can be due to the inaccuracy of the delineated drain site and its total area. Drain site delineations are provided by the projects which performed the drain flow monitoring and are based on a combination of knowledge about the actual tile drainage network and topographical delineations. For small drain sites and complex tile drainage networks, the uncertainty in the estimated area contributing to drain flow at the observation point could be significant. Moreover, measurement uncertainty in drain flow observations affects model performance as well.

Beven (1993) proposed the equifinality concept for hydrological modeling that states more than one solution (such as one parameter set) can have practically equal good model performance. Amongst others, Asadzadeh and Tolson (2013) and Anderson et al. (2015) also highlighted the non-uniqueness of groundwater models, that there could be more than one reasonable model with different combinations of parameter sets. We used multiple equally good calibration solutions to cover some model parameter uncertainty effects on generated drain flow. In this study, we incorporated the parameter uncertainty in simulated spatial DF by a limited selection of five solutions. Only five solutions are selected to reduce the computational time. Whether or not five solutions are enough to represent the uncertainty could be questionable. However, in the current study, we observed no difference in the correlation results when the spatial DF of 5 selected solutions is separately correlated with the physical control variables. Therefore, it is assumed sufficient to select five solutions.

Because of space and time-specific calibration, many groundwater models only apply to specific sites. Such site-specific calibrated models have low transferability as they cannot replicate the dynamics for other regions or climates (Montanari et al., 2013). The 10m resolution groundwater models are valid and transferable as spatially consistent parametrization combined with one joint calibration for all 26 drain sites. The natural variability in year-to-year climate is also covered by taking, on average, two years long drain flow data for calibration of most drain sites except Norsminde 9,10,11. Moreover, the time period of calibration is different for each drain site representing variation in precipitation in the model.

4.2 Applicability of correlation analysis

The pertinency of the correlation analysis and, thereby, the ability to generalize findings depends on the representation of the topographical and geological variability of Denmark's drained area by our 26 drain sites. Moller et al. (2018) developed a map for Denmark displaying the probability of an artificial drainage system. We used that probability map to limit the following analysis to areas across all of Denmark that likely are drained, using ≥33% probability as a cutoff for the existing drain. We evaluated whether the physical variables' ranges occurring across likely drained areas in all of Denmark are reasonably well covered by our 26 drain sites. Figure 13 shows the occurring ranges of values for the four most important variables for both the 26 drain sites and the likely drained areas across all of Denmark. The variability in all topographical and geological variables is covered in the 26 drain sites, while clay fraction below 10% is underrepresented. The regions with low clay fractions are not covered because we had limited drain flow data, but also, areas with clay fractions below 10% are less likely to be drained artificially. Therefore, our findings do not apply to regions with low clay fractions below 10%.

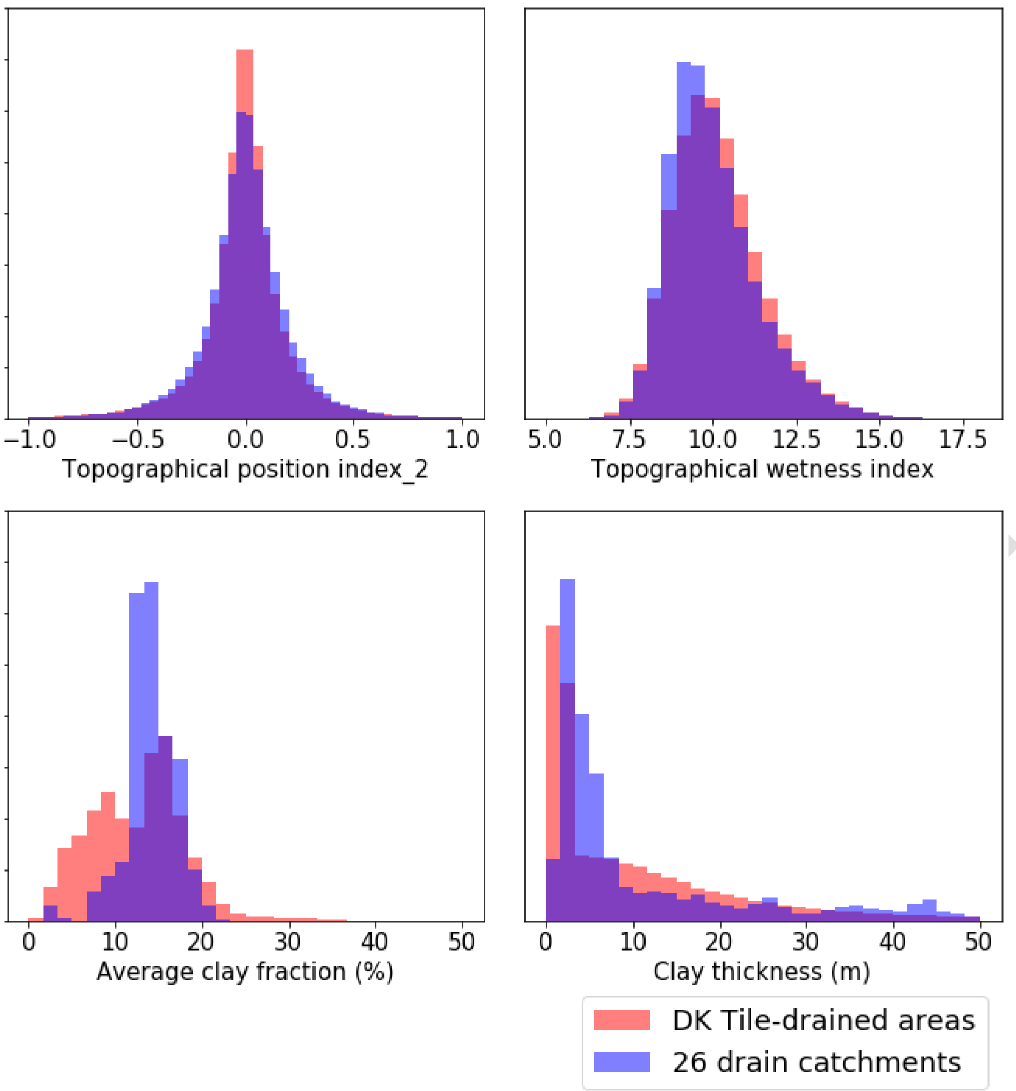

Figure 12 Histogram of distribution of geological and topographical variables. The purple area is overlap between the blue and pink region

4.3 Spatial DF

Due to a lack of piezometer head data, we have not calibrated the 10m resolution groundwater flow models against piezometer head observations. However, verifying simulated DF spatial patterns was crucial for this study. So, we validated the spatial patterns of simulated DF by inspecting the correlation between simulated DF spatial patterns and observed drain probability in winters for two drain sites with available groundwater level observations. Both sites showed an intermediate Spearman correlation of 0.5 (Norsminde3) and 0.68 (Gedved), assuring that simulated DF spatial patterns are reasonable and can be trusted.

Representation of tile drain site on 100m resolution does not provide an accurate picture of spatial variation within field scale. This study provides a more detailed spatial variation of DF within the field scale, allowing us to understand the behavior of drain flow with respect to field geology and topography. This lacked in the previously existing 100m DK-HIP model. Hydrological understanding of the spatial distribution of drain flow on the field scale is critical for water quality regulation in the agricultural sector. This study benefits the water managers in identifying areas of high drain flow and locates regions that feed agricultural water excess to surface water bodies.

4.4 Correlation Analysis

We found that among all topographical and geological variables, mostly relative differences in topography control the DF spatial distribution on the national scale. However, the correlations tend to get smaller with more significant aggregation levels. We do not observe a strong correlation with geological variables, and this might be because topographical variables are initially in 10m resolution, but geological variables are downscaled from 100m or 30m resolution to 10m resolution. When Denmark is aggregated into three parts based on similarities in geology, Fyn, Sjælland-Lolland, and Jylland, we observed a distinct behavior in Fyn. The drain flow pattern in Fyn is controlled mainly by clay fraction, whereas topographical variables do not influence the drain flow pattern. This distinct behavior in Fyn is because all available Fyn drain sites have relatively low differences in topography and a high clay fraction value (Table 2). Low hydraulic conductivity limits lateral groundwater flows, so the topographical influence is limited as water flows from peaks to depressions are prevented. We developed the hydraulic conductivity map from the clay fraction pedo-transfer function, and higher clay fractions are converted to lower hydraulic conductivity. Therefore, clay fraction is a limiting factor in Fyn's drain sites; thus, topographical variables do not show a significant correlation.

5 Conclusion

The study developed a groundwater flow model in 10m resolution that can well reproduce drain flow dynamics and spatial differences in drain flow fraction at the field scale after joint calibration of 26 drain sites. The achieved average KGE above 0.5 and |PBIAS| below 10% of drain flow affirm the robustness and accuracy of the groundwater model.

The study also demonstrates how spatial drain flow patterns correlate with physical variables of topography and geology to improve the understanding of drivers of the spatial distribution of drain flow. We found the TPI to be the most important physical covariate in regions where relative differences in topography exist. We also found clay fraction becomes a dominant factor when clay fraction percentage increase in relatively flat areas.

Even though the 10m resolution groundwater models can accurately produce drain flow patterns, developing a hydrological model-based DF map of Denmark in 10m resolution is impossible due to computational limitations. Nonetheless, the developed 10m resolution groundwater models for the selected drain sites can generate a training dataset for a data-driven algorithm study which could be explored in a future study.

6 Credit authorship contribution statement

Hafsa Mahmood: Conceptualization, Methodology, Software, Formal analysis, Investigation, Writing - original draft, Visualization, Validation, Writing - review & editing. Raphael Schneider: Supervision Conceptualization, Methodology, Investigation, Software, Formal analysis, Writing - review & editing. Simon Stisen: Supervision, Conceptualization, Methodology, Writing - review & editing. Rasmus Rumph Frederiksen: Supervision, Writing - review & editing. Anders Vest Christiansen: Supervision, Writing - review & editing.

7 Declaration of Competing Interest

None.

8 Acknowledgments

This study was part of T-Rex (Terrænnær redox og retentions-kortlægning til differentieret målrettet virkemiddelindsats indenfor ID15 oplande), funded by the Danish GUDP (Grønt Udviklings- og Demonstrationsprogram), project number 34009-18-1453 k and WATEC – Aarhus University Center for water technology in Denmark. The authors thank Bo V. Iversen for providing drain flow data and Carlos Duque Calvache for productive discussions. The authors would also like to thank Lars Troldborg for providing the clay and sand thickness maps.

9 Annexes

Drain catchment scale	TPI_2		TPI_20		Average clay fraction (%)		TWI		Clay thickness (m)		Slope		Plan curvature		Profile curvature	
	Mean	Std	Mean	Std	Mean	Std	Mean	Std	Mean	Std	Mean	Std	Mean	Std	Mean	Std
Fillerup	0.000	0.2	0.10	1.2	14.4	0.8	10.1	1.3	5.0	5.4	2.2	1.1	-0.002	0.2	-0.003	0.2
Ulvsborg	0.001	0.3	0.01	1.0	14.7	0.7	9.4	1.1	4.6	5.8	2.0	1.1	-0.006	0.3	-0.006	0.4
Vadum	-0.001	0.1	-0.03	0.2	3.6	2.2	11.1	1.1	0.2	0.5	0.3	0.2	0.003	0.1	0.003	0.1
Norsminde1	0.005	0.2	0.18	1.0	12.2	0.7	10.1	1.2	36.9	59.0	1.7	0.8	0.000	0.2	-0.004	0.2
Norsminde2	-0.002	0.3	-0.08	1.9	12.7	0.8	9.4	1.1	18.7	17.0	2.8	1.6	-0.003	0.3	0.001	0.3
Norsminde3	0.003	0.2	-0.01	0.9	12.2	0.6	9.8	1.2	329.4	231.5	1.7	1.0	-0.003	0.3	-0.006	0.3
Norsminde4	0.013	0.1	0.62	0.8	13.3	0.7	11.0	1.2	3.4	1.3	2.0	0.9	-0.006	0.2	-0.024	0.2
Norsminde5	-0.014	0.3	-0.04	2.0	13.3	0.7	9.5	1.1	38.2	12.0	3.1	1.8	-0.019	0.4	-0.002	0.5
Norsminde6	0.007	0.3	1.12	2.4	12.7	0.7	9.6	1.4	33.8	10.4	4.1	2.9	0.006	0.3	0.000	0.3
Norsminde7	0.090	0.2	2.60	1.8	12.7	0.5	9.5	0.7	12.2	14.3	3.3	2.7	0.023	0.1	-0.062	0.2
Norsminde8	0.019	0.5	1.55	2.4	12.4	0.8	9.0	1.3	1.6	0.7	3.9	2.1	0.036	0.5	0.012	0.6
Norsminde9	-0.009	0.3	0.02	2.2	12.4	0.8	9.6	1.3	389.0	196.4	3.6	2.6	-0.021	0.3	-0.004	0.4
Norsminde10	0.014	0.2	1.13	1.7	13.4	0.7	9.9	1.0	3.4	2.3	3.0	2.3	0.004	0.2	-0.008	0.2
Norsminde11	0.030	0.2	1.88	1.3	12.8	0.8	9.8	1.1	17.4	11.4	2.4	1.5	0.014	0.2	-0.017	0.2
Lillebaek1	0.020	0.1	0.54	0.6	18.9	0.5	10.0	0.9	14.2	5.9	1.1	0.6	0.008	0.2	-0.016	0.2
Lillebaek2	0.029	0.2	1.06	0.7	17.6	0.5	9.4	1.0	17.5	0.4	1.4	0.7	-0.023	0.4	-0.045	0.5
Lillebaek3	0.027	0.1	0.60	0.5	18.0	0.4	10.7	1.1	9.9	2.3	1.5	0.9	0.023	0.2	-0.014	0.2
Lillebaek4	-0.016	0.2	0.41	0.9	18.1	0.5	10.2	1.3	4.6	1.3	2.2	0.67	0.004	0.2	0.024	0.3
Lolland1	0.027	0.1	0.81	0.4	16.4	0.6	10.6	1.0	25.8	2.6	0.9	0.7	0.003	0.1	-0.027	0.1
Lolland2	-0.003	0.1	0.29	0.3	16.5	0.5	10.5	0.9	10.8	2.0	1.0	0.3	-0.008	0.1	0.006	0.1
Lolland3	0.066	0.1	2.03	1.6	15.7	0.3	9.4	0.7	17.9	1.3	2.6	1.0	0.023	0.2	-0.053	0.3
Lolland4	0.005	0.1	0.19	0.4	21.0	0.7	10.6	1.0	6.9	0.8	0.7	0.4	-0.005	0.2	-0.013	0.2
Gyldenholm1	-0.002	0.2	0.04	0.7	14.3	1.5	10.4	1.4	5.1	5.6	1.2	0.7	-0.004	0.2	0.004	0.2
Gyldenholm2	0.003	0.2	0.03	0.7	14.3	1.6	10.1	1.2	9.8	12.8	1.2	0.8	-0.003	0.2	-0.007	0.3
Gyldenholm3	-0.002	0.1	-0.03	0.5	16.9	1.5	10.5	1.3	7.1	6.0	0.9	0.7	0.001	0.2	0.005	0.2
Gyldenholm4	-0.007	0.2	-0.08	1.1	14.5	1.8	9.8	1.2	33.6	19.8	1.7	1.0	-0.009	0.2	-0.001	0.2

Table 2 Properties of drain sites

Table 3 Drain flow fraction statistics for different solutions

	1	2	3	4	5
mean	0.84	0.83	0.79	0.83	0.85
std	0.78	0.80	0.79	0.79	0.76
25%	0.46	0.43	0.42	0.44	0.46
50%	0.76	0.66	0.64	0.66	0.79
75%	0.94	0.99	0.93	0.98	0.97

Table 4 Drain flow fraction statistics for National, Regional, Catchment scale

	DK	Fyn	Jylland	Sjælland	Lillebæk	Lolland	Gyldenholm	Fensholt	Station	Fillerup	Ulvsborg	Vadum
Count	51924	1064	24481	26379	1064	1512	24867	12721	3563	3821	3487	889
Mean	3.51	1.59	3.54	3.56	1.59	2.92	3.60	3.56	3.30	3.73	3.47	3.82
Std	3.25	0.74	4.13	2.19	0.74	0.96	2.24	4.32	3.98	2.13	2.95	9.22
25%	1.92	1.02	1.25	2.60	1.02	2.58	2.60	0.98	1.14	2.43	1.70	0.03
50%	2.98	1.54	2.54	3.21	1.54	3.20	3.21	2.24	2.28	3.34	2.75	0.42
75%	4.08	2.25	4.26	4.06	2.25	3.48	4.12	4.36	3.87	4.49	4.05	3.20

Table 5 Drain flow fraction statistics for drain sites

	Vadum	Ulvsborg	FensholtD1	FensholtD2	FensholtD3	FensholtD4	FensholtD5	FensholtD6	FensholtD7	FensholtD8	Station31south	Station32in	Station33in
count	889.0	3487.0	3391.0	3279.0	2753.0	406.0	1193.0	724.0	353.0	622.0	1400.0	765.0	1398.0
mean	**3.8**	3.5	3.6	3.7	3.8	3.5	3.7	3.4	1.7	2.5	3.7	3.1	3.1
std	**9.2**	2.9	3.9	4.8	4.6	2.1	4.6	4.2	1.3	3.9	5.2	2.8	3.0
25%	**0.0**	1.7	1.1	0.7	1.0	2.1	1.6	0.9	0.8	0.2	0.6	1.5	1.4
50%	**0.4**	2.8	2.4	2.1	2.3	3.0	2.6	2.1	1.4	0.9	2.0	2.4	2.3
75%	**3.2**	4.1	4.4	4.5	4.8	4.3	4.3	4.5	2.4	3.0	4.4	3.7	3.6
	Fillerup	**Gyldenholm1**	**Gyldenholm2**	**Gyldenholm3**	**Gyldenholm4**	**LillebækD2**	**LillebækD4**	**LillebækD5**	**LillebækD6**	**LollandD103**	**LollandD105**	**LollandD106**	**LollandD107**
count	3821.0	4641.0	4870.0	12006.0	3350.0	446.0	100.0	256.0	262.0	578.0	264.0	204.0	466.0
mean	3.7	3.5	3.7	3.5	3.8	2.3	1.5	1.1	0.9	2.6	3.3	2.5	3.4
std	2.1	2.7	2.8	1.0	3.4	0.2	0.3	0.6	0.3	1.2	0.7	0.4	0.6
25%	2.4	2.2	2.3	2.9	1.9	2.2	1.3	0.6	0.6	1.9	3.0	2.4	3.4
50%	3.3	2.9	3.0	3.4	2.8	2.3	1.5	1.2	0.9	2.9	3.2	2.5	3.5
75%	4.5	4.2	4.4	4.1	4.5	2.5	1.7	1.4	1.1	3.4	3.5	2.7	3.6

10 References

Abbott, M. B., Bathurst, J. C., Cunge, J. A., Connell, P. E. O., & Rasmussen, J. (1986). An introduction to the European Hydrological System — Systeme Hydrologique Europeen, “SHE”, 1: History and philosophy of a physically-based, distributed modelling system. *Journal of Hydrology, 87*(1-2), 16. https://doi.org/https://doi.org/10.1016/0022-1694(86)90114-9

Adhikari, K., Kheir, R. B., Greve, M. B., Bocher, P. K., Malone, B. P., Minasny, B., McBratney, A. B., & Greve, M. H. (2013, May). High-Resolution 3-D Mapping of Soil Texture in Denmark. *Soil Science Society of America Journal, 77*(3), 860-876. https://doi.org/10.2136/sssaj2012.0275

Amado, A. A., Schilling, K. E., Jones, C. S., Thomas, N., & Weber, L. J. (2017, Sep). Estimation of tile drainage contribution to streamflow and nutrient loads at the watershed scale based on continuously monitored data. *Environmental Monitoring and Assessment, 189*(9). https://doi.org/ARTN 426
10.1007/s10661-017-6139-4

Anderson, M. P., Woessner, W. W., & Hunt, R. J. (2015). *Applied Groundwater Modeling: Simulation of Flow and Advective Transport*
Elsevier. https://doi.org/https://doi.org/10.1016/C2009-0-21563-7

Asadzadeh, M., & Tolson, B. (2013, Dec 1). Pareto archived dynamically dimensioned search with hypervolume-based selection for multi-objective optimization. *Engineering Optimization, 45*(12), 1489-1509. https://doi.org/10.1080/0305215x.2012.748046

Beven, K. (1993). Prophecy, reality and uncertainty in distributed hydrological modelling. *Advances in Water Resources, 16*(1), 41-51. https://doi.org/https://doi.org/10.1016/0309-1708(93)90028-E

Beven, K. J., & Kirkby, M. J. (1979). A physically based, variable contributing area model of basin hydrology. *Hydrolological Sciences Bulletin, 24*.

Boico, V. F., Therrien, R., Højberg, A. L., Iversen, B. I., Koganti, T., & Varvaris, T. (2022). Using depth specific electrical conductivity estimates to improve hydrological simulations in a heterogeneous tile-drained field. *Journal of Hydrology, 604*. https://doi.org/https://doi.org/10.1016/j.jhydrol.2021.127232

De Schepper, G., Therrien, R., Refsgaard, J. C., He, X., Kjaergaard, C., & Iversen, B. V. (2017, May). Simulating seasonal variations of tile drainage discharge in an agricultural catchment. *Water Resources Research, 53*(5), 3896-3920. <Go to ISI>://WOS:000403712100024

DHI. (2020). MIKE SHE - User Guide and Reference Manual. https://manuals.mikepoweredbydhi.help/2020/Water_Resources/MIKE_SHE_Print.pdf

EPA, D. (2020). *FOHM—Fælles Offentlig Hydrologisk Model* https://mst.dk/natur-vand/vand-i-hverdagen/grundvand/grundvandskortlaegning/kortlaegning-2016-2020/fohm-faelles-offentlig-hydrologisk-model/

Frederiksen, R. R., & Navarro, E. M. (2021). The importance of subsurface drainage on model performance and water balance in an agricultural catchment using SWAT and SWAT-MODFLOW. *Agricultural Water Management, 255*(107058). https://doi.org/https://doi.org/10.1016/j.agwat.2021.107058

Gallant, J. C., & Wilson, J. P. (2000). Primary topographic attributes. *Terrain Analysis: Principles and Applications*, 51-85. (Wiley, New York)

Hansen, A. L., Jakobsen, R., Refsgaard, J. C., Hojberg, A. L., Iversen, B. V., & Kjaergaard, C. (2019, Jan). Groundwater dynamics and effect of tile drainage on water flow across the redox interface in a Danish Weichsel till area. *Advances in Water Resources, 123*, 23-39. <Go to ISI>://WOS:000453714000003

Hansen, A. L., Storgaard, A., He, X., Hojberg, A. L., Refsgaard, J. C., Iversen, B. V., & Kjaergaard, C. (2019, Jan 30). Importance of geological information for assessing drain flow in a Danish till landscape. *Hydrological Processes, 33*(3), 450-462. <Go to ISI>://WOS:000456206300010

Henriksen, H. J., Kragh, S. J., Gotfredsen, J., Ondracek, M., van Til, M., Jakobsen, A., Schneider, R. J. M., Koch, J., Troldborg, L., Rasmussen, P., Pasten-Zapata, E., & Stisen, S. (2020). *Dokumentationsrapport vedr. modelleverancer til Hydrologisk Informations- og Prognosesystem*.

Henriksen, H. J., Voutchkova, D., Troldborg, L., Ondracek, M., Schullehner, J., & Hansen, B. (2019). *National Vandressource Model Beregning af udnyttelsesgrader, afsænkning og vandløbspåvirkning med DK model 2019*.

Hojberg, A. L., Hansen, A. L., Wachniew, P., Zurek, A. J., Virtanen, S., Arustiene, J., Stromqvist, J., Rankinen, K., & Refsgaard, J. C. (2017, Aug). Review and assessment of nitrate reduction in groundwater in the Baltic Sea Basin. *Journal of Hydrology-Regional Studies, 12*, 50-68. https://doi.org/10.1016/j.ejrh.2017.04.001

Horn, B. K. P. (1981). Hill shading and the reflectance map. *Proceedings of the IEEE, 69*, 14–47. https://doi.org/doi:10.1109/PROC.1981.11918

Højberg, A. L., Stisen, S., Olsen, M., Troldborg, L., Uglebjerg, T. B., & Jørgensen, L. F. (2015). *DK-model2014. Model opdatering og kalibrering. (Danmarks og Grønlands Geologiske Undersøgelse Rapport*.

Jacobsen, B. H., & Hansen, A. L. (2016, 2016). Economic gains from targeted measures related to non-point pollution in agriculture based on detailed nitrate reduction maps. *Science of the Total Environment, 556*, 265-275. https://doi.org/https://doi.org/10.1016/j.scitotenv.2016.01.103.

Jakobsen, P. R., Hermansen, B., & Tougaard, L. (2015). *Danmarks digitale jordartskort 1:25000 - Version 4.0,* .

Jeantet, A., Henine, H., Chaumont, C., Collet, L., Thirel, G., & Tournebize, J. (2021, Oct 14). Robustness of a parsimonious subsurface drainage model at the French national scale. *Hydrology and Earth System Sciences, 25*(10), 5447-5471. https://doi.org/10.5194/hess-25-5447-2021

Koch, J., Gotfredsen, J., Schneider, R., Troldborg, L., Stisen, S., & Henriksen, H. J. (2021, Sep 1). High Resolution Water Table Modeling of the Shallow Groundwater Using a Knowledge-Guided Gradient Boosting Decision Tree Model. *Frontiers in Water, 3*. https://doi.org/ARTN 701726 10.3389/frwa.2021.701726

Los Huertos, M., & Smith, D. (2013). *Wetland bathymetry and mapping* (Vol. 1). In Anderson JT.

Matott, L. (2017). *OSTRICH: An Optimization Software Tool, Documentation and User's Guide Version 17.12.19.* In University at Buffalo Center for Computational Research. http://www.civil.uwaterloo.ca/envmodelling/Ostrich.html

Mattivi, P., Franci, F., Lambertini, A., & Bitelli, G. (2019). TWI computation: a comparison of different open source GISs. *Open geospatial data, softw. stand, 6*(4).

Moller, A. B., Beucher, A., Iversen, B. V., & Greve, M. H. (2018, Jun 15). Predicting artificially drained areas by means of a selective model ensemble. *Geoderma, 320*, 30-42. https://doi.org/10.1016/j.geoderma.2018.01.018

Montanari, A., Young, G., Savenije, H. H. G., Hughes, D., Wagener, T., Ren, L. L., Koutsoyiannis, D., Cudennec, C., Toth, E., Grimaldi, S., Bloschl, G., Sivapalan, M., Beven, K., Gupta, H., Hipsey, M., Schaefli, B., Arheimer, B., Boegh, E., Schymanski, S., Di Baldassarre, G., Yu, B., Hubert, P., Huang, Y., Schumann, A., Post, D. A., Srinivasan, V., Harman, C., Thompson, S., Rogger, M., Viglione, A., McMillan, H., Characklis, G., Pang, Z., & Belyaev, V. (2013, Aug 1). "Panta Rhei-Everything Flows": Change in hydrology and society-The IAHS Scientific Decade 2013-2022. *Hydrological Sciences Journal-Journal Des Sciences Hydrologiques, 58*(6), 1256-1275. <Go to ISI>://WOS:000323242200003

Motarjemi, S. K., Moller, A. B., Plauborg, F., & Iversen, B. V. (2021, Aug). Predicting national-scale tile drainage discharge in Denmark using machine learning algorithms. *Journal of Hydrology-Regional Studies, 36*. https://doi.org/10.1016/j.ejrh.2021.100839

Prinds, C., Petersen, R. J., Greve, M. H., & Iversen, B. V. (2019, Jan 2). Locating tile drainage outlets and surface flow in riparian lowlands using thermal infrared and RGB-NIR remote sensing. *Geografisk Tidsskrift-Danish Journal of Geography, 119*(1), 94-105. https://doi.org/10.1080/00167223.2019.1573408

Riley, S. J., DeGloria, S. D., & Elliot, R. (1999). A terrain ruggedness index that quantifies topographic heterogeneity. *Intermountain Journal of Sciences, 5*(1-4), 23-27.

Salo, H., Warsta, L., Turunen, M., Nurminen, J., Myllys, M., Paasonen-Kivekas, M., Alakukku, L., & Koivusalo, H. (2017, May). Simulating 3-D water flow in subsurface drain trenches and surrounding soils in a clayey field. *Soil & Tillage Research, 168*, 20-32. <Go to ISI>://WOS:000394079800003

Scharling, M. (1999a). *Klimagrid Danmark - Nedbør, lufttemperatur og potentiel fordampning 20X20 & 40x40 km - Metodebeskrivelse*.

Scharling, M. (1999b). *Klimagrid Danmark Nedbør 10x10 km (ver. 2) - Metodebeskrivelse*.

Stenberg, M., Ulen, B., Soderstrom, M., Roland, B., Delin, K., & Helander, C. A. (2012, Sep 15). Tile drain losses of nitrogen and phosphorus from fields under integrated and organic crop rotations. A four-year study on a clay soil in southwest Sweden. *Science of the Total Environment, 434*, 79-89. https://doi.org/10.1016/j.scitotenv.2011.12.039

Stisen, S., Ondracek, M., Troldborg, L., Schneider, R. J. M., & van Til, M. J. (2019). *National Vandressource Model - Modelopstilling og kalibrering af DK-model 2019*.

Tang, C. S., Shi, B., Liu, C., Suo, W. B., & Gao, L. (2011, Apr). Experimental characterization of shrinkage and desiccation cracking in thin clay layer. *Applied Clay Science, 52*(1-2), 69-77. <Go to ISI>://WOS:000289610000010

Wang, F. Q., Franco-Penya, H., & Kelleher, J. D. (2017). An Analysis of the Application of Simplified Silhouette to the Evaluation of k-means Clustering Validity. *Conference: 13th International Conference on Machine Learning and Data Mining MLDM*. https://doi.org/DOI:10.1007/978-3-319-62416-7_21

Williams, M. R., King, K. W., & Fausey, N. R. (2015, Aug). Contribution of tile drains to basin discharge and nitrogen export in a headwater agricultural watershed. *Agricultural Water Management, 158*, 42-50. https://doi.org/10.1016/j.agwat.2015.04.009

Wilson, M. F. J., O'Connell, B., Brown, C., Guinan, J. C., & Grehan, A. J. (2007). Multiscale terrain analysis of multibeam bathymetry data for habitat mapping on the continental slope. *Marine Geodesy, 30*(1-2), 3-35. <Go to ISI>://WOS:000251571400002

Zevenbergen, L. W., & Thorne, C. R. (1987). Quantitative analysis of land surface topography. *Earth surface processes and landforms, 12*, 47–56.