Characterizing groundwater heat transport in a complex lowland aquifer using paleo-temperature reconstruction, satellite data, temperature–depth profiles, and numerical models

Alberto Casillas-Trasvina1,2, Bart Rogiers1, Koen Beerten1, Laurent Wouters3, and Kristine Walraevens2

1 Institute for Environment, Health and Safety, Belgian Nuclear Research Centre, SCK CEN, Boeretang 200, 2400 Mol, Belgium
2 Laboratory for Applied Geology and Hydrogeology, Department of Geology, Ghent University, Krijgslaan 281-S8, 9000 Ghent, Belgium
3 Long-term Management Department, ONDRAF/NIRAS, Belgian Agency for Radioactive Waste and Enriched Fissile Materials, Kunstlaan 14, 1210 Brussels, Belgium

Correspondence: Alberto Casillas-Trasvina (jesuscasillasmx@outlook.com)

Received: 16 November 2021 – Discussion started: 13 January 2022
Revised: 22 June 2022 – Accepted: 5 September 2022 – Published: 9 November 2022

Abstract. Heat is a naturally occurring, widespread groundwater tracer that can be used to identify flow patterns in groundwater systems. Temperature measurements, being relatively inexpensive and effortless to gather, represent a valuable source of information which can be exploited to reduce uncertainties on groundwater flow, and, for example, support performance assessment studies on waste disposal sites. In a lowland setting, however, hydraulic gradients are typically small, and whether temperature measurements can be used to inform us about catchment-scale groundwater flow remains an open question. For the Neogene Aquifer in Flanders, groundwater flow and solute transport models have been developed in the framework of safety and feasibility studies for the underlying Boom Clay formation as a potential host rock for geological disposal of radioactive waste. However, the simulated fluxes by these models are still subject to large uncertainties as they are typically constrained by hydraulic heads only. In the current study, we use a state-of-the-art 3D steady-state groundwater flow model, calibrated against hydraulic head measurements, to build a 3D transient heat transport model, for assessing the use of heat as an additional state variable, in a lowland setting and at the catchment scale. We therefore use temperature–depth (TD) profiles as additional state variable observations for inverse conditioning. Furthermore, a Holocene paleo-temperature time curve was constructed based on paleo-temperature reconstructions in Europe from several sources in combination with land surface temperature (LST) remotely sensed monthly data from 2001 to 2019 (retrieved from NASA’s Moderate Resolution Imaging Spectroradiometer, MODIS). The aim of the research is to understand the mechanisms of heat transport and to characterize the temperature distribution and dynamics in the Neogene Aquifer. The simulation results clearly underline advection/convection and conduction as the major heat transport mechanisms, with a reduced role of advection/convection in zones where flux magnitudes are low, which suggests that temperature is also a useful indicator in a lowland setting. Furthermore, the performed scenarios highlight the important roles of (i) surface hydrological features and withdrawals driving local groundwater flow systems and (ii) the inclusion of subsurface features like faults in the conceptualization and development of hydrogeological investigations. These findings serve as a proxy of the influence of advective transport and barrier/conduit role of faults, particularly for the Rauw fault in this case, and suggest that solutes released from the Boom Clay might be affected in similar ways.
1 Introduction

Heat is a naturally occurring, widespread groundwater tracer that can be used to identify flow patterns in groundwater systems (Anderson, 2005; Bense et al., 2017; Saar, 2011). Yet, it is often not evenly distributed in basins, as thermal heterogeneities are observed in the increase in temperature with depth (Dentzer et al., 2017). Recent work has broadened the use of heat in a quantitative way by incorporating it in formal solutions of the inverse problem to estimate hydraulic properties and groundwater flux (Hecht-Méndez et al., 2010; Jiang and Woodbury, 2006; Liu et al., 2019; Munz et al., 2017; Rau et al., 2010; Rodríguez-Escaselles et al., 2020). This is commonly done with numerical codes that at least enable a one-way coupling of the different processes, i.e. groundwater flow and heat transport.

Groundwater flow induces heat advection, which is a significant component of the total heat flux, especially in sedimentary basins, and thereby influences the subsurface temperature distributions for relatively deep groundwater flow (Anderson, 2005; Pollack et al., 1993; Saar, 2011). Due to this thermal signature transmitted by the movement of groundwater, an analysis of subsurface temperature distributions can yield quantitative insight into the groundwater flow systems behaviour (Saar, 2011). Additionally, when used as a tracer, groundwater temperatures are more sensitive to, for instance, the connectivity patterns and fault zones within an aquifer compared to hydraulic data alone, providing supplementary information on aquifer structure (Bense et al., 2017; Kurtz et al., 2014; Read et al., 2013). Temperature measurements, which are relatively inexpensive and effortless to gather, hence represent a valuable source of information which can be exploited to constrain groundwater flow (Anderson, 2005; Irvine et al., 2017; Kurylyk et al., 2018; Saar, 2011). However, as stated by Schilling et al. (2019), it is currently an underrepresented state variable in groundwater studies with respect to hydraulic head. Schilling et al. (2019) present a robust review on the use of what they call “unconventional” state variables, including temperature observations. They mention that the inclusion of temperature observations in combination with “conventional” observations (i.e. the hydraulic head) is beneficial for heat transport simulations, given that heat transport is not appropriately simulated on its own (Schilling et al., 2019). However, as demonstrated by Bravo et al. (2002), Kurtz et al. (2014), Irvine et al. (2015), and Delsman et al. (2016), when there are multiple unknowns on flux exchanges and material (i.e. thermal and hydraulic) properties, then the calibration using the hydraulic head and temperature observation as targets is likely to reproduce the temperature correctly in spite of a potential incorrect representation of the fluxes (Schilling et al., 2019).

Nonetheless, most of these studies (i.e. where temperature observations have been implemented to evaluate aquifer hydraulic characteristics) have been performed in shallow environments, mainly for surface water/groundwater interactions/exchanges at depths of around 20 m (Bartsch et al., 2014; Bravo et al., 2002; Delsman et al., 2018; Engelhardt et al., 2013; Kurtz et al., 2014; Liu et al., 2019; Ma et al., 2012; Munz et al., 2017; Rau et al., 2010; Read et al., 2013; Shanafield and Cook, 2014; des Tombe et al., 2018), with the exception of the works of Masbruch et al. (2014) and Irvine et al. (2015), who present works for deeper environments (i.e. a few hundreds of metres). Cases where temperature observations have been taken at relatively large depths (i.e. a few hundreds of metres) and used as observations for heat transport simulations are scarce (Masbruch et al., 2014). Several studies in which temperature profiles are measured are typically used for qualitative interpretations of the effects of anthropogenic stressors (Benz et al., 2018; Dong et al., 2018) and for the evaluation of deep (i.e. hundreds to thousands of metres) geothermal activities where substantial temperature variations occur (Dentzer et al., 2017; Majorowicz and Grasby, 2020; Marty et al., 2020; Sippel et al., 2013; Smith and Elmore, 2019). Little research has, however, been devoted to exploring the use of temperature in aquifers with depths in the range of a few tens of metres up to a few hundreds of metres in which the temperature range is rather limited (up to a maximum of ±10°C). Moreover, in a lowland setting, hydraulic gradients are typically smaller, and it is unclear whether the advection of heat at the catchment scale can be sufficiently large to enable the use of temperature.

For the Neogene Aquifer in Flanders, groundwater flow and transport models have been developed in the framework of safety and feasibility studies for the underlying Boom Clay formation as a potential host formation for the geological disposal of radioactive waste (Gedeon, 2008; ONDRAF/NIRAS, 2010, 2013; Rogiers et al., 2015; Vandersteen et al., 2013). However, the simulated fluxes by these models are still subject to large uncertainties as they are typically constrained by hydraulic heads only. While the evaluation of candidate host formations continues, this study investigates how heat transport is affected by groundwater flow in a sedimentary Neogene Aquifer, across the Nete catchment, in Belgium. Additionally, emphasis is put on the disturbances in heat transport by the presence of faults, highlighting the Rauw fault – a 55 km long normal fault. To this end, we use the state-of-the-art 3D groundwater flow model presented by Casillas-Trasvina et al. (2021) and gather temperature–depth (TD) profiles spread over the catchment, land surface temperature (LST) satellite image data (NASA’s Moderate Resolution Imaging Spectroradiometer, MODIS), and a paleo-temperature time series to set a transient temperature boundary condition. This case study will help in assessing the usefulness of temperature data in a catchment-scale lowland setting to characterize the magnitude and patterns of groundwater flow. This approach seems especially suitable within the framework of the disposal of radioactive waste as the idea is to learn as much as possible from measurements from low-to-non-invasive techniques. This work serves as a case study attempting to integrate the information provided.
by temperature data as an additional unconventional state variable at the catchment scale in a quantitative way, with the objective of further constraining the numerical models that serve as summaries of our system understanding.

2 Study area and hydrogeological setting

The Neogene Aquifer is located in the Campine area, in the northeast of Flanders, and is considered to be the most important groundwater reservoir in the region (Coetsiers and Walraevens, 2006). A full description of the study area has been presented elsewhere (Casillas-Trasvina et al., 2021). For clarity, only a brief summary is given here. The area of the Neogene Aquifer within the Nete catchment is studied in this work, as shown in Fig. 1. The area is characterized by a low relief, with altitudes ranging from \( \sim 5 \) to 70 m (above TAW, Tweede Algemene Waterpassing) along a west–east gradient. As a result, the hydrography is characterized by an east–west drainage system that belongs mainly to the Scheldt river basin (Van Keer et al., 1999; Rogiers et al., 2014). The aquifer is mainly composed of Neogene marine sand deposits that induce some variations in hydrochemical composition, although the groundwater is weakly mineralized (Coetsiers and Walraevens, 2008). The Oligocene (Rupelian) and Mio-Pliocene geology of the study area is presented in Sect. 2. The lithology mainly consists of fine-to-medium-grained sands, while the clay content is found to vary in certain units (e.g. Kasterlee, Lillo, and Diest formations), and basal gravels are sometimes present between the units (Laga et al., 2001). The sediments dip towards the north–northeast, with a gentle slope of about 1 \( \% \)–2 \( \% \), with some disturbances towards the west by different normal faults. In the eastern/northeastern part of the study area, faults occur that were formed as a consequence of the development of the Roer Valley Graben (RVG), which is the northwestern-most part of the Lower Rhine Graben (LRG; Verbeeck et al., 2017; Deckers et al., 2018). The most important of these faults outside the proper RVG, and in terms of Cenozoic offset, is the Rauw fault, which is proven to have been active during the Pleistocene (Verbeeck et al., 2017). The Rauw fault has a displacement of more than 7 m in the Quaternary, which increases with depth. It does, however, not have a clear surface expression. The Rauw fault consists of two separate branches (i.e. Rauw fault and Rauw east), around 700 m apart, for which the movement of the Rauw east fault stopped earlier, and Rauw fault has taken over the activity (Verbeeck et al., 2017). For a more detailed description, please refer to Verbeeck et al. (2017).

Quaternary deposits of varying texture overlie the Neogene units and constitute the upper few metres of the aquifer system. The hydrostratigraphic units occurring below the Quaternary are composed of Pleistocene and Pliocene sediments. These consists of the Malle, Merksplas, Mol, and Lillo sands, sitting on the Kasterlee Formation, a mixed clayey-sandy formation deposited in shallow marine to estuarine conditions. It is followed by the Diest Sands, overlaying the Lower Miocene Berchem Sands and Late Oligocene Voort Sands. The Boom Clay, an Oligocene marine sediment, forms the lower boundary of the system. For a more detailed description of the hydrostratigraphy of the area and the Boom Clay, please refer to Laga et al. (2001), Coetsiers and Walraevens (2008), Yu et al. (2013), and Vandenbergh et al. (2014). Notwithstanding their lithological differences, Patyn et al. (1989) concluded from hydrogeological observations that these sediments behave as a single aquifer. Similar to Casillas-Trasvina et al. (2021), in this work, the combined Quaternary deposits, Pleistocene, and Upper and Lower Pliocene aquifers, together with the Lower Miocene and Oligocene aquifer system, are referred to as the Neogene Aquifer.

Various subsurface and surface activities are taking place or being planned above and below the Neogene Aquifer, including, but not limited to, surface and geological disposal of low-level and short-lived and long-lived nuclear waste. The Nete catchment and the Neogene Aquifer have been subject to studies in the framework of geological disposal of nuclear waste (Beerten et al., 2010; Gedeon, 2008; Mallants, 2010; Mazeurek et al., 2009; ONDRAF/NIRAS, 2010, 2013).

3 Methodology

3.1 Data collection

3.1.1 Temperature–depth (TD) profiles

Temperature–depth (TD) profiles have been collected in the framework of various projects and campaigns. Historical data were compiled by Rogiers et al. (2015) from previous campaigns, as presented in Table 1. The quality of the TD profiles taken from 1988 to 1997 is unknown since the sensor resolution and accuracy, together with the performed logging speed (up to four logs per day), are not known, potentially inducing considerable bias in the obtained temperatures. The location of these TD profiles is presented in Fig. 2. The TD profiles collected during the site characterization for the cAt project (ONDRAF/NIRAS, 2010) were not considered during the assessment by Rogiers (2014) since they were believed to be influenced by drilling activities, and thermal equilibrium was not reached yet.

The four TD profiles reported by Rogiers (2014) were the only data for which the type and characteristics of the sensor and logging speed are known up to that date. For these wells, a correction was applied to the raw data to account for the logging error. The properties of the used sensor are presented in Table 2. Another five TD profiles were taken in a second campaign presented by Rogiers et al. (2015). A different temperature logger was used, for which the characteristics are presented in Table 2. The stop–go measurement
Figure 1. Geographical location of (a) the Nete catchment within Belgium with an indication of the faults in the Roer Valley Rift System, as in Vanneste et al. (2013), and (b) the land surface temperature averaged for January 2001 derived from satellite data (MODIS) for the study area within the Nete catchment.

Figure 2. Plan view of the study area as discretized in the second layer of the numerical model. It indicates the faults (emphasis on the highlighted Rauw fault), cross section, temperature–depth profile locations, and modelled formations derived from the hydrogeological 3D model in Deckers et al. (2019).

method by Harris and Chapman (2007) was followed, similar to the previous campaign.

The temperature probe used in this work was the Star-Oddi DST milli-TD (https://www.star-oddi.com/products/data-loggers/depth-sensor-water-level-data-logger-recorder-milli-TD, last access: 7 November 2022). This probe is typically used in TD logging by tagging fish in migration and behaviour studies. It was selected due to its low diameter, making it ideal to overcome a sort of bottleneck observed in a previous campaign. The dimensions of this probe are presented in Table 2, which summarizes the temperature sensors used.
Table 1. Compiled historical TD profile data.

<table>
<thead>
<tr>
<th>Campaign</th>
<th>Month</th>
<th>Year</th>
<th>TD profiles measured</th>
</tr>
</thead>
<tbody>
<tr>
<td>Belgian Geological Survey</td>
<td>March</td>
<td>1988</td>
<td>5</td>
</tr>
<tr>
<td>SCK CEN</td>
<td>August</td>
<td>1995</td>
<td>3</td>
</tr>
<tr>
<td>SCK CEN</td>
<td>February</td>
<td>1996</td>
<td>1</td>
</tr>
<tr>
<td>PHYMOL project</td>
<td>February</td>
<td>1997</td>
<td>2</td>
</tr>
<tr>
<td>Exploratory temperature logs (Rogiers, 2014)</td>
<td>February</td>
<td>2012</td>
<td>4</td>
</tr>
<tr>
<td>Rogiers et al. (2015)</td>
<td>August</td>
<td>2015</td>
<td>5</td>
</tr>
<tr>
<td>Current work</td>
<td>October</td>
<td>2019</td>
<td>15</td>
</tr>
</tbody>
</table>

Table 2. Summary of temperature sensors used for different campaigns.

<table>
<thead>
<tr>
<th></th>
<th></th>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td>Probe model</td>
<td>Schlumberger Mini-Diver</td>
<td>Solinst Levelogger Edge (M200)</td>
<td>Star-Oddi DST milli-TD</td>
</tr>
<tr>
<td>Dimensions (mm)</td>
<td>Ø 22 × 90</td>
<td>Ø 22 × 159</td>
<td>Ø 13 × 39.4</td>
</tr>
<tr>
<td>Weight (g)</td>
<td>70</td>
<td>129</td>
<td>9.2</td>
</tr>
<tr>
<td>Range (°C)</td>
<td>−20 to 80</td>
<td>−20 to 80</td>
<td>0 to 85</td>
</tr>
<tr>
<td>Accuracy (°C)</td>
<td>±0.1</td>
<td>±0.5</td>
<td>±0.1</td>
</tr>
<tr>
<td>Resolution (°C)</td>
<td>0.01</td>
<td>0.003</td>
<td>0.03</td>
</tr>
</tbody>
</table>

in this and in previous works. The measuring method described by Bense et al. (2017) and Kurylyk et al. (2018) was followed. In this work, 15 groundwater wells were measured to extend the number of observations spread across the Nete catchment. Finally, a total of 23 different locations were measured with a total of 35 TD profiles (some filters, i.e. R-12 e-f, R-15 b-f, R-16 c-d, R-1 b-c, and R-54 a-c-f, are in the same borehole), ranging from near the surface to roughly −450 m deep and distributed in time and space across the Nete catchment, are used as observations to constrain the heat transport model (see Appendix A1).

3.1.2 Air and soil temperature time series

Several meteorological stations exist within the region, and only three provide measured values of air (at 1.75 m above surface) and soil temperature (measured at 5 cm below surface) within or nearby the study area, i.e. Herentals, Overpelt, and Eindhoven stations. Data from the two stations in the Flemish region, belonging to the Flanders Environment Agency (VMM; https://www.vmm.be/, last access: 7 November 2022), were accessed through https://www.waterinfo.be/ (last access: 7 November 2022). Data from Eindhoven were obtained through the Royal Netherlands Meteorological Institute (KNMI; https://www.knmi.nl/home, last access: 7 November 2022). The temperature–time series for the data gathered for these stations from 1 January 2003 to 31 October 2016 is presented in Appendix A2.

3.1.3 Remote sensing data

Remote sensing data play an important role in the development of validated multi-scale Earth system models (Fick and Hijmans, 2017; Hazaymeh and Hassan, 2015; Tomlinson et al., 2011). Several satellite missions launched by NASA and the ESA (European Space Agency) collect data from the Earth’s crust up to the atmosphere on temperature, gravity variations, landscape characteristics, etc. Missions such as Landsat, ASTER, MODIS, and Copernicus collect, for example, land surface temperature (LST) values, which can be freely retrieved from their online platforms. Landsat, Copernicus, Meteosat Second Generation (MSG), and MODIS missions gather global LST values. However, MODIS gathers 1 km resolution images, whereas Copernicus images have a 5 km resolution. For this reason, LST images were retrieved from NASA’s MODIS Terra for the region, with 8 d and 8 night averages from 2001 to 2019. These were upscaled to monthly means for our purposes. The resulting raster stack was sampled for the 2001–2019 monthly period in two locations, corresponding to the VMM weather station soil temperatures measured at Herentals and Overpelt. These results are also included in Appendix A2. From this sampling, it is clear that the raster LST images correlate to the soil temperatures from the meteorological station, with the peak temperature values mostly during the summer months. Reinart and Reinhold (2008) discuss this effect on peak values, similarly, by comparing MODIS-retrieved temperature values and in situ measurements. They attribute these higher values to the so-called “skin effect”, since satellite radiome-
ters are only able to retrieve skin temperatures (Reinart and Reinhold, 2008) which are influenced by several other components, i.e. prevailing wind speed, time, and conditions on the day the measurement is taken (Fick and Hijmans, 2017; Tomlinson et al., 2011; Wan et al., 2002). A simple linear regression model was fitted to estimate the LST raster stack in relation to in situ soil temperature measurements from the weather stations (LST_{cor} = 0.6022 LST + 3.162; R^2 = 0.8; standard errors of slope 0.025 and intercept 0.46; see Appendix A3). Although there seems to be some seasonal hysteresis in the relation between the two, and one could potentially use, for example, the month as an additional regressor, we considered this to be sufficiently accurate for the current work. This temperature estimation was then applied to every LST raster (i.e. every month from 2001 to 2019). An example of the estimated monthly images for year 2001 is shown in Appendix A4. From the estimated monthly images, it is possible to observe the temporal but also the spatial distribution across the catchment.

### 3.1.4 Holocene paleo-temperature reconstruction

To advance the conceptualization of the transport model and to improve the initial boundary condition for the top of the model (i.e. temperature values), a long-term temperature–time input curve was built using data from several authors. The data were retrieved from the paleo-climatology data library of the National Centers for Environmental Information (NCEI; https://www.ncdc.noaa.gov/data-access/paleoclimatology-data, last access: 7 November 2022). Data from D’Arrigo et al. (2005, 2006), Casty et al. (2007), Buntgen et al. (2011), Tingley and Huybers (2013), Esper et al. (2014), Luterbacher et al. (2016), Langevin et al. (2017), Marsicek et al. (2018), Ljungqvist et al. (2019), and Glaser and Riemann (2009) were used for the estimation of the LST and temperature curve. The data from these authors present paleo-temperature reconstruction values for the Holocene in Europe in gridded format with coordinates for reference. Data from Mann et al. (2009) and Mann (2002) were used for comparison, since they represent a general and internationally accepted Northern Hemisphere temperature–time curve. The values from these authors were either taken from the gridded value or from values that contained Flanders within its/their boundaries. The obtained values from each author curve were then scaled to the relative Belgian average temperature for the period 1961–1990, as derived from Jones et al. (1999, 2012). These temperature–time curves are shown in Appendix A5. Inverse distance weighting (IDW: 1/d_{ij}^{0.5}) was performed on the data locations with reference to the core of the groundwater model domain. The resulting weights were used to determine an IDW average time series, for which the yearly (IDW – 1 year) average is displayed in Fig. 3. Recursive partitioning segmentation (RPS) was done using an implementation of the Lavielle (1999, 2005) method. This method was applied to different sections of the IDW-averaged temperature–time curve, aiming to obtain a segmentation from −8500 to 2000 (CE, common era) using average step sizes of 1 kyr (−8500 to −500), 100 years (−501 to 1500), 10 years (1501 to 1950), and 1 year (1951 to 2000). Monthly values for the last 19 years from remote sensing land surface temperature data were included at the end of the segmentation (see Fig. 3). A total of 90 time steps are then included to the steady-state solution to determinate the initial temperature, thus aiming to improve its efficiency.

### 3.2 Modelling framework

For the current work, the Neogene Aquifer model (NAM), a steady-state groundwater flow model constructed by Casillas-Trasvina et al. (2021) using MODFLOW-2005 (Harbaugh, 2005), is used. A heat transport model by Rogiers et al. (2015) has been updated using MT3D-USGS (Bedekar et al., 2016) and is used for simulations. The models were developed and post-processed using the RMODFLOW (Rogiers, 2015, 2016a) and RMT3DMS (Rogiers, 2015, 2016b) packages for R (R Core Team, 2020).

#### 3.2.1 Conceptual model

The NAM by Casillas-Trasvina et al. (2021) has a lateral boundary that coincides with the catchment boundaries of the Kleine and Grote Nete rivers. Similar to Gedeon (2008), the catchment is assumed to be laterally isolated, so no groundwater flow across the lateral boundary occurs. The top boundary is put at the ground surface elevation, while the bottom boundary coincides with the top of the Boom Clay formation (from a few metres below the surface up to roughly 450 m deep). The Neogene Aquifer becomes deeper in the northeasterly direction from the southwestern corner where the Boom Clay is present at the ground surface. The groundwater flow in the Neogene Aquifer is driven mainly by surface hydrological features (i.e. recharge and rivers) creating several local flow systems with local influences of abstraction wells. It is assumed to be in dynamic equilibrium, with no long-term trends in groundwater flow, which allows for a steady-state simulation (Gedeon, 2008). Casillas-Trasvina et al. (2021) included the conceptual model consideration that fault planes have hydraulic properties distinct from the surrounding sediments and do not penetrate up to the surface as they are buried by at least a thin Quaternary cover. Additionally, these planes are assumed to act as barriers if they behave in a significantly different way to the surrounding sediments, thereby affecting the horizontal component of the groundwater flow. Together with their buried nature, this causes upward flow at the upgradient side of the fault. On the down-gradient side of the fault, the vertical flow direction may also be affected, turning into a more parallel direction in respect to the fault. Finally, groundwater may be overflowing above the fault plane if it acts as a strong flow barrier, given the
fact that the fault is buried, creating a groundwater level step (head difference).

### 3.2.2 Groundwater flow model

The structural updates into NAM for heat transport follow the latest hydrostratigraphic 3D model for Flanders (H3D) by Deckers et al. (2019). Similar to NAM by Casillas-Trasvina et al. (2021), it assumes nine hydrostratigraphic units, i.e. non-tabular Quaternary (Quaternary cover above the Kempen aquifer system), Weelde, Malle, Merksplas, Mol, Lillo, Kasterlee, Diest, and Berchem and Voort, of which the geometry is based on Deckers et al. (2019). The model domain is discretized into 49 vertical layers that thin out closer to the surface to ensure smaller modelling cells close to surface hydrological features where groundwater gradients are the highest. The bottom of layers 1 and 2 were assigned equal to 50% and 30% of the ground surface elevation. The bottom of layer 3 was set equal to 0 m a.s.l. (above sea level; TAW). From layer 4 to layer 9, layer thicknesses of 5 m are used. From layer 10 to 49, the thickness used is 10 m. The modelled area was discretized into a regular model grid of 96 rows and 146 columns, resulting in cells with dimensions of 400 m × 400 m. A total of 23 faults were simulated with the horizontal flow barrier (HFB) package (Harbaugh et al., 2005), starting from the top of the second numerical layer (from around 12 to 18 m below surface) to the bottom of the modelled domain, given that the faults do not present a clear surface expression (Verbeeck et al., 2017). Rivers, lakes, canals, and abstraction wells are defined in the groundwater flow model. The groundwater abstractions range from a few cubic metres per day (m$^3$ d$^{-1}$) to more than 300 m$^3$ d$^{-1}$. Data from several sources were used to define these parameters, including the Flemish hydrographic atlas (Vlaamse Hydrografische Atlas, VMM 2017) and the IGN/NGI dataset (IGN/NGI 2017). Spatially distributed recharge is implemented with values obtained from DiCiacca (2020), which are derived from vegetation cover, soil texture, and spatial input layers of depth to the groundwater table, based on hydraulic head observations. A scaling factor was used during the model inversion for the calibration of the recharge initial value. A total of 1393 averages from hydraulic head time series are used in the NAM. These observations were obtained from the piezometric network monitored and maintained by SCK CEN for ONDRAF/NIRAS.
and from the subsurface database for Flanders (Databank Ondergrond Vlaanderen, DOV). The model performance of the NAM model had a root mean square error (RMSE) of 0.70 m, accounting for a total head loss of 46.5 m. Figure 4a–c show results from the groundwater flow model, including (a) the hydraulic head distribution over the Nete catchment, (b) the scatterplot of simulated equivalents vs. observed hydraulic head observations, and (c) a cross section (B–B’) with arrows indicating the flow magnitude results from the groundwater flow model per model cell. Appendix A6 shows a cross section with derived Péclet numbers ($Pe$), following Freeze and Cherry (1979; $Pe = \frac{Ve}{De}$), relating the rate of flow advection to the diffusion. In here, a value of $2 \times 10^{-9}$ m$^2$ s$^{-1}$ for the diffusion coefficient ($De$) is used, corresponding to the tritium diffusion coefficient in pure water (Gedeon, 2015), and a value of $L = 200$ m, corresponding to a rough average distance from the bottom of the aquifer to the surface, is used. It is notable that $Pe \gg 1$ is in the formations above Berchem and Voort sands, thus being more advection–dispersion-dominated areas. On the other hand, in the deepest areas of the aquifer (i.e. Berchem and Voort sands), $Pe \leq 1$, and therefore diffusion becomes an important transport mechanism. The values of $Pe \leq 1$ are very few, mostly near to 1, and found near the bottom of the aquifer where groundwater flow velocities are at the lowest. In these areas, the fluxes occur near the bottom no-flow boundary of the model. Research has been previously performed, as summarized by Vandersteen et al. (2013), which pointed out the low hydraulic conductivity values for the Berchem and Voort formations (as low as 0.02 m d$^{-1}$), which are thus not particularly acting as a barrier/clay but have modest advective/diffusive behaviour. A more detailed description of the NAM structure, parameters, and results is presented in Casillas-Trasvina et al. (2021).

### 3.2.3 Heat transport model development

#### Numerical modelling

In Liu et al. (2019) and Hecht-Méndez et al. (2010), it is demonstrated that MT3DMS is able to simulate heat transport, when temperature variations are limited, through the analogy with solute transport. If large variations in temperature are to be modelled (i.e. an aquifer thermal energy storage (ATES) and a ground source heat pump (GSHP) with large...
where $T$ (L (M L$^{-3}$)) is used to simulate heat transport, then these equivalent transport parameters are defined. The heat dispersion is the same as solute dispersion, which is determined by longitudinal and transversal dispersivities. For a more detailed description, the reader is referred to Anderson (2005), Langevin et al. (2008), Zheng (2010), Bedekar et al. (2016), des Tombe (2015), and Liu et al. (2019).

### Modelling set-up

A heat transport model of the original NAM (Gedeon, 2008) using MT3DMS (Zheng, 2010) has been constructed by Rogiers (2015). In this model, the top boundary condition uses a yearly average value for the temperature imposed on the ground surface. The transport simulation covered a period of 10519 years with yearly time steps, covering almost the entire Holocene period. The set temperature as a boundary condition at the surface accounted for the topographical influence and land cover classes, based on data presented by Leterme and Mallants (2012). A heat flow density of 0.06 W m$^{-2}$ was set to the bottom boundary condition, with the value being in accordance with the average heat flow density reported by Pollack et al. (1993) for Cenozoic and Mesozoic sedimentary and metamorphic crust (Rogiers et al., 2015). The specific heat capacity, $\varphi C_p$ fluid, and density, $\rho_w$, of the fluid (groundwater) were set to 4193 J kg$^{-1}$ K$^{-1}$ and 999.7 kg m$^{-3}$, respectively (Rogiers et al., 2015). Material properties defined in the model were (a) total porosity ($\theta$), (b) specific heat capacity ($\varphi C_p$ fluid), (c) density of the solid ($\rho_s$), (d) thermal conductivity ($\lambda_m$), and dispersivity ($\alpha_L$).

In this work, similar to Rogiers et al. (2015), we based the material properties on the reported literature values found in Hoes et al. (2005), Beerten et al. (2010), Chen et al. (2011), Govaerts et al. (2011), and the WTCB methodology reported in van Lysebetten et al. (2013). These values and a summary of the hydrologic properties required to parametrize the MT3D-USGS model used in the simulations are shown in Table 3. Note that the potential impact of geothermal heat flow

### Table 3. Overview of the hydrogeological unit properties used for the updated version of the NAM heat transport model in MT3D-USGS

<table>
<thead>
<tr>
<th>Model units</th>
<th>$n$ (%)</th>
<th>$C_s$ (J kg K$^{-1}$)</th>
<th>$\rho_s$ (g cm$^{-3}$)</th>
<th>$\lambda_m$ (W m K$^{-1}$)</th>
<th>$\alpha_L$ (m)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Non-tabular Quaternary</td>
<td>35.9</td>
<td>800</td>
<td>2.62</td>
<td>1.5</td>
<td>6</td>
</tr>
<tr>
<td>Weelde</td>
<td>35.9</td>
<td>800</td>
<td>2.66</td>
<td>2.47</td>
<td>6</td>
</tr>
<tr>
<td>Malle</td>
<td>40.2</td>
<td>800</td>
<td>2.66</td>
<td>2.47</td>
<td>6</td>
</tr>
<tr>
<td>Merksplas</td>
<td>40.2</td>
<td>800</td>
<td>2.66</td>
<td>2.47</td>
<td>6</td>
</tr>
<tr>
<td>Lillo</td>
<td>40.2</td>
<td>800</td>
<td>2.66</td>
<td>2.47</td>
<td>6</td>
</tr>
<tr>
<td>Mol</td>
<td>40.2</td>
<td>800</td>
<td>2.66</td>
<td>2.47</td>
<td>6</td>
</tr>
<tr>
<td>Kasterlee Clay</td>
<td>39.7</td>
<td>800</td>
<td>2.65</td>
<td>2.1</td>
<td>6</td>
</tr>
<tr>
<td>Diest</td>
<td>41.1</td>
<td>800</td>
<td>2.75</td>
<td>2.1</td>
<td>6</td>
</tr>
<tr>
<td>Berchem and Voort</td>
<td>39.5</td>
<td>800</td>
<td>2.83</td>
<td>2.1</td>
<td>6</td>
</tr>
<tr>
<td>Boom Clay</td>
<td>37.3</td>
<td>769</td>
<td>2.65</td>
<td>1.31</td>
<td>6</td>
</tr>
</tbody>
</table>

ranges), then a model that accounts for variable density and viscosity terms (e.g. SEAWAT, Langevin et al., 2008; SU-TRA, Voss and Provost, 2010; FEFLOW, Trefry and Muffels, 2007) should be applied. But, in the current setting, we do not consider this to be relevant. The heat and solute transport equations formulated in analogous forms show how MT3DMS can be used for heat transport simulations. For single species, the solute transport equation solved by MT3DMS (Langevin et al., 2008) as follows (Eq. 1):

$$
\left[1 + \frac{\varphi_b K_d}{\theta} \right] \frac{\delta(\theta C)}{\delta t} = \nabla \cdot \left[ \theta \left( D_m + \alpha \frac{q}{\theta} \right) \nabla C \right] - \nabla (q C) - q_s C_s, \tag{1}
$$

where $\varphi_b$ is the bulk density (mass of the solids divided by the total volume; M L$^{-3}$), $K_d$ is the distribution coefficient of the solute (L$^2$ M$^{-1}$), $\theta$ is the porosity (–), $C$ is the concentration of the solute (M L$^{-3}$), $\nabla$ is the Vector differential operator, $D_m$ is the molecular diffusion coefficient (L$^2$ T$^{-1}$), $\alpha$ is the dispersivity tensor (L), $q$ is the specific discharge vector (L T$^{-1}$), $C_s$ is the source concentration of the solute (M L$^{-3}$), and $q_s$ is the source or sink of the fluid (T$^{-1}$).

The analogous equation (Eq. 2) for the heat transport is defined as follows:

$$
\left[1 + \frac{1 - \theta}{\varphi C_p \text{ solid}} \right] \frac{\delta(\theta T)}{\delta t} = \nabla \cdot \left[ \theta \left( K_{\text{bulk}} + \alpha \frac{q}{\theta} \right) \nabla T \right] - \nabla (q T) - q_s T_s, \tag{2}
$$

where $T$ is the temperature (°C), $\rho_s$ is the density of the solid (M L$^{-3}$), $\varphi C_p \text{ fluid}$ is the specific heat capacity of the fluid (L$^2$ T$^{-2}$ °C$^{-1}$), $\varphi C_p \text{ solid}$ is the specific heat capacity of the solid (L$^2$ T$^{-2}$ °C$^{-1}$), $K_{\text{bulk}}$ is the bulk thermal conductivity of the aquifer (ML$^3$ T$^{-2}$ °C$^{-1}$), and $T_s$ is the source temperature (°C).

When the more recent MT3D-USGS code (Bedekar et al., 2016) is used to simulate heat transport, then these equivalent transport parameters are defined. The heat dispersion is

https://doi.org/10.5194/hess-26-5577-2022

Hydrol. Earth Syst. Sci., 26, 5577–5604, 2022

A. Casillas-Trasvina et al.: Characterizing groundwater heat transport in a complex lowland aquifer
of an increasing temperature with depth on viscosity and density of the fluid is not considered. This assumption is reasonable, as we consider a not so deep groundwater flow system (< 500 m) in which these temperature effects will be of minor importance (Bense and Person, 2006; Person et al., 1996) and where emphasis is not put on surface water-groundwater interactions when the impact on viscosity derived from temperature gradients may be considerable (Engelhardt et al., 2013; Kurtz et al., 2014; Liu et al., 2019; des Tombe et al., 2018).

Building on the previous work, the latest version of the heat transport model built by Rogiers et al. (2015) has been updated from MT3DMS (Zheng, 2010) to MT3D-USGS (Bedekar et al., 2016), following the new hydrostratigraphy and grid discretization used for the updated NAM groundwater flow model by Casillas-Trasvina et al. (2021). The dispersivity was based on a scale of 200 m, which is a rough measure for the depth of the aquifer, and thus the travel distance of heat. The linear dispersivity–scale relationship was derived from the literature data, as presented by Rogiers et al. (2013). For simplicity, the dispersivity is considered to be constant, with longitudinal dispersivity $\alpha_L$ equal to 6 m, and the ratios of transverse and vertical transverse dispersivity to $\alpha_L$ equal to 0.1 and 0.01, respectively.

Top boundary condition

As the first modelling step, a steady-state finite difference solution was obtained using temperature from $-8500 \, \text{ka}$ as the top boundary condition. This was used as the initial temperatures for transient simulations. The transient model is run for 10,519 years, from $-8500 \, \text{ka}$ to 2019. Every single time step is derived, as explained above in Sect. 3.1. A yearly average value from the LST raster stack was determined to be used as the spatial distribution for the Holocene paleo-temperature distribution implemented in the model as a top boundary condition. The temperature spatial distribution was derived from satellite data and used directly in the monthly part of the transient simulation. The 1 km LST rasters were downsampled to 100 $\times$ 100 using the (raster) package function disaggregate in R. Then, the mean values within a modelling cell was estimated for every 1 km LST raster that were sampled in the centre of each active model cell (MODFLOW model) with a 400 m $\times$ 400 m resolution. A total of 316 rasters were produced from $-8500 \, \text{ka}$ to 2000, and one for each month, from January 2001 to October 2019. Each of these rasters represent a time step for the transient transport simulation.

Paleo-temperature boundary condition test

To increase the accuracy of the simulated temperature and hence have a good simulated vs. observed temperature performance, the simulated temperature has to be able to stabilize (or reach a steady state). Thus, we require the estimation of initial conditions (initial temperature values for each stress period).

To show the importance of performing the paleo-temperature simulations to provide initial conditions for further stress periods, a test was performed. The (disturbed) simulations increasing by 1 °C in temperature at the initial top boundary condition at various time steps (i.e. 10 000, 9000, 8000, 7000, 6000, 5000, 4000, 3000, 2000, 1000, 900, 800, 700, 600, 500, 400, 300, 200, and 100 ka) and for the remaining of the simulation period were performed. The simulated temperature–depth (TD) profile from all these models was obtained for the same location (on well R-54f). The results of this test are shown in Fig. 5. In Fig. 5a, the simulated TD profile at the last time step (present time) from a normal (undisturbed) forward model run is shown. In Fig. 5b, the figure shows the differences between the disturbed and normal (undisturbed) simulated TD profiles, pointing out the time required for the model to reach a steady state, given a change of 1 °C in the top boundary temperature condition. This supports the use of a relatively long time series, which is actually a process that is not computationally intensive. The model has 337 stress periods in total and runs for around 40 min. For the nine time steps in the stress period between 10 000 and 2000 ka (in steps of 1000 years) it requires around 1 min (approximately 64 s) to compute a temperature stabilization of up to 30 % per temperature degree of change in the temperature top boundary condition. Having a higher-frequency temperature distribution is not within the scope of this work, as we are more interested in knowing what improvement the implementation of an additional state variable may bring to the conceptualization of the groundwater system rather than aiming for higher-frequency results for which monthly time steps seem reasonable.

Similar to what has been done in previous work in this research area (Casillas-Trasvina et al., 2021; Gedeon, 2008; Rogiers et al., 2015), and as indicated in Sect. 3.2.1, the aquifer is assumed to be in dynamic equilibrium with no long-term trends in groundwater fluxes, which allows us to simulate the groundwater flow in a steady state. For the purposes of our research, we find this assumption acceptable.

Model calibration

Automatic parameter optimization was implemented as a technique for model inversion. Several algorithms were used for global model optimization such as the Standard Particle Swarm Optimization (SPSO11; Clerc et al., 2012; Zambrano-Bigiarini et al., 2013; Zambrano-Bigiarini and Rojas, 2013) and differential evolution (Mullen et al., 2011). Temperature gradients, calculated by fitting a cubic smoothing spline with a knot distance of 1 m derived from each TD profile, and hydraulic head values were used as calibration targets during the joint inversion procedure. Results from the sps011 algorithm were selected for minimizing the RMSE despite requiring a slightly longer computational time.
3.2.4 Modelling strategy

Model 1: baseline

The updated version of the groundwater flow model by Casillas-Trasvina et al. (2021) and the heat transport model, including the Holocene temperature–time curve and satellite data for the last 19 years, is defined here as the baseline model (Paleo RPS). The heat transport simulation is jointly inverted with the groundwater flow model against temperature gradients and hydraulic head observations, respectively. The results from this model are compared against model case 2 and model case 3 (see below) to estimate the effect of thermal conduction and normal faults in the temperature distribution in the Neogene Aquifer across the Nete catchment. Additionally, a model without the effect of the paleo-temperature and including only the LST temperatures of the last 19 years (named the monthly LST model) and a steady-state transport model are also evaluated in comparison to the defined baseline model.

Model 2: thermal conduction

A model case accounting for only the thermal conduction is performed to enable the quantification of the effect of groundwater flow on the subsurface temperature distribution. The baseline simulation accounts for both the advection of heat via the groundwater flow and thermal conduction. For this second simulated model without groundwater flow, the same parameter set is used, but the advection of heat is switched off. The difference between the first and second simulation temperature distributions then gives us the changes in temperature induced by the groundwater flow.

Model 3: heat transport without faults

An exploration of the temperature distribution in the aquifer without faults (i.e. no HFB included in the flow model by Casillas-Trasvina et al., 2021) is also performed. For the heat transport model, parameters are set to be the same as for the baseline model. Both this case and the baseline model account for both the advection of heat via groundwater flow and...
thermal conduction. The difference between both of these cases allows the quantification of the effect of faults on the temperature distribution and provides an idea of the parameter sensitivity of the hydraulic conductivity of such structures.

4 Results and discussion

4.1 Paleoclimate effect on temperature distribution

Every measured TD profile is compared with its respective simulated temperature based on its spatial and temporal (i.e. month and year) location. A scatterplot of the model-simulated temperature gradient values vs. the observed measurement is shown in Fig. 6a, where a comparison is made between the monthly LST model (see Sect. 3.2.4; model 1 – baseline) and the baseline model of the current work (Paleo RPS), separating the observations with a z value below and above –5 m a.s.l. Note that temperature/geothermal gradients are more important than absolute temperature values as they provide more insights to groundwater fluxes. The gradient is calculated by fitting a cubic smoothing spline with a knot distance of 1 m. A steady-state transport solution result is also included for reference. From this scatterplot, both transient heat transport simulations have a bias towards lower gradient values, indicative of flux magnitudes being potentially overestimated. Additionally, most of the simulated values above the $z = -5$ m a.s.l. limit show large fluctuations. This may be explained by the imposed temperature value as the top boundary condition (at the first layer of the model), the potentially heterogeneous properties of the materials near the surface, and the space–time discretization of the model (i.e. dimensions of the cells and time step sizes). The spatial variations in the imposed temperature (i.e. scaling of the thermal images/model cell size vs. the scale where TD profiles were measured), together with the local flows occurring near the surface of the model affected in very local scales (e.g. surface/groundwater exchanges), participate in these fluctuations that seem to stabilize at around –5 m a.s.l. depth. At these shallow depths, temperature variations can drive fluctuations in heat transport related to the fluid viscosity (Hecht-Méndez et al., 2010; Liu et al., 2019; des Tombe et al., 2018), which are limited to these few metres (~20 m) below the surface. Additionally, although yearly and monthly time steps were defined in the model, seasonal/diurnal temperature changes may also affect the propagation of heat into the subsurface (Benz et al., 2018; Dong et al., 2018). Nevertheless, the simulated results present an acceptable compromise between observed and simulated values. The paleo-temperature reconstruction model results are more clustered and less disperse, despite having some values spreading at a lower simulated equivalent gradient, in comparison with the LST transient model (see Fig. 6c), thus indicating that this simulation is, overall, slightly more accurately representing the observed situation (RMSE = 0.03°C m−1; 1.15°C), though several inaccuracies are observed in deeper parts of the aquifer (Fig. 6c).

4.2 Current temperature distribution

The flow and transport parameter composite scaled sensitivities (CSSs) with respect to hydraulic head and temperature gradients, respectively, are shown in Appendix A7, with the parameter names shown in Table A1. It can be seen that the
flow parameters have a much larger effect on the model results than the temperature parameters, implying their relevance to the produced results. In terms of temperature, the Diest Formation thermal conductivity presents the highest sensitivity, given that most of the temperature profiles were taken at these depths. On the contrary, the Berchem Formation (Berchem and Voort sands), which does not have a large number of observations, presents a relatively high sensitivity as well. Given that, at these depths, the magnitude of groundwater flow is very low (around \(1 \times 10^{-7} - 1 \times 10^{-5} \text{ m d}^{-1}\)), the transport of heat is less advective and more driven by thermal conduction. Nevertheless, in general, flow model parameter sensitivities are far larger than those for the heat transport model, given that thermal properties have less variation than the hydraulic ones (Saar, 2011).

The temperature distribution at the end of the calibrated heat transport model simulation is shown in Fig. 7. In Fig. 7a, a cross section B-B’ is shown, as indicated in Fig. 7b. Figure 7a shows a distribution that closely relates to groundwater flow, as seen in Fig. 4c, indicating the influence of advection. At the east of the Rauw fault, colder water infiltrates, travels downwards, and mixes with warmer groundwater from the bottom of the aquifer. Groundwater with a slightly higher temperature is transported upwards before penetrating the Rauw fault; however, most groundwater flows across the fault at around \(-150 \text{ m a.s.l.}\). Then, upstream of the Rauw fault, a portion of the groundwater flowing upwards overflows above it to then continue its flow direction downwards, mixing again with colder infiltrated water at the western end of the Rauw fault. Groundwater to the west of the Rauw fault then is mixed with colder and younger infiltrated water, reducing its temperature along the approximate flow direction (i.e. to the west) until the end of the fault block, and then continues flowing laterally to the west, merging with another more local flow system. This can be seen in Fig. 7b, which displays the model layer 10, where groundwater temperature to the west of the Rauw fault zone is higher than to the east. Groundwater in this zone is largely influenced by the surface water features carrying groundwater from deeper in the aquifer that has already crossed the Rauw fault. Groundwater to the west of the Rauw fault seems to have a lower temperature, given that mixing with warmer groundwater occurs deeper, near the bottom of the aquifer.

In the shallower parts of the aquifer (\(z > -100 \text{ m a.s.l.}\); Fig. 7a), some “isolated zones” with lower temperatures can be identified in the vicinity of other faults (i.e. from east to west; represented as black dashed lines). These isolated zones correspond to the groundwater flow distribution producing local flow systems in the area formed by the relatively large density of rivers and canals. Even though the faults in the vicinity have a relatively low sensitivity (see Appendix Fig. A7), these might still have an effect on the distribution of temperature. In their review, Irvine et al. (2012) mention that, although thermal properties might have low variations, as stated by Saar (2011), these variations might have a reasonable effect on the overall temperature distribution in sedimentary environments (Chang and Yeh, 2012; Constantz et al., 2003; Hidalgo et al., 2009), such as the Neogene Aquifer. On the other hand, in the deeper parts of the aquifer (near the bottom of the aquifer; Fig. 7a), higher temperatures are present (close to the main source of heat to the aquifer). At these depths, groundwater fluxes are low (\(\sim 1 \times 10^{-7} \text{ m d}^{-1}\)), which would indicate that the heat transport here would be mostly conduction driven. The discharge

\[\text{https://doi.org/10.5194/hess-26-5577-2022}\]
zone of the aquifer downstream to the Rauw fault location seems to have some influence from the surface water network driving deep groundwater to the surface. Groundwater abstractions in this zone seem to have an impact, added to the effect of the surface water network, driving heat from deeper parts of the aquifer into more superficial groundwater flow systems.

At the eastern side of the Rauw fault, colder temperatures are found from infiltrating water up to \( z \sim -150 \text{ m a.s.l.} \). At approximately this elevation, infiltrating colder water mixes with warmer groundwater from the bottom of the aquifer (Fig. 7a; green arrow), delineating an area where the mixing of younger and older groundwater seems to occur that appears to be defined by the flow velocities characteristic of the hydrostratigraphic formations (\( z \sim -200 \text{ m a.s.l.} \), at around the limit of the Diest Formation with Berchem and Voort formations). On the other hand, at the western side of the Rauw fault zone, an intercalation of the local flow systems is seen, following the east-to-west flow direction (Fig. 7a, cyan arrows, from east to west), as (i) the infiltration area with colder waters, (ii) an upwards groundwater flow driven by the surface water network and groundwater abstractions, (iii) the infiltration area with cold water, and (iv) an upwards groundwater flow. In the section where the aquifer becomes thinner at around 100 m thick (around the centre of the aquifer) towards the west, several even narrower local flow systems occur (as seen in Figs. 4c and 7a). Although these patterns are difficult to discern in the current figure, the majority of the temperature anomalies indicate upward fluxes, which can be explained by the relatively larger density of rivers and drains in the area.

Selected locations along the approximate flow direction where TD profiles were measured are shown in Fig. 8a. Their measured and calculated TD profiles for the three modelled cases (i.e. Paleo RPS, only conduction, and without faults) are shown in Fig. 8b (all measured TD profiles are shown in the Appendix). From east to west, the measured sites are R-13d (to the east of the Rauw fault), R-54c, R-11c, R-15f, R-2c, R-34c, R-1b, R-51c, and R-43c (to the west of the Rauw fault). All measurements show a C-shaped profile, which is an indication of transient conditions at the surface (Anderson, 2005; Bense et al., 2017). From Fig. 8b, the simulated and observed temperature values show that the closer the TD profiles were taken to the fault (R-54c being the closest one), the larger the absolute temperature error is for either of the three simulated cases. The simulated temperature at R-54c is lower than observed, which can be explained by its close proximity to the Rauw fault (of around 10 m apart) and thus being largely influenced by the flux across it, which might be overestimated, as indicated by the lower RMSE by the simulation case accounting only for the thermal conduction. The results from the simulation only accounting for the thermal conduction show the overestimation of the groundwater fluxes in few locations, which are mostly those located from the centre to the east of the studied area. Although the disparity in the model performance is relatively small, it shows the influence that the dense network of rivers and lakes in this zone has on the temperature distribution. Towards the west, where the density of rivers is lower and the aquifer becomes thinner, these differences become nearly negligible. The complete temperature–depth profile series simulated by the paleo-transport model is included (grey lines), indicating the range of potential temperature values that this model with these specific boundary conditions and heat source can achieve. These paleo-temperature simulated values, starting low, seem to stabilize after several transport steps, achieving an apparent relatively steady condition in the deeper parts of the aquifer (\( z < -150 \text{ m a.s.l.} \)). Similarly, as presented by Dong et al. (2018) and Benz et al. (2018), long-term temperature variations in the top boundary condition are potentially driven by external factors such as climate change or changes in land use/cover. These factors seem to have an influence on temperature observations at depths of around \(-50 \text{ m a.s.l.} \) or shallower, according to Fig. 8b. In Fig. 8a, the complexity of the river and canal network is visualized, and several lakes close to the measured sites can be observed. These lakes might explain anomalies occurring relatively close to the surface, but given that groundwater is mainly driven by the surface water network, they represent a likely driver for these shallow variations, potentially driving deeper/warmer waters upwards. Given that the simulated and measured temperature gradients are in fair agreement, the surface water network might have some influence on the temperature top boundary condition for which LST temperature images and the paleo-temperature time curve might not be accounting.

### 4.3 Temperature distribution cases

A comparison between the simulated temperature fields by the baseline (paleo-temperature reconstruction) model versus (a) thermal conduction only (no advection) and (b) without faults as HFB is shown in Fig. 9 and described here.

#### 4.3.1 Thermal conduction case

**Local flow systems patterns**

Removing the advection transport mechanism would allow temperature to be freely conducted without being transported by the groundwater flow. It is clear that advective groundwater flow has a large impact on the simulated temperature fields, and thus it is one of the main mechanisms for heat transport in the Neogene Aquifer (see Fig. 4). Given that advection is not being considered (no temperature convection), Fig. 9b clearly shows a temperature pattern that mimics the surface water network. This strongly supports the statement that heat transport in the aquifer may indeed be mainly driven by the advection in local flow systems, where temperatures increase near rivers and streams at the surface by
Figure 8. (a) Map showing selected TD profile borehole locations (magenta circles) along the approximate flow direction for R-13d, R-54c, R-11c, R15f, R-2c, R34c, R-1b, R51c, and R-43c. The background map source: © OpenStreetMap contributors, 2020. Distributed under the Open Data Commons Open Database License (ODbL) v1.0. (b) Temperature–depth profiles of selected boreholes and their filters (in lowercase letters) showing simulated values for the paleo-temperature model (baseline) and for simulation cases 2 and 3 (only thermal conduction and without faults, respectively). Modelled TD profiles for every time step of the baseline case are also included (in grey, paleo modelled).
Figure 9. Cross sections and plan view (model layer 10) showing the differences between the baseline model minus both cases, with panels (a, b) for the thermal conduction only case and panels (c, d) for the temperature distribution without faults.

deep groundwater discharge, and by large advection related to groundwater abstractions. The intention of this case is not to determine which case is more or less accurate than the other, as it is known that advective transport exists and not including it would not be realistic, but to highlight the spatial importance or dominance of the advective transport mechanism across the aquifer.

The interfluves are shown to be where most negative values are located, clearly indicative of higher temperatures in the purely conductive model. These areas are indicative of downward flows serving as recharge areas feeding the local flow systems which, in turn, are being discharged in more downstream areas by the surface water drainage effect. The effect of these local flow systems seems to depend on the depth of the aquifer. In the northeastern part of the study area, where the aquifer is thickest, the largest effects are present, with respect to the pure thermal conduction scenario, while in the southwest, where the aquifer is thin, there is in fact very little difference.

Depth of the aquifer vs. local flow system

Groundwater on the downstream side of the Rauw fault, in the upper part of the aquifer at around \( z = -50 \text{ m a.s.l.} \) (Fig. 9a), can be seen as warmer when advection is considered due to the Rauw fault overflowing effect and the areas below rivers. However, in the deeper parts of the aquifer, higher groundwater temperatures are shown for the thermal conduction case, suggesting that conduction may potentially be the dominant mechanism of heat transport (due
to the lower hydraulic conductivity; see above). Large temperature differences can be found in several locations where abstraction wells are located. The withdrawals made by these abstraction wells drive groundwater to flow upwards at several locations (see Fig. 4c), hence transporting heat at various depths from the bottom of the aquifer to shallower levels. The effect of the local groundwater flow systems driving heat from the bottom of the aquifer at several locations along the apparent flow direction is relatively widespread across the aquifer (as seen in Fig. 9a). However, although groundwater flow magnitudes are the lowest (~ 1 x 10^{-7} m d^{-1}) at the bottom of the aquifer, groundwater abstractions represent an important driver for the upwelling of deep warmer groundwater at these depths. By this groundwater upwelling, parcels of groundwater flow upwards in the surroundings of the abstraction well, which are then spread over the aquifer by the natural circulation of local groundwater flow systems.

### 4.3.2 Heat transport without faults

In Fig. 9c, the delta temperature distribution of the baseline model minus the heat transport without faults model is shown. It indicates that a considerable variation in the temperature to the western side of the Rauw fault zone would occur. Groundwater would flow freely and without needing to find a way across, either by slowing down or being forced upward east of the Rauw fault. Groundwater with a higher temperature at the bottom of the aquifer would be driven in the upwards and downstream directions, due to having a relatively broad spread of heat. Adjacent to the Rauw fault on the eastern side, at around z ∼ −100 m a.s.l., a higher temperature zone can be observed. This indicates the approximate depths at which groundwater begins to flow upwards, and heat is being transported across the fault in the top last few metres, crossing and creating the occurrence of higher temperatures adjacent to the Rauw fault at the western side. These relatively higher variations in temperature occur in the vicinity of the Rauw fault (Fig. 9c). The results suggest that, further downgradient, a much lower or even negligible impact exists (Fig. 9d).

### 5 Conclusions and recommendations

Our study provides highly detailed spatial and temporal temperature data (i.e. the LST raster stack and a paleo-temperature time series) for the implementation of boundary conditions for reconstructing the past and current temperature distribution with the use of a transient heat transport model. Limitations that are typically associated with applying a specified temperature across the top of the model domain are tackled here by constructing a Holocene temperature–time curve to define a surface temperature boundary for the subsurface spatially distributed temperature to reduce the uncertainty in the temperatures at the

Water table. Long-term transient heat transport simulations were done using a steady-state groundwater flow field and calibrated on the basis of TD profile measurements in the Neogene Aquifer across the Nete catchment. The use of a relatively long time series is actually a process that is not computationally intensive, as the time steps in the starting stress periods are quite spaced. The implementation of the paleo-temperature time series in the stress period between 10 000 and 2000 ka (in steps of 1000 years) requires around 1 min to compute and offers a considerable temperature stabilization across the aquifer of up to 30 % per temperature degree of change in the temperature top boundary condition. For this reason, the use of a paleo-temperature time series is a robust and accessible approach for determining initial conditions.

We collected several state variable observations (i.e. TD profiles) from the literature and executed different temperature logs ourselves as well, for the purpose of testing the process model and performing calibration. This study builds on previous efforts, providing an upgraded example of the combined impact of conductive and advective phenomena associated with paleo-climatic fluctuations to characterize the temperature distributions observed in the Neogene Aquifer. The use of a fully 3D numerical model for flow and transport in groundwater is widely supported for this type of problem, as is it relatively easy to implement, and their ability to solve the partial differential equations governing these phenomena for which an analytical solution would perhaps present complications for accounting for the spatiotemporal stresses (e.g. vertical/horizontal in/out fluxes across time and space).

The model provides an acceptable representation of the groundwater system since all of the simulated TD profiles are within 2°C (0.03°C m^{-1}) of the simulated values (with a maximum difference of 1.7 °C). With an RMSE of 0.03 °C m^{-1} (1.15 °C), the paleo-temperature heat transport model outperforms the model performed in steady state and the model where only the LST were considered by 0.04 °C m^{-1} (1.6 °C) and 0.08 (3.22 °C), respectively. A total of 70 % of the simulated TD profiles by the paleo-temperature model are within 1°C (0.015°C m^{-1}) of the observed values, and 50 % of them are within 0.5°C. In some areas, the modelled behaviour deviates from the observed temperature–depth values, and this is regarded as being due to groundwater flow over-/underestimations which drive the advection of heat in addition to conduction. Although some of these deviations may be relatively considerable in a few areas, they are indicators of local processes and/or drivers that may be isolated or combined, which the current models do not account for. In general, absolute temperatures at very shallow depths (z > −5) and in the deepest parts of the aquifer are more difficult to reproduce due to, for example, the imperfectness of the boundary condition and/or local heterogeneities; however, the temperature gradients have shown a good fit for most of the aquifer central and deep areas.
Because groundwater temperatures are highly affected by the magnitude of groundwater flow, it is clear that advection and convection play a major role as mechanisms of heat transport in the Neogene Aquifer ($Pe \gg 1$). This behaviour changes with depth, where, deeper in the aquifer, groundwater fluxes are lower, and thermal conduction seemingly becomes dominant ($Pe \leq 1$). Assumptions have been made on particularly (i) homogeneous thermal properties and (ii) homogeneous heat flux source. By these assumptions we neglect the variability in the presumably heterogeneous thermal properties (mainly in clay-rich formations) and heat source flux, which may potentially introduce local variations in the temperature distribution where areas of strong heterogeneity occur. Despite this, and given that advection/convection is seemingly the main mechanism of heat transport in the system, these uncertainties might only be relevant in some parts of the system, for instance, (i) at the surface derived by the hysteresis produced by the LST vs. soil temperature measurements, (ii) where low groundwater fluxes occur, for instance, right at the bottom of the aquifer, (iii) and possibly in clay-rich formations, such as the Kasterlee Formation. It is therefore suggested that future research focus on these topics. Nonetheless, given the acceptable agreement between observed and simulated temperature values and gradients, and the distribution of temperature across the aquifer, a good approximation to the real thermal properties and heat flux source has been implemented.

With the developed baseline model, we evaluated different conceptual models to quantify the separate effects of advection and faults acting as a barrier, and we see the following:

i. how good groundwater flow is at spreading heat across the whole model domain, thereby accentuating the drainage effect of the surface water network, corresponding local flow systems, and groundwater withdrawals, and

ii. the role of subsurface features, specifically the Rauw fault in this particular case, which act as barriers/conduits disturbing temperature distribution, although this mainly occurs locally.

These aspects are both of the utmost importance, particularly in the context of geological disposal of radioactive waste, but also advance our process understanding in terms of using of temperature observations as calibration targets and their implementation in a joint calibration. For future investigations near the surface, a refined grid cell size (e.g. a few to scores of metres) would improve the simulation accuracy, especially in areas of hydrogeological or thermal complexity, such as in the surroundings of the Rauw fault. Currently, the relatively coarse grid cells generalize relevant local-scale complexities (i.e. surface water/groundwater exchanges) that can potentially affect local flow paths and temperature gradients.

All these temperature variations serve as a proxy of the influence of advective transport and the role of faults as barriers or conduits, such as the Rauw fault, and suggest that solutes released from Boom Clay might be affected in similar ways. While our study used temperature variations to characterize subsurface transport at the catchment scale and in the vicinity of the Rauw fault zone, in the future, other environmental tracers will be evaluated. These unconventional state variable observations are to be implemented for further examination of the groundwater flow and age distribution of the Neogene Aquifer and to be used in a joint inversion procedure to constrain numerical models for groundwater flow and solute transport.
Figure A1. Measured and simulated temperature–depth profiles for all locations across the Neogene Aquifer within the Nete catchment.
Figure A2. (a) Temperature–time series of air and soil measured in different stations within and nearby the Nete catchment. Corresponding raw and corrected average land surface temperature (LST) based on remote sensing data are presented as well. (b) Inset of panel (a).

Figure A3. In situ measurements at weather stations vs. sampled LST values. The linear model that was used to correct raster LST values is included.
Figure A4. Corrected MODIS monthly average LST with a 1 km resolution for the year 2001 within the Neogene Aquifer model (NAM) boundary.

Figure A5. Temperature–time curve for each author in the legend scaled to the 1961–1990 Belgian average for the past 10 519 years.
Figure A6. Pécler numbers ($Pe$) derived from groundwater model results presented by Casillas-Trasvina et al. (2021). Note the considerably lower values ($Pe \leq 1$) in the deepest parts of the Neogene Aquifer corresponding to the Berchem and Voort sands.

Figure A7. Composite scaled sensitivities for (a) groundwater flow parameters, where HK means horizontal hydraulic conductivity for the hydrogeological formations, VA is vertical anisotropy, HFB and CON represent the hydraulic conductance for the horizontal flow barriers and drains, and (b) heat transport parameters, where TC means thermal conductivity for each hydrogeological formation.
Table A1. Groundwater flow and transport models parameter names, as shown in Fig. A7.

<table>
<thead>
<tr>
<th>Model</th>
<th>Parameter name</th>
<th>Parameter description</th>
<th>Units</th>
</tr>
</thead>
<tbody>
<tr>
<td>Groundwater flow model</td>
<td>RECH_SPA</td>
<td>Spatially distributed recharge</td>
<td>mm yr$^{-1}$</td>
</tr>
<tr>
<td></td>
<td>CON_RIVER</td>
<td>River conductance</td>
<td>m d$^{-1}$ m$^{-1}$ (cond. coeff.)</td>
</tr>
<tr>
<td></td>
<td>DIES_HK</td>
<td>Horizontal conductivity of the Diest Formation</td>
<td>m d$^{-1}$</td>
</tr>
<tr>
<td></td>
<td>MOL_HK</td>
<td>Horizontal conductivity of the Mol Formation</td>
<td>m d$^{-1}$</td>
</tr>
<tr>
<td></td>
<td>KEMP_HK</td>
<td>Horizontal conductivity of the non-tabular Quaternary formation</td>
<td>m d$^{-1}$</td>
</tr>
<tr>
<td></td>
<td>KastClay_HK</td>
<td>Horizontal conductivity of the Kasterlee Formation</td>
<td>m d$^{-1}$</td>
</tr>
<tr>
<td></td>
<td>KastClay_VA</td>
<td>Vertical anisotropy of the Kasterlee Formation</td>
<td>Va</td>
</tr>
<tr>
<td></td>
<td>LILLO_HK</td>
<td>Horizontal conductivity of the Lillo Formation</td>
<td>m d$^{-1}$</td>
</tr>
<tr>
<td></td>
<td>HFB_RAUW</td>
<td>Horizontal flow barrier for the Rauw fault</td>
<td>m d$^{-1}$</td>
</tr>
<tr>
<td></td>
<td>MALLE_HK</td>
<td>Horizontal conductivity of the Malle Formation</td>
<td>m d$^{-1}$</td>
</tr>
<tr>
<td></td>
<td>WEELDE_HK</td>
<td>Horizontal conductivity of the Weelde Formation</td>
<td>m d$^{-1}$</td>
</tr>
<tr>
<td></td>
<td>MERKS_HK</td>
<td>Horizontal conductivity of the Merksplas Formation</td>
<td>m d$^{-1}$</td>
</tr>
<tr>
<td></td>
<td>HFB_POPPEL</td>
<td>Horizontal flow barrier for the Poppel Fault</td>
<td>m d$^{-1}$</td>
</tr>
<tr>
<td></td>
<td>CON_DRAIN</td>
<td>Drain conductance</td>
<td>m d$^{-1}$ m$^{-1}$ (cond. coeff.)</td>
</tr>
<tr>
<td></td>
<td>BERCH_HK</td>
<td>Horizontal conductivity of the Berchem Formation</td>
<td>m d$^{-1}$</td>
</tr>
<tr>
<td>Heat transport model</td>
<td>TC_Diest</td>
<td>Thermal conductivity of the Diest Formation</td>
<td>W m K$^{-1}$</td>
</tr>
<tr>
<td></td>
<td>TC_Berchem</td>
<td>Thermal conductivity of the Berchem Formation</td>
<td>W m K$^{-1}$</td>
</tr>
<tr>
<td></td>
<td>TC_Mol</td>
<td>Thermal conductivity of the Mol Formation</td>
<td>W m K$^{-1}$</td>
</tr>
<tr>
<td></td>
<td>TC_Kasterlee</td>
<td>Thermal conductivity of the Kasterlee Formation</td>
<td>W m K$^{-1}$</td>
</tr>
<tr>
<td></td>
<td>TC_Malle</td>
<td>Thermal conductivity of the Malle Formation</td>
<td>W m K$^{-1}$</td>
</tr>
<tr>
<td></td>
<td>TC_Lillo</td>
<td>Thermal conductivity of the Lillo Formation</td>
<td>W m K$^{-1}$</td>
</tr>
<tr>
<td></td>
<td>TC_Merksplas</td>
<td>Thermal conductivity of the Merksplas Formation</td>
<td>W m K$^{-1}$</td>
</tr>
<tr>
<td></td>
<td>TC_Campine</td>
<td>Thermal conductivity of the non-tabular Quaternary formation</td>
<td>W m K$^{-1}$</td>
</tr>
<tr>
<td></td>
<td>TC_Weelde</td>
<td>Thermal conductivity of the Weelde Formation</td>
<td>W m K$^{-1}$</td>
</tr>
<tr>
<td></td>
<td>TC_BoomClay</td>
<td>Thermal conductivity of the Boom Clay formation</td>
<td>W m K$^{-1}$</td>
</tr>
</tbody>
</table>

**Code and data availability.** Model codes and data are available upon request from the corresponding author.

**Author contributions.** ACT, BR, and KB designed the experiments and collected and processed data in different field campaigns. ACT carried the modelling experiments and performed simulations with contributions from BR, BR, KB, and KW provided supervision during the research period. LW (ONDRAF/NIRAS) provided part of the financial support and resources in this work. ACT prepared the paper, with contributions to writing, reviewing, and editing from all co-authors.

**Competing interests.** The contact author has declared that none of the authors has any competing interests.

**Disclaimer.** Publisher’s note: Copernicus Publications remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.

**Acknowledgements.** This work has been performed in co-operation with, and with the financial support of, ONDRAF/NIRAS, the Belgian Agency for Radioactive Waste and Fissile Materials, and the SCK CEN Academy by SCK CEN, the Belgian nuclear research centre, as part of the programme on geological disposal of high-level/long-lived radioactive waste that is carried out by ONDRAF/NIRAS. The authors would like to acknowledge Victor Bense, for his advice and support during the first temperature-depth measurement campaign of 2019.

**Financial support.** This research has been supported by NIRAS/ONDRAF and SCK CEN Academy in the framework of Alberto Casillas-Trasvina doctoral grant agreement.

**Review statement.** This paper was edited by Monica Riva and reviewed by two anonymous referees.
References


Clerc, M., Knight, N., Sousa, P., Barrett, J. L., and Atran, S.: Standard particle swarm optimisation, hal-00764996, HAL open science, https://hal.archives-ouvertes.fr/hal-00764996/document (last access: 7 November 2022), 2012.


R Core Team: R: A Language and Environment for Statistical Computing, https://www.r-project.org/ (last access: 7 November 2022).


Rogiers, B.: MODFLOW: Pre- and post-processing of MODFLOW files in R, GitHub [code], https://github.com/rogiersbart/RMODFLOW (last access: 7 November 2022), 2016a.

Rogiers, B.: RMT3DMS: Pre- and post-processing of MT3DMS files, GitHub [code], https://github.com/rogiersbart/RMT3DMS (last access: 7 November 2022), 2016b.


