Three-dimensional hydrostratigraphy and groundwater flow models in complex Quaternary deposits and weathered/fractured bedrock: evaluating increasing model complexity

Susanne Charlotta Åberg 1 · Annika Katarina Åberg 1 · Kirsti Korkka-Niemi 1

Received: 17 June 2020 / Accepted: 18 December 2020 / Published online: 17 February 2021
© The Author(s) 2021

Abstract
Greater complexity in three-dimensional (3D) model structures yields more plausible groundwater recharge/discharge patterns, especially in groundwater/surface-water interactions. The construction of a 3D hydrostratigraphic model prior to flow modelling is beneficial when the hydraulic conductivity of geological units varies considerably. A workflow for 3D hydrostratigraphic modelling with Leapfrog Geo and flow modelling with MODFLOW-NWT was developed. It was used to evaluate how the modelling results for groundwater flow and recharge/discharge patterns differ when using simple or more complex hydrostratigraphic models. The workflow was applied to a study site consisting of complex Quaternary sediments underlain by fractured and weathered crystalline bedrock. Increasing the hydrostratigraphic detail appeared to improve the fit between the observed and simulated water table, and created more plausible groundwater flow patterns. Interlayered zones of low and high conductivity disperse the recharge/discharge patterns, increasing the vertical flow component. Groundwater flow was predominantly horizontal in models in which Quaternary sediments and bedrock were simplified as one layer per unit. It appears to be important to define the interlayered low-conductivity units, which can limit groundwater infiltration and also affect groundwater discharge patterns. Explicit modelling with Leapfrog Geo was found to be effective but time-consuming in the generation of scattered and thin-layered strata.

Keywords Groundwater flow · Heterogeneity · Unconsolidated sediments · Fractured rocks · Numerical modelling

Introduction

The main aim of hydrogeological studies is often to model groundwater flow, in which the three-dimensional (3D) geological structure is simplified; however, variation in hydraulic conductivity and the geological structure has a considerable impact on groundwater flow and groundwater discharge patterns, as presented in the study of Freeze and Witherspoon (1967). The construction of a 3D hydrostratigraphic model prior to flow modelling is essential to accurately characterize groundwater flow patterns (e.g. Artimo et al. 2004; Wycisk et al. 2009; Cox et al. 2013; Huntington et al. 2013), especially when the hydraulic conductivity of geological units varies considerably. In addition, variations in hydraulic conductivity in 3D groundwater flow models are usually based on model calibration, probabilities or uncertainty analysis (Refsgaard et al. 2012; Feinstein et al. 2012); however, the interpolation of scattered data can cause significant misinterpretation, as areas having highly dispersed aquifers, confining units and deposits are formed in a complex and unpredictable manner. For example, modelling of the hydrostratigraphy of an area that has formed during several glaciations with weak glacial erosion is challenging due to the local and scattered nature of the aquifers and confining units (Ross et al. 2005; Howett et al. 2015; Åberg et al. 2019). The hydraulic conductivity ($K$) of glacial till is orders of magnitudes lower than that of sands and gravels ($10^{-5}–10^{-2}$ m s$^{-1}$, e.g. Freeze and Cherry 1979), and understanding of the continuity of the units is thus important. In addition, weak glacial erosion has enabled the preservation of weathered bedrock, which commonly occurs in areas where intrusive or metamorphosed bedrock is exposed (Wright 1992; Hall et al. 2015).

Susanne Charlotta Åberg
susanne.aberg@helsinki.fi

1 Department of Geosciences and Geography, University of Helsinki, P.O. Box 64, Gustaf Hällströminkatu 2, FI-00014 Helsinki, Finland
The chemically weathered top of crystalline bedrock can be clayey, and its hydraulic conductivity thus tends to be low (10^{-9}–10^{-7} m s^{-1}, Welby 1981; Schoeneberger and Amoozegar 1990). By comparison, less weathered sandy or till-like weathered bedrock (gruss) can be more conductive (~10^{-6} m s^{-1}, Maréchal et al. 2004; Schoeneberger and Amoozegar 1990). It is also common that the fractured upper part of bedrock underlying the weathered zone has even higher hydraulic conductivity (10^{-6}–10^{-5} m s^{-1}, Maréchal et al. 2004), and it can behave like an aquifer depending on the fracturing rate and connections between the fractures. Highly simplified single or two-layered groundwater flow models are quick to produce but usually not sufficient, especially in the case of transport modelling—for example, Artimo (2002) found that a two-layer model, one layer with unconsolidated Quaternary deposits and the other with weathered/fractured bedrock, was not adequate for transport modelling. Freeze and Cherry (1979) also illustrated the effect of a high conductivity zone on groundwater flow paths, demonstrating the importance of K variation in transport models.

The ore potential of northern Finland is high, promoting intensive exploration activities (Eilu 2012). Northern Finland has been an area of intensive exploration for gold, nickel, copper and platinum group elements (PGE) since the 1970s (Eilu 2012) and glacial stratigraphy and geochemistry have been methods applied (Hirvas et al. 1977; Gustavsson et al. 1979) in many investigations. Extensive data collection has also enabled the study of Quaternary glaciation history (e.g. Kujansuu 1976; Hirvas et al. 1977; Hirvas 1991), which has revealed the presence of complex sediment arrangements of alternating glacial and interstadial deposits. Due to the close proximity of former ice-divide zones, complex sediment packages are a recurring feature in ore-prospective areas in Finnish Lapland (Hirvas et al. 1991; Helmens et al. 2000; Salonen et al. 2014; Howett et al. 2015; Lunkka et al. 2015; Åberg et al. 2017a). Previous studies have also revealed that weathered bedrock is a prominent feature in Central Lapland (Hirvas 1991; Hall et al. 2015), and has been preserved due to weak glacial erosion (Pulkkinen 1983; Hirvas 1991).

In this study, the Sakatti mining development project in northern Finland (Brownscombe et al. 2015) was used as a study area, since a large amount of detailed information on Quaternary sediments and their hydraulic properties already exists (Hirvas et al. 1977; Gustavsson et al. 1979; Hirvas 1991; Åberg et al. 2017a; Åberg et al. 2019). The project is located in an area of complex hydrostratigraphy consisting of peat, alternating till and sorted sediment units and weathered/fractured bedrock overlying more competent bedrock (Åberg et al. 2017a, 2019). In this mining development project, hydrology and hydrogeology have an important role in environmental impact assessment (EIA), the mine design and infrastructure planning and it has been investigated with three-dimensional (3D) flow modelling. Understanding of the connections between the main hydrostratigraphic units, i.e. peat, glacial and fluvial sediments, weathered/fractured bedrock, competent bedrock and nearby river, is needed. The surficial aquifer-aquitard system consists of relatively thin (ca. 10 m on average) but complex unconsolidated Quaternary deposits. Aquifers that mainly consist of glacigenic and fluvial deposits can be disconnected due to interlayered tills. In addition, underlying fractured bedrock can host a bedrock aquifer, in which connections with the Quaternary aquifer-aquitard system are poorly known.

Recently, a 3D geological model of the Quaternary deposits (Åberg et al. 2017a) and a 3D groundwater flow model (Åberg et al. 2017b, 2019) were constructed for the Sakatti mining development site. In this study, a four times as large 3D geological model was constructed. Based on additional data, a model with more detailed hydraulic properties of sedimentary units and weathered and fractured crystalline bedrock was constructed. Concurrently, a workflow for generating a 3D flow model using Leapfrog Geo (Seequent Ltd. 2020) was described. The groundwater flow model was used to examine the groundwater flow, recharge and discharge patterns of the western part of the Viiankaapaa mine. Detailed understanding of the flow patterns is needed to assess possible groundwater dependency of the ecosystems in mire areas as well as to understand the possible interaction between surface water, shallow groundwater, and water in weathered bedrock or bedrock fractures.

The aim of this study was to analyse how the modelling results for groundwater flow and recharge/discharge patterns differ when simple or more elaborate hydrostratigraphic models of the study site are used. A workflow originally applied by Åberg et al. (2019), which combines 3D modelling in Leapfrog Geo using the hydrogeology tool of Leapfrog and the graphical user interface (GUI) ModelMuse (Winston 2009 USGS) with the flow modelling code MODFLOW-NWT (Niswonger et al. 2011), was described and further developed. The workflow was used to examine how hydrostratigraphic model complexity affects the results. It was hypothesized that a simple hydrostratigraphic structure is inadequate for modelling site-scale groundwater flow and recharge/discharge patterns in complex formerly glaciated sedimentary environment, while greater complexity in a hydrostratigraphic model yields more detailed and plausible groundwater flow patterns. The hypothesis was evaluated with four different versions of hydrostratigraphic models.

This paper discusses the effects of K heterogeneity on site-scale flow patterns, groundwater discharge in mire areas, groundwater flow model calibration, and 3D modelling methods, including simplification of the geological model in the hydrostratigraphic model.

**Geological and hydrological background**

The study area is located in the municipality of Sodankylä in northern Finland (Fig. 1) and lies in the valley of the River
Kitinen, which is surrounded by vast mire areas. The topography within the valley is relatively flat, lying 180–200 m above sea level. The valley is surrounded by inselbergs rising 10–100 m above the flat-lying areas.

The study area is situated in the boreal and subarctic regions of the Northern Hemisphere, with cold winters and cool summers. The annual mean temperature is −0.4 °C and the average annual precipitation varies between 500 and 600 mm (Ilmatieteen laitos, 2018).

The Proterozoic crystalline bedrock within the study area mainly comprises mafic volcanic rocks, gabbros, quartzite and mica schists (Tyrväinen 1980, 1983). The upper part of the crystalline bedrock is highly fractured (10–100 m) and the top of the upper part is partly chemically weathered (Åberg et al., 2019). According to Hall et al. (2015), the weathering indices indicate a variable degree of weathering within the study area. Gruss-type weathering appears to dominate in the study area, while clay-type weathering and weakly weathered areas have patchy occurrences.

The rugged bedrock surface is covered with Quaternary sediments (Åberg et al. 2017a). Multiple gliation cycles have deposited alternating glacial tills and glaciolfluvial and fluvial sediments (Sarala et al. 2015; Åberg et al. 2017a), which have been preserved due to weak glacial erosion (Hirvas 1991; Johansson and Kujansuu 2005) within the study area. In addition, aeolian Holocene/deglacial deposits have been observed in the study area. A diverse sediment package that consists of at least four till units and four sorted deposit units, mainly gravel and sand, creates a highly heterogeneous aquifer-aquitard system in Kärväsniemi (Fig. 1). Comparable sediment packages have also been encountered in the river banks and beneath the mire (Åberg et al. 2017a).

According to Åberg et al. (2019), the annual water budget is positive since the mean annual precipitation rate of 550 mm exceeds the mean annual evaporation rate of 300 mm, and groundwater recharge dominates within the valley of the River Kitinen. Major groundwater recharge occurs from April to September. The highest recharge rate in spring results from the spring thaw and the secondary maximum results from intense rainfall in summer and autumn. The River Kitinen is a gaining river, and the groundwater originates from the surrounding river banks and the vast mire areas.

The study of Åberg et al. (2019) indicated that the hydraulic conductivities of the Quaternary sediments are highly variable, ranging from $2.2 \times 10^{-7}$ to $4.3 \times 10^{-3}$ m s$^{-1}$ (AA Sakatti Mining Oy, unpublished data, 2019; Åberg et al. 2019). Groundwater is mainly unconfined, and scattered groundwater reservoirs exist in the river banks and beneath the Viiankiaapa mire. However, local confined or perched groundwater was observed at least in Kersilöngangs area (Fig. 1), where till may act as a confining unit. The relatively high hydraulic conductivity of the fractured bedrock, from $2.7 \times 10^{-8}$ to $6.5 \times 10^{-5}$ m s$^{-1}$, indicates the presence of a bedrock aquifer (SRK, unpublished data, 2019; AA Sakatti Mining Oy, unpublished data, 2019; Åberg et al. 2019).

Minerogenic Quaternary deposits are overlain by peat deposits. The mire located on the western side of the River Kitinen is called Vantioaapa. The eastern side of the River Kitinen is covered by the Viiankiaapa mire, which hosts endangered groundwater-dependent ecosystems and is a Natura 2000 conservation area (Hjelt and Pääkkö 2006). The Viiankiaapa mire is a patterned mire-fen complex, which started to develop about 10,000 years ago (Suonperä 2016), soon after the Scandinavian Ice Sheet retreated from the study area (Johansson and Kuivansuu 2005; Stroeven et al. 2016). Groundwater discharge in the mire is probably related to the underlying fluvial sediments (Åberg et al. 2019).

Materials and methods

General overview

A workflow was generated starting with 3D geological modeling using Leapfrog Geo (Seequent Ltd. 2020) followed by 3D hydrostratigraphic modeling and groundwater flow modeling using MODFLOW-NWT (Niswonger et al. 2011). In this study, two 3D geological models, GM1 and GM2, were generated using Leapfrog Geo. In addition, four hydrostratigraphic models, HSM1, HSM2, HSM3 and HSM4, were generated from the GM models to test how model complexity affects groundwater flow patterns and the distribution of recharge and discharge zones (Table 1). The structures of HSM1, HSM2 and HSM3 were based on GM1, whereas the structure of HSM4 was based on GM2.

Leapfrog Geo is designed for rapid implicit modelling. The 3D mesh generation in Leapfrog Geo is based on contact surfaces that define the upper and lower limit of the unit. The interpreted stratigraphy, which defines the relative age of the units, affects the behaviour of the defined contacts. Leapfrog Geo offers three types of contact surfaces: sediment contact, erosive contact and intrusion contact (Seequent Ltd. 2020). The contact surfaces of the 3D geological model were generated with the first two types. The FastRBF interpolation method of Leapfrog Geo, which resembles the kriging method (Oliver 1990), was used to generate the contact surfaces.

Geological models GM1 and GM2

The geological models were constructed based on drilling data from the Kersilö database (see Åberg et al. 2017a for the drilling data references) and data obtained from base-of-till (BOT) drilling by AA Sakatti Mining Oy (unpublished data, 2017). In addition, the detailed till geochemistry and bedrock observation databases of the Geological Survey of Finland
Fig. 1 The study area on a surficial deposit map. The locations of cross-sections are presented with letters A–B and W–E. The coordinate system is EUREF-FIN TM35. Surficial deposits are modified and reproduced after the Geological Survey of Finland. The administrative borders of Finland and the gravel intake area are modified after the National Land Survey of Finland. Contours and Natura 2000 National Land Survey of Finland. Administrative borders of Sweden, Norway and Russia, source GADM 2020.
GTK) were used. Moreover, ground penetrating radar (GPR) profiles from Åberg et al. (2017a) were utilized.

The geological models GM1 and GM2 were extended and modified from the geological model presented by Åberg et al. (2017a). The extent of the models was set to $6 \times 8 \text{ km}^2$ (48 km$^2$) to cover the study area presented in Fig. 1. The GM models consisted of ten units (3D solids), including the River Kitinen, peat, four sorted sediment units, four till units, weathered bedrock and fractured bedrock (Table 1). The geological models were constructed with both explicit and implicit modelling using the methods of Åberg et al. (2017a). The horizontal resolution of both GM models was $20 \times 20 \text{ m}$, while the vertical resolution was not applicable in Leapfrog. The River Kitinen was modelled as an erosive contact. However, in GM2, the river bottom was slightly edited compared to GM1 based on the bedrock observations database of GTK.

The peat bottom contact was generated with ArcMap (ESRI) as depth contours. Next, a triangulated irregular network (TIN) surface was generated from the peat contours with ArcMap. After this, the TIN surface that depicted the peat bottom depth was converted into a raster, and the raster was converted to the altitude above sea level by subtracting it from the topography. The converted raster was used as an erosion-type contact defining the peat bottom in a geological model in Leapfrog. However, the peat bottom contact in GM2 was reedited based on new GPR profiles in spring 2019. During contour generation, 0.5-m-thick contours were created 2 m inwards from the border of the mire areas to prevent the peat close to the borders from being too thin. In addition, areas with no peat were defined with 20-m-thick contours 5 m outwards from the border of the mire areas. A random negative height rising above the topography was selected.

Minerogenic Quaternary sediment units (presented in Table 1) were modelled in an explicit manner as sediment contacts with the polyline tool. In GM2, revisions to the Quaternary sediment units were made compared to GM1, including changes to the layer structure of the Sahankangas area (Fig. 1) and expansion of the lowest till unit (Fig. 2).

The bedrock surface contact for both GM models was generated as an erosive contact with implicit modelling. It was adjusted with the polyline tool in those areas where drilling did not reach the bedrock using the methods described in Åberg et al. (2017a).

In both GM models, the bedrock was modelled with two layers: the upper depicted the chemically weathered zone, whereas the lower zone depicted fractured bedrock, as in Durand et al. (2017) and Wright (1992), among others. In GM1, chemically weathered bedrock was modelled with a deposit-type contact above the fractured bedrock surface, as in Åberg et al. (2017a). In GM2, a more precise and extensive chemically weathered zone was generated based on Hall et al. (2015), Raatikainen (2019), AA Sakatti Mining Oy, unpublished data, 2019; and GPR surveys in 2015. The weathering zone was extended to cover the whole model area (Fig. 2), contrary to the restricted coverage of GM1. It was generated with Leapfrog as an offset surface from the bedrock surface with a minimum thickness of 3 m. A few dummy points and polylines were added to offset the surface to manually improve the interpolated bottom contact. In GM1, everything below the bedrock surface was assumed to be fractured bedrock, and the bottom was set to 140 m asl. In GM2, the bottom of the fractured zone was generated as an offset surface from the bedrock surface with a minimum thickness of 5 m (Fig. 2). The offset surface was also adjusted with a few dummy points and polylines. The fractured zone delineated the base of the model.

**Table 1** The hydrostratigraphic units used in 3D geological models (Fig. 2) and 3D groundwater flow models

| Modelled geological unit | GM 1–2 | HSM1 | HSM2 | HSM3 | HSM4 |
|--------------------------|--------|------|------|------|------|
| River Kitinen            | x      | x    | x    | x    | x    |
| Peat                     | x      | x    | x    | x    | x    |
| Top deposits             | x      | –    | –    | x    | x    |
| Upper till               | x      | –    | –    | x    | x    |
| Middle sorted deposits   | x      | –    | –    | x    | x    |
| Middle till              | x      | –    | –    | –    | –    |
| Lower sorted deposits    | x      | –    | –    | –    | –    |
| Lower till assemblage    | x      | –    | –    | x    | x    |
| Lowest sands and gravels | x      | –    | –    | x    | x    |
| Lowest till              | x      | –    | –    | –    | x    |
| Weathered bedrock        | x      | –    | –    | –    | x    |
| Bedrock                  | x      | x    | x    | x    | x    |

In *HSM1* and *HSM2*, the Quaternary deposit layers, except peat, were combined into one layer.

**Hydrostratigraphic HSM models**

Model complexity increased from HSM1 to HSM4 (Table 2). HSM1 had a uniform Quaternary sediment package and uniform fractured bedrock, whereas HSM2 had a uniform Quaternary sediment package but variable fractured bedrock. HSM3 had multi-layered Quaternary sediments and a single layer of variable fractured bedrock, whereas HSM4 had multi-
Fig. 2  An exploded 3D view and a 3D view of GM2. The Quaternary sediments and the weathered and fractured bedrock were reedited compared to GM1 (see Figs. 3 and 4)
layered Quaternary sediments and two-layered variable bedrock.

HSM1 and HSM2 were simple models that consisted of four layers including the units the River Kitinen, peat, Quaternary sediments and fractured bedrock. Simplification from the geological model to the hydrostratigraphic model was performed in Leapfrog by removing contact surfaces between the sediment units (Table 1).

In the complex models, HSM3 and HSM4, most of the units from the GM models were retained (Tables 1, 3 and 4) and the Quaternary sediments consisted of several units, contrary to one combined unit in HSM1 and HSM2. HSM3 had eight layers, whereas HSM4 had ten layers including the units in Table 1. In HSM3, the middle till, lowest till and weathered bedrock units were removed due to their originally assumed small coverage, and the lower sorted and middle sorted units were combined into one unit. In HSM4, middle sorted deposits and lower sorted deposits were likewise combined into one unit, and middle till was removed from the model. Contrary to the construction of HSM3, the weathered bedrock and lowest till units were retained in HSM4 due to the expansion of their coverage.

### Hydrogeology tool and conversion to a groundwater flow model grid

The hydrogeology tool of Leapfrog was used to convert the 3D hydrostratigraphic models into MODFLOW grids (Figs. 3 and 4). The units of the 3D hydrostratigraphic model were projected within the grid using the evaluation tool of Leapfrog. The resolution of the horizontal cell size was $20 \times 20$ m, and the minimum allowed cell size was 0.5 m for vertical resolution (Fig. 4c), except in HSM4, in which it was 0.0017 m (Fig. 4d). This minimum thickness appeared between the scattered units as “pinched” cells (e.g. Feinstein et al. 2012), while the basic MODFLOW codes only allow continuous layers. The bases of the MODFLOW grids were edited in HSM1–3 to reach 35 m (assumed average thickness of the fractured bedrock) (Fig. 4c) downwards from the bedrock surface, instead of 140 m as in GM1 (as in Fig. 4a). In HSM4, the irregularly elevated base of the model (Fig. 4b,d), corresponding to the solid bedrock surface, was used as the base extent as such. The river top was reduced to 0.5 m above the bottom of the river in all HSM models, as described in Åberg et al. (2019). Layer 1 in the mire area was modified in complex models HSM3 and HSM4 to represent two-thirds of the mire thickness, as presented in Åberg et al. (2019), to correspond to the mire acrotelm. Layer 2 in the mire area corresponded to the mire catotelm. Acrotelm is the less decomposed upper part of peat with higher hydraulic conductivity and catotelm is more decomposed peat below the acrotelm (Holden and Burt 2003). In the simple models HSM1 and HSM2, layer 1 was unmodified from GM1 and peat was uniform.

### Parametrization and boundary conditions of the groundwater flow models

The 3D units of the HSM models were used as zones for horizontal hydraulic conductivity (HK) parameters and vertical anisotropy (VANI) parameters in ModelMuse. The HK parameters are shown in Table 2.
and VANI parameters of the HSM models were based on earlier slug test and grain-size analyses, as described in Åberg et al. (2019), and more recent grain-size analyses (SRK, unpublished data, 2019; AA Sakatti Mining Oy, unpublished data, 2019). The HSM models were set as steady-state, which was assumed to be adequate for the generic understanding of groundwater flow patterns since annual water-table variation in most continuous monitoring wells is less than 1 m (Åberg et al. 2019).

The precipitation rate was 560 mm year\(^{-1}\) for the hydrological year lasting from November 2013 to November 2014 to correspond to a single snow-covered period as in Åberg et al. (2019), and this configuration was used for recharge estimations. The recharge rate was interpolated from recharge estimates for two zones called “open water and mire area” and “forested area”, as described in Åberg et al. (2019). The recharge rate of the mire area was assumed to be uniform for simplicity. The recharge rates of the remaining areas were interpolated with multiplier points, as described in Åberg et al. (2019), based on the results of the episodic master recession (EMR; Nimmo et al. 2015) and water-table fluctuation (WTF; Meinzer 1923) methods. All the models were surrounded by constant head boundaries in locations where the mire contacted the model border. Head-dependent flux boundaries were added in sediment boundary areas with the general head boundary (GHB) package. Their heads were calculated from the closest open water bodies, and conductance was defined as in Åberg et al. (2019). The drain package (DRN) was used to simulate discharge within the open mire area and the gravel intake area (Fig. 1). The River Kitinen was modelled with the river package (RIV) to simulate the river discharge.

### Definition of fractured and weathered bedrock zones in the HSM models

In the simplest model, HSM1, the fractured bedrock was defined with a uniform initial hydraulic conductivity of \(5.3 \times 10^{-6}\) m s\(^{-1}\). In models HSM2, HSM3 and HSM4, the fractured bedrock was defined similarly as in Åberg et al. (2019). The hydraulic conductivity of the fractured bedrock was interpolated from hydrogeological testing in boreholes, including packer, spinner and constant head tests (SRK, unpublished data, 2019), along with two slug tests (Golder Associates, unpublished data, 2012). The thrust faults and brittle faults

| Unit                          | Thickness (m): mean, SD, min–max | Volume (m\(^3\)) | % of surficial deposits/bedrock |
|-------------------------------|----------------------------------|------------------|--------------------------------|
| Peat                          | 2.0, 1.0, 0–6.8                  | 5.12 \times 10^7 | 13.6                           |
| Sorted sediments              | 2.0, 1.9, 0–20.5                 | 1.30 \times 10^8 | 34.7                           |
| Tills                         | 2.3, 1.9, 0–38.5                 | 1.94 \times 10^8 | 51.7                           |
| Quaternary sediments, all     | 7.8, 4.4, 0–48.3                 | 3.75 \times 10^6 | 100.0                          |
| Weathered bedrock             | 6.0, 2.6, 0–41.3                 | 2.87 \times 10^6 | 19.0                           |
| Fractured bedrock             | 26.0, 8.2, 0.1–77.9              | 1.24 \times 10^9 | 81.0                           |

Fig. 3 A MODFLOW grid generated (HSM3) from the simplified geological model GM1 with Leapfrog. The Leapfrog evaluation tool defined the parameter zones. The colour scheme of the zones is depicted in Fig. 2. Vertical exaggeration of the model is 10x.
Fig. 4 Cross-sections of the Leapfrog models vs. MODFLOW grids from complex models HSM3 and HSM4 with initial hydraulic conductivities. The location of cross-section W–E is depicted in Fig. 1. a Cross-section of geological model GM1 used as a basis for HSM3. Everything below the bedrock surface was assumed to be fractured bedrock, and the bottom was set to 140 m asl. b Cross-section of GM2 used as a basis for HSM4. c Cross-section of model HSM3 presented as a modified MODFLOW grid in ModelMuse. d Cross-section of model HSM4 presented as a modified MODFLOW grid in ModelMuse. The thin “pinched” parts connect the scattered layers as a whole. The parameter zones of the HSM models were defined with the Leapfrog evaluation tool. The river was modelled as a separate unit, as it penetrated multiple layers.
including fractures were defined as their own parameter zones (Table 2).

In HSM4, the type of weathered bedrock zone was defined as gruss type, clay type or nonweathered based on the Weathering Index of Parker (WIP, Parker 1970; Hall et al. 2015). The clay and the gruss zones were delineated based on WIP calculations of till geochemistry data (weathered bedrock samples and fine fraction of till) in Hall et al. (2015) and Raatikainen (2019). Gruss and clay-zone polygons were drawn and used as parameter zones (Fig. 5b). The remaining least weathered areas were included in the same parameter zone as fractured bedrock.

**Calibration of the HSM groundwater flow models**

A total of 165 observations, including head observations from groundwater monitoring wells (n = 34), water-table measurements from flarks and other open water bodies (n = 126) on the LiDAR DEM (light detection and ranging digital elevation model) and spring observations (n = 5), were used as the calibration points in all the HSM models. The head observations of the wells were added to the layer that corresponded the depth of the screen of the well (Table 6). The observations from LiDAR DEM were added to layers 1 or 2 depending on the expected water table. The hydraulic conductivities were first manually defined with interpolated HK parameter multipliers in ModelMuse, as presented in Åberg et al. (2019). Then, the hydraulic conductivity of the peat acrotelm, thrust faults, brittle faults and fractures, the gruss zone and clay zone, and RIV and DRN conductance were calibrated with PEST (Doherty 2015), depending on the model version (see Table 5). Parameters that had low sensitivity or indicated ill-posedness were set as fixed parameters. The variation of Marquardt lambda and singular value decomposition were enabled during parameter estimation, which helped to find the best improvement of the objective function of the head observation (see Doherty 2015). The limits for parameter values were set as follows: the maximum allowed value was 0.01 for all calibrated parameters except brittle faults and fractures, which had a maximum of 1, and the minimum allowed value was 1 × 10⁻10 for all parameters to maintain the convergence of the model. The unit was m s⁻¹ or m² s⁻¹ depending on the parameter. Starting values for RIV and DRN were 1 × 10⁻⁵ and 1 × 10⁻⁶ m² s⁻¹, respectively, based on test calibrations. The starting K of peat, Quaternary sediments and fractured bedrock were 5.4 × 10⁻⁵, 1 × 10⁻⁵ and 5.3 × 10⁻⁶ m s⁻¹ (Table 5), respectively, based on the definition in Åberg et al. (2019). Starting K for gruss and clay were 1 × 10⁻⁵ and 1 × 10⁻⁸ m s⁻¹ (only used in HSM4), respectively. The initial VANI parameters were defined as in Åberg et al. (2019) as 10 for Quaternary deposits, 100 for peat and 1 for all bedrock parameters. The VANI parameters remained fixed due to their low sensitivities (see e.g. Hill 1998; Poeter et al. 2014). Root mean square error (RMSE), mean absolute error (MAE) and bias were calculated from residuals of the head observations. Linear uncertainty was calculated for calibrated parameters with PEST. Since all calibrated parameters used were log-transformed, the uncertainty variances were calculated to logarithms of parameters (Doherty 2018).

**Verification and validation of groundwater flow models**

Model evaluation was performed with particle tracking with measured mean δ¹⁸O, δD and d-excess isotope values of the water samples in the Viiankaapa mire based on data presented in Bigler (2019) and measured in September or October 2016. Particle tracking was performed in forward and backward directions with MODPATH (Pollock 1994) to describe the discharge and recharge locations as well as flow routes of the water having certain isotope composition. Isotope compositions were assumed to reflect the evaporation state of water in the recharge area. Neither altitude effect (altitude varies from 180 to 200 m) nor groundwater transport were supposed to affect the isotope values (Gat 2010). The δ¹⁸O, δD and d-excess values were inserted to modelled flow paths with ArcMap to study the flow paths of different isotope compositions in 3D with Leapfrog. Moreover, the isotope values of the particle flow paths were compared to measured isotope values of springs on the shore of the River Kitinen presented in Åberg et al. (2019) measured in August 2015. Firstly, the isotope values of the springs were grouped into five groups based on their locations. Secondly, the flow paths close to the defined five groups were selected and studied in 3D view. Then, the fit between isotope values of the selected flow paths and spring observations were calculated and defined based on the following criteria: the deviation from spring observation is ±1, ±5, and ±1.5% in δ¹⁸O, δD and d-excess values, respectively. The difference is assumed to be acceptable if the value is within one-sixth part of the ranges −14.8 to −9.9, −108.8 to −77.0, and −4.4 to 8.6‰ in δ¹⁸O, δD and d-excess values, respectively. If all isotope values of flow paths coincided with the spring observations within acceptable limits, the fit of the group was good. If only some of the spring values coincided with flow path values, the fit of the group was moderate. If all flow path values deviated from spring observations, the fit was poor. If no particle paths crossed the spring observation, the group was categorised as “no-match”. To evaluate the models, overall points were calculated to HSM models with values 2, 1, 0 and empty for good, moderate, poor and no-match fits of the groups, respectively. The final decision of the group match was based on the fit of d-excess, which has a less seasonal variation (Gat 2010).
Results

General distribution of Quaternary sediments and fractured/weathered bedrock

Based on the geological and hydrostratigraphic models, the study area consists of 51.7–59.0% till, 26.7–34.7% sorted sediments (silt, sand or gravel) and 13.6–14.4% peat, and their volumes are ca. 0.2, 0.1 and 0.05 km³, respectively. The Quaternary sediment package is 8 m thick on average, and it consists of till, sorted sediment and peat units that are 2 m thick on average (Tables 3 and 4). The altitude of the bedrock topography varies from 135 to 206 m asl, with an average altitude of 180 m asl (Fig. 5). According to the most complex models, GM2 and HSM4, the thickness of the fractured bedrock is highly variable (0–77.9 m), with the average thickness being 26.0 m. The average thickness of weathered bedrock is 6 m.
The units of RMSE, MAE and bias are in metres. Cal denotes calibrated models. Calibrated parameters are shown in italic. K and the RCH1 multiplier (recharge) are in units of m s\(^{-1}\). All K represent horizontal hydraulic conductivity. Thrust faults, brittle faults and fractures are presented in Fig. 5. Conductance in RIV and DRN has units of m\(^2\) s\(^{-1}\). The Conduct factor that determines the volume on transmitting interface of boundary condition (see Harbaugh 2005) was assumed to be 400 in RIV and DRN conductance in all models due to the grid size of 20 × 20 m. The actual conductance value is the product of the parameter value and the Conduct factor.

**Groundwater flow budget and recharge/discharge patterns**

Flow modelling budgets (see Appendix for the second through fifth tables therein) after calibration indicated that the percentage of groundwater discharge to the River Kitten was a lower proportion of the overall outflow (21.8–29.9%) in the simple models HSM1 and HSM2 than in the complex models HSM3 and HSM4 (34.7–35.5%). However, the percentage of mire discharge in the overall outflow appeared to be higher in the simple models (65.7–74.6%) than the complex models (57.2–59.1%).

In the simple models HSM1 and HSM2, the recharge rate was estimated to be 25% of the precipitation in forested areas. According to the complex models HSM3 and HSM4, the recharge rate in forested areas varied from 28 mm year\(^{-1}\) (5% of precipitation) to 380 mm year\(^{-1}\) (68% of precipitation; Fig. 6). In the mire area, a constant recharge rate of 100 mm year\(^{-1}\) (19% of precipitation) was used in all HSM models based on Åberg et al. (2019). The calibrated hydraulic conductivity of fractured bedrock was 1.7 \times 10^{-6} m s\(^{-1}\) in HSM1, whereas K of bedrock units varied from 8.4 \times 10^{-9} to 1.3 \times 10^{-3} m s\(^{-1}\) in models HSM2, HSM3 and HSM4 (Table 5). In HSM4, the calibrated hydraulic conductivity for gruss and clay zones was 2.2 \times 10^{-7} and 8.4 \times 10^{-9} m s\(^{-1}\), respectively. The hydraulic conductivity of tills varied from 5.0 \times 10^{-8} to 1.3 \times 10^{-4} m s\(^{-1}\) and that of sorted sediments from 1.2 \times 10^{-6} to 6.1 \times 10^{-4} m s\(^{-1}\) in models HSM3 and HSM4 (Table 11).

The root mean square error (RMSE) of the calibrated HSM models varied from 0.38 to 0.33 m, and the residual errors between the simulated and observed water table were within ±1.6 m (Fig. 7) with Pearson correlation coefficient (SPSS ver. 24) of 0.995–0.997 (p value <0.01). The RMSE appeared to decrease when model complexity increased from model HSM1 → HSM2 → HSM3. Overall MAE and bias were similar in all models; however, groundwater wells in Table 12 indicated a notable decrease of RMSE, MAE and bias from HSM1 to HSM4.

The calculated uncertainties with 95% confidence limits indicated that parameters brittle faults and Quaternary sediments (simple models) have similar confidence limits and uncertainty variance (0.05–0.1) regardless of the model. However, some parameters like thrust fault and river parameter had different variance. In HSM3, the variance of the thrust fault was smaller (~0.5) than in HSM2 and HSM4 (>1.5). In HSM 2, the variance of the river parameter was two compared to the smaller variance of ~0.5 in other models. Brittle fault seemed to have the lowest variance in HSM2–4 (<0.1) indicating higher certainty. The variance in drain parameter was high (1.3–1.4) in both simple models. Gruss parameter had

### Table 5 Parameters and RMSE, MAE and bias of the HSM models

| Parameter | HSM1 | HSM1 cal | HSM2 | HSM2 cal | HSM3 | HSM3 cal | HSM4 | HSM4 cal |
|-----------|------|---------|------|----------|------|----------|------|---------|
| RMSE      | 0.40 | 0.38    | 0.37 | 0.36      | 0.35 | 0.37      | 0.34 | 0.37    |
| MAE       | 0.22 | 0.22    | 0.22 | 0.22      | 0.23 | 0.22      | 0.24 | 0.23    |
| Bias      | −0.08| −0.04   | −0.07| −0.03     | −0.06| −0.04     | −0.08| −0.05   |
| Peat layer 1 K\(^a\) | 5.4 \times 10^{-5} | 2.8 \times 10^{-4} | 5.4 \times 10^{-5} | 5.9 \times 10^{-5} | 5.4 \times 10^{-5} | 7.7 \times 10^{-5} | 5.4 \times 10^{-5} | 8.8 \times 10^{-5} |
| Peat layer 2 K\(^b\) | 5.4 \times 10^{-5} | 2.8 \times 10^{-4} | 5.4 \times 10^{-5} | 5.9 \times 10^{-5} | 1.0 \times 10^{-6} | 1.0 \times 10^{-6} | 1.0 \times 10^{-6} | 1.0 \times 10^{-6} |
| Quaternary sediments K | 1.0 \times 10^{-5} | 3.2 \times 10^{-5} | 1.0 \times 10^{-5} | 1.5 \times 10^{-5} | 5.0 \times 10^{-6} | 5.0 \times 10^{-6} | 5.0 \times 10^{-6} | 5.0 \times 10^{-6} |
| Gruss type weathered bedrock K | – | – | – | – | – | – | – | – |
| Clay type weathered bedrock K | – | – | – | – | – | – | 1.0 \times 10^{-5} | 2.2 \times 10^{-5} |
| Fractured bedrock K | 5.3 \times 10^{-6} | 1.7 \times 10^{-6} | 2.7 \times 10^{-8} | 6.5 \times 10^{-5} | 2.7 \times 10^{-8} | 6.5 \times 10^{-5} | 2.7 \times 10^{-8} | 6.5 \times 10^{-5} |
| Brittle faults and fractures K | – | – | 1.0 \times 10^{-3} | – | 1.0 \times 10^{-3} | – | 1.5 \times 10^{-5} | 1.0 \times 10^{-3} |
| Thrust faults K | – | – | 1.0 \times 10^{-5} | 1.2 \times 10^{-5} | 1.0 \times 10^{-5} | 3.9 \times 10^{-4} | 1.0 \times 10^{-5} | 1.3 \times 10^{-3} |
| RIV | 1.0 \times 10^{-5} | 4.4 \times 10^{-7} | 1.0 \times 10^{-5} | 1.0 \times 10^{-5} | 1.6 \times 10^{-6} | 1.0 \times 10^{-5} | 7.1 \times 10^{-7} |
| DRN | 1.0 \times 10^{-5} | 2.7 \times 10^{-6} | 1.0 \times 10^{-6} | 2.8 \times 10^{-6} | 1.0 \times 10^{-6} | 1.0 \times 10^{-6} | 1.0 \times 10^{-6} | 1.0 \times 10^{-6} |
| RCH1 multiplier | 0.25 | 0.25 | 0.25 | 0.25 | 0.05–0.68 | 0.05–0.68 | 0.05–0.68 | 0.05–0.68 |

The variance in drain parameter was 2.2 \times 10^{-5} ms\(^{-1}\), which has units of m \times s\(^{-1}\). All models have similar confidence limits and uncertainty variance (0.05–0.1) regardless of the model. However, some parameters like thrust fault and river parameter had different variance. In HSM3, the variance of the thrust fault was smaller (~0.5) than in HSM2 and HSM4 (>1.5). In HSM 2, the variance of the river parameter was two compared to the smaller variance of ~0.5 in other models. Brittle fault seemed to have the lowest variance in HSM2–4 (<0.1) indicating higher certainty. The variance in drain parameter was high (1.3–1.4) in both simple models. Gruss parameter had

\( a \) Acrotelm in models HSM3 and HSM4

\( b \) Catotelm in models HSM3 and HSM4
low variance (0.1) and clay had high variance (1.7). Peat parameter had high (>1.2) variance in all models except HSM1 with the variance of 0.4.

**Groundwater flow patterns**

The flow model results indicated that groundwater mainly flows towards the River Kitinen (Fig. 7). In addition, horizontal flow was dominant in the simple HSM models (Fig. 8a,b), while the downward and upward flow component increased in the complex HSM models (Fig. 8c,d).

The variation in hydraulic conductivity appeared to be an important factor determining groundwater flow patterns. When a high-conductivity zone, e.g. a sand unit, lay below the peat catotelm, vertical flow areas in peat layers expanded. Subtill sorted sediments appeared to cause a similar effect.
Fig. 7 The residual of the observed minus the simulated water table and the simulated water-table contours in calibrated HSM models (a–d). Negative values indicate that the water table is higher compared to the observation and positive values indicate that the water table is lower, respectively. Location points A, B and C indicate visible differences in residuals. The LiDAR DEM is modified and reproduced after the National Land Survey of Finland.
Moreover, the low hydraulic conductivity of the clay type in weathered bedrock caused a partial increase in the hydraulic gradient in HSM4 (Fig. 7 location C). In addition, the existence of low hydraulic conductivity fractures or faults increased the altitude of the water table (see e.g. Bense et al. 2003).

Flow model results indicated that groundwater discharge appeared to occur in river shoreline and mire areas (Fig. 9). The discharge areas were interpreted from drain cells or flooding cells, as presented in Åberg et al. (2019). The complex models provided more evenly distributed groundwater discharge patterns, reflecting the string and flark pattern of the mire (Fig. 9, location E). The Matarakoski hydroelectric power plant created a recharge area in the river due to the higher river stage compared to the water table on the surrounding shore areas (Fig. 9, location F).

Particle tracking indicated that the groundwater flow in surficial deposits occurred mainly in high hydraulic conductivity zones. The K of brittle faults and fractures seem to affect the particle paths indicating their importance to model results in HSM2–4. The deepest flow paths also seemed to mimic the model bottom indicating the importance of the topography of more competent bedrock, which was used as the model bottom. Moreover, particle tracking indicated groundwater recharge areas in river banks and in the Viiankiaapa mire, as well (Fig. 10).

Discussion

Site-scale flow patterns

Flow patterns vs. complexity

In the simplest model, HSM1, the small variation in hydraulic conductivity between units led to the dominance of horizontal flow (Fig. 8a). However, the more prominent variation in hydraulic conductivity created complexity in flow patterns due to the refraction of groundwater flow lines in locations of conductivity change (see Freeze and Cherry 1979), which can be seen in the more advanced models HSM2–4. The refraction of flow lines due to hydraulic conductivity variation also appears to be a factor that affects the occurrence of discharge areas (Freeze and Witherspoon 1967). High variation leads to higher refraction, increasing the vertical flow component.

In addition to refraction, the order of high and low conductivity layers appears to affect the site-scale flow patterns. Horizontal flow seemed to dominate when lower conductivity bedrock appeared beneath higher conductivity surficial deposits in HSM1 and HSM2 (Fig. 8a,b). The dominance of horizontal flow when a lower conductivity unit (e.g. till) was overlain by a high conductivity zone was already observed by Freeze and Witherspoon (1967). However, if the hydraulic conductivity of the bedrock was higher than that of Quaternary deposits, as in HSM2, the vertical flow component slightly increased (Fig. 8b, location D). The expansion in vertical flow in a mire due to high conductivity subpeat sediments was also observed by Reeve et al. (2000). Based on particle tracking the recharge areas seem to be quite similar in all models. However, modifications in high K units affected recharge areas of the flow paths, as seen in HSM4 in the southeastern corner in Fig. 10d. Simultaneously, the discharge locations vary in the shore of the River Kitinen and the mire, as well.

Groundwater discharge in mires

The mire discharge areas in the simple models HSM1 and HSM2 were tilted compared to the complex models HSM3 and HSM4, in which discharge zones matched the flark and string patterns more closely (Fig. 9). As higher variation in hydraulic conductivity increased the upward component in flow (Fig. 8c,d), discharge was more widespread in the complex models (Fig. 11). However, the lower percentage of discharge (see Appendix for the second through fifth tables therein) in DRN in the complex models appears to be related to the lower hydraulic conductivity of the peat catotelm, which limits the groundwater discharge (Peat layer 2 K in Table 5). Even though model HSM1 had the most restricted distribution of discharge (Fig. 9) in the mire area, it had the highest discharge rate due to highest overall K in peat (peat layer 1 K in Table 5).

Åberg et al. (2019) used the “flow lower” dataset of MODFLOW from a cell-by-cell output file (see Harbaugh 2005; Niswonger et al. 2011), which describes groundwater in or outflow in every cell bottom, as a definition of groundwater discharge. The flow lower dataset indicated similar patterns to those in the DRN outflow cells in every model: upward flow (presented as a negative value) in flow lower datasets in layer 2 (peat bottom) resembled the areas of DRN outflow in the mire (Fig. 11). This means that water flowing upward from sediment discharges out of the model in flarks if the head rises above the model surface. However, in the simple models, the water table was observed to be tilted in flarks, and the discharge rate was two orders of magnitude higher in an upstream direction compared to a downstream direction (Fig. 9, location G). In contrast, in the complex models,
the water table in flarks was smoother. The observed difference can be explained due to the lower refraction of flow lines and more horizontal flow direction in peat layers due to lower variation in hydraulic conductivity between the peat and the sediment beneath it. These findings indicate that the hydraulic conductivity of peat is the limiting factor for groundwater discharge. It is apparent that DRN may overestimate the discharge rate, as it assumes that all water that rises above the water table in DRN cells is discharge. Moreover, using DRN in open mires has its limitations, as discussed in Åberg et al. (2019) and in more detail in Batelaan and De Smedt (1998).

Groundwater flow model fit vs. model reality

The overall slight decrease of RMSE from HSM1 to HSM3 was not statistically significant. In contrast, the difference in...
Fig. 9 Groundwater discharge areas defined with DRN outflow cells and the river leakage dataset in the calibrated HSM models (a–d). *Presented as \(-\log_{10}\) to distinguish recharge (reddish colours) from discharge (greenish colours). The arrow at location G indicates the groundwater flow direction. F indicates the discharge area formed due to the Matarakoski hydroelectric power plant. Errors in the LiDAR DEM altitude are also visible in the lake area and horizontal unconformity surfaces (X). The River Kitinen (RIV cells) and mire borders (DRN cells) are modified and reproduced after the National Land Survey of Finland.
Fig. 10 3D particle tracking in calibrated HSM models in 3D views (a–d) and cross-sections (e–h). The range of each cross-section is 500 m towards the northeast and the southeast, and the length is ca. 4,200 m. Balls and flow paths coloured with $d$-excess values present the starting point of forward and backward tracking. Recharge point indicates an ending point of backward tracking and discharge points an ending point of forward tracking.
RMSE and MAE of groundwater well heads (Table 12) was statistically significant ($\rho < 0.05$) between simple and complex models according to the Mann-Whitney U test, indicating that model fit increases when more complex multi-layered models are used. However, the differences in residuals between models HSM3 and HSM4 can be related to the local structural revisions rather than the overall increase in complexity. Further changes to the structures of bedrock and surficial deposits between HSM3 and HSM4 indicate that only relying on obtaining a good fit is not adequate in model development. HSM4 was the only model with a weathered bedrock unit, and it had a higher overall RMSE than HSM3, indicating that model complexity also added more uncertainties to the structure. In the Kokkolampi area (Fig. 7), indicating that model complexity depends on the research questions.

The areas of the groundwater flow model with a poor fit could indicate defects in the layer structure, the definition of hydraulic conductivity, the recharge rate, or the existence of perched groundwater—for example, the residuals of heads indicated that a constant area of poor fit exists in the Kärväslampi area in all models (Fig. 7, location A). Moreover, stable $\delta^{18}$O isotope studies (Åberg et al. 2019) indicated that Kärväslampi probably represents a perched water table. These differences in residuals between models indicate that the calibration results could be used for model development by identifying the poorly fitting areas needing reconsideration of the hydrostratigraphic/geological structure or the recharge rate.

In groundwater flow modelling, flooding is usually considered as a model defect; however, in some cases, for example close to river cells (Feinstein et al. 2012), it can be considered as a simplification of groundwater discharge. In the complex models, low hydraulic conductivity units caused unrealistic flooding with moderate slopes, whereas in the simple models, unrealistic flooding was present when a smooth or high water table crossed the topography. Flooding also existed in areas where low conductivity features that were probably too thick appeared close to the model top, simultaneously with an overall high water table. In addition, if the estimated recharge is too high in relation to hydraulic conductivity, this can cause flooding. While the definitions of HK and the recharge multiplier were based on interpolation, some unexpected combinations might appear with overall low conductivity and too high recharge. Moreover, as noticed from the flow patterns, ignoring the clay zone (in HSM1–3) causes unrealistic connections between surficial deposits and the bedrock in clay weathering areas (Fig. 5b), while the refraction of flow lines and the limitation of infiltration due to a low conductivity zone are omitted.

In water balance studies, simple models can give as reasonable results as complex models (e.g. Hudon-Gagnon et al. 2015); however, HSM models indicated that details are important in studies related to recharge/discharge patterns since too simple models led to the underestimation of the vertical flow component and $K$ variation between layers. Especially in mires, locations of groundwater discharge are notably affected by vertical flow component; thus, the importance of model complexity depends on the research questions.

**Effect of hydraulic conductivities in the HSM models**

**Effects of low conductivity zones and transmissivity**

The variations in hydraulic conductivity were found to be an important factor in site-scale flow patterns. The low hydraulic conductivity units such as lens-like till units, clay-type weathering zones or bedrock elevations, were observed to behave similarly to low hydraulic conductivity faults or fractures. The existence of a low conductivity zone created a barrier to groundwater flow, raising the water table in the upstream direction and lowering it in the downstream direction. However, the water table already started to lower in the middle of the blocking unit due to the steepened hydraulic gradient; thus, the dimension of the lens-like low conductivity zone appears to be the important factor in complex aquifer-aquitard systems.

During model development, manual sensitivity analysis (see Reilly and Harbaugh 2004) indicated that changing the hydraulic conductivity by two orders of magnitude in sediment layers changed the water level by about 1 m, whereas a one order of magnitude change only had a minor effect. In addition to hydraulic conductivity, transmissivity depends on the layer thickness, which varied between 0 and 50 m for Quaternary sediments and between 0 and 80 m for bedrock, causing prominent variation in internal transmissivity. Differences in flow patterns between HMS3 and HSM4 could be due to variation in the layer thicknesses.

**Effect of interlayered tills on recharge**

While variation in hydraulic conductivities existed in the complex models HMS3 and HSM4, the recharge could be estimated in more detail. Moreover, the layer structure should be known when adjusting the recharge rates, as an impeding layer between sorted sediments may reduce the infiltration rate. The existence of interlayered tills (Åberg et al. 2017a) causes a reduction in the recharge rate. The upper till represented in the models was mainly
Brittle faults and fractures appear to affect the river discharge in location I. The River Kitinen (RIV cells) and mire borders (DRN cells) are modified and reproduced after the National Land Survey of Finland.

Fig. 11 Flow lower dataset in layer 2 presenting the upward flow cells (originally negative value) in calibrated models HSM1–4 (a–d). In the mire area, coloured cells indicate flow from sediment to the peat layer.
Within the study area, the bedrock is partially highly fractured to depths of 30–100 m and can be considered as an aquifer or aquitard (Åberg et al. 2019; Salminen 1975; Lestinen 1980; Hall et al. 2015). Two combinations of weathering profiles have been observed in Finnish Lapland: tripartite clay-gruss-saprock and bipartite gruss-saprock (Hall et al. 2015). The gruss-saprock type dominates in the study area, but the clay type also exists (Fig. 5b). In HSM4, a simplification was made concerning the clay type: the existence of plausible gruss-type weathering beneath the clay type was omitted due to its unknown distribution. This simplification possibly led to an overestimation of the clays weathering zone thickness, which was found to be the reason for flooding in the NW parts, including the Kersilö and Vanttioaapa areas (Fig. 7, location C).

Modelling of the weathered bedrock zone could be improved by generating separate gruss and clay units with Leapfrog. However, according to the drilling data (AA Sakatti Mining Oy, unpublished data, 2019), thicknesses of weathered and fractured bedrock zones are highly variable within short distances (<50–100 m) and interbedded fresh bedrock exists within fractured media (Fig. 5b), causing challenges in interpolation.

Hydraulic connection between surficial deposits and weathered/fractured bedrock

Based on Åberg et al. (2019) and model results of this study, there is groundwater discharge from sediments and fractured bedrock into the river; however, the existence, distribution and magnitude of a hydraulic connection between groundwater in Quaternary sediments and the bedrock is still somewhat an open question. According to modelling, there is groundwater recharge in the mire. This can be seen in stable isotope composition ($\delta^{18}O, \delta D$) of groundwater and surface waters, as well (Åberg et al. 2019; J. Karhu et al., University of Helsinki, unpublished report, 2020). The evaporated component in groundwater was observed in sediments and upper bedrock in an area located in western part of Viankiapa mire, Pahanlaaksonmaa and Kersiliönkangas areas. Based on flow modelling, the evaporated component in groundwater originates from the open water in the mire area. The thickness and composition of unconsolidated sediment affects the connection of water in the surficial sediment system and the bedrock. Subpeat sediments consist of mainly sand and gravel overlying till on top of the bedrock. Till deposits can be scattered, and the hydraulic conductivity of till can be variable ($5.0 \times 10^{-8} - 1.3 \times 10^{-4} \text{ m s}^{-1}$ SRK, unpublished data, 2019; AA Sakatti Mining Oy, unpublished data, 2019) enabling groundwater to fractured bedrock. Kværner and Snilsberg (2011) also observed that there is hydraulic connection from mires to bedrock fractures through thin subpeat till deposits (<0.5 m) in Norway.

Calibration of the HSM groundwater flow models

Calibration of complex multi-layered groundwater flow models

As noted by Åberg et al. (2019), the calibration of complex groundwater flow models that have scattered Quaternary sediments and fractured bedrock is challenging due to the low sensitivities of the parameters in thin layers. Calibrating simple models such as HSM1 and HSM2 is more straightforward than complex models (HSM3 and HSM4), as their layers are thicker and their sensitivity to changes in the water table is higher. Åberg et al. (2019) also noted that the existence of perched groundwater causes uncertainties in calibration, and the observation points representing perched groundwater should be omitted. The known challenge in the previous model and HSM models is that the head observations from LiDAR DEM are at a higher level compared to the heads in groundwater monitoring wells (head variation can be as high as 0.3 m), leading to the overestimation of groundwater discharge in the mire area (see Åberg et al. 2019). Observations from monitoring wells were based on GPS altitudes fitted to the N2000 altitude system. In Finland, there is land uplift of 7 mm year$^{-1}$ due to recovery from the last glaciation (Vestøl et al. 2019). Thus, the difference in altitudes between N2000 measures and LiDAR DEM produced in the 2010s is about 10 cm. In addition, LiDAR DEMs were produced in May 2010 and 2015 when the mire water table is the highest due to spring thaw.

The influence of hydraulic conductivities of the thrust faults, brittle faults and fractures was found to be spatially limited due to their slab-like nature, and the calibration results depended on the observations nearby. Therefore, the differences in calibration results between models (Table 5) and the high uncertainty variance indicated that the hydraulic conductivity of thrust faults is difficult to estimate, and it requires further measurements and nearby observations. However, the calibration results for brittle faults were quite similar with the low uncertainty variance in models HSM2, HSM3 and HSM4, indicating some certainty. The direction and distribu-
tion of the trust faults, brittle faults and the fractures appeared to affect the calibration results: if a brittle fault or a thrust fault was more perpendicular to the groundwater flow direction, the calibration was more likely to be successful than if it was parallel to the groundwater flow direction. Multiple thrust faults were parallel to the flow direction, which probably caused high uncertainty in the thrust fault parameters. Moreover, the hydraulic gradient was also important, and many thrust faults occurred in low-gradient areas where calibration is generally challenging due to low sensitivities as also noticed by Reilly and Harbaugh (2004) in their earlier studies.

The benefit of using a transient model instead of a steady-state model could be more detailed modelling of the groundwater recharge during the spring thaw. Model fit of the transient model would most probably be poorer in simple models due to the conflict between oversimplified Quaternary sediments and high variation of the head in monitoring wells presented in Åberg et al. (2019). In contrast, in complex transient models, the model fit could be better but it would include probably more misinterpretations due to the higher number of parameters depending on the detailed subsurface structure.

**Calibration of boundary conditions**

Previous studies such as Doherty (2015), have indicated the importance of defining the boundary conditions; however, the variation in hydraulic conductivities and recharge appeared to have a higher impact on the results in the HSM models. During model development, it was observed that in relatively extensive models with a generally low hydraulic gradient, changing the conductance of the GHB boundary condition affects the water table near the boundary. GHB conductance changed the water table by more than 10 cm, while the distance from the model border was less than 700 m if the change in the GHB conductance value was more than two orders of magnitude.

The RIV conductance had a considerable effect on the calibration results due to the observed steep hydraulic gradient towards the River Kitinen (Fig. 7). The calibrated uniform RIV conductance values of the different HSM models varied from $4.4 \times 10^{-7}$ to $1.2 \times 10^{-4}$ m$^2$ s$^{-1}$ (Table 5), reflecting the hydraulic conductivity of sediments and brittle faults below (Fig. 9 location H, Fig. 11, location I). However, as Åberg et al. (2019) noted, the bottom sediments of the River Kitinen are poorly known and require further understanding for better estimation of the RIV conductance value and also estimation of the discharge area in the river. In the complex models HSM3 and HSM4, DRN conductance was found to be challenging to calibrate, as its sensitivity varied considerably, and the uncertainty variance was high. Therefore, DRN conductance was set to a constant value of $1 \times 10^{-5}$ m$^2$ s$^{-1}$. The calibration could be improved by adding flow observations, since only head observations were used in the HSM models. Flow observations could improve the uniqueness of the parameters (see Hill 1998; Poeter et al. 2014), and thus enhance the calibration results. Altogether, the calibration results were only easy to evaluate when the hydraulic gradient was steep enough. Calibration and the uncertainty of the calibrated parameters, both in the river and drain parameters, could be enhanced with Markov chain Monte Carlo (MCMC) uncertainty analysis to avoid possible local minima in parameters estimation (Poeter et al. 2014). However, confidence limits of linear uncertainty analysis provided reasonable values ($10^{-8}$–$10^{-2}$ m$^2$ s$^{-1}$) representing the most probable variation of hydraulic conductivity in the river bottom and peat layers.

**Validation with particle tracking**

The simulated flow paths based on the $d$-excess value dataset from Bigler (2019) were verified with the selected $d$-excess verification groups (1–5) based on a dataset from Åberg et al. (2019; Fig. 12). The verification groups and the simulated flow paths had the best match of $d$-excess values in the noncalibrated HSM4 model (Fig. 12, see the last table of the Appendix). Groups 2 and 3 had good fit, and groups 1 and 4 had mainly intermediate or poor fit in all models (Table 13). The intermediate fit of group 1 can be due to the lack of observations in the southwestern area (Fig. 12). In HSM1, group 1 had no match since the southernmost particles discharged already before the river (Fig. 12). Group 5 seemed to have a poor fit or no match in all models. The isotope values of group 5 are probably related to infiltration of evaporated water in the Kärväslampi Pond. Similarly, highly evaporated water in group 4 is probably infiltrated water from the mire (Fig. 12). Since MODPATH considers only advection, the mixing of different isotope compositions was not computed. In addition, evaluation of no match group is challenging since it can be related to a lack of measurements instead of a poor match. However, fairly good overall match with the highly variable $d$-excess composition in relatively short distances indicated that the flow paths from different sources can be characterized with simple methods like MODPATH.

**Converting the geological structure to the flow model grid**

**Construction of a MODFLOW-NWT grid based on the Leapfrog model**

With MODFLOW-NWT, it was possible to use relatively thin cells of 0.5 and 0.0017 m (Fig. 4c,d) as “pinched” cells in the HSM models, since the upstream weighting (UPW) package
co-used with the Newton solver in the code prevented cells from drying, enabling model convergence (Niswonger et al. 2011). This means that the water table is also simultaneously calculated for “dry cells”, making solver convergence easier for highly nonlinear models. It also allows a more realistic approach for recharge related to precipitation, which can pass through unsaturated zones to the saturated cells, since all cells in MODFLOW-NWT are kept active (Niswonger et al. 2011).

As for the geological point of view, the hydrogeology tool creating the parameter zones in Leapfrog makes the variation of the different units more realistic. Leapfrog assumes that thin cells belong to the geological unit that covers the cells the most; therefore, differences in discretization or cell resolution can generate different results. However, the variable cell thicknesses can cause errors in flow patterns due to the staggered nature of the cell heights (Gao 2011), while the flow only occurs on the faces of the six neighbouring cells when the rectangular prism grid is used (Harbaugh 2005).

Adding more vertical discretization within a hydrostratigraphic unit would enhance model accuracy (see Gao 2011), while head variation within a hydrostratigraphic unit could be taken into account in groundwater flow modelling. However, in the locations of “pinched” cells, the discontinuity of the layer can cause more instability and errors in the flow calculations if more layers are added. In addition, the block model in Leapfrog could be used to create parameter zonation for multi-layered scattered units, and the block model could be imported into ModelMuse. However, using the block model in discretization with Leapfrog instead of the hydrogeology tool was more challenging and time consuming—in particular, adding the boundary conditions to the correct layer was more challenging.

In the complex HSM models, the peat unit had two layers, whereas other geological units had only one layer—for example, Reeve et al. (2000) used multiple layers in both peat and subpeat sediment units in cross-sectional groundwater flow models to obtain more accurate flow patterns. McDonald and Harbaugh (1988) also suggested that in low conductivity units, it is preferable to use many layers with layer discretization to yield more realistic flow patterns. In the simple models, the upper layer of the mire was 0.5 m thick and the hydraulic conductivity was uniform for the whole peat thickness, while the complex models had an acrotelm of varying thickness on top of a lower conductivity catotelm.

Fig. 12 Forward particle tracking (flow paths) with isotope $d$-excess values from Bigler (2019) (starting points) were compared to selected verification groups 1–5 (squares) (Åberg et al. 2019): a calibrated HSM1 and b calibrated HSM4. Ortho images modified and reproduced after National Land Survey of Finland.
Explicit vs. implicit modelling

Leapfrog is designed for fast implicit modelling, but it can also be used in explicit modelling. Implicit modelling requires a dense drill-hole network if a realistic representation of the geological structure is desired, whereas in explicit modelling, less data is sufficient, since geological knowledge can be used to construct the continuum of the strata. Explicit modelling with the polyline tool enabled the correction of unrealistic interpolated surfaces to more closely resemble the natural shape of geological structures.

Interpolation challenges with leapfrog

FastRBF interpolation generated bullseyes for the interpolated surface of the bedrock topography when the snap option was used for drill holes and polylines (Fig. 5). The snap option forces the interpolated surface to follow the exact altitude of an observation (e.g. drill holes, polylines), enhancing the undulation of the surface, which is useful in areas with a rugged bedrock surface (as in Fig. 4a,b). To improve the interpolation of the bedrock surface, the W–E orientation of schistosity could have been taken into account in the trend or anisotropy. Additionally, to obtain a smoother interpolated surface (as for the peat bottom), external software (ArcGIS) could be used.

Modelling of scattered units with contact surfaces and the polyline editing tool

The construction of Quaternary sediment units with Leapfrog and selected contact types depends on the geological structure. Moreover, the contact surfaces can be used in a creative way differing from its original purpose—for example, in this study, the bottom contact of the peat layer was generated using the erosion type. This type is suitable if the shape of the unit is exported from external software such as ArcGIS and the aim is to preserve the shape. In addition, using the erosional contact type is useful when the invasion of a lower unit is not intended, and the penetration of an older unit is avoided.

Leapfrog Geo appears to lack a predefined option or tool to model lens-like structures other than intrusions (Seequent Ltd. 2020) suitable, for example, to model a braided river environment, eskers or drumlins. The most difficult scenario in Leapfrog is when two different adjacent sediment units are observed in drill-hole data such as fill and bar-like structures (Fig. 13). This can be corrected with explicit modelling by adding additional dummy polylines or points bending the contacts, thereby avoiding the invasion of inappropriate units (Fig. 13a). Fill-like structures can be created by raising the borders of the bottom contact above overlying younger contacts (Fig. 13b). Bending the surface of a deposit-type contact beneath an older contact surface (either deposit or erosion) creates lens-like structures (Fig. 13c). The configuration in Fig. 13b could be used in fill structures such as fluvial or glaciofluvial fills, whereas the configuration in Fig. 13c could be used in braided river bars, drumlins or eskers. Cross-section-based modelling (e.g. GSIS, see Ming et al. 2010) could be more effective for constructing lens-like structures.

Alternative options to construct models with complex quaternary sediments and weathered and fractured bedrock

In this study, the variation in hydraulic conductivities was generated within the predefined hydrostratigraphic units by interpolation with points. Alternatively, the units could be separated into smaller units, either with geostatistical tools such as T-PROGS (Carle 1999) or by defining smaller areas with polygons. There are also mathematical codes and stochastic methods that generate complex hydraulic conductivity fields based on pseudo-genetic models such as braided river bar structures (e.g. Felletti et al. 2006; Pirot 2015).

MODFLOW, which was used in this study, assumes that all units used in the model are equivalent to continuous porous media (CPM), including the fractured bedrock layer. However, Darcy’s law is only valid when the fracture network is dense enough and behaves like a porous medium (Singhal and Gupta 1999). In addition, if the hydraulic conductivity of a fracture is too high, turbulent flow appears and Darcy’s law is not valid. Thus, to improve the modelling accuracy for fractured bedrock, software such as ConnectFlow (Jackson et al. 2000) could be used to generate grids, applying both CPM and a discrete fractured network (DFN) (Hartley and Joyce 2013). Furthermore, ConnectFlow algorithms can be used to calculate K within the Quaternary sediment system underlain by fractured bedrock (Hartley et al. 2009; Hartley and Joyce 2013).

In most of the MODFLOW series codes, including MODFLOW-NWT, horizontally continuous layers are assumed, and the pinched cell approach is the only way to create lens-like structures. MODFLOW 6 (Hughes et al. 2017) could be an option for creating scattered units using unstructured grids, where the number of cells between layers can vary. Unfortunately, the hydrogeology tool of Leapfrog cannot currently create unstructured grids; however, Leapfrog is able to create uneven cell sizes in X and Y directions, enabling denser gridding for river banks, for example, and looser gridding for poorly known or homogeneous sediment areas.

To further develop the modelling workflow used in this study, multiple layers in bedrock units should be added to the MODFLOW grid to create fractures with varying dips. Additionally, fractures could have been generated with the
deposit or erosion type of contact in Leapfrog and used as such for the hydrogeology tool, although this would be complicated. Moreover, Leapfrog Geo version 4.4 and older did not allow the use of the vein and intrusion type as contact surfaces in the hydrogeology tool. To model the hydraulic properties of fractures in more detail, the MODFLOW package Horizontal Flow Barrier (HFB) could be used to create low conductivity zones at 90° surrounding the faults and fractures. In addition, in groundwater flow modelling with the code FEFLOW, it is possible to generate discrete features such as conduits or fault structures inside the cell (Diersch 2014). Leapfrog could generate a FEFLOW grid with the hydrogeology tool, and it would therefore be possible to test a corresponding workflow.

Conclusions

The hydrostratigraphic model and its corresponding groundwater flow model for complex Quaternary sediments underlain by fractured and weathered crystalline bedrock was possible to construct with the tested workflow. A workflow, starting with 3D geological modelling using Leapfrog Geo and followed by 3D hydrostratigraphic modelling and groundwater flow modelling with MODFLOW-NWT and ModelMuse, was developed and tested for models differing in complexity. An increase in hydrostratigraphic details appears to improve the fit of the modelled and observed heads in groundwater wells at different depths. The appropriate model complexity is dependent on the research question. In water balance studies, simple models can give as reasonable results as complex models; however, if detailed understanding of groundwater flow patterns is needed, more detailed models are recommended. This is suggested especially in studies related to mire areas or groundwater dependent ecosystems.

In hydrostratigraphic modelling, Leapfrog is easy to use, and the polyline editing tool is advantageous for explicit modelling when the input data are not sufficient to generate realistic models by implicit modelling or when generating thin (<2 m) and undulating sediment units. However, using contact-based 3D modelling software in areas with unconsolidated sediments consisting of glacigenic and fluvial material seems to be challenging, since they often form lens-like structures and scattered layers. The lens-like structures can be explicitly created if dummy points and the contact surfaces are used innovatively. Cross-section-based software could enhance the automatic modelling of lens-like structures. The main advantage of using Leapfrog is that Leapfrog Geo with the hydrogeology tool interacts well with ModelMuse and MODFLOW-NWT. However, changing the geological 3D model in an explicit manner to test the effect of 3D flow model results was found to be time consuming.

MODFLOW-NWT code is suitable for groundwater flow modelling with complex but thin layered Quaternary strata having multiple unsaturated units. Scattered hydrostratigraphic units can be modelled with this code as continuous layers when the “unnecessary thin cells” between scattered units are defined using the Leapfrog evaluation tool.
Details of the hydrostratigraphic model can easily be added in ModelMuse with polygons and interpolations between points, which adds detail to the groundwater flow patterns and groundwater discharge areas, and simultaneously reduces the time for working with Leapfrog.

Modelling results indicated the importance of variation in hydraulic conductivity: low conductivity units behave like low conductivity faults or barriers, regardless of their shape. Thus, for example, scattered till units with low hydraulic conductivity or clay-type weathered bedrock should be taken into account in model construction. Interlayered low and high conductivity zones disperse the recharge/discharge patterns in areas with a flat topography. In addition, high conductivity units beneath a mire add vertical flow. It is important to estimate the thickness of the upper fractured zone of bedrock, as the thickness of this layer affects its transmissivity. Model complexity generated in Leapfrog adds details to groundwater flow paths and discharge patterns and improves model fit compared to simple models when high variation in hydraulic conductivity is present.

The structure of a 3D geological model can be evaluated with groundwater flow modelling results, and the areas that may need reconstruction of the recharge rate, hydraulic conductivity or layer structure can be detected. Areas with a flat low hydraulic gradient, such as extensive mire areas, are difficult to calibrate, whereas high-gradient areas such as river banks, are easier to calibrate due to their higher sensitivity. In areas with a high hydraulic gradient, changing the grid was found to have a large local effect on the groundwater flow modelling results. Overall, it is worth constructing a detailed geological model if high variation in hydraulic conductivities between geological units is present and the target is to understand groundwater discharge and recharge patterns accurately. However, there should be enough data and time available for modelling.

Acknowledgements We thank AA Sakatti Mining Oy for collaboration, especially Anne Rautio, Jukka Jokela, Perti Lamberg, Ulla Syrjälä, Joanna Kuntonen-van’t Riet, and Janne Siikaluoma. We thank the personnel of the University of Helsinki Veli-Pekka Salonen, Seija Kultti, and Emilia Koivisto for their contribution and assistance, and the students of the University of Helsinki Jyri Laakso, Mimmi Takalo, and Harri Turttainen for their contribution during fieldwork. We would also like to express our gratitude to Richard Winston for guidance with ModelMuse and John Doherty for guidance with PEST.

Funding Open Access funding provided by University of Helsinki including Helsinki University Central Hospital. Additional funding was provided by the Doctoral Programme in Geosciences (GeoDoc). The funding of the project was provided by the K.H. Renlund Foundation and Maa-ja vesitekniikan tuki ry.

Appendix

Table 6  Monitoring and observation wells used in model calibration

| Groundwater well | Mean observed WT (m asl) | Observation N | Observation time | Screen depth (m) | Corresponding unit in GM2 |
|------------------|--------------------------|---------------|------------------|-----------------|--------------------------|
| GA102            | 186.4                    | 356           | 1 Jan 2014–31 Dec 2014 | 1–3             | Peat                     |
| GA103            | 186.0                    | 356           | 1 Jan 2014–31 Dec 2014 | 1–2             | Peat                     |
| GA200            | 185.7                    | 356           | 1 Jan 2014–31 Dec 2014 | 4.5–6.5         | Middle sorted/lower sorted |
| GA201            | 186.8                    | 356           | 1 Jan 2014–31 Dec 2014 | 3–5             | Lower till                |
| GA301            | 187.2                    | 356           | 1 Jan 2014–31 Dec 2014 | 7.5–9.5         | Lowest sands and gravels |
| GA302            | 186.3                    | 356           | 1 Jan 2014–31 Dec 2014 | 8–10            | Middle sorted/lower sorted |
| GA303            | 185.7                    | 356           | 1 Jan 2014–31 Dec 2014 | 10.5–12.5       | Lower till                |
| GA304            | 186.1                    | 356           | 1 Jan 2014–31 Dec 2014 | 12–14           | Lowest till               |
| GA305            | 182.9                    | 356           | 1 Jan 2014–31 Dec 2014 | 8–10            | Lower till                |
| GA401            | 187.2                    | 356           | 1 Jan 2014–31 Dec 2014 | 9.5–12.5        | Weathered bedrock        |
| GA402            | 186.4                    | 356           | 1 Jan 2014–31 Dec 2014 | 11–13           | Weathered bedrock        |
| GA403            | 185.8                    | 356           | 1 Jan 2014–31 Dec 2014 | 15–17           | Weathered bedrock        |
| GA404            | 183.6                    | 356           | 1 Jan 2014–31 Dec 2014 | 8.5–10.5        | Weathered bedrock        |
| GA405            | 182.8                    | 356           | 1 Jan 2014–31 Dec 2014 | 11.5–13.5       | Weathered bedrock        |
| MOS8049          | 185.5                    | 5             | 14 Jun 2017–22 Oct 2017 | Open bedrock well | Fractured bedrock        |
| MOS8120          | 184.1                    | 6             | 14 Jun 2017–17 Oct 2017 | Open bedrock well | Fractured bedrock        |
| MOS8125          | 185.4                    | 6             | 4 Jun 2017–17 Oct 2017  | Open bedrock well | Fractured bedrock        |
| HYD009           | 185.2                    | 6             | 4 Jun 2017–18 Oct 2017  | 2.6–6.6         | Lower till                |
| HYD010           | 185.9                    | 6             | 14 Jun 2017–18 Oct 2017  | 1.4–3.4         | Middle sorted/lower sorted |
| HYD013           | 181.1                    | 6             | 14 Jun 2017–17 Oct 2017  | Open bedrock well | Fractured bedrock        |
Table 6 (continued)

| Groundwater well | Mean observed WT (m asl) | Observation N | Observation time | Screen depth (m) | Corresponding unit in GM2 |
|------------------|--------------------------|---------------|------------------|------------------|---------------------------|
| HYD014           | 182.2                    | 6             | 13 Jun 2017–18 Oct 2017 | Open bedrock well | Fractured bedrock          |
| HYD015           | 189.2                    | 6             | 4 Jun 2017–10 Oct 2017  | 3–5, 7–9        | Lower till                 |
| HYD016           | 186.5                    | 6             | 13 Jun 2017–18 Oct 2017 | 3–7             | Middle sorted/lower sorted |
| HYD018           | 185.3                    | 3             | 6 Sep 2017–18 Oct 2017  | 4.4–8.4         | Middle sorted/lower sorted |
| HYD019           | 189.3                    | 6             | 4 Jun 2017–10 Oct 2017  | 1.2–3.2, 5.2–7.2| Middle sorted/lower sorted |
| HYD020           | 187.6                    | 6             | 4 Jun 2017–18 Oct 2017  | 4–8             | Middle sorted/lower sorted |
| HYD021           | 184.2                    | 5             | 7 Aug 2017–18 Oct 2017  | 8–12            | Middle sorted/lower sorted |
| HYD022           | 184.4                    | 6             | 4 Jun 2017–18 Oct 2017  | 8.2–12.2        | Lowest sands and gravels   |
| GA104            | 185.4                    | 6             | 13 Jun 2017–17 Oct 2017 | 2–4             | Peat                      |
| GA204            | 184.8                    | 6             | 13 Jun 2017–17 Oct 2017 | 2–4             | Upper till                |
| GA205            | 183.7                    | 5             | 13 Jun 2017–22 Sep 2017 | 2–4             | Top deposit               |
| GA300            | 187.1                    | 6             | 14 Jun 2017–17 Oct 2017 | 3–4             | Top deposit               |
| GA400            | 187.5                    | 6             | 14 Jun 2017–17 Oct 2017 | 5–7             | Weathered bedrock         |

Locations are presented in Fig. 1. WT is the water table

*a Depth from ground surface

Table 7 Water budget in the calibrated groundwater flow model HSM1

| Component                        | Cumulative volume (in) | Cumulative volume (out) |
|----------------------------------|------------------------|-------------------------|
|                                  | Value (m³)             | Percentage              | Value (m³)             | Percentage              |
| Storage                          | 0                      | 0.0%                    | 0                      | 0.0%                    |
| Constant head                    | 1,757,501.125          | 23.5%                   | 187,426.1875           | 2.5%                    |
| Drains                           | 0                      | 0.0%                    | 5,573,878.5            | 74.6%                   |
| River leakage                    | 28,112.7949            | 0.4%                    | 1,628,083.375          | 21.8%                   |
| Head-dependent boundaries        | 45,869.7266            | 0.6%                    | 77,384.7422            | 1.0%                    |
| Recharge                         | 5,635,287              | 75.5%                   | 0                      | 0.0%                    |
| Total                            | 7,466,770.5            | 100.0%                  | 7,466,772.5            | 100.0%                  |

Table 8 Water budget in the calibrated groundwater flow model HSM2

| Component                        | Cumulative volume (in) | Cumulative volume (out) |
|----------------------------------|------------------------|-------------------------|
|                                  | Value (m³)             | Percentage              | Value (m³)             | Percentage              |
| Storage                          | 0                      | 0.0%                    | 0                      | 0.0%                    |
| Constant head                    | 1,795,203.75           | 23.2%                   | 253,755.9062           | 3.3%                    |
| Drains                           | 0                      | 0.0%                    | 5,076,190              | 65.7%                   |
| River leakage                    | 189,945.875            | 2.5%                    | 2,307,309              | 29.9%                   |
| Head-dependent boundaries        | 106,359.4219           | 1.4%                    | 89,552.9062            | 1.2%                    |
| Recharge                         | 5,635,287              | 72.9%                   | 0                      | 0.0%                    |
| Total                            | 7,726,796              | 100.0%                  | 7,726,808              | 100.0%                  |
Table 9  Water budget in the calibrated groundwater flow model HSM3

| Component                        | Cumulative volume (in) | Cumulative volume (out) |
|----------------------------------|------------------------|-------------------------|
|                                  | Value (m³)              | Percentage              | Value (m³)               | Percentage  |
| Storage                          | 0                      | 0.0%                    | 0                        | 0.0%        |
| Constant head                    | 1,695,598.625           | 23.4%                   | 375,258.5938              | 5.2%        |
| Drains                           | 0                      | 0.0%                    | 4,270,219.5               | 59.1%       |
| River leakage                    | 97,152.6953             | 1.3%                    | 2,506,017.25              | 34.7%       |
| Head-dependent boundaries        | 108,514.8594            | 1.5%                    | 79,736.3906               | 1.1%        |
| Recharge                         | 532,9917.5              | 73.7%                   | 0                        | 0.0%        |
| Total                            | 7,231,184               | 100.0%                  | 7,231,231.5               | 100.0%      |

Table 10  Water budget in the calibrated groundwater flow model HSM4

| Component                        | Cumulative volume, in  | Cumulative volume, out |
|----------------------------------|------------------------|------------------------|
|                                  | Value (m³)              | Percentage              | Value (m³)               | Percentage  |
| Storage                          | 0                      | 0.0%                    | 0                        | 0.0%        |
| Constant head                    | 1,515,531.25            | 21.6%                   | 414,071.8438              | 5.9%        |
| Drains                           | 0                      | 0.0%                    | 4,018,178.75              | 57.2%       |
| River leakage                    | 75,602.5078             | 1.1%                    | 2,496,457.5               | 35.5%       |
| Head-dependent boundaries        | 90,433.2812             | 1.3%                    | 99,574.1641               | 1.4%        |
| Recharge                         | 5,346,703               | 76.1%                   | 0                        | 0.0%        |
| Total                            | 7,028,270               | 100.0%                  | 7,028,282                 | 100.0%      |

Table 11  The variation in $K$ for the minerogenic units and the River Kitinen used in the HSM models

| GM/HSM solids         | $K$ (m s$^{-1}$) |
|-----------------------|-----------------|
| Kitinen               | $1.0 \times 10^{-2}$ |
| Top deposits          | $1.2 \times 10^{-6} - 6.1 \times 10^{-4}$ |
| Upper till            | $2.4 \times 10^{-7} - 1.3 \times 10^{-4}$ |
| Middle sorted         | $3.3 \times 10^{-6} - 4.3 \times 10^{-3}$ |
| Middle till           | Combined into 'middle sorted' |
| Lower sorted          | Combined into 'middle sorted' |
| Lower till assemblage | $5.0 \times 10^{-8} - 8.3 \times 10^{-5}$ |
| Lowest sands and gravels | $1.5 \times 10^{-4}$ |
| Lowest till           | $1.0 \times 10^{-5}$ |
Hydrogeol J (2021) 29:1043–1074

Table 12 RMSE, MAE and bias (units in metres) of head observations were calculated separately for groundwater well (GW), surface water observation (SW), and spring observations. Cal indicates the calibrated model.

| Parameter | HSM1 | HSM1 cal | HSM2 | HSM2 cal | HSM3 | HSM3 cal | HSM4 | HSM4 cal |
|-----------|------|----------|------|----------|------|----------|------|----------|
| Bias SW   | −0.01| 0.02     | 0.01 | 0.03     | −0.03| −0.01    | −0.05| −0.03    |
| Bias GW   | −0.41| −0.30    | −0.38| −0.27    | −0.25| −0.20    | −0.26| −0.19    |
| Bias spring | 0.25 | 0.32     | 0.06 | 0.17     | 0.43 | 0.27     | 0.43 | 0.31     |
| MAE SW    | 0.10 | 0.13     | 0.13 | 0.13     | 0.15 | 0.16     | 0.17 | 0.17     |
| MAE GW    | 0.60 | 0.53     | 0.54 | 0.51     | 0.46 | 0.43     | 0.44 | 0.41     |
| MAE spring | 0.50 | 0.52     | 0.40 | 0.43     | 0.57 | 0.46     | 0.56 | 0.53     |
| RMSE SW   | 0.21 | 0.23     | 0.22 | 0.24     | 0.23 | 0.24     | 0.27 | 0.27     |
| RMSE GW   | 0.75 | 0.65     | 0.68 | 0.63     | 0.56 | 0.54     | 0.54 | 0.50     |
| RMSE spring | 0.60 | 0.62     | 0.47 | 0.50     | 0.68 | 0.57     | 0.70 | 0.63     |

Table 13 Comparison of d-excess in flow paths of particle tracking to spring observations in groups 1–5. Cal indicates the calibrated model. Point sum indicates the overall match of d-excess values summed per model. Model point sum is the sum of noncalibrated and calibrated models. Group points is the sum of points per group.

| Model | Group 1 | Group 2 | Group 3 | Group 4 | Group 5 | Point sum | Model point sum |
|-------|---------|---------|---------|---------|---------|-----------|----------------|
| HSM1  | −       | 0       | 2       | 0       | 0       | 2         | 6             |
| HSM1 cal | −     | 1       | 2       | 1       | 0       | 4         |               |
| HSM2  | 1       | 1       | 2       | −       | −       | 4         | 8             |
| HSM2 cal | 1    | 1       | 2       | −       | −       | 4         |               |
| HSM3  | 1       | 1       | 2       | 0       | 0       | 4         | 8             |
| HSM3 cal | 1   | 1       | 2       | −       | 0       | 4         |               |
| HSM4  | 1       | 1       | 2       | 1       | −       | 5         | 9             |
| HSM4 cal | 1   | 1       | 2       | 0       | −       | 4         |               |
| Group points | 6 | 7 | 16 | 2 | 0 | 2 | 4 |

Explanation of numerical notations: 2 good match with observation, 1 moderate match, 0 poor match, − no match

Open Access This article is licensed under a Creative Commons Attribution 4.0 International License, which permits use, sharing, adaptation, distribution and reproduction in any medium or format, as long as you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons licence, and indicate if changes were made. The images or other third party material in this article are included in the article’s Creative Commons licence, unless indicated otherwise in a credit line to the material. If material is not included in the article’s Creative Commons licence and your intended use is not permitted by statutory regulation or exceeds the permitted use, you will need to obtain permission directly from the copyright holder. To view a copy of this licence, visit http://creativecommons.org/licenses/by/4.0/.

References

Åberg SC, Korkka-Niemi K, Rautio A, Salonen V-P, Åberg AK (2019) Groundwater recharge/discharge patterns and groundwater–surface water interactions in a sedimentary aquifer along the River Kitinen in Sodankylä, northern Finland. Boreal Environ Res 24:155–187

Artimo A (2002) Application of flow and transport models to the polluted Honkala aquifer, Säkylä, Finland. Boreal Environ Res 7(2):161–172

Artimo A, Salonen V-P, Pietiäi S, Sarapäär S (2004) Three-dimensional geologic modeling and groundwater flow modeling of the Tollinenä Aquifer in the Hitura nickel mine area, Finland; providing the framework for restoration and protection of the aquifer. Bull Geol Soc Finl 76:5–17

Batelaan O, De Smedt F (1998) An adapted DRAIN package for seepage problems. In: MODFLOW’98 Proceedings, Citeseer, Golden, CO, October 1998, pp 555–562

Bense V, Van Balen R, De Vries J (2003) The impact of faults on the hydrogeological conditions in the Roer Valley Rift System: an overview. Neth J Geosci 82:41–54. https://doi.org/10.1016/j.earscirev.2013.09.008

Bigler P (2019) Hydrogeology and hydrogeochemistry of the western margin of the Viiankaapa mire in Sodankylä: factors affecting the distribution of endangered species. MSc Thesis University of Helsinki, Helsinki

Brownsocombe W, Ihlenfeld C, Coppard J, Hartshorne C, Klatt S, Sikanalauka JK, Herrington RJ (2015) The Sakatti Cu-Ni-PGE sulfide deposit in northern Finland, chap 3.7. In: Maier WD, Lahitten R, O’Brien H (eds) Mineral deposits of Finland. Elsevier,
Amsterdam, pp 211–252. https://doi.org/10.1016/B978-0-12-410438-9.00009-1

Carle SF (1999) T-PROGS: transition probability geostatistical software: user’s manual. University of California, Davis, CA. http://gmsdocs.aquaveo.com/t-progs.pdf. Accessed 20 Nov 2019

Cox ME, James A, Hawke A, Raiber M (2013) Groundwater Visualization System (GVS): a software framework for integrated display and interrogation of conceptual hydrogeological models, data and time-series animation. J Hydrol 491:56–72. https://doi.org/10.1016/j.jhydrol.2013.03.023

Diersch, HJG (2014) FEFLOW: finite element modeling of flow, mass and heat transport in porous and fractured media. Manual. Springer, Heidelberg, Germany

Doherty J (2015) Calibration and uncertainty analysis for complex environmental models. Watermark, Brisbane, Australia

Doherty J (2018) PEST: model independent parameter estimation—user manual, 7th edn. Watermark, Brisbane, Australia

Durand V, Léonardi V, de Marsily G, Lachassagne P (2017) Quantification of the specific yield in a two-layer hard-rock aquifer model. J Hydrol 551:328–339. https://doi.org/10.1016/j.jhydrol.2017.05.013

Eilu P (ed) (2012) Mineral deposits and metallogeny of Fennoscandia. Special Paper 53, Geological Survey of Finland, Espoo, Finland, 401 pp

Felletti F, Bersezio R, Giudici M (2006) Geostatistical simulation and kriging methods [Report on geochemical mapping methods]. Geological Survey of Finland, Espoo, Finland, 189 pp

Felletti F, Bersezio R, Giudici M (2006) Geostatistical simulation and kriging methods. Geological Survey of Finland, Espoo, Finland

Felletti F, Bersezio R, Giudici M (2006) Geostatistical simulation and kriging methods. Geological Survey of Finland, Espoo, Finland

Feinstein DT, Fienen MN, Kennedy JL., Buchwald CA, Greenwood MM (2012) Development and application of a groundwater/surface-water flow model using MODFLOW-NWT for the Upper Fox River Basin, southeastern Wisconsin. US Geol Surv Sci Invest Rep 2012-5108

Feinberg BM, Greenwood MM (1992) Development and application of a groundwater/surface-water flow model using MODFLOW-NWT for the Upper Fox River Basin, southeastern Wisconsin. US Geol Surv Sci Invest Rep 2012-5108

Freeze R, Cherry J (1979) Groundwater, Prentice Hall, Englewood Cliffs, NJ

Freeze RA, Witherspoon PA (1967) Theoretical analysis of regional groundwater flow: 2, effect of water-table configuration and subsurface permeability variation. Water Resour Res 3:623–634. https://doi.org/10.1029/WR003i002p00623

GADM (2020) GADM maps and data. http://gadm.org. Accessed 20 Nov 2019

Gao H (2011) Groundwater modeling for flow systems with complex geological and hydrogeological conditions. Proceed Earth Planet Sci 3:23–28. https://doi.org/10.1016/j.eprogs.2011.09.061

Gat J (2010) Isotope hydrology: a study of the water cycle. Imperial College Press, London, 189 pp

Gustavsson N, Tomasen P, Tanskanen H (1979) Seloste geokemiallisen kartoituksen tutkimusmenetelmiä [Report on geochemical mapping methods]. Geological Survey of Finland, Espoo, Finland

Hall AM, Sarala P, Ebert K (2015) Late Cenozoic deep weathering patterns on the Fennoscandian shield in northern Finland: a window on ice sheet bed conditions at the onset of Northern Hemisphere glaciation. Geomorphology 246:472–488. https://doi.org/10.1016/j.geomorph.2015.06.037

Harbaugh AW (2005) MODFLOW-2005, The U.S. Geological Survey modular ground-water model: the ground-water flow process. US Geol Surv Tech Methods 6-A16.http://pubs.usgs.gov/tm/2005/tm6A16/. Accessed 20 November 2019

Hartley L, Joyce S (2013) Approaches and algorithms for groundwater flow modeling in support of site investigations and safety assessment of the Forsmark site, Sweden. J Hydrol 500:200–216. https://doi.org/10.1016/j.jhydrol.2013.07.031

Hartley L, Hoek J, Swan D, Roberts D, Joyce S, Follin S (2009) Development of a hydrogeological discrete fracture network model for the Olkiluoto Site Descriptive Model 2008. Posiva working report 2009–61, Posiva Oy, Eurajoki, Finland. http://www.posiva.fi/files/1114/WR_2009-61.web.pdf. Accessed 28 October 2019

Helmens KF, Räsänen ME, Johansson PW, Jungner H, Korjonen K (2000) The last interglacial-glacial cycle in NE Fennoscandia: a nearly continuous record from Sokli (Finnish Lapland). Quatern Sci Rev 19:1605–1623. https://doi.org/10.1016/S0277-3791(00)00044-4

Hill MC (1998) Methods and guidelines for effective model calibration, with application to UCODE, a computer code for universal inverse modeling, and MODFLOWP, a computer code for inverse modeling with MODFLOW. US Geol Surv Water Resour Invest Rep 98-4005. http://pubs.er.usgs.gov/publication/wri984005. Accessed 28 October 2019

Hiras H (1991) Pleistocene stratigraphy of Finnish Lapland. Geol Surv of Fin Bull 354:1–123

Hiras H, Allthn A, Pulkkinen E, Puranen R, Tynn R (1977) Raportti malminetsintää palvelevasta maaperätutkimuksesta Pohjois-Suomessa vuosina 1972–1976 [A report on glacial drift investigations for ore prospecting purposes in northern Finland 1972–1976]. Report of Investigation no. 19, Geological Survey of Finland. http://tupa.gtk.fi/julkaisu/tutkimusraportit/tr_019.pdf. Accessed 20 November 2019

Hjelt A, Pääkkö E (2006) Viiankiaavan hoito- ja käyttösuunnitelma [Treatment and use plan of Viiankaipaa]. Metsähallituksen luonnon suojelujulkaisuja, Sarja C [Metsähallitus’ nature conservation publications, series C], vol 11, Metsähallitus, Vantaa, Finland, 51 pp

Holden J, Burt TP (2003) Hydrological studies on blanket peat: the significance of the acrotelm-catotelm model. J Ecol 91(1):86–102. https://www.jstor.org/stable/3599540. Accessed 20 Nov 2019

Howett PJ, Salonen V-P, Hyttinen O, Korkka-Niemi K, Moreau J (2015) A hydrostratigraphical approach to support environmentally safe siting of a mining waste facility at Rautavaara, Finland. Bull Geol Soc Finl 87:51–66. https://doi.org/10.17741/bgsf/87.2.001

Hudson-Gagnon E, Chesaux R, Cousineau PA, Rouleau A (2015) A hydrostratigraphic simplification approach to build 3D groundwater flow numerical models: example of a Quaternary deltaic deposit aquifer. Environ Earth Sci. 74:4671–4686. https://doi.org/10.1007/s12665-015-4439-y

Hughes JD, Langevin CD, Banta ER (2017) Documentation for the MODFLOW 6 framework: US Geological Survey Techniques and Methods, book 6, chap A57, 40 pp. https://doi.org/10.3133/tm6A57

Huntington JL, Niswonger RG, Rajagopal S, Zhang Y, Gardner M, Hjelt A, Pääkkö E. (2006) Viiankiaavan hoito- ja käyttösuunnitelma [Treatment and use plan of Viiankaipaa]. Metsähallituksen luonnon suojelujulkaisuja, Sarja C [Metsähallitus’ nature conservation publications, series C], vol 11, Metsähallitus, Vantaa, Finland, 51 pp

Ilmatieteen laitos (2018) Vuosittilastot [Annual statistics]. http://ilmatieteenlaitos.fi/vuosittilastot. Accessed 4 December 2020

Jackson CP, Hoch AR, Todman SJ (2000) Self-consistency of a heterogeneous continuum porous medium representation of a fractured medium. Water Resour Res 36(1):189–202. https://doi.org/10.1029/1999WR900249

Johannson P, Kujansuu R (2005) Pohjois-Suomen maaperä: maaperäkarttojen 1:400 000 selitys. [Quaternary deposits of northern Finland: explanation to the maps of Quaternary deposits 1:400 000]. Erikoisjulkaisut [Special Publications], Geological Survey of Finland, Espoo, Finland
Welby CW (1981) A technique for evaluating the hydraulic conductivity of saprolite. UNC-WRRI-81-164, Water Resources Institute, Raleigh, NC

Winston RB (2009) ModelMuse—a graphical user interface for MODFLOW–2005 and PHAST. U.S. Geological Survey Techniques and Methods 6–A29, p 52. http://pubs.usgs.gov/tm/tm6A29. Accessed 20 Nov 2019

Wright EP (1992) The hydrogeology of crystalline basement aquifers in Africa. Geol Soc Lond Spec Publ 66:1–27

Wycisk P, Hubert T, Gossel W, Neumann C (2009) High-resolution 3D spatial modelling of complex geological structures for an environmental risk assessment of abundant mining and industrial megasites. Computers Geosci 35:165–182. https://doi.org/10.1016/j.cageo.2007.09.001

Publisher’s note Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.