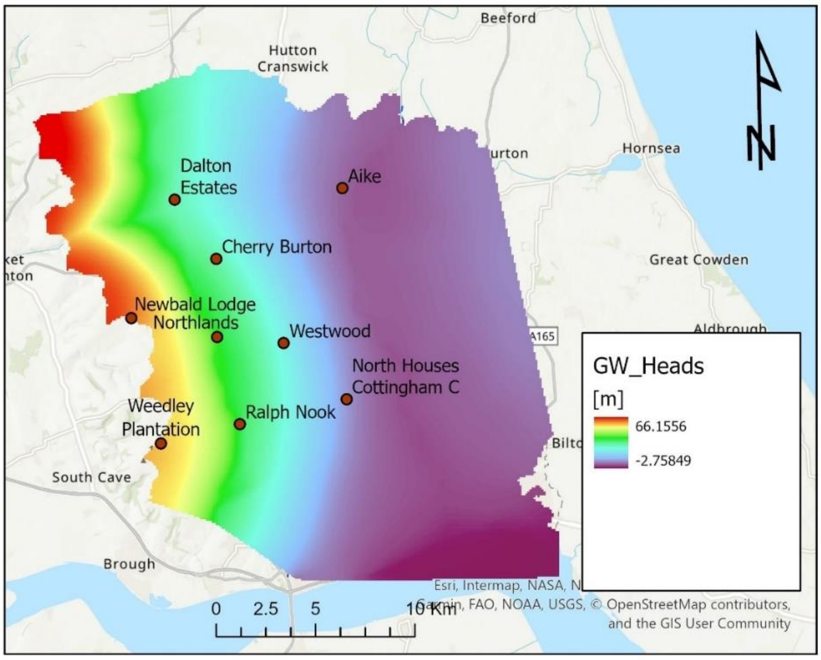
CAESAR-GW a modified version of CAESAR-LF with improved groundwater model: An application to the Hull digital twin FLOODTWIN

Environmental change, adaptation and resilience programme

Internal report OR/25/023



If

BRITISH GEOLOGICAL SURVEY

ENVIRONMENTAL CHANGE, ADAPTATION AND RESILIENCE PROGRAMME

INTERNAL REPORT OR/25/023

CAESAR-GW a modified version of CAESAR-LF with improved groundwater model: An application to the Hull digital twin FLOODTWIN

M Mansour, J Scheidegger, A Barkwith, T Coulthard

Keywords

Modelling; Flooding; Digital twin; CEASAR, CLiDe platform.

Front cover

Groundwater head contours at Hull and East riding area.

Bibliographical reference

Mansour, M, Scheidegger, J, Barkwith, A, Coulthard. 2025. Addition of a hydrogeological model component to CAESAR-LF: An application to the Hull digital twin FLOODTWIN. *British Geological Survey Internal Report*, OR/25/023. 27pp.

Keyworth, Nottingham British Geological Survey 2025

BRITISH GEOLOGICAL SURVEY

The full range of our publications is available from the BGS shop at Nottingham and Cardiff (Welsh publications only). Shop online at https://shop.bgs.ac.uk/

The London Information Office also maintains a reference collection of BGS publications, including fossils, for consultation.

We publish an annual catalogue of our maps and other publications; this catalogue is available online or from the BGS shop.

The British Geological Survey carries out the geological survey of Great Britain and Northern Ireland (the latter as an agency service for the government of Northern Ireland), and of the surrounding continental shelf, as well as basic research projects. It also undertakes programmes of technical aid in geology in developing countries.

The British Geological Survey is a component body of UK Research and Innovation.

British Geological Survey offices

Nicker Hill, Keyworth,
Nottingham NG12 5GG
Tel 0115 936 3100

BGS Central Enquiries Desk
Tel 0115 936 3143
email enquiries@bgs.ac.uk

BGS Sales
Tel 0115 936 3241
email sales@bgs.ac.uk

The Lyell Centre, Research Avenue South,
Edinburgh EH14 4AP
Tel 0131 667 1000
email scotsales@bgs.ac.uk

Natural History Museum, Cromwell Road,
London SW7 5BD
Tel 020 7589 4090
Tel 020 7942 5344/45
email bgslondonstaff@bgs.ac.uk

Cardiff University, Main Building, Park Place,
Cardiff CF10 3AT
Tel 029 2167 4280

Maclean Building, Crowmarsh Gifford,
Wallingford OX10 8BB
Tel 01491 838800

Geological Survey of Northern Ireland, 7th Floor,
Adelaide House, 39-49 Adelaide Street, Belfast, BT2 8FD
Tel 0289 038 8462
www2.bgs.ac.uk/gsni/

Natural Environment Research Council, Polaris House,
North Star Avenue, Swindon SN2 1EU
Tel 01793 411500 Fax 01793 411501
www.nerc.ac.uk

UK Research and Innovation, Polaris House,
Swindon SN2 1FL
Tel 01793 444000
www.ukri.org

Website: https://www.bgs.ac.uk
Shop online: https://shop.bgs.ac.uk/

Foreword

BGS has previously modified the CAESAR-LF landscape evolution platform to include an improved approach to water partitioning and a simple representation of groundwater (Barkwith and Coulthard, 2015). This new platform, CLiDE, performed well in homogeneous geological settings, but was ineffective under more complex geology. CLiDE is also computationally expensive to run and difficult to keep updated due to fundamental changes in the CAESAR code that the platform is based on.

This report describes improvements to the water partitioning and groundwater modules CLiDE platform and realignment with CAESAR-LF code, as implemented in the Hull and East Riding digital twin demonstrator FLOODTWIN (NERC NE/Z503411/1). To reflect the realignment with CAESAR, the new platform is called CAESAR-GW. This report covers the development of CAESAR-GW, testing of the model and calibration of the groundwater components to reproduce the observed groundwater time series recorded at nine observation boreholes in the study area.

Contents

FIGURES

TABLES

1 Introduction

A Digital Twin is a dynamic virtual copy of a physical asset, process, system or environment that looks like and behaves identically to its real-world partner. Unlike a digital model, which is calibrated to reproduce an observed process behaviour, the digital twin must adapt to any alteration made to the real-world physical system while maintaining a good representation of the processes it simulates (Blair, 2021).

Stakeholder engagement with data and models to better understand the environment, change impacts and risk is becoming more prominent with the advances in digital technology. The availability of earth observation (EO), field sensors and networks all integrated in a modelling platform facilitate stakeholder engagement and provide a decision-making tool that can be invaluable to address potential risks such as flooding. The aim of FLOODTWIN is to build a Digital Twin for water-related hazard forecasting and decision-making in the Hull and the East Riding of Yorkshire, UK, a region heavily impacted by several hydrometeorological hazards.

FLOODTWIN project involves four research institutes including the University of West England, the University College London, the British Geological Survey, and is led by the University of Hull. The project involves four work packages designed to deliver a demonstrator digital twin. The first work package deals with the integration of observations, the second work package focuses on the hydrological model digital twin, mainly the groundwater simulator which is described in this report. The third work package addresses the requirements to develop a real-time digital hub, and the fourth work package explores stakeholder engagement.

There are multiple ways to numerically represent hydrology within a model. The most appropriate method dependents on the scale, resolution and complexity requirements of any particular study. At the regional scales we're interested in here, cellular automata models provide a reasonable compromise between computational expense and process representation. CAESAR-LF is a landscape evolution model that makes use of the 2D hydrodynamic LISFLOOD model (Burek et al., 2013) to represent surface water flow. In 2013 BGS modified CAESAR-LF, replacing TOPMODEL (hydrological component that drives water flow) with a water partitioning model and a groundwater model. The resulting platform (CLiDE) has been used to simulate hydrological and hydrogeological processes in the UK, Belgium, France, Switzerland and India.

CLiDE integrates distributed water partitioning / recharge, groundwater flow, surface water, and storm surge inundation to provide an almost complete representation of the hydraulic processes causing flooding at the study area. The groundwater / surface water fluxes are exchanged using two-way couplings at runtime in a framework (see Barkwith et al., 2015). The water partitioning and recharge calculation is based on the hydrological SLiM methodology (Wang et al., 2012) and the groundwater utilises a simple Darcian approach most suited to unconfined, homogeneous aquifers. The timestep changes required to provide a more accurate daily water balance along with the addition of the water partitioning component and groundwater model massively increased the runtime compared to the original CAESAR-LF model. The timestep changes also increased the resources needed to keep the CLiDE up-to-date with the latest CAESAR-LF releases.

Hull represents a complex hydrological setting and application of CLiDE proved to be problematic, generating unrealistic representations of groundwater. Additionally, the groundwater flow simulation in CLiDE is undertaken using an explicit approach, which depends on the use of a high temporal resolution (computationally expansive) to maintain stability and numerical accuracy. In this report we update the CLiDE model to address these two issues to improve the performance of the model and enable the digital twin to address the main objectives of informing the users / stakeholders within a reasonable time. This new model is more closely aligned with the CAESAR-LF, with minimal changes to code outside new water partitioning and groundwater modules. The new model is called CAESAR-GW.

2 Summary of the existing CLiDE platform

The CLiDE platform (Barkwith and Coulthard, 2013) consists of three modules that represent the movement of water in the soil store, within the saturated ground, and on the ground surface as illustrated in Figure 1.

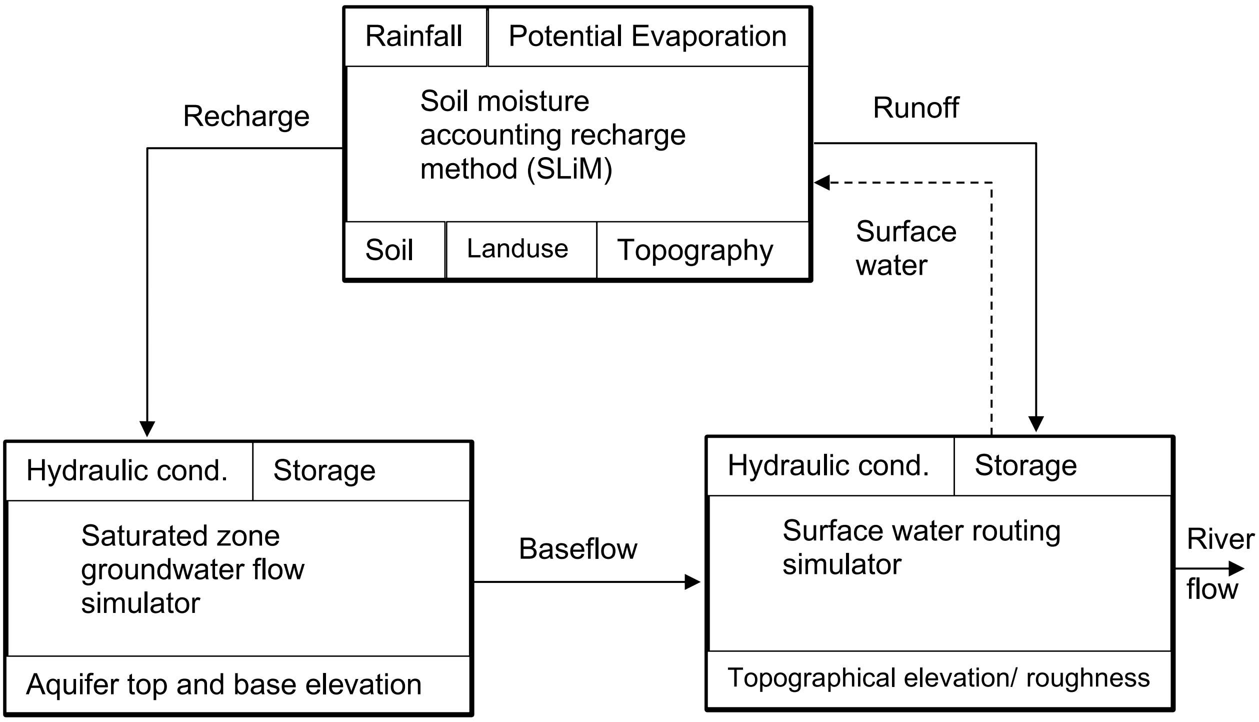


Figure 1 Schematic showing the linkages between the CLiDE platform hydrological components (Barkwith and Coulthard 2013).

2.1 WATER PARTITIONING MODULE (SLIM)

The calculations start at the water partitioning module that splits the rainfall into actual evaporation, soil moisture, runoff and recharge using a methodology referred to as SLiM and described in detail by Wang et al. (2012). This is a soil moisture accounting method that considers plant interception. If the rainfall intensity exceeds the capacity available in the soil storage, the evapotranspiration, and losses due to plant interception, the remaining water is considered as excess water that is divided into recharge and runoff components based on baseflow index parameter (BFI). This baseflow index parameter depends on the permeable nature of the catchment and is equal to one minus the runoff coefficient (ROC) a hydraulic term that is used more in hydrological models.

In the SLiM methodology (Figure 2), the BFI value is not only linked to the hydrological characteristics of the ground surface but also to the topographical gradient. An average gradient value is calculated for the topography across the hydrological domain within the study area. The BFI value at one location is then adjusted based on the topographical gradient at that location compared to the average gradient. The recharge volume for the given time step duration is calculated by multiplying the volume of excess water for that time step by BFI. The runoff volume, on the other hand is calculated by multiplying the volume of excess water by one minus BFI value.

The input files to SLiM are:

- Rainfall (gridded files or rainfall stations)
- Potential evaporation (gridded files or rainfall stations)
- Landuse type (gridded)
- Soil data (gridded)

This module runs using a sub-daily time step defined by the user for the rainfall data input. For each time step it calculates the amount of actual evaporation based on the value of the soil moisture deficit. The soil moisture deficit value can range between zero and a maximum value calculated using the root depth of the crop multiplied by the soil moisture at full saturation minus the soil moisture at wilting point. Evapotranspiration may be equal to potential evaporation if the soil moisture deficit value is close to zero; however, it is zero if the soil moisture deficit is at its maximum value.

The model domain is defined by the DEM, i.e. where values are different than -9999. Model resolution is also defined by the cell size specified in the DEM gridded ascii file.

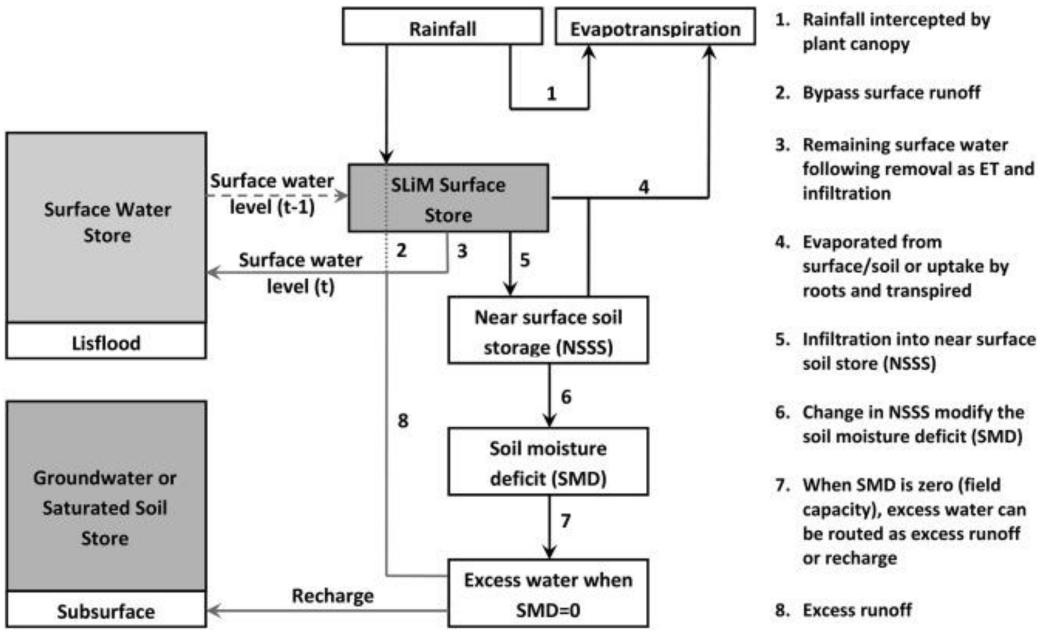


Figure 2 Flow chart for partitioning of water at the soil surface using the SLiM soil water balance method. Connections with the other water components are included (reproduced from Barkwith et al., 2015).

2.2 THE GROUNDWATER MODULE

Groundwater flow is simulated using a two-dimensional grid, constructed using square cells linked through a hydraulic conductance based on the hydraulic conductivity value. Groundwater head is calculated at a grid node for a time-step using Darcy's law to calculate the water fluxes between the cell and its neighbouring cells.

$$q_{gw} = \frac{2T_iT_j}{T_i+T_j}(h_i - h_j) \quad \text{Equation 1}$$

Where T_i is the transmissivity that can be approximated by multiplying hydraulic conductivity by aquifer depth. Some controls have been introduced as a flux limiter to improve stability when large outliers in transmissivity are encountered. The total flux for a cell is calculated including fluxes from or to neighbouring cells, source and sink terms, and recharge fluxes. The aquifer heads are then updated at each point in the domain simultaneously for each time step (dt) using the mass balance equation:

$$h_{t+\Delta t} = h_t + \frac{1}{S_y}\frac{q_{gw}}{\Delta x^2}\Delta t$$ Equation 2

The outer boundary conditions can be defined as specified aquifer head (Dirichlet) boundary conditions or no-flow (Neumann B) boundary conditions. The base of the aquifer is defined as impermeable; however, leakage from the base of the modelled aquifer is included in the flux algorithm as a secondary sink term. The boundary conditions across the top of the aquifer allows groundwater to be returned to the surface component as baseflow where aquifer head is greater than the topography.

To ensure model stability, the time step is defined based on the Reynolds number calculated at each node for each time-step. Reduction of the Reynolds number is possible by reducing the time-step or increasing the cell size. The volume of water returned to rivers is dependent on the specific yield of the groundwater node and the thickness and hydraulic conductivity of the riverbed.

2.3 THE SURFACE WATER MODULE

The surface water module of CAESAR, which routes surface runoff and channel flow, is controlled by a stripped-down version of the LISFLOOD-FP model (as used by Coulthard et al., 2013). LISFLOOD-FP is a one-dimensional inertial model that is applied in the x and y directions to simulate flow in two dimensions over a raster grid.

Routing is based on the one-dimensional Saint-Venant equations, as modified by Bates et al. (2010). The flow equation considers acceleration, water surface gradient and friction properties. This equation is given by:

$$q_{t+\Delta t} = \frac{q_{t-gh_f S\Delta t}}{1+gh_f\Delta t q_t n^2/h_f^{3.33}}$$ Equation 3

Where h_f is the flow depth, S is the hydraulic gradient, n is Manning's coefficient, and g is the gravitational constant.

To maintain stability, the LISFLOOD component calculates the maximum allowable length of the next time step (Δt_{max}) for which a converging solution is highly likely to be maintained. The calculation is based on the shallow water Courant-Friedrich-Levy (CFL) condition, in which:

$$\Delta t_{max} = \alpha\, \Delta x/\sqrt{gh}$$ Equation 4

The coefficient α is typically defined as being between 0.3 and 0.7 (Bates et al., 2010).

The surface water module receives runoff from the water partitioning module and from the groundwater module on a daily time step. It split the daily time step into sub time steps based on Δt_{max} calculated above.

The depth of the surface water calculated by the surface water module at the end of the one day is passed back to the water partitioning module and is added to rainfall. That means that the surface water routing affects the soil moisture / soil moisture deficit value simulated by the water partitioning module indirectly.

3 Changes made to the CLiDE groundwater module

The first runs undertaken for this study to simulate compound flooding in Hull and the wider county highlighted the need to modify the code to improve model performance. The main issues relate to the contrast between the topographical characteristics of the study area to the west, hilly where the chalk aquifer outcrops, and these characteristics to the east of the area, flat where the chalk aquifer is covered with superficial deposits. Much of the solution within the resulting CAESAR-GW model was found by aligning the groundwater module features with the UK national scale model BGWM (British Groundwater Model: Bianchi et al., 2024).

The hydraulic properties of the Chalk aquifer have previously been optimised within the BGWM model and these were used to inform the CAESAR-GW model calibration for this study. However, this technique required some modifications to the original CLiDE model. Firstly, BGWM is driven with recharge values calculated using the distributed recharge model ZOODRM (Mansour et al., 2018), which applies a recharge calculation method that is slightly different from SLiM methodology in CLiDE. ZOODRM applies a simplified version of the FAO recharge calculation method, which is similar to SLiM but less complex. This results in different modelled recharge values in CLiDE and BGWM despite the same inputs. Secondly, the introduction of topographical gradient within SLiM (used to adjust the runoff coefficient) results in recharge values that can't be calibrated to an acceptable level. This problem stems from the fact that most of the area is relatively flat except the part where the Chalk outcrops to the West, which produce relatively steep topographical gradients. Within CLiDE, the correlation of the baseflow index value (one minus the runoff coefficient value) to the topographical gradient greatly increases the baseflow index over the flat areas, causing significant recharge to occur over the areas considered relatively impervious. In reality, the steep gradient at the Chalk outcrop lowers the baseflow index value leading to the estimated recharge being insignificant compared to elsewhere. Finally, the map used to calculate the baseflow index values for recharge calculation is also used to control the amount of groundwater returning to the surface. This setting prevented the model from converging to a solution. It was necessary to simplify the water partitioning method and to uncouple the input data that control the water partitioning module from those used to control the groundwater flow simulation.

The approach used to solve for groundwater heads in CLiDE has proved to be inefficient when the model is used to reach steady state conditions, or when the model is used to simulate transient conditions over a period of time longer than a couple of months. The explicit nature of the solution necessitated the use of a very small timestep to maintain numerical stability and accuracy, which caused the model to take long time to reach a solution and consequently prevented the calibration of the model to refine the hydraulic parameter values. In addition to the efficiency problem, the groundwater module in CLiDE considers the base of the aquifer as horizontal flat surface at an elevation of 0.0 m AOD. This introduces problems as the simulated groundwater heads approach a value of zero (for example around the coast) and the numerical solution fails.

To overcome these problems, both the water partitioning module and the groundwater simulators in CLiDE were modified as described below.

3.1 UPDATING THE WATER PARTITIONING MODULE

The SLiM methodology for the calculation of recharge and runoff used in CLiDE was replaced with the recharge calculation method implemented in the distributed recharge model ZOODRM (Mansour and Hughes, 2004; Mansour et al., 2018) in CAESAR-GW. This method is based on an approach that is very similar to SLiM but removes the dependence of recharge and runoff estimation on the topographical gradient. In addition, this approach does not allow for plant interception and surface water ponding is not added to rainfall intensity in the next timestep. Despite its simplicity (compared to SLiM), it can successfully reproduce the groundwater heads and river flows when used to drive BGWM (Bianchi et al., 2024). This method is referred to as the FAO-Basic hereafter.

The FAO-Basic recharge calculation method is based on the EA-FAO method, as modified by Griffiths et al. (2006). It uses the moisture content of the soil at field capacity (θ_{fc} [L^3 L^{-3}]) and at wilting point (θ_{wp} [L^3 L^{-3}]) together with the root constant of the landuse (Z_r [L]) to calculate two hydraulic parameters (TAW and RAW) used to determine the amount of evapotranspiration.

The TAW and RAW parameters stand for total available water and readily available water respectively and given by:

$$TAW = Z_r(\theta_{fc} - \theta_{wp}) \quad \text{Equation 5}$$

$$RAW = p * TAW \quad \text{Equation 6}$$

Where p [-] is referred to as a depletion factor.

Griffiths et al. (2006) suggest calculating the evapotranspiration e_s [L] as a function of the potential evaporation and an intermediate soil moisture deficit value calculated as:

$$s_s^* = s_s^{t-1} - r + e_p \quad \text{Equation 7}$$

Where s_s^{t-1} [L] is the soil moisture deficit calculated at the previous time step, r is the rainfall [L], and e_p is the potential evaporation [L].

e_s is then calculated from:

$$e_s = e_p \left[\frac{s_s^*}{TAW-RAW}\right]^{0.2} \quad \text{for } s_s^* < RAW$$

$$e_s = e_p \quad \text{for } s_s^* \geq RAW \quad \text{Equation 8}$$

$$e_s = 0 \quad \text{for } s_s^* \leq TAW$$

The current time step soil moisture deficit is then calculated from:

$$s_s = s_s^{t-1} - r + e_s \quad \text{Equation 9}$$

Recharge and overland flow are only generated when the calculated soil moisture deficit becomes zero. The volume of water remaining from precipitation after accounting for soil moisture deficit and evapo-transpiration, the excess water, is then split into recharge and overland flow using a run-off coefficient.

3.2 UPDATING THE GROUNDWATER MODULE

The explicit approach used to solve the finite difference form of the basic flow equation is replaced with an implicit approach. In two dimensions, the numerical form of the basic flow equation is given by:

$$T_x \frac{H_{i+1,j} - 2H_{i,j} + H_{i-1,j}}{\Delta x^2} + T_y \frac{H_{i,j+1} - 2H_{i,j} + H_{i,j-1}}{\Delta y^2} = S \frac{H_{i,j} - H_{i,j}^*}{\Delta t} \quad \text{Equation 10}$$

Where T_x and T_y denote the transmissivity values of the saturated aquifer in the x and y directions only at node (i, j) respectively. $H_{i,j}$ and $H_{i,j}^*$ are the groundwater heads calculated at node (i, j) for the current and previous time steps respectively.

The difference between the explicit approach and the implicit approach to solve Equation 10 is to assume that the heads to the left-hand side of the equation are those being calculated at the previous time step, i.e., they are all superscripted with the character '*', or those to be calculated for the current time step respectively. The explicit form of Equation 10 allows the calculation of $H_{i,j}$ directly as a function of the groundwater heads at the surrounding nodes calculated at the previous time step. The implicit form of Equation 10, on the other hand, has the previous time step groundwater head $H^*_{i,j}$ as the only known groundwater heads while all other groundwater heads need to be determined for the current time step. An iterative solution is required in this case to reach a solution for the implicit approach.

The implicit approach for solving the numerical form of the basic flow equation involves a system of linear equation that can be represented in a matrix form as $\boldsymbol{A} \,.\, \boldsymbol{u} = \boldsymbol{f}$, where $\boldsymbol{u}$ denotes a vector of groundwater heads, $\boldsymbol{A}$ is a matrix of the hydraulic conductance values between the nodes, and $\boldsymbol{f}$ is a vector of the values of the known parts of the numerical equation. In linearised unconfined groundwater flow domains, the matrix A has a sparse structure that can be divided into three matrices $\boldsymbol{D}$ representing the diagonal elements, $-\boldsymbol{U}$ is the upper triangular elements and $-\boldsymbol{L}$ is the lower triangular element. The linear system of equation can then be written as:

$$(\boldsymbol{D} - \boldsymbol{L} - \boldsymbol{U}) \,.\, \boldsymbol{u} = \boldsymbol{f} \qquad \text{Equation 11}$$

Let $\boldsymbol{v}$ denote an approximate solution to $\boldsymbol{u}$. Substituting in Equation 11 and using the components of the approximation as soon as they are updated during an iteration (The Gauss-Seidel method). The iteration method in this case is written as:

$$\boldsymbol{D}\boldsymbol{v}^{n+1} - \boldsymbol{L}\boldsymbol{v}^{n+1} - \boldsymbol{U}\boldsymbol{v}^{n} = \boldsymbol{f}$$

$$(\boldsymbol{D} - \boldsymbol{L})\boldsymbol{v}^{n+1} = \boldsymbol{U}\boldsymbol{v}^{n} + \boldsymbol{f} \qquad \text{Equation 12}$$

$$\boldsymbol{v}^{n+1} = (\boldsymbol{D} - \boldsymbol{L})^{-1}\boldsymbol{U}\boldsymbol{v}^{n} + (\boldsymbol{D} - \boldsymbol{L})^{-1}\boldsymbol{f}$$

The correction for the Gauss-Seidel iteration is given by the difference between $\boldsymbol{v}^{n+1}$ and $\boldsymbol{v}^{n}$; let Δ denote this difference. When this correction is added to $\boldsymbol{v}^{n}$, the new value of $\boldsymbol{v}$ is expected to become closer to the exact solution. In addition, if this correction is enlarged by multiplying it by a factor ω that has a value greater than one, it will result in a value of $\boldsymbol{v}^{n+1}$ that is even closer to the exact solution $\boldsymbol{u}$. $\boldsymbol{v}^{n+1}$ can therefore be determined from of $\boldsymbol{v}^{n+1} = \boldsymbol{v}^{n} + \omega\Delta$, where ω is called the relaxation factor.

From Equation 12 let $\boldsymbol{v}^{n+1} = \boldsymbol{R} = (\boldsymbol{D} - \boldsymbol{L})^{-1}\boldsymbol{U}\boldsymbol{v}^{n} + (\boldsymbol{D} - \boldsymbol{L})^{-1}\boldsymbol{f}$, then the new approximation of $\boldsymbol{v}$ can be written as: $\boldsymbol{v}^{n+1} = \boldsymbol{v}^{n} + \omega(\boldsymbol{R} - \boldsymbol{v}^{n})$. The last equation can be rearranged to take the final form of the SOR scheme as follows: $\boldsymbol{v}^{n+1} = \omega\boldsymbol{R} + (\boldsymbol{I} - \omega)\boldsymbol{v}^{n}$. The value of ω ranges between 1.0 and 2.0 and a value of 1.6 is typically adopted in the numerical models.

Reverting to the application of the conventional finite difference form of the basic flow equation with the implicit approach removes the limitation of having the base of the aquifer defined as horizontal flat surface at an elevation of 0.0 m AOD. In the modified approach, both the base and top elevations of the aquifer are defined at every grid node. The aquifer saturation is calculated, therefore, based on the defined base elevation whether below or above the sea level, and the simulated groundwater heads. If the simulated groundwater heads reach values that are higher than the top elevations of the aquifer, the groundwater flows discharge through the top face of the nodes. The discharge flux is calculated at a node using a hydraulic conductance value defined at the node multiplied by the difference between the groundwater head and the elevation of the top of the aquifer at that node. If the value of the hydraulic conductance is set to zero, no groundwater flows will discharge through the upper face of the node causing confining conditions to take place at this node. It must be noted however, that the model does not switch from specific storage to specific yield or vice versa depending on the confining hydrological conditions, rather storage coefficient value specified in the input file will be used in all cases. Knowing the location of the confining conditions and specifying the storage coefficient values accordingly is a process that the user must undertake before initialising the simulation.

An additional update was made to this module allowing the base elevation of the aquifer to be specified at any depth rather than being fixed at an elevation of zero m AOD. In addition, the numerical solution was modified to prevent nodes dewatering, which resulted in a more stable approach and allowed the simulation of more complex scenarios.

Figure 3 shows an updated diagram illustrating the connection of the different modules within the modified CLiDe platform together with the update methodology description.

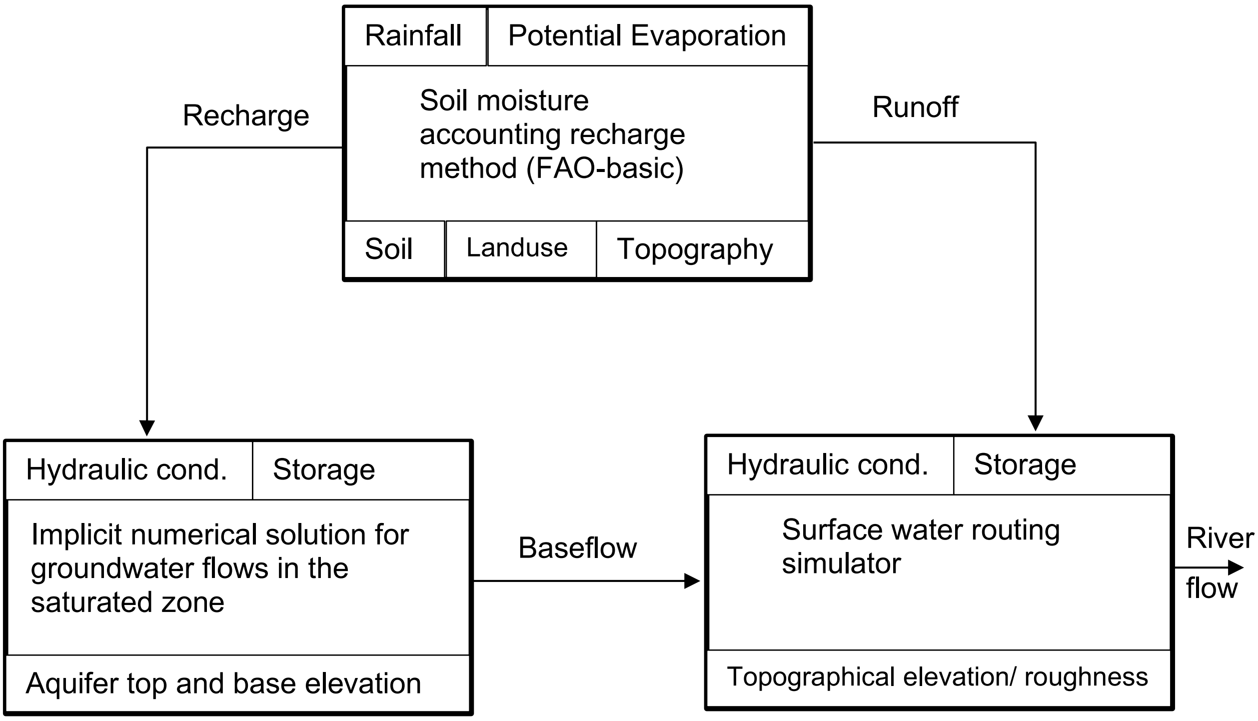


Figure 3 Schematic showing the linkages between the updated CLiDe platform hydrological components.

The interface used in CLiDE was updated in CAESAR-GW to allow for the specification of parameter values required by the implicit solver implemented in the modified groundwater module. Figure 4 shows the four edit boxes added to the “Groundwater” tab of the interface to: specify the values of the over relaxation parameter ω; the maximum permissible flow inaccuracy at a node; the maximum number of iterations the solver undertake before moving to the next time step if the error stays above the maximum permissible error; and finally the number of cycles over which the hydraulic conductance values between the nodes are updated in one time step to improve the accuracy of the unconfined solution.

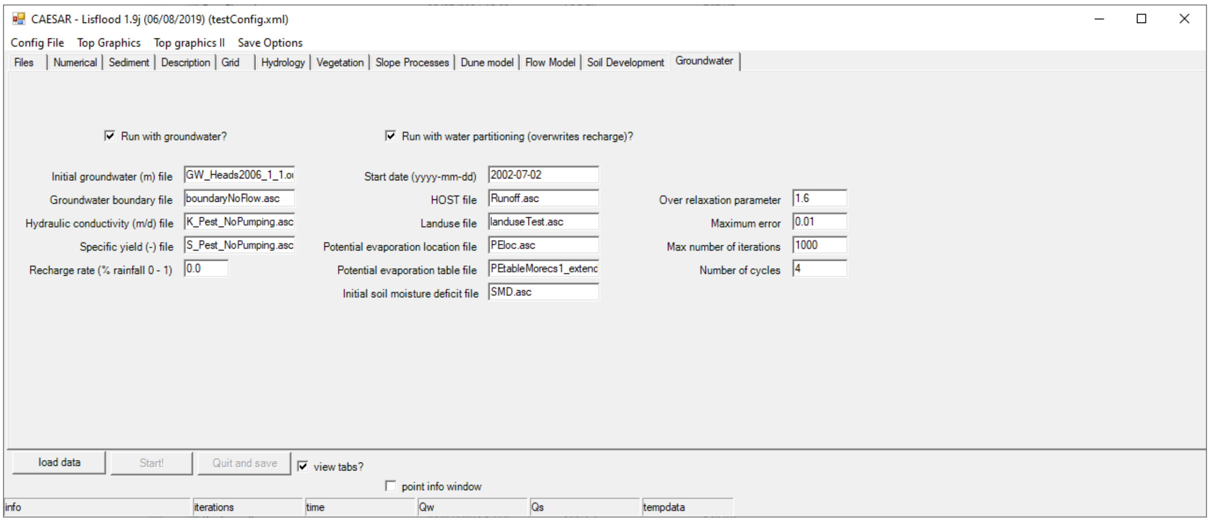


Figure 4 The updated “Groundwater” tab of the CLiDe GUI.

4 Study area

The pilot area used to build the hydro Digital Twin is the Hull and the East Riding of Yorkshire (Figure 5) in the northeast of England. The region is characterised by a gently undulating rural topography, much of which lies below the high-tide level of the Humber estuary. Flood protection at the mouth of the principal river (River Hull) relies on a tidal barrier that is utilised during storm surges. Surface water flooding is alleviated through pumping surface water across and out of the catchment during heavy rainfall. A series of drainage channels and pumping stations drain much of the surrounding countryside into the Hull.

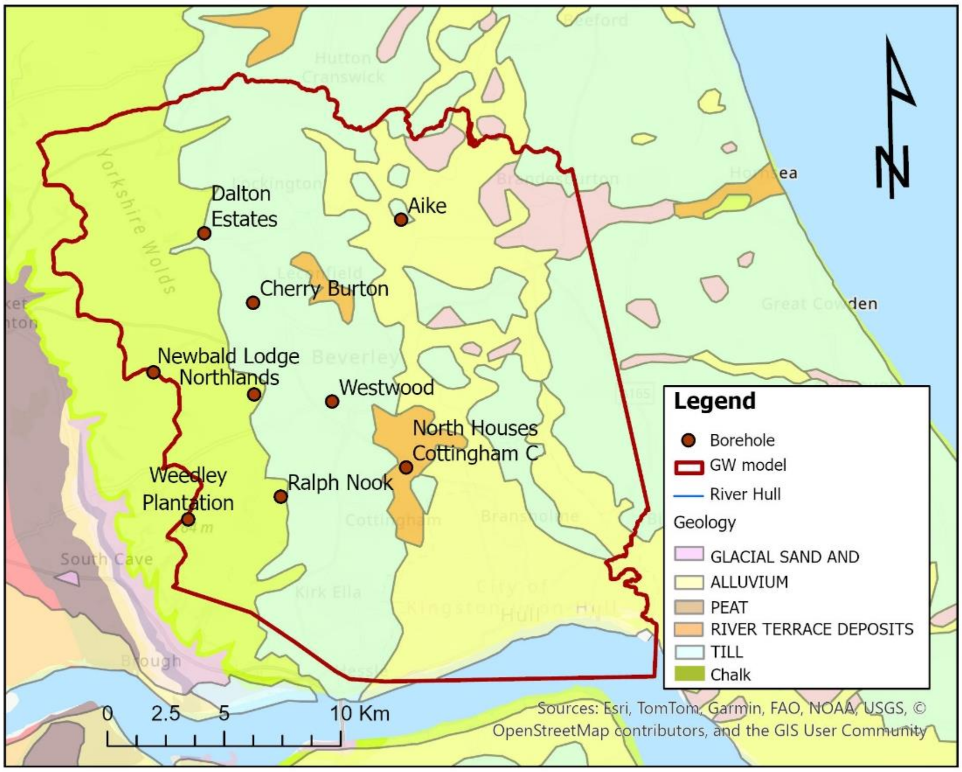


Figure 5 Study area represented by the digital twin

The geology of the region consists of a Permo-Triassic sequence with the Sherwood Sandstone group to the west and the Yorkshire Chalk aquifer to the East. The chalk sections are heavily developed and supply Hull with water, which has adverse impact on spring flows in the area. Pumping for water supply causes saline intrusion from the coastal area into the aquifer. Quaternary deposits that cover the chalk are structurally complex, they mostly confine the chalk aquifer underneath due to their low permeability but can also form locally important aquifers.

The low permeability of the surface deposits increases overland runoff, limiting the amount of water that infiltrates into the Chalk aquifer. Most of the water that does reach the aquifer at its outcrop eventually discharges into the River Hull. Combined with storm surges, this makes the Hull and East Riding of Yorkshire area highly vulnerable to flooding from rainfall, groundwater, and the sea. Consequently, the natural River Hull valley is one of the most at-risk developed floodplains in England.

5 Data files used with CAESAR-GW for the FLOODTWIN project

Within the CLiDE user guide (Barkwith and Coulthard, 2014), a full list of the input files for the original model is provided. In this section, we list the files that are required in the updated water partitioning and groundwater modules in CAESAR-GW (Table 1). For further detail the reader is referred to the user guide document for complete description of the philosophy behind the use of the CLiDE platform, how the model is setup, and the steps used to run the model.

There are seven files that control the water partitioning calculations. Five of these files give spatially gridded information and two that give information in time series format. The names of these files can be flexibly defined by the user through the CAESAR-GW GUI, as shown in the set of edit boxes in the middle of Figure 4.

The files that provide variable inputs over time are precipitation, potential evaporation and tidal stage. Within the GUI, the potential evaporation file name is specified in the “Groundwater” tab (Figure 4) and the precipitation and tidal stage files specified in the “Hydrology” tab. Table 1 shows the names of these file as used in the model of the FLOODTWIN project, but these can have different names (note, tidal stage not used).

The are six files that are used by the groundwater module to define the hydraulic characteristics of the domain and its physical properties as well as the initial conditions. The names of three of these files can be also flexibly specified in the “Groundwater” tab of the CAESAR-GW GUI using the edit boxes to the left of Figure 4. These files provide spatially distributed information for the initial groundwater heads, the hydraulic conductivity and the specific yield. The names of these file shown in Table 1 are those used in the FLOODTWIN model, but these can be named differently. The remaining files have their names hardcoded and they must be specified with the names as listed in Table 1. These files define the top and base elevations of the aquifer and the vertical conductance values that control the discharge of groundwater flows out of the aquifer. In unconfined aquifers, the top elevation of the aquifer is usually the same as the elevation of the ground surface. The topographical information is specified in a file, the name of which is given in the “Files” tab. The use of two files that provide the same information may not seem efficient; however, the availability of a file that defines the top elevation of the aquifer gives the flexibility to simulate groundwater flows in confined aquifers, the top elevation of which may be different from the topography.

Table 1 Input files used by the water partitioning and groundwater modules

Input File name	Description	Note
DEM.asc	Defines the ground elevations across the study area. Also dictates the extent of the study area and the resolution of the numerical grid used to discretise the area.	Name specified through GUI
Runoff.asc	Defines the runoff coefficient ID spatially across the area. The runoff coefficient value is derived from an array of values hardcoded in the model and based on this ID.	Name specified through GUI
landuseTest.asc	Defines the landuse ID spatially across the area. This ID is used to derive the root constant and depletion factor values from arrays of values that are hardcoded in the model.	Name specified through GUI
Soil.asc	Defines the soil ID spatially across the area. This ID is used to derive the soil hydraulic properties, mainly saturation at field capacity and saturation at wilting point, from arrays of values that are hardcoded in the model.	File name is fixed
SMD.asc	Defines the initial values of the soil moisture deficit spatially across the area.	Name specified through GUI
PEloc.asc	Defines the PE zone ID spatially across the area. This ID is used to define the column from where PE is read from in the PE table file (PEtableMorecs1.dat).	Name specified through GUI
PEtableMorecs1 _extended.dat	A table that gives the potential evaporation values for each month of the year but in the units of mm/day. Header is: Year Month ID1 ID2 ID3 etc. Start and end years are independent from but must cover the simulation period.	Name specified through GUI
Raingrid.asc	Defines the rainfall zone ID spatially across the area. This ID is used to define the column from where rainfall is read from in the rainfall table file (RainfallStations.txt).	Name specified through GUI
RainfallStations.txt	A table that gives the rainfall values for each time step length specified in the CLiDe GUI. The file has no header and no time column. The data must be synchronised with the simulation start date as also specified in the CLiDe GUI.	Name specified through GUI
GW_Heads_ini.out	Gives the initial groundwater head values spatially across the area.	Name specified through GUI
K_Pest_NoPumping.asc	Specifies the hydraulic conductivity values spatially across the area.	Name specified through GUI

S_Pest_NoPumping.asc	Specifies the storage coefficient values spatially across the area.	Name specified through GUI
BaseElevation.asc	Specifies the base elevation values of the aquifer spatially across the area.	File name is fixed
TopElevation.asc	Specifies the top elevation values of the aquifer spatially across the area.	File name is fixed
VerCond.asc	Specifies the vertical hydraulic conductance which controls the groundwater discharge out of the aquifer values spatially across the area.	File name is fixed
boundaryNoFlow.asc	A file that specifies the hydrological boundaries as described in Section 3.6.3 in Barkwith and Coulthard (2013)	Name specified through GUI

6 Application of the groundwater model to reproduce the observed groundwater heads

The water partitioning and groundwater modules are used to simulate the groundwater flows in the study area (Section 2). The surface water module is switched off to maximise the performance of the model. The numerical model consists of one layer and has the extents as defined by the polygon shown in Figure 5. The boundary conditions are defined as impermeable everywhere. No pumping is assumed to take place, and recharge is the only source of water driving the groundwater flows. The area is drained mainly by the river Hull, which runs approximately in the centre of the study area. Groundwater can also discharge through the top of the aquifer when the groundwater heads become above the ground surface and the hydraulic conductance values are set to allow this.

The characteristics of the study area and the driving data are provided to the model as time series text files or as gridded ascii files. These include data released under the Open Government Licence; namely, rainfall data (https://environment.data.gov.uk/hydrology/explore; © Environment Agency copyright and/or database right) and topsoil moisture data (https://mapapps2.bgs.ac.uk/ukso/home.html; Henrys et al. (2014), © UK Centre for Ecology & Hydrology. Contains British Geological Survey materials © UKRI 2014. Contains Ordnance Survey data © Crown copyright and database right 2007). Additional data include potential evaporation data (Hough & Jones (1997); available under the CC BY-NC-SA 2.5 licence), digital elevation data (Lehner et al. (2008); from HydroSHEDS version 1 database © World Wildlife Fund, Inc. (2006-2022)), the BGS landuse map, and aquifer top and base elevations obtained from geological models prepared by BGS.

Figure 5 shows that the Chalk aquifer outcrops to the west of the area and dips beneath the superficial deposits to the east. The superficial deposits are characterised as low permeability materials and they may lead to the aquifer confinement. The hydraulic parameter values are defined differently over these two zones, for example, Figure 6 shows the hydraulic conductivity values specified over these two zones and Figure 7 the vertical hydraulic conductance. The specific storage used in the model was almost the same everywhere in the model with a value of approximately 0.00005 (m^{-1}), however, a distribution of storage coefficient values derived by multiplying the specific storage value by the distributed aquifer thickness is used in the model.

There are nine observation boreholes with time series of groundwater levels available in the area the locations of which are also shown in Figure 5. The run (model setup) that produced the best match between the observed and simulated groundwater heads has hydraulic parameter values shown in the labels of Figure 6 and Figure 7. An example of groundwater head distribution across the study area is shown in Figure 8. It must be noted that the depths of these boreholes are not known, and it is difficult to confirm if these boreholes are registering the groundwater heads in the Chalk aquifer or in the superficial deposits above. Table 2 shows the quality of the match between the observed and simulated time series using the root mean square error (RMSE), the mean average error (MAE), and the Nash Sutcliffe efficiency (NSE) measures. Figure 9 shows plots of the observed and simulated groundwater head timeseries at these boreholes. The model fails to produce the fluctuations at Aike and North Houses Cottingham C, where there is a good agreement on the long-term groundwater levels, but the simulated heads are too flashy compared to the observed groundwater head fluctuations, which are more subdued. Also, the model fails to produce a good match at Newbald Lodge and Weedley Plantation, where the simulated groundwater fluctuations compare well with the observed ones; however, there is a bias difference with the simulated heads being lower than the observed ones. The model produces a good match between the observed and simulated groundwater heads at five of these boreholes including: Dalton Estates, Cherry Burton, Northlands, Westwood, and Ralph Nook as illustrated in the efficiency measure reported in Table 2.

Table 2 Efficiency measures calculate at the boreholes in the study area.

Borehole name	RMSE	MAE	NSE
Dalton Estates	3.91	3.11	-0.86
Cherry Burton	3.56	2.63	0.48
Aike	0.68	0.53	-6.35
Westwood	2.43	1.77	-0.07
Northlands	4.62	3.39	0.56
Newbald Lodge	9.76	8.18	-3.44
North Houses Cottingham C	1.65	1.26	-20.26
Ralph Nook	3.95	2.95	0.54
Weedley Plantation	20.85	19.43	-37.10

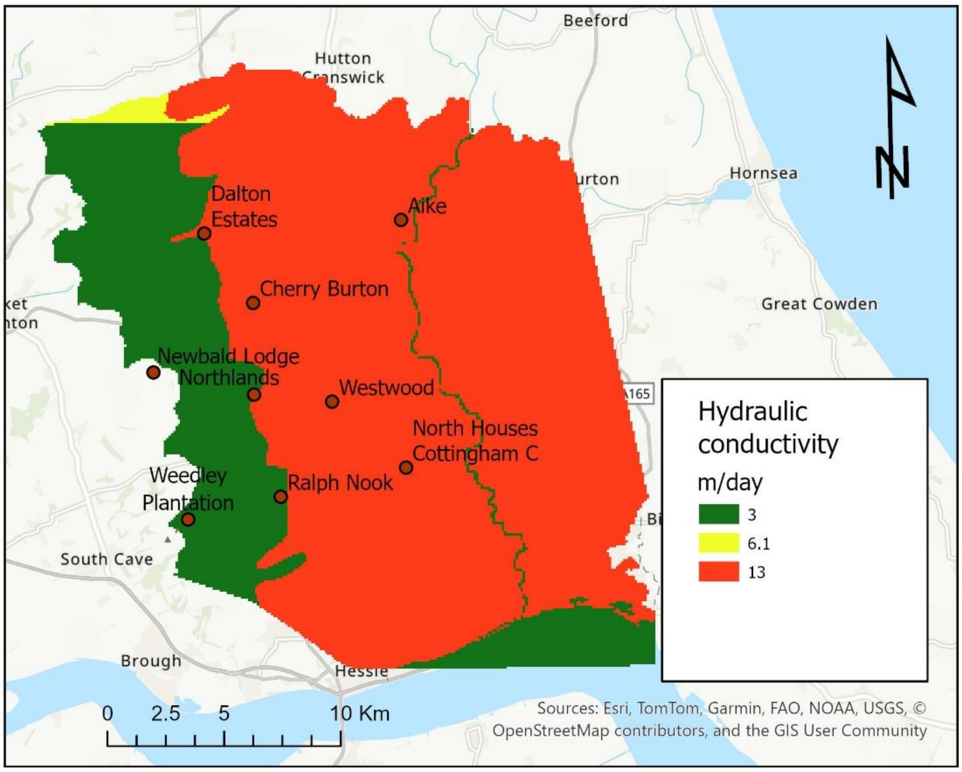


Figure 6 Hydraulic conductivity values used over the unconfined (west) and confined (east) parts of the aquifer

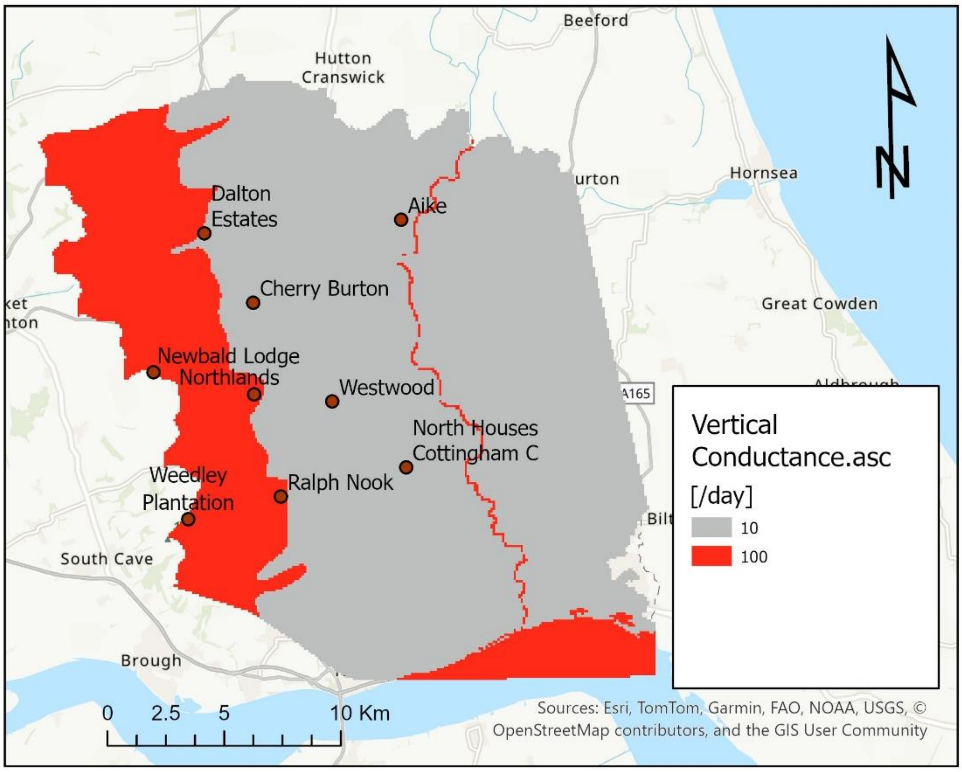


Figure 7 Vertical hydraulic conductance values used over the unconfined (west) and confined (east) parts of the aquifer

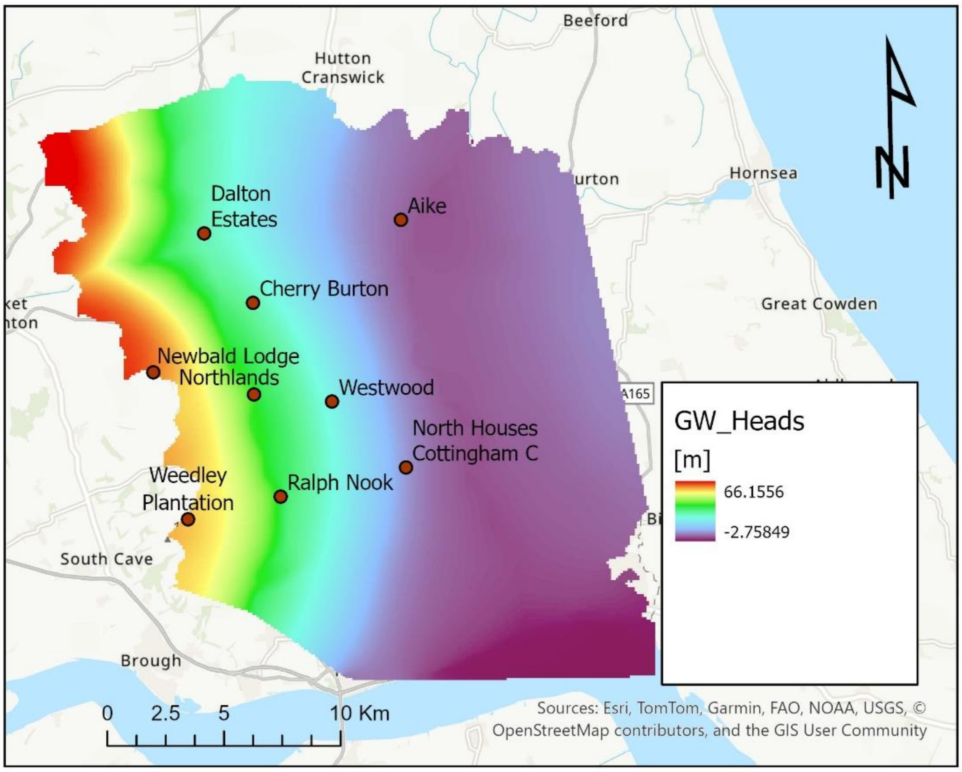


Figure 8 Peak groundwater heads simulated on 23/11/2023

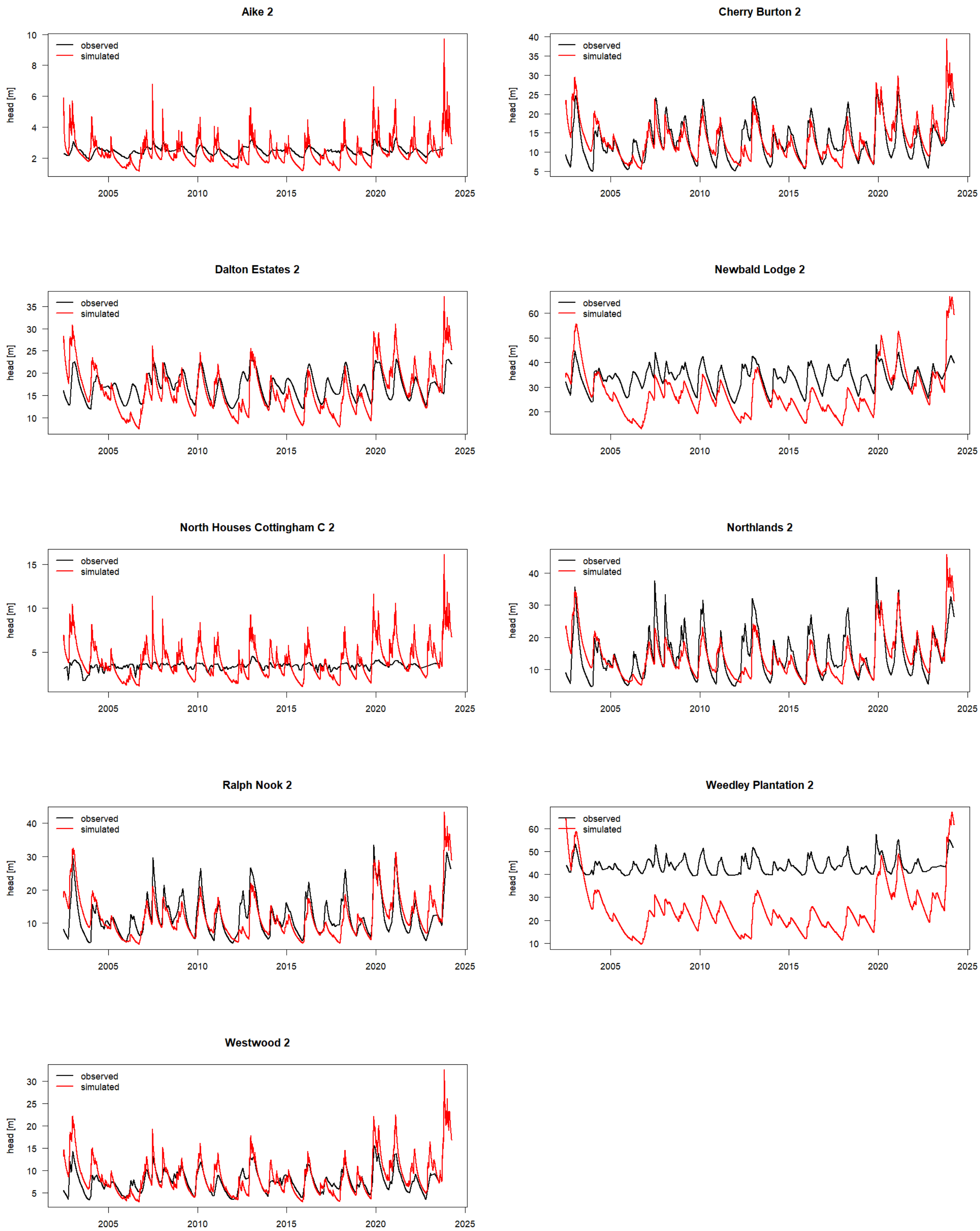


Figure 9 Observed ((https://environment.data.gov.uk/hydrology/explore) versus simulated groundwater heads at the observation boreholes

As the groundwater heads become above the ground elevation, the groundwater discharges on the ground surface in the form of baseflow. This baseflow is added to the runoff calculated by the water partitioning module and routed downstream following the topographical gradients. Figure 10 shows the surface water depth simulated on the 20th of November 2023 as a result of the runoff calculated by the water portioning module and the groundwater baseflow.

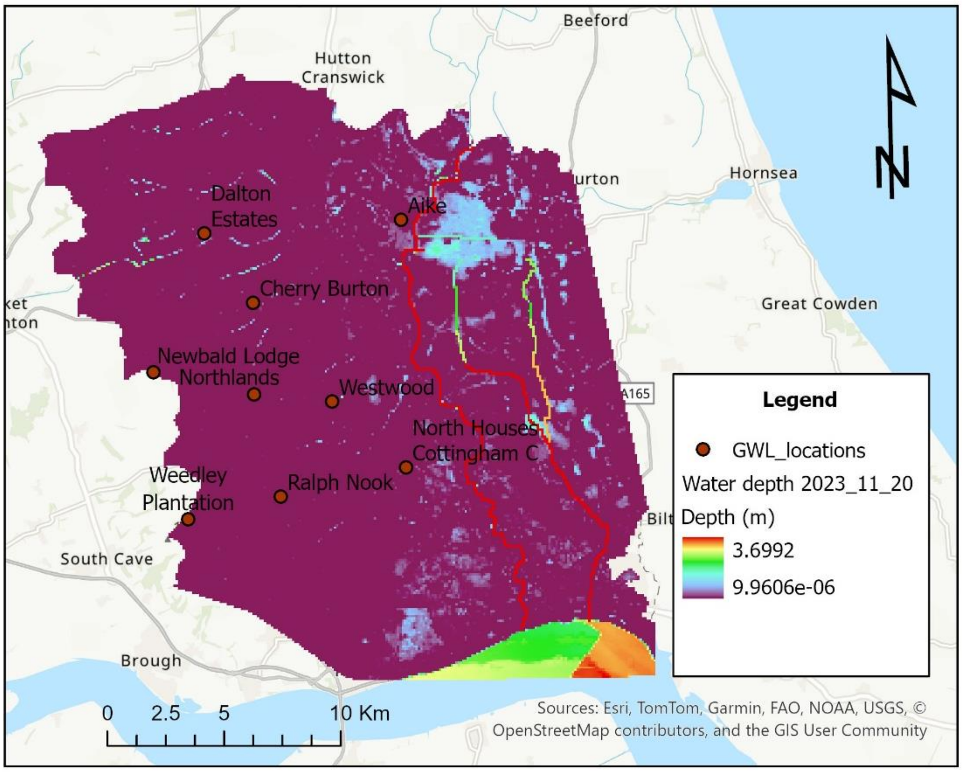


Figure 10 Simulated surface water depth, which is the result of runoff and baseflow simulated on 20/11/2023.

7 Summary and further work

This report describes the modifications undertaken to the CLiDE modelling platform and subsequent update to CAESAR-GW. These changes focussed on the water partitioning and groundwater modules, using an implementation in the Hull and East Riding as a digital twin demonstrator for the FLOODTWIN project.

The water partitioning model was simplified as the SLiM (surface water partitioning) methodology, that includes the topographical gradient in the calculation of recharge and runoff, does a poor job of estimating the recharge in complex hydrological settings. The use of the method of one file that describes multiple processes has added to these difficulties. SLiM was replaced with a simplified FAO method, previously implemented in a national scale distributed recharge model of the UK (BGWM). As a result, the recharge estimates in CAESAR-GW are more aligned with this national model.

The groundwater module was updated and an implicit approach for solving the finite difference form of the basic groundwater flow solution used. This method is more stable than the explicit method originally used in the CLiDE platform. It allows the specification of any time step length without affecting the stability or the accuracy of the solution. However, the larger the timestep size, the longer is the time required to reach a solution. An additional important update was undertaken, allowing the base elevation of the aquifer to be specified at any depth rather than being fixed at an elevation of zero m AOD. In addition, the numerical solution was modified to prevent node dewatering which resulted in a more stable approach and allowed the simulation of more complex scenarios.

While the hydraulic parameter values used with the groundwater module are specified through gridded ascii files, some of the parameter values used with the water portioning module are hard coded within CAESAR. It is recommended that the values of these parameters are specified in text files that are used as input to the model. This reduces the possibility of mistakes related to lack of user knowledge in the code, it also gives the user flexibility to assign zone numbers that relate to real-world parameter values.

Currently, the baseflow fluxes returned at the ground surface can moved overland through the routine implemented in the surface water part of the model. However, these fluxes do not contribute to the precipitation that takes place in the next time step. The importance of this feature on the simulated recharge and runoff values must be investigated and the possibility of turning this option on or off must be made as an option to reflect the hydrologic conditions of the study area being investigated.

This report describes the modifications and application of the CAESAR-GW platform to the region of Hull and East Riding as a digital twin demonstrator for FLOODTWIN. To improve this digital twin, further work is required to integrate surface water management options, including surface water pumps; stormwater reservoirs; and operation of tidal barriers. When integrated, and kept up to date with current climate, tidal data and soil moisture derived from earth observation, it could be used as a flood forecasting and early warning tool. The user could then test out different management scenarios to manage flooding in Hull.

References

The British Geological Survey holds most of the references listed and copies may be obtained via the library service subject to copyright legislation (contact libuser@bgs.ac.uk for details). The library catalogue is available at https://of-ukrinerc.olib.oclc.org/folio/

Barkwith, A, Coulthard, T J. 2013 CLiDE version 1.0 user guide. Nottingham, UK, British Geological Survey, 54pp. (OR/14/011) (Unpublished)

Barkwith, A, Hurst, M D, Jackson, C R, Wang, L, Ellis, M A, Coulthard, T J. 2015 Simulating the influences of groundwater on regional geomorphology using a distributed, dynamic, landscape evolution modelling platform. *Environmental Modelling and Software*, 74. 1-20. https://doi.org/10.1016/j.envsoft.2015.09.001

Bianchi, M, Scheidegger, J, Hughes, A, Jackson, C, Lee, J, Lewis, M, Mansour, M, Newell, A, O'Dochartaigh, B, Patton, A, Dadson, S. 2024 Simulation of national-scale groundwater dynamics in geologically complex aquifer systems: an example from Great Britain. Hydrological Sciences Journal, 69 (5). 572-591. 10.1080/02626667.2024.2320847

Blair, G S. 2021 Digital twins of the natural environment. Patterns. https://doi.org/10.1016/j.patter.2021.100359.

Burek, P, Van Der Knijff, J, De Roo, A. 2013. LISFLOOD - Distributed Water Balance and Flood Simulation Model - Revised User Manual. EUR EUR 26162. Luxembourg (Luxembourg): Publications Office of the European Union; 2013. JRC78917

Henrys, P.A.; Keith, A.M.; Robinson, D.A.; Emmett, B.A. (2014). Model estimates of topsoil moisture [Countryside Survey]. NERC Environmental Information Data Centre. https://doi.org/10.5285/8db84900-5fdb-43be-a607-e56c843d9b87

Hough, M. N. & Jones, R. J. A. 1997. The United Kingdom Meteorological Office rainfall and evaporation calculation system: MORECS version 2.0 – an overview. Hydrology and Earth System Sciences, 1, 227–239. https://doi.org/10.5194/hess-1-227-1997

Lehner, B., Verdin, K., Jarvis, A. (2008). New global hydrography derived from spaceborne elevation data. Eos, Transactions, American Geophysical Union, 89(10): 93–94. https://doi.org/10.1029/2008eo100001

Mansour, M M, Wang, L, Whiteman, M, Hughes, A G. 2018 Estimation of spatially distributed groundwater potential recharge for the United Kingdom. *Quarterly Journal of Engineering Geology and Hydrogeology*, 51 (2). 247-263. 10.1144/qjegh2017-051

Mansour, M M, Hughes, A G. 2004 User's manual for the distributed recharge model ZOODRM. British Geological Survey, 61pp. (IR/04/150) (Unpublished)