MSc Graduation Research
Earth Sciences | Earth Surface and Water

The global distribution of modern groundwater derived from high-resolution 3-D flow path simulations

Bas van Dijk
Student number: 3801063
Email: b.vandijk4@students.uu.nl

March 2018
Utrecht, The Netherlands

B. van Dijk

BSc Earth Sciences

MSc student Earth Surface and Water

Student number: 3801063

Email: b.vandijk4@students.uu.nl

Utrecht University, the Netherlands

University of Victoria, BC, Canada

First Supervisor

Dr. L.P.H. van Beek

Assistant Professor of Physical Geography

Utrecht University

Second Reader

Prof. dr. ir. M.F.P. Bierkens

Professor of Physical Geography

Utrecht University

External Research Supervisor

Dr. T. Gleeson

Associate Professor of Civil Engineering

University of Victoria, BC, Canada

Preface

This research was partially carried out at the University of Victoria in British Columbia, Canada under the supervision of dr. Tom Gleeson. A major part of his scientific research on groundwater from the last few years has served as both theoretical information, as well as input data and material for comparison and validation. The remainder of this research has been carried out at the University of Utrecht under the supervision of dr. Rens van Beek. Model input data and the concept of the model partially originate from dr. ir. Inge de Graaf's research, who is a former PhD student at the University of Utrecht under van Beek's supervision.

First of all I would like to express my gratitude towards Rens and Tom. Rens, thank you for your inspiration, your help with all the programming, and especially your boundless guidance. Tom, thank you for letting me be part of your groundwater science and sustainability group at UVIC, the countless life lessons, and all your feedback both in person as well as during the many skype calls. Secondly, I would like to thank Inge for her assistance with the kick-off of my research, and Kevin Befus from the University of Wyoming and Peter Vermeulen from Deltares for helping me out with several modelling issues.

Furthermore I am grateful to my Canadian friends; Kristyna, Mikhail, Tara and Xander, who reached out to me and made my stay in Victoria feel like home. Finally, I would like to thank my parents, my sisters, my girlfriend, and the rest of my family and friends for their endless support.

Bas van Dijk

Utrecht, March 2018

Abstract

Research on groundwater and its sustainable management has become increasingly significant due to growing water requirements and over-abstraction of groundwater resources. The time that water spends in the subsurface is an essential feature for the understanding of the dynamics of groundwater systems and patterns. Modern groundwater, which only represents the water in the subsurface recharged after above-ground thermonuclear testing fifty years ago, is of major interest as it is most vulnerable to global change, but also able to indicate renewable aquifers. The primary aim of this study was to improve recent estimates of the global volume and the distribution of modern groundwater using high-resolution 3-D flow path simulations. In order to analyse the results, they were compared to the research by Gleeson et al. (2015). The total groundwater volume in the continental crust is estimated at 4.2 million km^3, of which approximately 147 thousand km^3 or 3.5% is younger than fifty years. The volume of modern groundwater is equal to a depth of 1.1 m if it was extracted and pooled at the earth's land surface like a flood; almost 3 times smaller than presented by Gleeson et al. (2015). The model's spatial distribution of modern groundwater is mainly controlled by mountainous and coastal areas while river systems are underrepresented. The relatively modest groundwater volumes are assumed to be largely dependent on the aquifer thickness and the resolution of input data.

Contents

List of figures

List of tables

1 Background

1.1 Concepts of groundwater flow and travel times

Groundwater is an important fresh water source, as it is not only the primary global drinking water supply, but also a cornerstone of ecosystems and biodiversity (de Graaf et al., 2015). Research on groundwater and the sustainable management of it have become a subject of increasing significance to society due to growing water requirements and over-abstraction of groundwater resources (Erskine & Papaioannou, 1997; Gleeson et al., 2012).

The groundwater age is fundamental for the understanding of the dynamics of groundwater systems, which can be defined as the period of time groundwater travels between recharge and the point of measurement. This must however not be mistaken with the residence time that is defined as the groundwater age specifically at the point of discharge (Bethke & Johnson, 2008).

The age of groundwater can be directly linked to above-ground thermonuclear testing during the late 1950s until the early 1960s, which led to a significant peak in tritium concentrations of groundwater (Alley et al., 2002). Water in the subsurface recharged after these events (i.e. less than fifty years old) is defined as modern groundwater and of major interest as it is most vulnerable to global change (Gleeson et al., 2015). Furthermore a relatively high ratio of modern groundwater might characterise an aquifer or region as renewable, whereas one with mostly "old" water, will presumably not renew within a time period that is useful for human society (Bethke & Johnson, 2008). Globally measured groundwater ages can vary between a couple of months for humid regions to centuries, millennia, or in exceptional cases millions of years in deserts (Gleeson et al., 2015).

Groundwater flow and transport can most simply be conceptualised as a constant unit of water that infiltrates the surface (e.g. rainfall), following a specific flow path through an aquifer until the point of measurement that can either be in the subsurface or at the point of discharge (Zimmerman et al., 1966; Kazemi et al, 2006). This two-dimensional flow characterisation, known as the piston flow model (Figure 1.1), which considers the water packet to act simply as a piston (i.e. a cylinder fitting closely within a tube), does not take into account the exchange of water molecules between the aquifers and aquitards however. A local, intermediate, and regional flow type system is preferred as proposed by Toth (1963), as the piston flow model falls short in a scale distinction, which is claimed to be all the more important for higher topographies, but also expected to be fundamental for the distinction of modern groundwater.

Bethke & Johnson (2008) have set forth that since the beginning of the 21st century the hydrology discipline has gotten familiar with an alternative view towards the definition and computation of groundwater age. In order to achieve the most accurate model, the water-carrying subsurface must not be interpreted as a closed system. It is fundamental that groundwater should not be regarded as sealed packets, as each individual water molecule in the sediment can follow an abundance of possible tortuous flow paths, which is defined as hydrodynamic dispersion

(Bethke & Johnson, 2008). The mean age of each water molecule in a sample, i.e. the average period it has been in the subsurface, can be used to quantify groundwater age, but this has been criticised since the same mean age can be derived from a variety of water distributions (Jasechko, 2016). The impact on the spatial distribution of groundwater by hydrological phenomena such as phases of recharge or the mixing of water from different sources, makes the adaption of this definition all the more crucial (McGuire & McDonnell, 2006; Bethke & Johnson, 2008).
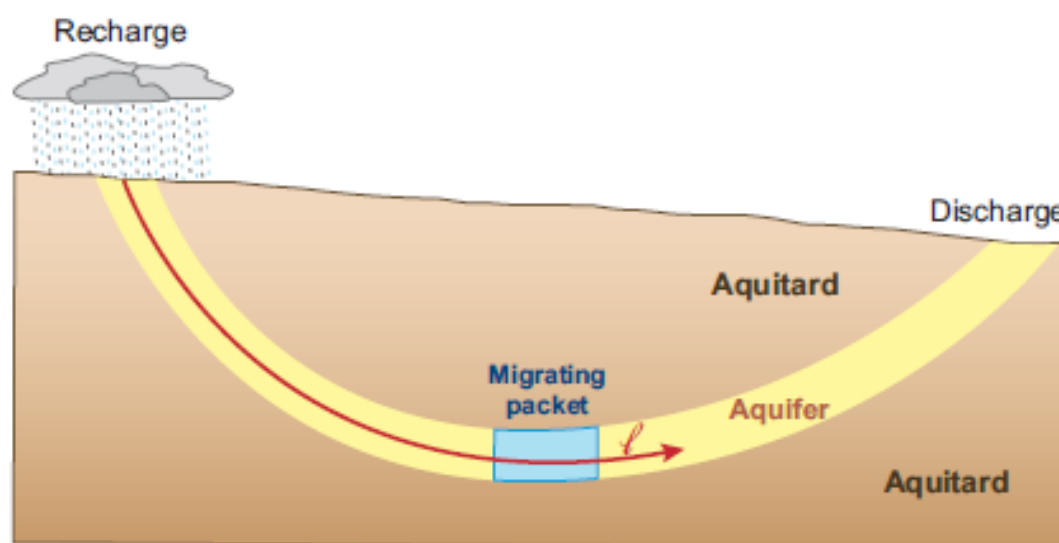
Figure 1.1: The piston flow model (Bethke & Johnson, 2008).

1.2   Previous work on groundwater simulations

Even though others have researched (modern) groundwater (Garmonov et al., 1975; Chahine, 1992; Bethke & Johnson, 2008), until recently the volume and distribution of it have always been unknown (Gleeson et al., 2015). Interest in groundwater age has been stimulated again as a consequence of developments in river monitoring and simulation programmes, which can be used to compute the volume of water in large hydrological systems (McGuire & McDonell, 2006). However assumptions and limitations associated to various modelling approaches have led to uncertainties, which are difficult to assess. Traylor & Zlotnik (2016) claimed that particularly the lack of available input data and uncertainties have always been the restraints of accurate groundwater modelling and age simulations.

Gleeson et al. (2015) recently determined the volume of modern groundwater on earth derived from two different techniques. The first method uses a water-table-driven flow model based on an abundance of available spatial watershed data including more than 40,000 porosity measurements globally (Gleeson et al., 2014), worldwide water table depths (Fan et al., 2013), and lithology data (Hartmann & Moosdorf, 2012). Using these the one-dimensional groundwater equivalent, which is the water level if it were to be abstracted from the ground and pooled on the land surface like a flood, can be computed. By integration of the specific porosity for every

soil type, and weighing its volumetric fraction. Subsequently the total groundwater storage can be calculated by multiplying the groundwater equivalent with the studied surface area (Gleeson et al., 2015).

The modern groundwater equivalent was derived from the spatial groundwater distribution in a cross-sectional watershed model assuming steady-state groundwater flow. The assumption was made that one cross section is representative of a whole watershed and one aquifer consists of multiple watersheds. Also, the depth of an aquifer is derived from the length scale, which is in direct correlation to the width of the watershed. The total modern groundwater volume is ultimately derived by means of global integration of the individual watersheds.

The second method is isotope-based, and primarily founded on almost 4,000 groundwater measurements of tritium (^{3}H), which is the radioactive isotope of hydrogen. From this data in combination with the radioactive decay rate of tritium, Gleeson et al. (2015) have managed to derive the mass proportion of modern groundwater, which is expressed by the parameter $R_{modern,3H}$ and dependent on depth. Together with aquifer-specific porosity values and modelled groundwater volumes, the modern groundwater volume could be calculated.

In order to assess the consistency between the two methods, Gleeson et al. (2015) compared the differences between the outputs of the hydrogeological 2-D model and the tritium observations. As there is a lack of available global concentration data of the latter, only the thirty aquifers that contain the largest amount of accessible tritium data are considered, which are presented in the inset map of Figure 1.2. Greenland and Antarctica were excluded, as there is no data available, as well as the fact that the groundwater flow can be assumed to be negligible there. Gleeson et al. (2015) state that ruling out these areas adds uncertainty to the global groundwater volume estimate. Furthermore it must be mentioned that the considered aquifers are actually convex hull zones (hereafter referred to as aquifer areas), restricted by the tritium data points with a buffer zone of 10 km as presented in the main map covering Central Asia (Gleeson et al., 2015). It can be observed that as a consequence of the tritium sample locations, differences in the aquifer area sizes can be substantial.

The comparison between the results of the two methods for these thirty aquifer areas considered by Gleeson et al. (2015) is graphically presented in Figure 1.3, where the agreement is expressed by the parameter $V_{storage}$ ratio, which defines the modelled storage estimate divided by the tritium based estimate. As can be observed, about two out of three aquifer areas have similar volumes for the two methods, when taking the uncertainty range in account (representing both uncertainty within the tritium calculations, but also for the permeability of the 2-D simulations), indicated by the grey bars.

Gleeson et al. (2015) state that it is prohibitive to model the groundwater flow and age distribution separately for every watershed on the world. They circumvented this by binning

Figure 1.2: Global inset map presenting the locations of all thirty aquifer areas that contain the most accessible tritium data (yellow), and the main map featuring the majority of the Asian aquifer areas, including the original tritium sample points (red) and the circumscribing convex hull areas (light orange).
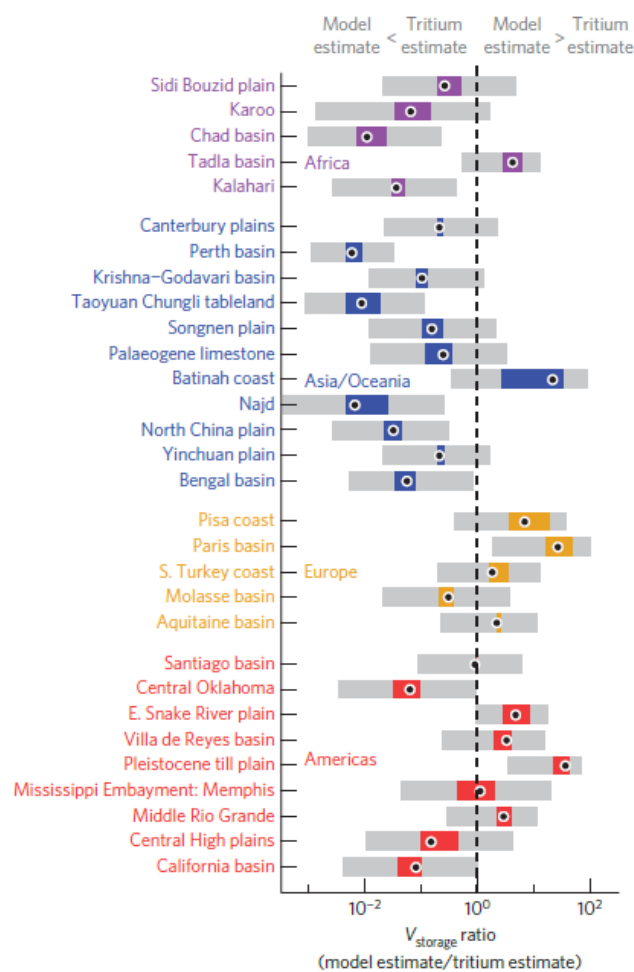
Figure 1.3: Comparison of the modern groundwater storage estimates derived from the hydrologic 2-D model and the isotope-based method for the thirty aquifer areas with the most available tritium data. The black dots present the ratio of the average model estimate divided by the median tritium estimate. The coloured bars represent the uncertainty range based on solely the tritium analysis, while the grey bars define a combined uncertainty range. (Gleeson et al., 2015).

watershed properties, i.e. grouping similar watersheds, which was mainly based on equivalent water tables and watershed sizes. By virtue of running some extra model simulations, Gleeson et al. (2015) have been able to make a robust approximation of both the total volume of groundwater, as well as the modern groundwater in the uppermost 2 km of the subsurface. The total volume was estimated at 22.6 million km^3, with an absolute range from 16 to 30 million km^3 allowing for porosity uncertainties. The part of this which can be defined as modern groundwater is estimated to be 0.35 million km^3 (1.5% of total) with a range of 0.10 to 5.00 million km^3 accounting for uncertainties including input parameters such as the watershed length scale and various hydraulic gradients.

Although Gleeson et al. (2015) established the first estimate of the global volume and spatial distribution of modern groundwater, the research is limited by significant assumptions:

- The limited geological variety in locations where tritium samples were compiled induces bias. The majority of samples originate from watersheds with a relatively high permeability, which must be considered to overestimate the modern groundwater volume to an unknown extent.
- The mixing of water that originates from multiple sources, thus also multiple flow paths, is considered to affect the tritium concentration at any point, therefore the spatial distribution of groundwater.
- Considering the hydrogeological model, only one sediment type is assumed per watershed, which biases the permeability. This should be accounted for in the uncertainty range. Gleeson et al. (2015) state that it is an important topic of ongoing research.
- Using a 2-dimensional model, instead of a 3-dimensional model for the sake of computational expediency, the resulting artefacts in the simulated lateral flow might affect the spatial groundwater distribution.

Gleeson et al. (2015) therefore suggest an alternative method for the estimate of modern groundwater storage; one where geospatial data and flow path simulation are incorporated in the modelling of the spatial groundwater distribution. A 3-D model of the spatial distribution is preferred, as this would come with the inclusion of subsurface processes, such as lateral and vertical groundwater flow, and also a more accurate location-specific visualisation, which makes it simpler to analyse the correlation between the groundwater age and local conditions.

De Graaf et al. (2015) recently developed a global 3-D MODFLOW aquifer, which includes lateral groundwater flow. On account of the shortage of complete hydrogeological data, large-scale hydrological models generally have not been able to encompass a lateral groundwater flow component, until Sutanudjaja et al. (2011) produced an experimental groundwater flow model for the Rhine-Meuse basin, based on global lithology and elevation data sets. The model composed in MODFLOW was coupled to PCRaster Global Water Balance (PCR-GLOBWB), which is a model developed by van Beek et al. (2011) that is able to accurately simulate

hydrological processes in the subsurface, consisting of both soil stores, as a linear groundwater store. Coupling of these models started off with the computation of the average annual net recharge and groundwater discharge over the period of 1960 to 2010. Subsequently the recharge and surface water levels, which can be derived from the discharge, serve as drivers for the MODFLOW model. Rather than using the linear groundwater store of PCR-GLOBWB, de Graaf et al. (2015) substituted this by a single MODFLOW layer, which represents a global unconfined aquifer (Schmitz et al., 2009). As the thickness of this layer has always been uncertain, its extent is determined by the available terrain attribute data, which can be extrapolated for all types of environments. Assuming that the extent of the global aquifer corresponds with sediment basins beneath fluvial valleys, mountain ranges could be associated with an insignificant aquifer thickness, primarily made up of hard rock, while sediment basins correlate to a relatively large aquifer thickness (de Graaf et al., 2015). During groundwater flow simulations they have restricted themselves to a steady-state, i.e. the global aquifer storage is constant, serving as a try-out for transient simulations in further research in prospect.

The 3-D model presented by de Graaf et al. (2015) computes the water balance for each grid cell at a certain time step regulated by a climate simulation, which includes the influence of precipitation, evapotranspiration, and the temperature (de Graaf et al., 2016). By means of the resulting global groundwater head distribution, flow path simulations were carried out using MODPATH (Pollock, 1994), from the point of infiltration to the point of discharge. This provides the groundwater travel times of a particle at the edge of each grid cell.

De Graaf et al. (2015) conducted large-scale flow path simulations for two continents as presented in Figure 1.4, where the movement of one particle from one cell to another is computed from the point of recharge until the point of discharge (e.g. a river or sea). As expected, the figure shows that particularly dry areas (e.g. the Sahara and the Kalahari desert) in Africa can be associated to long flow paths and old groundwater, while the majority of Europe primarily contains modern groundwater. It must be noted that regarding the travel times quantitatively, the travel times of flow paths spreading over the continents in purple (less than 1 month), but also the maximum values (rarely more than 100-200 years) seem considerably underestimated, and in contradiction to the results presented by Gleeson et al. (2015).

The 3-D groundwater model by de Graaf et al. (2015) potentially overcomes some of the weaknesses that are inherent to the 2-D modelling by Gleeson et al. (2015), in particular the exclusion of subsurface hydrological processes such as later groundwater flow, and the standardising of hydrologic properties within watersheds. However, as a large-scale model exercise it still has to deal with incomplete or uncertain data and therefore it is important not only to understand the model but also its assumptions and limitations. From a preliminary evaluation of the model and its global input files, the following limitations became clear:

- Some maps have missing values, and are therefore incomplete (e.g. transmissivity).
- On some locations the thickness of the aquifer is equal to zero (i.e. DEM = bottom model layer), with no groundwater flow as a consequence.
- Standard porosity values of 0.01 and 0.3 for respectively the aquitard and aquifer have been applied, which might bias the simulated travel times of the flow paths.
- The original geographic coordinate system (GCS) of the input maps is the world geodetic system 1984, better known as WGS84, which has degrees as its angular unit. The corresponding resolution of 5 arc minutes, i.e. 0.0833 degrees, can be associated to a cell size of 9.27 km at the equator, where the distance between longitudes approaches zero moving towards earth's poles. Therefore the length of the flow paths, and thus the travel times, becomes distorted. A single conversion from the angular unit to a linear one was applied by de Graaf et al. (2015), however this ignores the inherent warping of flow lines across a geographic grid and travel times are approximate at best. A more rigorous treatment of the earth's surface is therefore in order.
- The original resolution of the spatial input files is ~10 km (5 arc minutes). A finer resolution is definitely preferred, as particles will not be generated correctly considering the frequency of data points per particle, but also the possible misrepresentation of hydrological attributes.
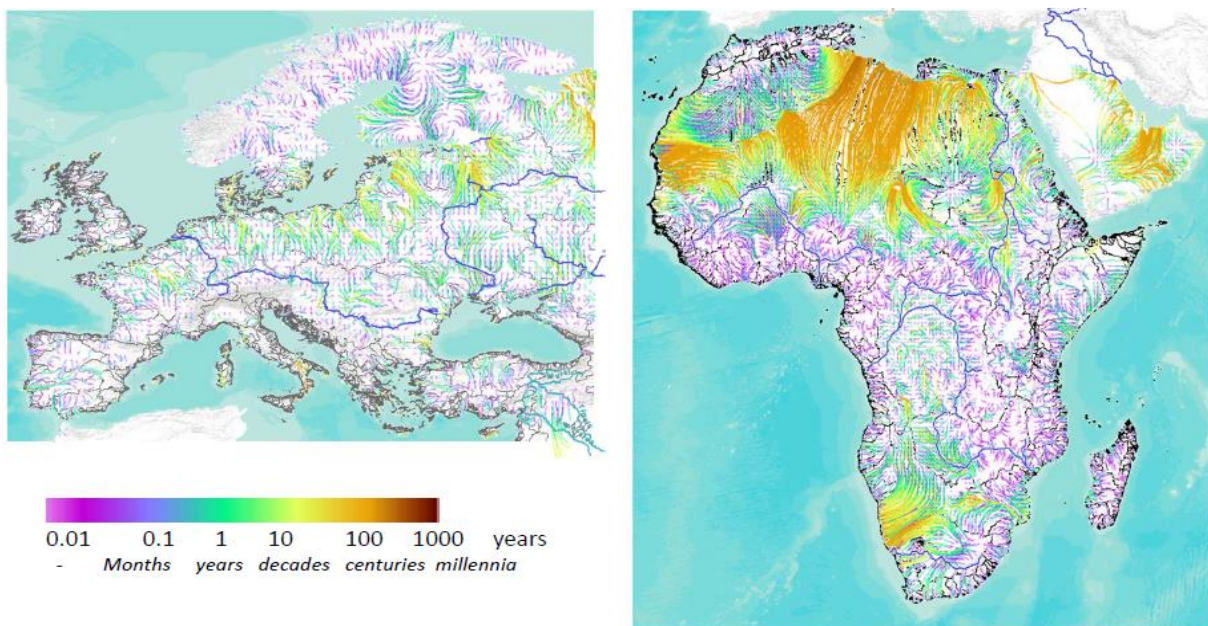
Figure 1.4: Flow paths simulated for Europe and Africa using an unconfined aquifer one-layer model (de Graaf et al., 2015).

1.3 Objectives and research questions

The primary research objective is to improve the estimate of the global volume and the distribution of modern groundwater. Following the suggestion of Gleeson et al. (2015) and building on the simulations by de Graaf et al. (2015), flow paths will be simulated on the global scale with MODPATH (Pollock, 1994) in 3-D and used to provide a more robust distribution of

simulated groundwater ages than that obtained by Gleeson et al. using 2-D hydrological simulations. The underlying assumption is that the presence of modern groundwater below the modelled aquifer layer is negligible, which is not unrealistic as the aquifers included in the simulation by de Graaf et al. (2015) are typically deep.

In order to achieve the primary objective, a number of research questions have been formulated to structure the research process:

1. The model data inputs used by de Graaf et al. (2015) can be regarded to be of insufficient quality in some aspects. What other data should be used to complement these, and what projection and higher spatial resolution should be used in order to achieve the most accurate flow path simulations?
2. How does the volume of modern groundwater output, which is calculated from a global hydrogeological 3-D model, compare to tritium based and the 2-D model based observations for the thirty aquifer areas with the most tritium data as presented by Gleeson et al. (2015), and what controls the differences in these calculations?
3. What is the overall age distribution of groundwater as derived from the global hydrogeological 3-D model, and where are lateral groundwater flows most important to the distribution of modern groundwater?

In order to achieve the introduced primary objective, particle tracking will applied with MODPATH (Pollock, 1994). The second research question will not only be a sensitivity analysis serving as a form of authentication of the data and used methods, but will also allow comparison to the results of Gleeson et al. (2015) as presented in Figure 1.2.

The results from the groundwater simulations might be able to identify areas where lateral groundwater flows are significant, and provide an understanding of location specific differences in groundwater movement, but will also make it possible to produce an approximation of groundwater travel times. In contrast to the study by Gleeson et al. (2015), results are expected to present the importance of the inclusion of lateral flow. Lastly the global groundwater volume and the proportion that is modern can be derived from the simulated travel times and once again compared to the ones presented by Gleeson et al. (2015).

2   Methodology

Chapter 2 addresses the development of the global hydrogeological 3-D model that is used in this research. The global spatial distribution of groundwater, as well as the total volume of modern groundwater is determined from the model simulations. The model, its settings and the assumptions made throughout the research are described as well. Lastly it will be presented how the groundwater volumes can be extracted from the model simulations and how they can be analysed.

2.1   Simulation of flow paths

In order to simulate groundwater flow paths, an interactive modelling programme (iMOD) that allows for particle tracking has been used. There is an abundance of available numerical groundwater flow models, of which some have considerable overlap. Deltares has anticipated on this by means of the development of iMOD, which is heavily influenced by MODFLOW, to minimise the costs of the redundant development of individual models over and over again (Vermeulen et al., 2016). iMOD supports very large high-resolution groundwater flow models, and offers both the possibility to accurately evaluate effects on a local, as a global scale. It is therefore suited for this research, as not only the modern groundwater volume of individual aquifer areas will be calculated, but also the volume globally.

It is possible to track particle flow paths and to approximate the coherent travel times of groundwater flows by means of MODPATH (Pollock, 1984). This is a particle-tracking post-processing program for MODFLOW and incorporated in iMOD, which is able to compute the flow rate components of inflow and outflow for every direction in every grid cell per time step. The groundwater flow equation used to simulate the flow of groundwater through aquifers is:

$$\frac{\partial}{\partial x}\left(K_{xx}\frac{\partial h}{\partial x}\right) + \frac{\partial}{\partial y}\left(K_{yy}\frac{\partial h}{\partial y}\right) + \frac{\partial}{\partial y}\left(K_{zz}\frac{\partial h}{\partial z}\right) + W = S_S\frac{\partial h}{\partial t} \qquad \text{(McDonald \& Harbaugh, 1988)}$$

where
- K is the hydraulic conductivity for all three dimensions
- h is the groundwater head
- W is the additional out and in flow by respectively sinks or sources
- S_S is the specific storage of the porous material
- t is time

This partial-differential equation of groundwater flow is considered to be standard code for aquifer simulation. In combination with boundary and initial conditions, which are defined in a run file (Appendix A), it is possible to accurately track flow paths (Harbaugh et al., 2000) and to construct the accompanying groundwater velocity distribution. By means of iMODFLOW budget terms and the streamflow can be established, which will be used as the steady-state model

input. Subsequently the flow paths are simulated from the location where a particle infiltrates the soil up to the moment of discharge. The MODPATH output provides the coordinates of a particle and the elapsed time in years since the moment of release, i.e. the groundwater age. The full modelling process is explained step-by-step in Appendix B.

2.2   Model settings & assumptions

2.2.1   Input data and boundary conditions

The basis of the global groundwater simulations in this study stems from the work of de Graaf et al. (2015). Their work was based on two datasets, the high-resolution global lithological map (GLiM) developed by Hartmann & Moosdorf (2012) and global permeability estimates as introduced by Gleeson et al. (2011). From the former, hydrogeological units were derived and the relevant properties (e.g. permeability and porosity) assigned from the latter. This information provides the basis for the twelve spatial input files that are required in order to simulate the groundwater head and travel times with iMOD. These files are listed in Table 2.1.

From the global digital elevation model (DEM), presented in the HydroSHEDS dataset by Lehner et al. (2006), de Graaf et al. (2015) delineated sedimentary basins. Additionally they made the questionable assumption that for areas that have not been identified as aquifers, the bottom of the layer is precisely 100 meters lower than the DEM. Alternatively the bottom of the layer for the uncharted areas is complemented by means of a linear interpolation with the depth to bedrock as presented by Shangguan et al. (2016). The accompanying python script can be found in Appendix C. As a result, the depth of the aquifer areas differs from that used by Gleeson et al. (2015), which may influence the comparison with the 2-D modelled results and the tritium values that are evaluated over the same convex hull zones.

Similar to the bottom of the layer, the hydraulic conductivity, which is in direct relation to the permeability, is rendered by the transmissivity as put forward by Gleeson et al. (2014).

The spatial boundaries on the surface of the model are set by a landmask. This file defines the extent of active flow and gives a value greater than zero to every location where particle tracking is allowed, and a value equal to zero for inactive groundwater flow areas, such as the oceans. It must be mentioned that the original landmask (de Graaf et al., 2015) did not completely match with the other maps, which induced unrealistic values for several coastal areas. As a solution a new landmask has been adopted from the DEM.

The underlying layer is considered as an aquitard for modelling purposes, and thus presumed to be impermeable. De Graaf et al. (2015) have therefore set the porosity at standard values of 0.01 and 0.3 for respectively the aquitard and aquifer. The latter has been considered as a global average that might substantially underestimate travel times of flow paths. As a consequence it has been replaced by the storage coefficient map as currently presented by Gleeson et al. (2014).

Table 2.1: Overview of all used raster maps, a short description, their units, and original source(s).

Map type	Description and processing	Unit	Source
DEM	Digital elevation model of the earth's surface from the HydroSHEDS dataset.	m	Lehner et al. (2006)
Bottom of layer	Lower boundary of the modelled global aquifer complemented by depth to bedrock data.	m	de Graaf et al. (2015) & Shangguan et al. (2016)
Transmissivity	Groundwater flow rate, derived from the aquifer thickness multiplied by the hydraulic conductivity.	m^2/d	Gleeson et al. (2014)
Landmask	Mask derived from the DEM distinguishing land from water, i.e. active and inactive cells.	-	
Porosity	Storage coefficient of the soil.	-	Gleeson et al. (2014)
Water head	Simulated groundwater heads validated by global piezometer data.	m	de Graaf et al. (2015)
Recharge	Steady-state groundwater recharge attained from PCR-GLOBWB.	m^3/d	de Graaf et al. (2015) & Beek et al. (2011)
Baseflow	Water flow between river streams and the aquifer.	m^3/d	de Graaf et al. (2015)
Additional drainage	Other drainage representing locations where the groundwater head exceeds the DEM.	m^3/d	de Graaf et al. (2015)
River head	Average river head derived from long-term river discharge computed by PCR-GLOBWB.	m	de Graaf et al. (2015) & Beek et al. (2011)
Bottom river bed	Bottom elevation of the river bed calculated with the channel depth assuming a rectangular channel.	m	de Graaf et al. (2015)
River condition	Value defining the rate of interaction between groundwater and surface water.	-	de Graaf et al. (2015)

Before MODFLOW can generate an equilibrium water table, a starting head is needed, which has been derived from a combination of groundwater simulations and global piezometer data by de Graaf et al. (2015). The head for surface water is defined by the groundwater recharge and boundary conditions including the hydraulic (total) head in rivers. These are based on local PCR-GLOBWB averages of the period 1960-2010 (de Graaf et al., 2015).

At the same time there must be a form of discharge, which will act as the boundary or sink of a particle's flow path. This is incorporated in the model in the shape of a drainage package and a river package, which will determine the interaction between the groundwater and surface water. The first is represented by two maps, i.e. the baseflow between streams and the aquifer, and additional drainage defining locations where the groundwater level is higher than the surface

e.g. springs. This has been applied due to the inaccuracy that comes with the 5 arc minutes resolution of the model. Lastly the river package consist of three maps that specify the river head, the approximated bottom of the river bed, and the so-called river condition, which is a measure indicating the degree of interaction between the groundwater and surface water derived from the width of the rivers and streams (de Graaf et al., 2015).

2.2.2 Projection & resolution

In order to improve the quality in the flow path simulations and remove any bias that arises from approximate cell dimensions when using geographical projections as used by de Graaf et al. (2015) in the form of WGS 1984, this study employed Lambert Equal-Area projections.

While the Lambert Equal-Area projection cannot represent angles faithfully, which leads to a distortion of shapes (Figure 2.1), it represents gradients between cells truthfully. Hence, the projection was adopted by the United States Geological Survey (Steinwand et al., 1995) as part of the Hydro1k dataset to represent Digital Elevation Models and the derived maps of flow directions and upstream area per continent. This study adopted the same projections and subdivision of the world into six continents. Prior to the MODFLOW simulation all input data taken from de Graaf et al. and corrected when necessary (subsection 2.2.1), was converted by invoking GDAL (Geospatial Data Abstraction Library). The accompanying python script is presented in Appendix C.

The original resolution of 5 arc minutes (~10 km) might be insufficient for this study, considering the frequency of data points per particle and the possible misrepresentation of hydrological attributes. As a solution for the deficient resolution, spatial interpolation (inverse distance) could have been used to fill in the gaps between the time values originating from the flow path simulations, which would produce an age distribution map. However the discharge of a certain particle, which can occur on a finer resolution than the original grid, would then be neglected. Moreover this approach is assumed to be deficient considering groundwater flows (with different travel times) crossing each other.

Instead, spatial interpolation is used to resample the spatial input maps to a finer resolution (this step is also incorporated in the python script presented in Appendix C), which induces a higher number of data points for a particle, and therefore a higher number of data points per square kilometre. During this research different resolutions will be explored, inspired by the fact that iMOD otherwise does not generate the particles correctly. In case of divergent flow for instance, there will be areas that are initially not covered by particle data points.

Different resampling methods have been used for the spatial interpolation, dependent on the type of map. Nearest neighbour has been used for the majority of input maps, except for the landmask, and the three maps representing the river package. For the first a boolean resampling method was used, while the latter three were subjected a bilinear interpolation, so the correct river beds would stay intact.
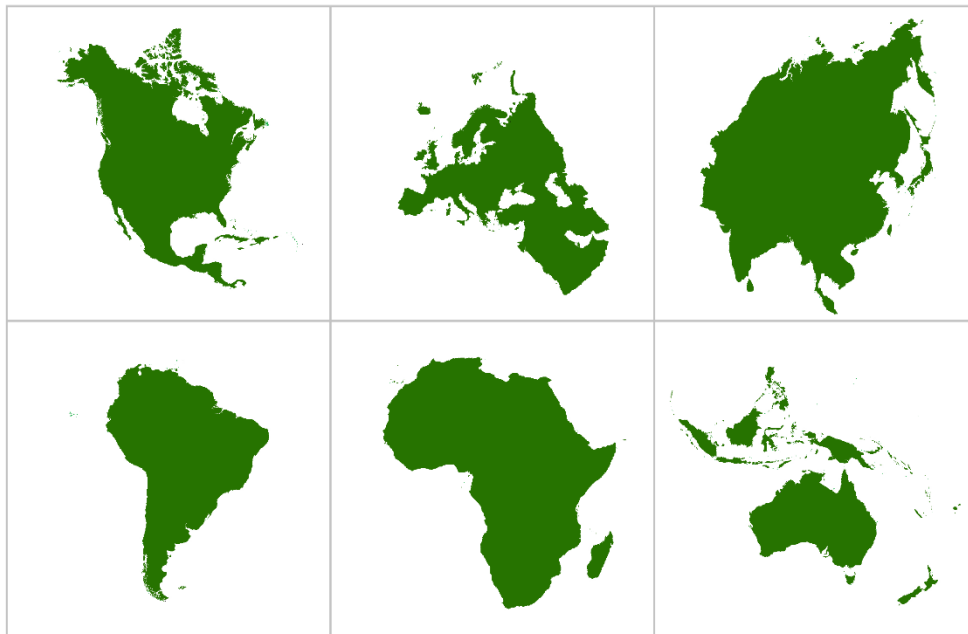
Figure 2.1:Global distribution of the continents in WGS84 as used by de Graaf et al. (2015), and continental maps projected in the Lambert equal-area projection as defined in the Hydro1k dataset.

2.3   Extraction of travel times and modern groundwater volumes

Once the model settings have been established, the 3-D flow path simulations can be executed, i.e. MODPATH generates the coordinates of tracked particles and the accompanying groundwater ages for every crossed cell. F_{modern} can be derived, which is the amount of particles from the flow path simulations with a travel time values less than fifty years divided by the total amount of particles that can be computed per entity, i.e. cell, aquifer area, or continent. This is not conform with R_{modern}, presented by Gleeson et al. (2015) that actually presents the mass proportion of modern groundwater, nor is it an indisputable definition of the age distribution of groundwater as each flow path line is assumed to be proportionate to one another, since the particle tracking solely models the travel time of a flow path, not the correlated volume of water.

Combining F_{modern} with the total groundwater volumes, the modern groundwater volume can be computed, as explained hereafter. The processes from the script presented in Appendix C implemented to achieve a global modern groundwater estimate and a spatial distribution map, are discussed in this section step by step as presented in a flow diagram in Figure 2.2.

1. First an array is built that will provide the information to generate a map of groundwater ages at the considered resolution (10K, 5K, 2K and 1K respectively). Each entry along the rows of this array represents a single location or grid cell that is characterized by its unique x- and y-coordinates (X_c, Y_c). To each of these cells an ID is assigned that corresponds to the convex hulls for aquifer areas identified by Gleeson et al. (2015). Areas outside a convex hull are given the value of 0 and used to derive continental values. The column headers of this array correspond to twenty-four travel times (TT) bins, ranging from 0 to a maximum of 100 million years, which are presented in Table 2.2.

Table 2.2: Class boundaries of travel times (upper limit) in years.

0.001	1	2.5	5	10	25	50	100
250	500	1,000	2,500	5,000	10,000	25,000	50,000
100,000	250,000	500,000	1,000,000	2,500,000	5,000,000	10,000,000	100,000,000

2. As particles are seeded in MODPATH at every node, each cell can be crossed by multiple flow paths that stem from the upstream area, where the groundwater flow is simulated by means of the partial-differential equation presented in section 2.1 (McDonald & Harbaugh, 1988). These flow paths are described by vertices (x, y) with a corresponding value of the local flow velocity and the accumulated travel time. To capture this, all vertices within a certain cell are identified (using the shortest distance to the cell centre) and their travel time used to update the number of entries in the bins.

3. From the obtained numbers per bin, frequencies can be computed. These can be computed per entry, i.e. per cell, or using the IDs for an aquifer area or an entire continent by the following equation:

$$f_t = \frac{N_t}{\sum_{t=t0}^{t=T} Nt}$$

where
- f_t is the frequency for the present entity
- N_t is the total number of particles within the current bin with travel time t
- $t0$ and T are the travel time of the first and last bin respectively

Within each cell, the groundwater volume can be computed on the basis of the depth of the water table and the effective porosity using the following equation:

$$V_t = f_t h_w n_{eff} A$$

where
- V_t is the groundwater volume of a cell
- h_w is the water level above the base of the aquifer
- n_{eff} is the effective porosity
- A is the cell area

This can then be multiplied with the travel time frequencies of a cumulative density function (CDF) per cell to obtain the groundwater distribution per age. Section 2.3 will elaborate on this.

4. The frequencies of travel times and groundwater volume per age can be used to create CDF per entity (cell, aquifer area, or continent). Alternatively, the location of each cell can be used to create maps that provides a more visual insight of the spatial distribution of ages and groundwater volumes. The percentage of travel time values that is younger than fifty years can be derived per cell using n_c, which is the data points sample size of a cell. This fraction $F_{modern,c}$, linked to X_c, Y_c, can be converted to a fraction map presenting the amount of modern groundwater for every cell.

5. The data input for the last step is the total groundwater volume for each cell. This approach is on a cell-scale, as differences in thickness and/or the porosity across an aquifer must be taken into account.

Just as in steps 3 and 4, the CDFs and fraction map for the absolute volumes can be calculated. Lastly for the map with the modern groundwater equivalent in m, the total groundwater volume for a cell must be divided by the area of a cell and multiplied with the linked $F_{modern,c}$.

2.3   Sensitivity analysis & data processing

In order to answer the first research question a sensitivity analysis will be executed to determine what resolution is adequate for this research, taking in account both the accuracy of the results, as well as the computation time it requires. For the eight European aquifer areas with the most available tritium data (this includes the four in the Middle-East following the Lambert projection as presented in Figure 2.1), groundwater flow path lines have been simulated at a resolution of 10, 5, 2 and 1 km; from here on defined as 10K, 5K, 2K and 1K respectively.

				TT$_1$	TT$_2$	TT$_3$
ID# (X$_1$,Y$_1$)						
ID# (X$_2$,Y$_2$)						
ID# (X$_3$,Y$_3$)						

1.

Empty array:
- Rows – X$_c$,Y$_c$ and ID
- Columns – 24 travel time bins

Data from 3-D flow path simulations

X$_i$,Y$_i$	TT (yrs)
X$_a$,Y$_a$	13
X$_b$,Y$_b$	185
X$_d$,Y$_d$	67

2.
- Data points (X$_i$,Y$_i$) are linked to cell locations (X$_c$,Y$_c$).
- Counts per bin are filled in array.

	TT$_1$	TT$_2$	TT$_3$
ID# (X$_1$,Y$_1$)	0	1	1
ID# (X$_2$,Y$_2$)	1	0	2
ID# (X$_3$,Y$_3$)	0	3	1

3.
- Select IDs; aquifer area or continent.
- Counts per bin are summed, and sample size (n) determined.

	TT$_1$	TT$_2$	TT$_3$	
ID# (X$_1$,Y$_1$)	0	1	1	
ID# (X$_2$,Y$_2$)	1	0	2	
ID# (X$_3$,Y$_3$)	0	3	1	
Σ	1	4	4	n

Travel time CDFs

4.
- Bins are summed for every cell (n_c).
- CDFs are produced for every cell.
- Ratio of modern GW ($F_{modern,c}$), which is linked to X$_c$,Y$_c$, is derived.

Modern GW distribution map (%)

GW volume per cell (km^3)

5.

Groundwater volume of a cell
*
Counts per bin of a cell
/
Total counts of a cell =

Groundwater volume of a cell
/
Area of a cell
*
$F_{modern,c}$ =

GW volume CDFs

Modern GW distribution map (m)

Figure 2.2: Flow diagram presenting the process of extracting the travel times and modern groundwater distribution from the 3-D flow path simulations.

CDFs presenting the time values of every European aquifer area will be produced and examined for each resolution in order to decide which of these four is the most adequate for global flow path simulations. Multiple aspects should be taken into account during the sensitivity analysis; besides the time values, the number of data points is of significance as well; whether or not cells are sufficiently covered represents an important reliability factor.

First the age distributions will be analysed in CDFs for each considered resolution. These four distribution curves present the proportion of the total amount of particles within the aquifer area for a certain age. This interpretation is however not possible for the comparison between the 3-D modelled results and the ones presented by Gleeson et al. (2015), as these numbers are not available for all three methods.

Alternatively the second part of the sensitivity analysis compares the absolute groundwater volumes. It must be mentioned that the results are not expected to converge as the set aquifer boundaries are not similar. Ternary plots are considered as a highly applicable way to visually present this comparison, as these two-dimensional triangular diagrams have the ability to outline data sets on three axes. Conventionally they are used to present the ratios of three variables that sum up to either 1 or a 100%, e.g. the proportions of sand, silt and clay in a soil sample (Graham & Midgley, 2000).

For this research the modern groundwater volumes of the thirty aquifer areas from the tritium based and the 2-D modelled methods presented by Gleeson et al. (2015) will be examined in relation to the 3-D modelled results following from this study itself. In its current form these values are neither dependent on each other, nor adding up to 1. Therefore alternative assumptions must be made.

Fictional aquifers A, B, C and D are considered in order to understand how the diagram works. The modern groundwater volumes are presented in km^3 for each method in Table 2.3. In order to remove the bias between methods, the modern groundwater volumes are standardised across the columns, which allows for a more direct comparison and a better understanding of the modern groundwater distribution over the aquifer areas within one method, using the following equation:

$$z = \frac{X - \mu}{\sigma}$$

where
- z is the standardised value
- X is the observation
- μ is the mean
- σ is the standard deviation

The standardised values present the difference compared to the mean in terms of the standard deviation, which can result in both positive and negative values. Ternary plots are however not able to project the latter. As a solution the minimum standardised value per method (underlined in Table 2.4) is subtracted from every standardised value for that method, which results in values greater than or equal to zero. These three minimums must be proportionate in order to have the smallest deviation.

The adjusted values are subsequently converted into ratios, where the sum of all three modern groundwater volumes for a convex hull zone across a row is assumed as 1. The ratios are presented in Table 2.5, and plotted in Figure 2.3. The plot, also serving as a legend, shows dark blue areas, which represent a bias towards one specific method considering the standardised modern groundwater volume for an aquifer area, middle blue areas, which represent an overlap of two methods, and a lighter blue area, which represents (moderate) similarity for all three methods. It can be observed that, for instance, data point A (upper left) can be associated to a convex hull area that contains relatively more modern groundwater for the tritium method in comparison to the other two methods.

When analysing thirty aquifer areas at the same time, a ternary plot will not only allow for an effective visual comparison, but these diagrams can also add another level of information if the size or colour of a data point is used to represent an extra variable.

Table 2.3: The modern groundwater volumes of fictional aquifers A, B, C and D.

Aquifer	Tritium (km³)	2-D modelled (km³)	3-D modelled (km³)
A	160	24	7
B	39	5	71
C	19	63	24
D	44	17	18

Table 2.4: The standardised values of the modern groundwater volumes of fictional aquifers A, B, C and D, where the column minimums, which will be subtracted from every value in a column, are underlined.

Aquifer	Tritium (-)	2-D modelled (-)	3-D modelled (-)
A	1.14	-0.23	<u>-0.81</u>
B	-0.44	<u>-0.87</u>	1.12
C	<u>-0.70</u>	1.09	-0.30
D	-0.38	-0.46	-0.48

Table 2.5: The ratios derived from the adapted standardised modern groundwater volumes for fictional aquifers A, B, C and D.

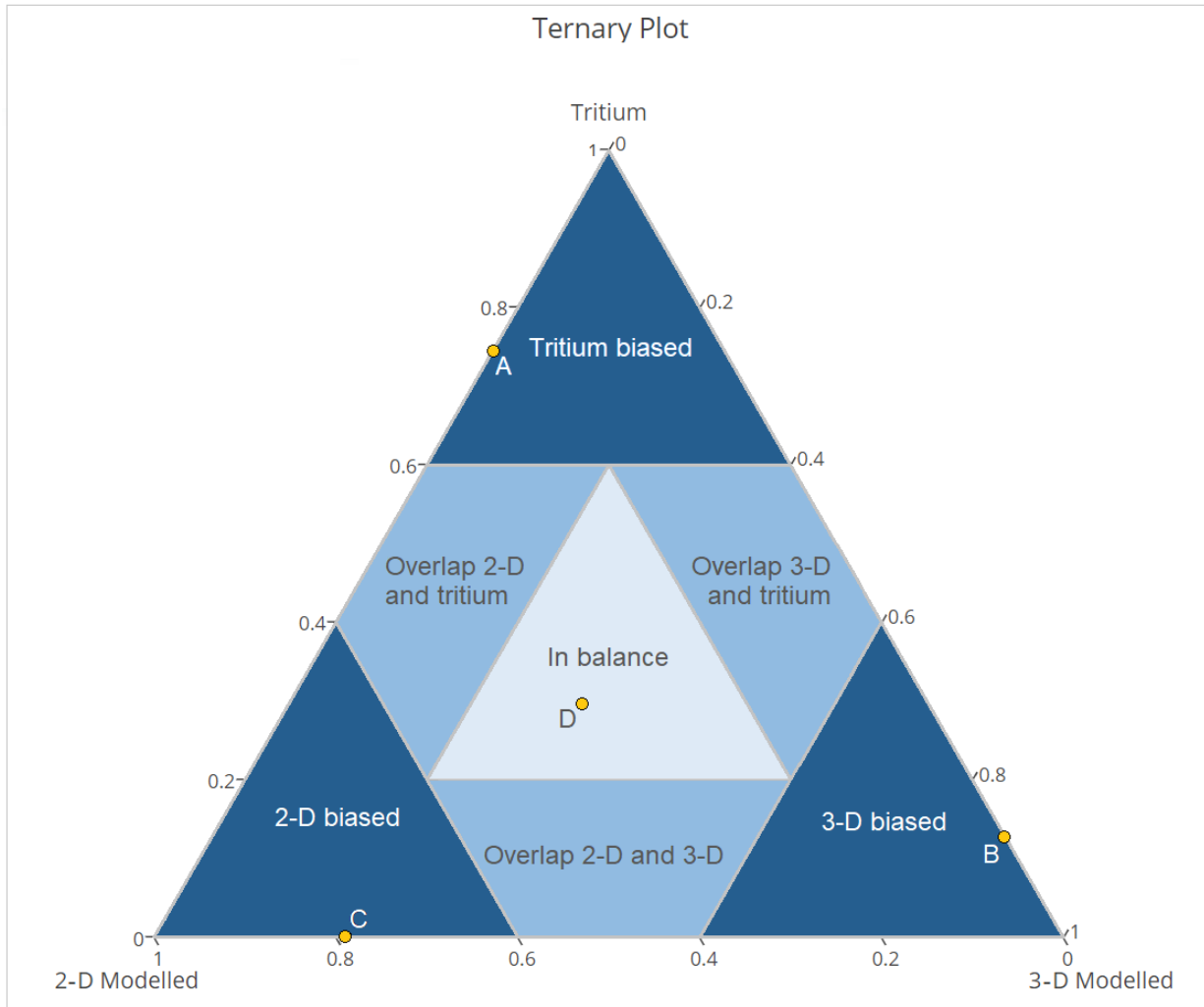| Aquifer	Tritium (km³)	2-D modelled (km³)	3-D modelled (km³)
A	0.74	0.26	0.00
B	0.12	0.00	0.88
C	0.00	0.79	0.21
D	0.31	0.38	0.31

Figure 2.3: Ternary plot exhibiting the converted modern groundwater volumes for fictional aquifers A, B, C and D. Coloured areas define the following: dark blue; a bias towards one specific method, middle blue; an overlap or a bias of/towards two methods, light blue; a (moderately) balanced outcome for all three methods.

3   Results

3.1   Sensitivity analysis

The primary goal of the sensitivity analysis is to determine what resolution will be used for the global simulations. With increasing resolution, flow paths can be simulated more truthfully and realistically. However this effect will be limited by the fact that the original data and groundwater simulations by de Graaf et al. (2015) were performed at 5 arc minutes (~10 km) and have been merely subsampled when deriving the flow paths with MODPATH. Moreover, from a practical perspective the computation time should be taken into account as well; a coarser resolution may prove to be of satisfactory quality for this purpose of this study.

Travel times were extracted for the eight aquifer areas in Europe at 10K, 5K, 2K, and 1K resolution (section 2.3). Figures 3.1a-h present the cumulative distribution functions (CDFs) for each aquifer area and all resolutions. The original data can be found in appendix D; the comprehensive results for all aquifer areas per resolution, including significant values such as the number of data points per km^2 and the total amount of modern groundwater in km^3.

It can be easily derived from figures 3.1a-h that a resolution of 10K will not suffice for this research. In not a single graph its distribution intersects with the dashed line, suggesting no modern groundwater at all.

As for the remaining three resolutions, overall the curves look fairly similar. Although the absolute modern groundwater volumes seem to be convincing for the 5K, the average 0.59 data points per km^2 make these results less reliable, as cells are probably not sufficiently covered.

Remarkably the 2K results show more modern groundwater than the 1K results, respectively 19.57 km^3 and 8.50 km^3 on average. Although the 1K resolution might be preferred considering the presumed higher accuracy of the simulations, a cell resolution of 2 kilometres appears to be fine enough to get an adequate representation of the pathways, and provides ample data points per km^2. Taking into account that an estimated eight billion data points would cover earth's surface for the 1K resolution, and the correlated extensive computation times, the cell resolution of 2K will be used for both the other aquifer areas and the global flow path simulations.

3.2   Controls on groundwater age

A comprehensive overview of the results of all thirty aquifer areas, simulated at 2K, can be found in Appendix E, including the modern groundwater volume for every individual aquifer area. In order to compare the results with the ones presented by Gleeson et al. (2015), the absolute volumes have been standardised, converted into ratios, as discussed in section 2.4, and are plotted in Figure 3.2. Table 3.1 presents the accompanying aquifer area names and numbers. The absolute modern groundwater volumes of all thirty aquifer areas for the three methods next to each other are presented in Appendix F.
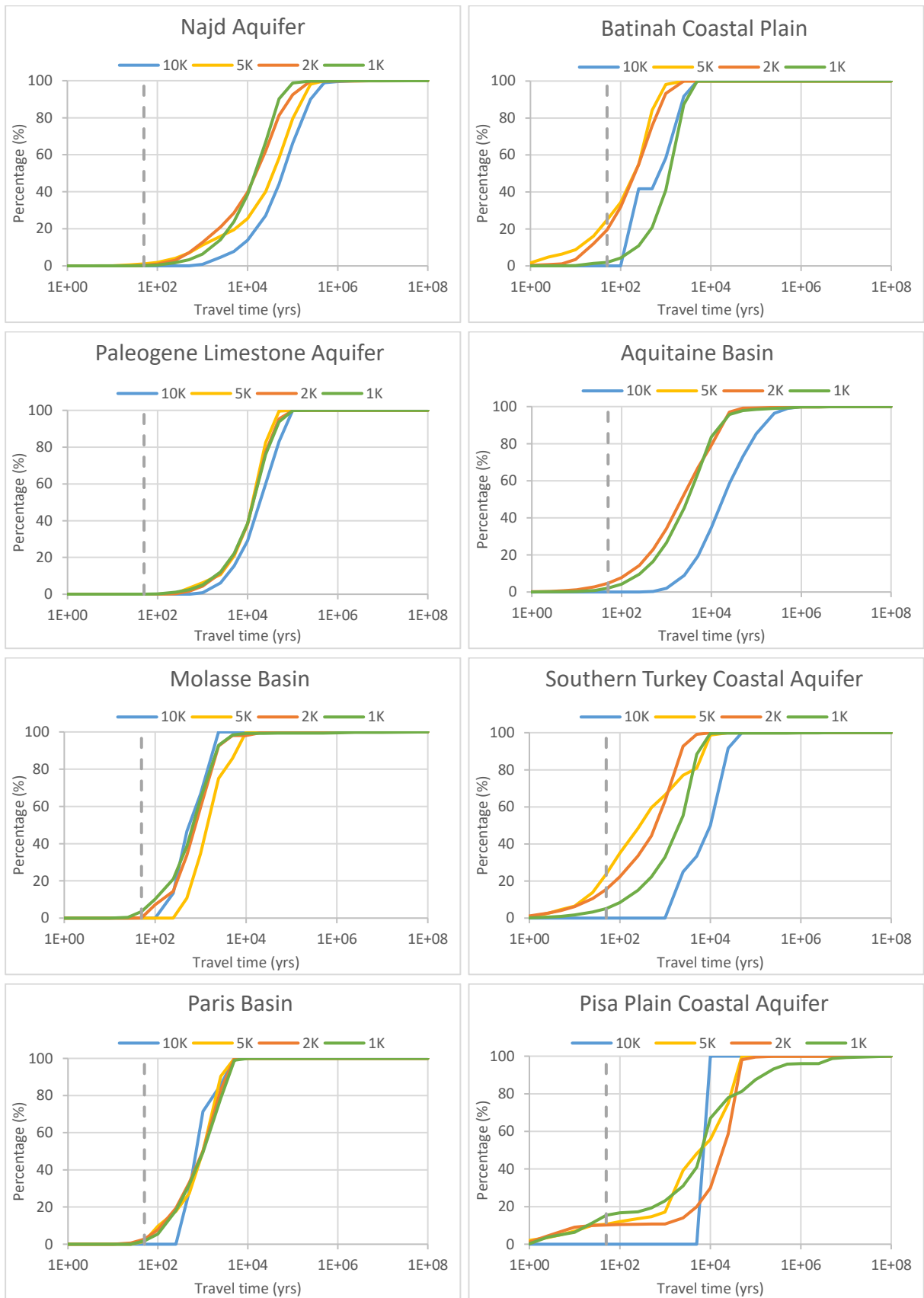
Figures 3.1a-h: Travel time cumulative distribution functions (CDFs) for all four resolutions for the eight considered European aquifer areas (dashed grey line is located at fifty years, indicating the limit of modern groundwater).

In Figure 3.2 ternary plots are shown. The upper left plot shows the agreement in modern groundwater volume. All aquifer areas seem to bias towards one or two methods (with the exception of an overlap of 3-D and tritium), where only the California basin (#1) approaches a balance between all three methods.

For the other three ternary plots in Figure 3.2, the size of the data point varies to represent additional information. For the upper right plot, the surface area was chosen as this is consistent among all methods, where the volume is unavailable for the tritium measurements and it varies with the assumed depth, which differs for the two model approaches. Where a few large aquifer areas including the (by far) largest (i.e. #18; Kalahari Desert Ntane Sandstone) indicate a strong bias towards the tritium method, multiple smaller aquifer areas seem particularly biased towards the 3-D modelled method. For the bottom two ternary plots the size of the data point defines the average porosity (left) and the average transmissivity (right). Aquifer areas that bias towards the 3-D modelled method, including four located in Asia, show a variety for both variables. Remarkably aquifer areas with small values for both porosity and transmissivity, which include four located in North America, seem to bias towards the 2-D modelled method.

3.3   Global groundwater storage

In the sensitivity analysis the resolution for the global flow path simulations has been set at 2K. Thereupon it is now possible to run the hydrogeological 3-D model for every continent, and answer the third research question; what is the age distribution of global groundwater?

Similar to the CDFs presenting the travel time distribution for each of the four resolutions of all European aquifers, Figure 3.3 shows the travel time distribution for each continent, which is the result of nearly 1.8 billion simulated flow path data points. Since in this research the groundwater age is considered as a function of particles only, which are not explicitly tied to the volume of groundwater flow across a cell, it must be noted that 100% on the y-axis represents the total amount of data points for every separate continent, and therefore does not indicate the absolute groundwater volume. As can be observed, South-America is the continent relatively containing more young water, where Australia and Africa hold more old water. Climate is presumably the controlling factor, which indirectly points at the recharge and the groundwater flow volume.

Under the assumption of a proportional distribution of water volumes with the tracked ages within a cell (or aquifer area), as explained in section 2.3, the total groundwater volume and the absolute modern groundwater volume can be computed as well. Figure 3.4 presents the groundwater volume for each continent in a stacked area chart. A similar distribution of young and old groundwater as presented in Figure 3.3 can be recognised, where in order Africa, Asia and Australia hold the most, but also the oldest groundwater, which is also evident from the stacked bar chart in Figure 3.5.

Table 3.1: The numbers and names of the thirty aquifer areas considered by Gleeson et al. (2015), where the colours of the numbers represent the continent, as indicated in the ternary plots.

Cont.	#	Aquifer area name	Cont.	#	Aquifer area name
NA	1	California basin		16	Paris Basin
	2	High Plains Aquifer		17	Pisa Plain Coastal Aquifer
	3	Middle Rio Grande Basin	AF	18	Kalahari Desert Ntane Sandstone
	4	Mississippi embayment Memphis aquifer		19	Tadla basin
	5	Pleistocene till plain		20	Chad Sedimentary basin
	6	Villa de Reyes basin		21	Karoo Aquifer
	7	Eastern Snake River Plain		22	Sidi Bouzid plain
	8	Central Oklahoma alluvium and terrace aquifer	AS	23	Bengal Basin
SA	9	Santiago Basin		24	Yinchuan plain
EU	10	Najd Aquifer		25	North China Plain
	11	Batinah Coastal Plain		26	Songnen Plain aquifer
	12	Paleogene Limestone Aquifer		27	Taoyuan Chungli Tableland
	13	Aquitaine Basin		28	KrishnaGodavari alluvial basin
	14	Molasse Basin	AU	29	Canterbury Plains
	15	Southern Turkey Coastal Aquifer		30	Perth Basin
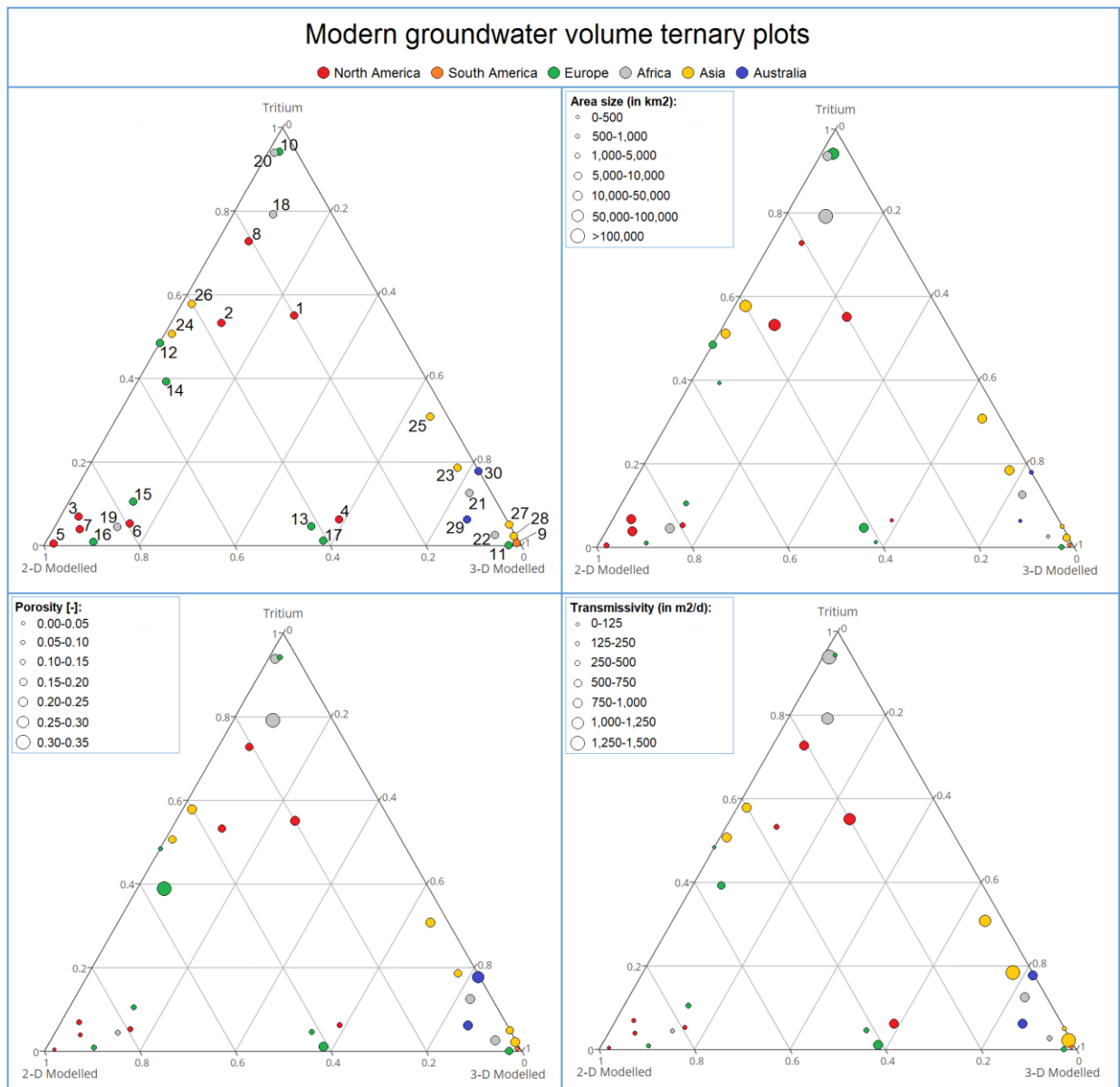
Figure 3.2: Ternary plots presenting the correlation between the three methods considering the aquifer area's modern groundwater volume, where from upper left to bottom right the aquifer area numbers, the area size, the average porosity, and the average transmissivity are considered.

A comprehensive overview of both continental and global groundwater results can be found in Appendix G, including the global groundwater volume and the modern groundwater volume. The first, which can also be derived from Figure 3.4 is estimated at 4.227 million km^3. Although simulations resulted in an F_{modern} of 1.040%, the proportion of modern groundwater on earth relative to the global groundwater volume exceeds this considerably, which is a consequence of the cell-scaled volume computation. This ratio is determined at 3.477%, which represents a global modern groundwater volume of 0.147 million km^3. Asia, Africa and Europa have the largest share in this; respectively 26%, 23% and 22%.

Following from the model results, global distribution maps have been created; Figures 3.6a-b present from left to right the total groundwater equivalent expressed in m, the dimensionless fraction of modern groundwater relative to the total groundwater, and the modern groundwater equivalent in m, which is a result of combining the first two. High values seem to be systematically located at mountainous and coastal areas.
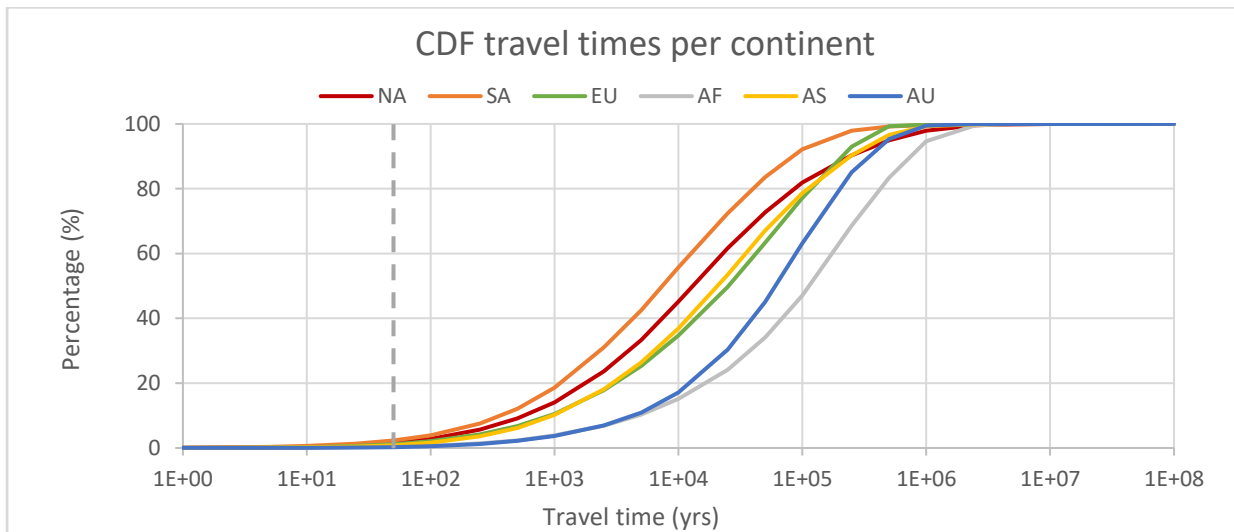
Figures 3.3: Travel time CDF for all six continents (dashed grey line indicates boundary of modern groundwater).
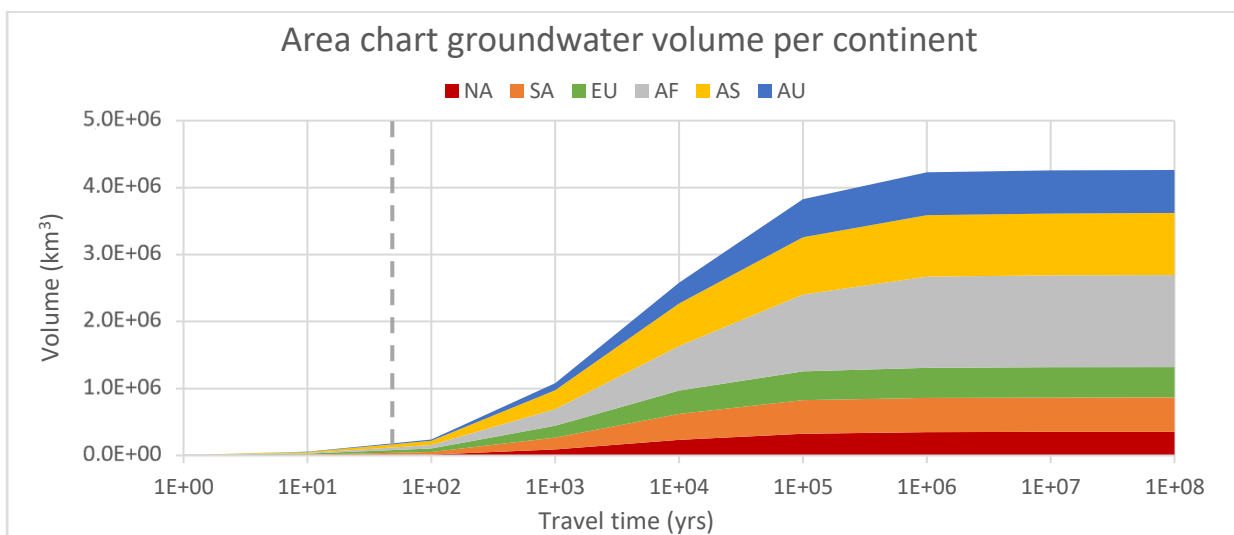
Figure 3.4: Stacked area chart of the total groundwater volume of each continent plotted against the travel time in years (dashed grey line indicates boundary of modern groundwater).
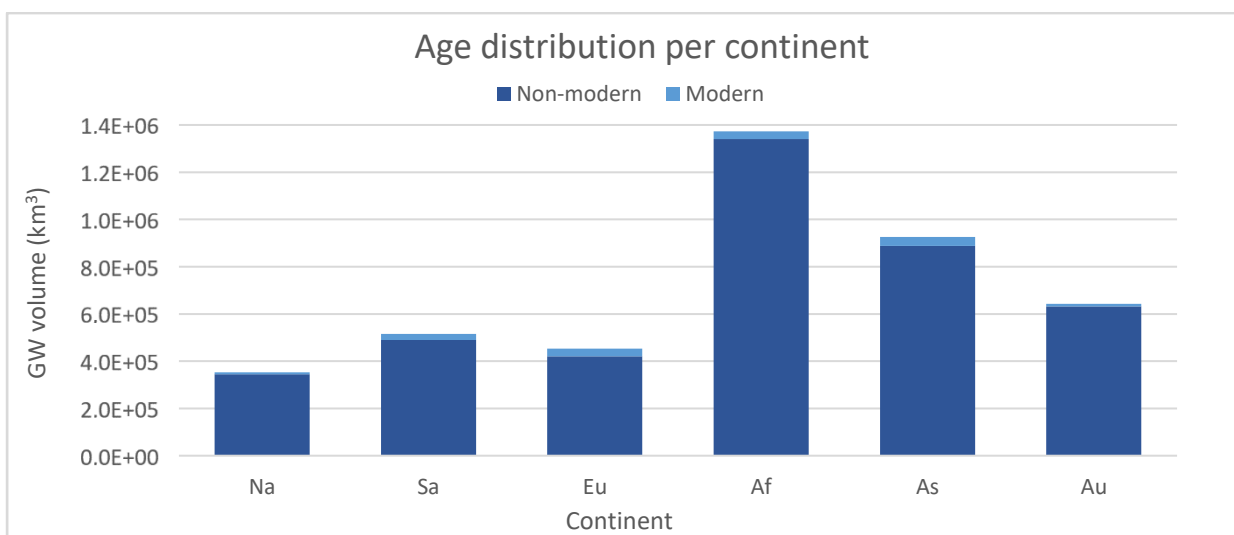
Figure 3.5: Stacked bar chart presenting both the total and modern groundwater volume per continent.

Total groundwater equivalent (m)	Fraction of modern groundwater (-)	Modern groundwater equivalent (m)

☐	0 – 2.50	☐	0 – 0.001	☐	0 – 0.25
☐	2.50 – 5.00	☐	0.001 – 0.010	☐	0.25 – 0.50
☐	5.00 – 10.0	☐	0.010 – 0.100	☐	0.50 – 1.00
☐	10.0 – 25.0	☐	0.100 – 0.500	☐	1.00 – 2.50
☐	25.0 – 50.0	☐	0.500 – 1.000	☐	2.50 – 5.00
☐	50.0 – 100			☐	5.00 – 10.0
☐	> 100			☐	> 10.0

Figure 3.6a: The global distribution of the total groundwater equivalent in m (left), the fraction of modern groundwater, F_{modern} (middle) and the modern groundwater equivalent in m (right) for the northern continents. Legend displays boundaries from and including the lower limit to the upper limit.

Total groundwater equivalent (m)	Fraction of modern groundwater (-)	Modern groundwater equivalent (m)
0 – 2.50	0 – 0.001	0 – 0.25
2.50 – 5.00	0.001 – 0.010	0.25 – 0.50
5.00 – 10.0	0.010 – 0.100	0.50 – 1.00
10.0 – 25.0	0.100 – 0.500	1.00 – 2.50
25.0 – 50.0	0.500 – 1.000	2.50 – 5.00
50.0 – 100		5.00 – 10.0
> 100		> 10.0

Figure 3.6b: The global distribution of the total groundwater equivalent in m (left), the fraction of modern groundwater, F_{modern} (middle) and the modern groundwater equivalent in m (right) for the southern continents. Legend displays boundaries from and including the lower limit to the upper limit.

4   Discussion

4.1   Limitations of the model

The first steps of this research have focussed primarily on establishing proper input files and an adequate set up of the hydrogeological model within iMOD. First of all it must be mentioned that this research has used a single model layer with no vertical decay of any hydrogeological variable, which is certainly unrealistic and less advanced than the analysis by Gleeson et al. (2015). Also, the DEM is assumed as the top of the saturated layer instead of the water table, which is a simplification also used by Gleeson et al. (2015); in that perspective it might be a more accurate comparison between the methods themselves. This assumption might induce deeper groundwater flow and overestimated recharge rates (Gleeson et al, 2015). Although the model of Gleeson et al. (2015) can be assumed to be more detailed in terms of hydrogeological characterisation, this research is certainly of added value as it considers the third dimension, and therefore includes the topographical control on convergent and divergent groundwater flow.

The evaluated models differ in the depth of the aquifer. The 3-D model used here, using the geomorphology-based aquifer presented by de Graaf et al. (2015), happens to be thinner than the p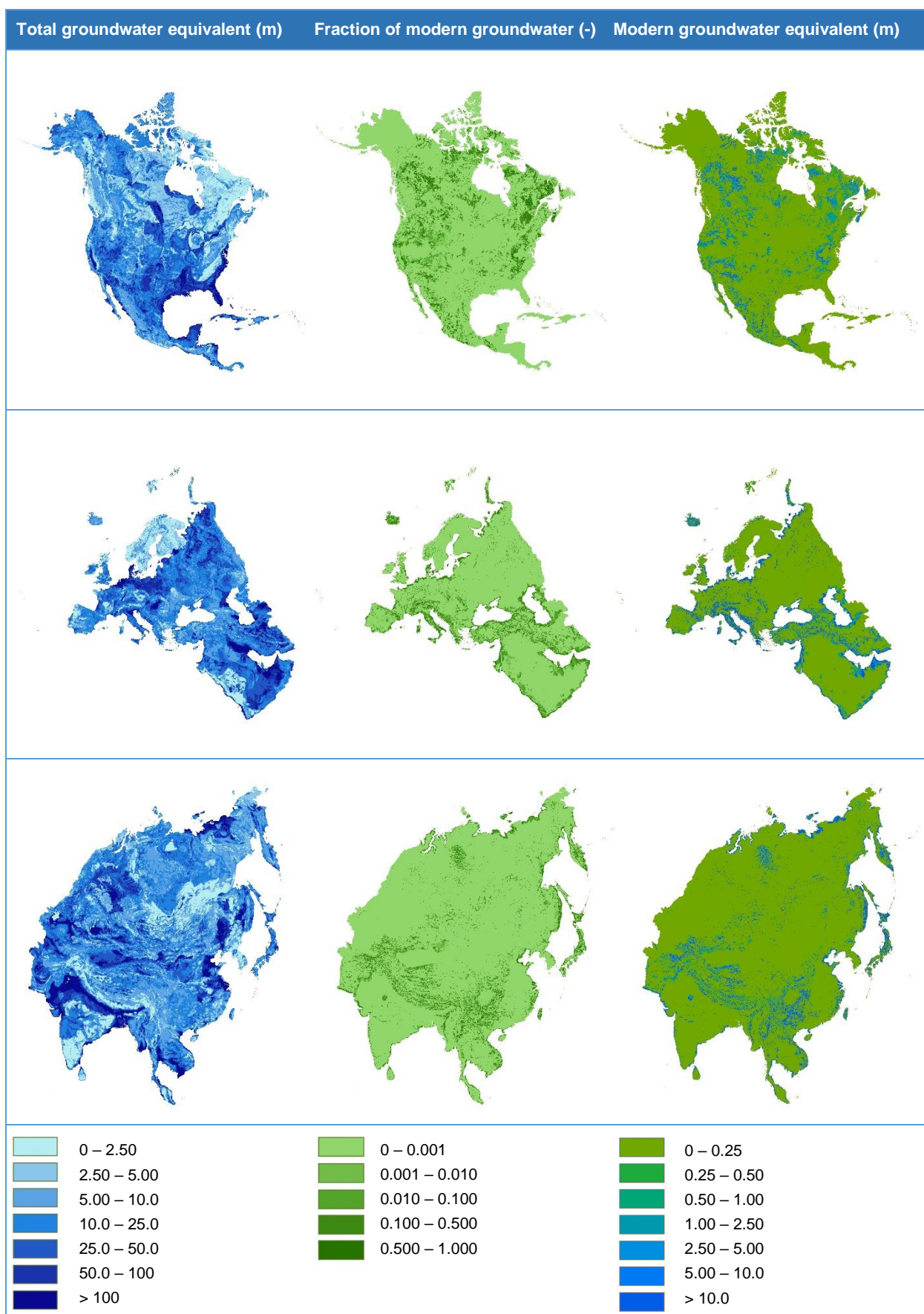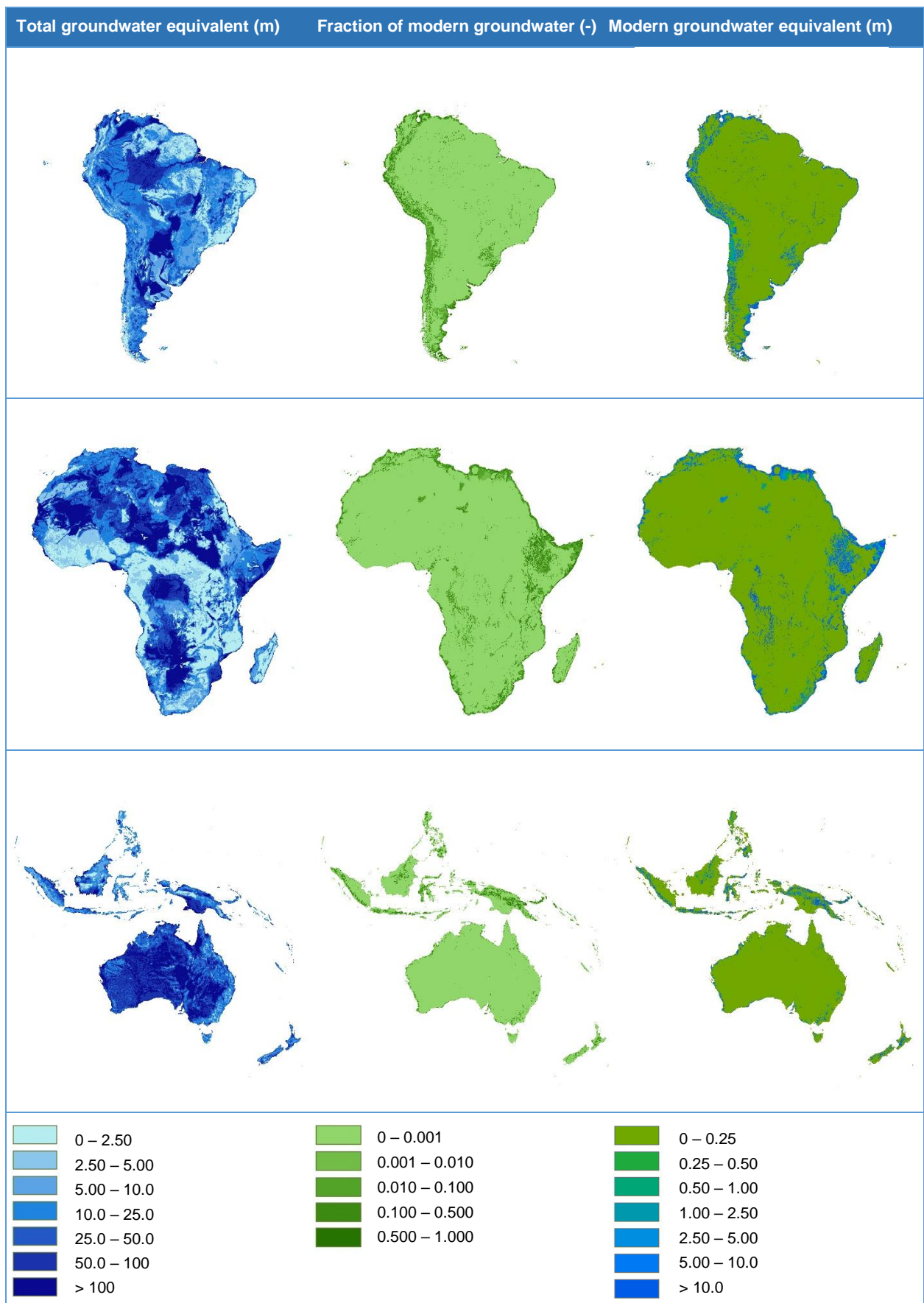ermeable layer from Gleeson et al. (2015) almost everywhere. This may lead to an underestimation of groundwater volumes. Yet, the impact on the relative estimate of modern groundwater is expected to be minimum, as this is predominantly located in the first hundreds of meters below the surface and its presence decreases with depth, where less recharge occurs (Gleeson et al., 2015).

To improve upon the research by de Graaf et al. (2015) and to remove the distortion that is the consequence of the WGS84 projection used, the Lambert equal-area projection was adopted for this research. Each continent is projected separately relative towards a shifted pole, and results in an equidistant-grid where angles become distorted (Figure 2.1).

In spite of the modelled 2K resolution, the majority of the data had an original resolution of 10K. A variation of spatial interpolation methods have been used to resample the data over a finer grid that goes hand in hand with a few complications. The increased resolution secures a near total coverage of data points, and a fuller registration of flow paths, which is all the more important for research on modern groundwater. However, increasing the resolution does not add detail; for instance the DEM will be too smooth over complex terrains that may underrepresent the control of rivers on the groundwater head.

Just as de Graaf et al. (2015) stated, it is impossible to capture all the small-scale flow paths in shallow groundwater systems with the current model resolution. One might propose a resolution of 100 m for instance, which would in theory approach reality the best. Sadly there is no input data to support this, nor is it warranted to disaggregate the current data to this spatial resolution.

An alternative form of disaggregating the model would be setting out more particles per cell, and thereby create more flow paths in total, which would result in better statistics for every resolution. Multiple particles in one cell will follow the exact same path however, if not for the inclusion of weak sinks in the model, which indicates that the flow of a particle stops when entering a cell where the discharge to sinks is larger than a specified fraction of the total cell inflow. It is impossible for the model itself to decide truthfully whether a particle should stop or not, nor is it easy to determine the correct fraction that represents reality adequately. The length of flow paths could be heavily influenced by weak sinks, and therefore the fraction of modern groundwater as well. In the case that it might be combined with a finer resolution, it becomes all the more important, as the occurrence of weak sinks will increase (Vermeulen et al., 2016). Its influence could be assessed by comparing the results with this study's modern groundwater volume estimate.

Considering the flow path simulations, the most significant limitation on the derived age distribution is probably the assumption that every particle represents the same amount of groundwater flow, i.e. the model simulates flow paths without taking the volume of flow into account. Where the impact on the estimate of the modern groundwater volume is difficult to determine, it is safe to say that the total groundwater volume estimate is not affected by this deficiency. This restriction could possibly be corrected for, by setting out more particles per cell as discussed earlier, which would produce a more comprehensive and balanced time distribution per cell, however computation times must be kept in mind as well.

4.2   Discussion of the results

This study looked into how to compute the global volume and the distribution of modern groundwater by means of high-resolution 3-D flow path simulations in iMOD. As a research with a similar objective has been executed before (Gleeson et al., 2015), results have been compared on aquifer-scale and global-scale.

By means of the sensitivity analysis it was possible to determine what higher spatial resolution should be used for this research to achieve the most accurate flow path simulations, where computation times must be kept in mind as well. Modern groundwater volumes have been compared for selected aquifer areas at a 10K, 5K, 2K and 1K resolution. Maybe against the expectations, overall the 2K results showed a higher volume of modern groundwater than the 1K results for the European aquifer areas. Remarkably the 5K results seem to be even more equivalent to the 2K results. It is more than likely that this can be imputed to the disaggregating process. This might have a certain threshold value that probably lies in between a resolution of 1 and 2 km, based on the average modern groundwater volumes from the thirty selected aquifer areas (10K: 0.00 km^3, 5K: 18.97 km^3, 2K: 19.57 km^3 & 1K: 8.50 km^3).

According to the aquifer area results, the average modern groundwater volume derived from tritium samples (680 km^3) considerably exceeds the 2-D and the 3-D simulated numbers (respectively 175 km^3 and 28.7 km^3). The ternary plots however, which consider the distribution

over the thirty aquifer areas within each method (due to the standardisation), set forth a different relation. With overall relatively large aquifer area sizes and a high average porosity and transmissivity it can be stated that wide, deep sedimentary (i.e. permeable non-consolidated material) basins tend to bias towards the tritium method, and in lesser extent towards the 3-D modelled method. In particular multiple aquifer areas located in Asia cause to favour the latter, which can presumably be explained by a relatively large groundwater recharge generated by the local environment (e.g. monsoon and mountainous areas), and quick drainage to rivers (e.g. the Ganges and the Yellow River), leading to short flow path lines, and therefore relatively more modern groundwater. The 2-D modelled method is overrepresented by less productive regions with a low permeability, i.e. often hard-rock reservoirs in highlands.

Furthermore, as tritium samples are taken from wells, which can be all on the same location for some aquifer areas, differences can be quite significant as some aquifer areas are nothing more than a circular buffer zone with a radius of 10 km, and therefore will not cover a lot of flow path data points for the modelled methods in comparison to larger areas. Moreover the locations of wells globally are generally biased towards lower altitudes, which often act as discharge areas. This could explain for a local bias for modern groundwater, however not a bias towards a specific method.

The aquifer dimensions (Lehner et al., 2006; de Graaf et al., 2015; Shangguan et al., 2016) combined with the effective porosity (Gleeson et al., 2014) have resulted in a total groundwater volume estimate of 4.2 million km^3, as presented in Figure 3.4; similar to a groundwater equivalent of 33 m. This is considerably lower than the total groundwater volume estimate by Gleeson et al. (2015) of 22.6 million km^3 (16-30 million km^3 considering the uncertainty in porosity), similar to a groundwater equivalent of 180 m. Therefore the total groundwater volume assumed in this research is almost four times smaller than the minimum estimate by Gleeson et al. (2015). The substantial contrast is presumably due to the different spatial boundaries for the bottom of the aquifer per model, which will affect the absolute amount of modern groundwater as well.

Following from global estimates derived from nearly 1.8 billion simulated flow path data points, the percentage of groundwater younger than fifty years is estimated at 3.5%, which is equal to a global volume of 147 thousand km^3 and an effective global depth of 1.1 m, as presented in Table 4.1. Although the tritium based method by Gleeson et al. (2015) results in a ratio of modern groundwater that is only 2.1% higher, the volume and the modern groundwater equivalent are almost one order of magnitude larger. Although the modern groundwater fraction derived from the 3-D model exceeds the small ratio of 1.5% from the 2-D modelled results, it still produces a volume and a modern groundwater equivalent more than double the 3-D based results. This contradictory correlation is also a consequence of the difference in layer thickness. The modern groundwater estimate of this research does however fall within the overall uncertainty-based range set by Gleeson et al. (2015), i.e. 0.1–5.0 million km^3. Although it can be inferred from

Figure 1.4, that the simulations presented by de Graaf et al. (2015) presumably could result in more modern groundwater than the estimate from this study, a numerical comparison is impossible as the necessary data are unavailable.

Table 4.1: The global amount of modern groundwater expressed as a percentage of the total groundwater, the cubic volume and the depth if it was extracted and pooled at the earth's land surface like a flood, for the two methods by Gleeson et al. (2015) and the 3-D modelled results from this research.

Method	Ratio modern (%)	Modern groundwater volume (10^5 km^3)	Modern groundwater equivalent (m)
Tritium	5.6	12.7	10.1
2-D modelled	1.5	3.5	2.7
3-D modelled	3.5	1.5	1.1

As modern groundwater is said to be climate-controlled, a comparison with groundwater recharge cannot be omitted. Global hydrological models by Döll & Fiedler (2008) and Wada et al. (2010) respectively present an average annual global groundwater recharge of 12,700 km^3 per year and 15,200 km^3 per year. This is similar to a total groundwater recharge ranging from 635,000 km^3 to 760,000 km^3. Hence, this study's global modern groundwater volume estimate only represents 19.3% – 23.1% of the water that has entered the subsurface during fifty years. The remainder is assumed to have already discharged, which includes baseflow. The tritium-based modern groundwater estimate (1,270,000 km^3) by Gleeson et al. (2015) exceeds the total recharge of fifty years by far, which can only be justified by the fact that recharge is a poorly defined quantity, while the 2-D modelled estimate represents 46.1% –55.1% of the recharged groundwater during fifty years.

The distribution of the total groundwater, presented in Figures 3.6a-b on the left, seems to be dependent on multiple factors; primarily the permeability and the river network. Parts of rivers clearly stand out (e.g. the Niger, the Po, and the Nile) whereas for other rivers (e.g. the Amazon, the Tigris and the Ganges) the entire basin is in contrast with the surroundings. The modern groundwater fraction maps and the modern groundwater distribution maps, presented in Figures 3.6a-b respectively in the middle and on the right, show a lot less variation in their distribution patterns. They both seem to be almost directly correlated to the coast and mountains. At the coast, flow paths will be shorter due to early groundwater discharge in the sea and river deltas. Mountain ranges stand out due to factors such as the impermeability, the aquifer thickness and the topography. The modern groundwater distribution does not show a correlation with climate; a primary factor proposed by Gleeson et al. (2015). This can be ascribed to the fact that the 3-D modelled results are considerably less heterogeneous.

Surprisingly the stream density does not seem to have much impact (i.e. the number of streams per km^2), which should be much higher in more humid areas. It could be expected to be more important, as this basically determines the length of flow paths, and therefore the travel times.

This is presumably an effect of the disaggregated DEM that does not represent the drainage density correctly.

The smoothed modern groundwater distributions are considered as a direct consequence of the disaggregated DEM; topography has become smoother and some river streams have dissolved. In order to improve the model results it is a priority to use an alternative DEM; for instance the one from the Hydro1k dataset, which has an original resolution of 1 km. One would start with a finer DEM, then aggregate it, and subsequently incise the rivers, so that the river network would not be mispresented again (flows can be blocked by the DEM of higher topographies), which is expected to be of great influence on the drainage levels. This alternative method with a higher resolution drainage network cannot be assumed to be ideal however, as the DEM might not be in agreement with the groundwater heads simulated by de Graaf et al. (2015).

5   Conclusion

The primary objective of this study was to improve the estimate of the global volume and the distribution of modern groundwater. The previous results by Gleeson et al. (2015) and the flow path simulations from de Graaf et al. (2015) have been used as a foundation on which a global high-resolution hydrogeological 3-D model was built. The basis of this simulation was the adaptation of an equidistant projection to represent the flow paths more truthfully, replacing the original geographic coordinate system WGS84 to an equal-area Lambert projection, and increasing the spatial resolution from the original 5 arc minutes (~10 km) to 2 km (2K). For this new model, which combines spatial input data from various sources, flow paths have been simulated in order to provide a potent projection of the groundwater age distribution.

Gleeson et al. (2015) have determined the volumes of modern groundwater (younger than 50 years) for thirty aquifer areas with the most available tritium data. The two data sets, tritium based and 2-D model based, considerably exceed this study's 3-D simulated estimates. The controls of these differences have been examined by means of ternary plots, which consider the standardised modern groundwater distribution. A high porosity and transmissivity induce a relatively large volume of modern groundwater for this study and the tritium based method, which can be coupled to deep sedimentary basins, while lower values for these variables, representing hard-rock reservoirs, bias towards the 2-D model.

The 3-D model estimates the total global groundwater volume at 4.2 million km^3, of which 147 thousand km^3 can be defined as modern groundwater. The first is substantially lower than the estimate by Gleeson et al. (2015), i.e. a total groundwater volume between 16 and 30 million km^3. The latter falls just within the range of modern groundwater, which is 0.1 to 5.0 million km^3. The difference is for a substantial part due to the dissimilar aquifer thickness. Furthermore the flow path simulations from the 3-D model might underestimate the modern groundwater volume due to disaggregating the original input data.

One of the main goals of this research was to find out what the distribution of modern groundwater as derived from the global hydrogeological 3-D model looks like. This could be particularly significant in order to determine where lateral groundwater flows are most important to the distribution of modern groundwater; something that other models have not yet been able to do. The somewhat unsatisfying results are unable to provide a better understanding of location specific differences in groundwater movement however. A recommended first step of improvement would be using a DEM with a higher original resolution, so the synergy between groundwater and rivers will be no longer disparaged. Nonetheless can it be concluded that this approach of groundwater modelling is promising for the comprehension of the groundwater distribution and for producing a robust estimate of the global volume.

6   References

Alley, W.M., Healy, R.W., LaBaugh, J.W. & Reilly, T.E. (2002). Flow and storage in groundwater systems. Science, 296, 1985-1990.

Bethke, C.M., & Johnson, T.M. (2008). Groundwater age and groundwater age dating. Annual Review of Earth and Planetary Sciences, 36, 121–152.

Chahine, M.T. (1992). The hydrological cycle and its influence on climate. Nature, 359, 373-380.

de Graaf, I.E.M., Sutanudjaja, E.H., van Beek, L.P.H. & Bierkens, M.F.P. (2015). A high-resolution global-scale groundwater model. Hydrology and Earth System Sciences, 19, 823-837.

de Graaf, I.E.M., van Beek, L.P.H., Gleeson, T., Moosdorf, N., Schmitz, O., Sutanudjaja, E.H., Bierkens, M.F.P. (2016). A global-scale two-layer transient groundwater model: development and application to groundwater depletion. Hydrology and Earth System Sciences Discussions, 121, in review

Diem, S., Renard, P., & Schirmer, M. (2014). Assessing the effect of different river water level interpolation schemes on modeled groundwater residence times. Journal of Hydrology, 510, 393-402.

Döll, P., & K. Fiedler (2008). Global- scale modeling of groundwater recharge, Hydrol. Earth Syst. Sci., 12, 863–885.

Erskine, A.D., & Papaioannou, A. (1997). The use of aquifer response rate in the assessment of groundwater resources. Journal of Hydrology, 202, 373-391.

Fan, Y., Li, H. & Miguez-Macho, G. (2013). Global patterns of groundwater table depth. Science 339, 940-943.

Garmonov, I.V., Konoplyantsev, A.A. & Lushnikova, N.P. (1974). The World Water Balance and Water Resources of the Earth, Hydrometeoizdat, 48-50.

Gleeson, T., Smith, L., Moosdorf, N., Hartmann, J., Dürr, H.H., Manning, A.H., van Beek, L.P.H., & Jellinek, A.M. (2011). Mapping permeability over the surface of the Earth. Geophys. Res. Lett., 38.

Gleeson, T., Wada, Y., Bierkens, M.F.P., and van Beek, L.P.H. (2012). Water balance of global

aquifers revealed by groundwater footprint. Nature, 488, 197–200.

Gleeson, T., Moosdorf, N., Hartmann, J., & van Beek, L.P.H. (2014). A glimpse beneath earth's surface: GLobal HYdrogeological MaPS (GLHYMPS) of permeability and porosity. Geophys. Res. Lett., 41, 3891–3898.

Gleeson, T., Befus, K.M., Jasechko, S., Luijendijk, E., & Bayani Cardenas, M. (2015). The global volume and distribution of modern groundwater. Nature Geoscience, 9, 161-167.

Graham, D.J. & Midgley, N.G. (2000). Graphical representation of particle shape using triangular diagrams: an Excel spreadsheet method. Earth Surface Processes and Landforms 25(13): 1473-1477.

Harbaugh, A.W., Banta, E.R., Hill, M.C., & McDonald, M.G. (2000). MODFLOW-2000, The U.S. geological survey modular ground-water model-user guide to modularization concepts and the ground-water flow process, U.S. Geological Survey, 10.

Hartmann, J. & Moosdorf, N. (2012). The new global lithological map database GLiM: a representation of rock properties at the Earth surface. Geochem. Geophy. Geosy., 13.

Jasechko, S. (2016). Partitioning young and old groundwater with geochemical tracers. Chemical Geology 427, 35–42.

Kazemi, G.A., Lehr, J.H., & Perrochet, P. (2006). Groundwater Age. John Wiley and Sons, Hoboken, New Jersey, USA, 325.

Lehner, B., Verdin, K., & Jarvis, A. (2006) HydroSHEDS Technical Documentation v1.0, United States Geological Survey, 1-27.

Manning, A.H., Clark, J.F., Diaz, S.H., Rademacher, L.K., Earman, S., Plummer, L.N. (2012). Evolution of groundwater age in a mountain watershed over a period of thirteen years. Journal of Hydrology, 460-461, 13–28.

McDonald, M.G., and Harbaugh, A.W. (1988). A modular three-dimensional finite-difference ground-water flow model: U.S. Geological Survey Techniques of Water-Resources Investigations, book 6, chap. A1, 586 p.

McGuire, K.J., & McDonnell, J.J. (2006). A review and evaluation of catchment residence time modelling. Journal of Hydrology, 330, 543-563.

Pollock, D. (1994). User's Guide for MODPATH/MODPATH-PLOT, version 3: A particle tracking post-processing package for MODFLOW, the US Geological Survey finite-

difference groundwater flow model. Tech. rep., US Geological Survey Open-File Report, 94-464.

Shangguan, W., Hengl, T., de Jesus, J. M., Yuan, H. and Dai, Y. (2016). Mapping the global depth to bedrock for land surface modeling. J. Adv. Model. Earth Syst.

Schmitz, O., Karssenberg, D., van Deursen, W., & Wesseling, C. (2009). Linking external components to spatio-temporal modelling framework: coupling MODFLOW and PCRaster. Environ. Modell. Softw., 24, 1088–1099.

Steinwand, D.R., Hutchinson, J.A., and Snyder, J.P. (1995). Map projections for global and continental data sets and an analysis of pixel distortion caused by reprojection: Photogrammetric Engineering and Remote Sensing, v. 61, 1487-1497.

Sutanudjaja, E.H., van Beek, L.P.H., de Jong, S.M., van Geer, F.C., & Bierkens, M.F.P. (2011). Large-scale groundwater modeling using global datasets: a test case for the Rhine-Meuse basin, Hydrol. Earth Syst. Sci., 15, 2913-2935.

Tóth, J. (1963). A theoretical analysis of groundwater flow in small drainage basins, J. Geophys. Res., 68(16), 4795–4812.

Traylor, J.P., & Zlotnik, V.A. (2016). Analytical modeling of irrigation and land use effects on streamflow in semi-arid conditions. Journal of Hydrology, 533, 591-602.

van Beek, L.P.H., Wada, Y., & Bierkens, M.F.P. (2011). Global monthly water stress: 1. Water balance and water availability. Water Resour. Res., 47.

Vermeulen, P.T.M., Burgering, L.M.T., & Minnema, B. (2016). iMOD; User Manual. Deltares, 265-320.

Wada, Y., van Beek, L.P.H., van Kempen, C.M., Reckman, J.W.T.M., Vasak, S., & Bierkens, M.F.P. (2010). Global depletion of groundwater recourses, Geophys. Res. Lett., 37, 2.

Wood, E.F., Roundy, J.K., Troy, T.J., van Beek, R., Bierkens, M., Blyth, E., de Roo, A., Döll, P., Ek, M., Famiglietti, J., Gochis, D., van de Giesen, N., Houser, P., Jaffe, P., Kollet, S., Lehner, B., Lettenmaier, D.P., Peters-Lidard, C.D., Sivapalan, M., Sheffield, J., Wade, A. J., & Whitehead, P. (2012). Reply to comment by Keith, J. Beven and Hannah, L. Cloke on "Hyperresolution global land surface modeling: Meeting a grand challenge for monitoring Earth's terrestrial water". Water Resour. Res., 48.

Zimmerman, U., Münnich, K. O., Roether, W., Kreutz, W., Schubach, K., & Siegel, O. (1966). Tracers determine movement of soil moisture and evapotranspiration. Science, 152, 346–347.

Appendices

Appendix A

'C:\iMOD34\Results\Lambert_As_2k'
1   1   1   0   1   0   0   0   0   0   0   0
     1   0   0   0   0   0   1.0
          100   1250   0.001000   0.1   0.98 0.98
          -3700000 -3900000   4300000   4300000 2000 1.0 0.0
ACTIVE MODULES
1,1,0 (bnd)
1,1,0 (shd)
1,1,0 (kdw)
1,1,0 (drn)
1,1,0 (riv)
1,1,0 (rch)
'C:\iMOD34\Input_data\As_2k\landmask.IDF'
MODULES FOR EACH LAYER
1,(bnd)
  1,1.0,0.0,'C:\iMOD34\Input_data\As_2k\LANDMASK.IDF'
1,(shd)
  1,1.0,0.0,'C:\iMOD34\Input_data\As_2k\gw_head.IDF'
1,(kdw)
  1,1.0,0.0,'C:\iMOD34\Input_data\As_2k\transmissivity.IDF'
PACKAGES FOR EACH LAYER AND STRESS-PERIOD
1,0.0,STEADY-STATE,1
1,(drn)
  1,1.0,0.0,'C:\iMOD34\Input_data\As_2k\kq3.IDF'
  1,1.0,0.0,'C:\iMOD34\Input_data\As_2k\base_s3_used.IDF'
1,(riv)
  1,1.0,0.0,'C:\iMOD34\Input_data\As_2k\riv_cond_comb.IDF'
  1,1.0,0.0,'C:\iMOD34\Input_data\As_2k\riv_head_comb.IDF'
  1,1.0,0.0,'C:\iMOD34\Input_data\As_2k\riv_bot_comb.IDF'
  1,1.0,0.0,1.0
 1,(rch)
  1,1000.0,0.0,'C:\iMOD34\Input_data\As_2k\rch_md.IDF'

Flow path simulations in iMOD

In order to compute the streamflow, the runfile can simply be operated through iMODFLOW, which is used in combination with iMODPATH, by means of the windows command prompt. Before the path line simulations can be started, the starting locations of particles, i.e. the so-called start points, have to be defined. All points of the area that will be considered are located inside a polygon, specified in three-dimensional coordinates. The lateral distribution of the start points is 2 km, which is the distance between every flow path particle in both the X and Y-direction. The depth distribution of the particle start points can be given theoretical numerical values, but are defined by the DEM for the top-level and the bottom of the layer for the base in this case.

The location of the set of particles is defined, which means the particles can be tracked through time. The path lines can be simulated by means of the flow field computed by iMODFLOW, which will be used as the model input as they are transported by advection (Vermeulen et al., 2016). Before running the simulations, iMOD presents the results that are available and the amount of model layers. It will show only the budget flow for the right face (x) and front face (y), as there is no particle movement to a lower face (z) because this model only assumes a singular aquifer layer.

Further input variables concern the boundary settings for instance, which is defined by the landmask. This file rules the extent of active flow that gives a value greater than zero to every location where particle tracking is allowed, and a value lower or equal to zero to inactive groundwater flow areas, such as the oceans. The layer boundaries are defined by the DEM and bottom file, and the porosity by the by the storage coefficient as presented by Gleeson et al. (2014).

As for the time settings, iMOD is able to either run the model for a specific period selection, or a maximal trace time of a certain amount of years. The latter is best applicable for this research, as a steady-state is assumed.

Lastly the weak and strong sinks can be defined (Figure B1). A weak sink can be described as a cell where a minimum of one flow component is directed outwards, while a strong sink will take all the water and is not associated to outflow at all (Vermeulen et al., 2016). Although they are both purely model-based properties, which do not exist in nature in, they could be of significance for this model. During the groundwater flow simulation particles will be stopped when they enter an adjacent model cell, of which the discharge is larger than a specific ratio. Unfortunately the flow path algorithm is not designed to determine whether a particle should stop, or flow on to the next cell. However it is unknown how particles passing through cells with weak sinks will behave. Furthermore it is impossible for the model itself to decide truthfully whether a particle

should stop or not, nor is it easy to determine the correct fraction that represents reality adequately. Naturally the approach towards sinks affects the flow paths substantially.
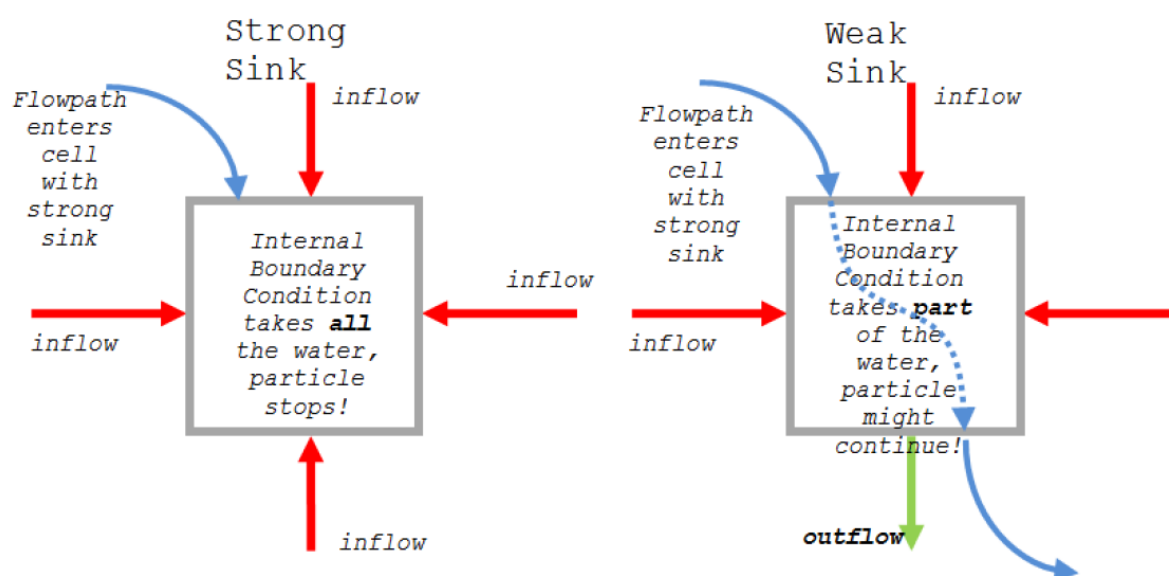
Figure B1: The concept of a strong sink (left) and a weak sink (right) within a flow model (Vermeulen et al., 2016).

The flow paths can be simulated now, this is possible in either a forward or backward trace direction. The first will be computed, as this is the natural direction of flow. The simulation results, in the form of an interchange file format (*.iff), will provide the following data:

- Particle number (total particle number will be dependent on the grid size)
- Model layer number (i.e. will be 1 for our entire model)
- X coordinate of the particle
- Y coordinate of the particle
- Z coordinate of the particle (not to be confused with the budget flow lower face (z), which is only active in case of multiple model layers)
- Elapsed time in years since the moment of release of the particle, i.e. groundwater age or travel time
- Velocity in meters per day of the particle on that location
- Row number of location
- Column number of location

Only the coordinates and travel time are relevant for the goals of this research. Subsequently the data is processed as discussed in section 2.3.

Appendix C

Used python scripts are available on Google drive through this link:

https://drive.google.com/drive/folders/1x-FLXJZCsdtdRxNCyX-1yT7EuQ4Q0nbF

It contains the following scripts:

- compute_iMOD_data.py: used to compute adjusted data values for the bottom of the layer, the porosity, and the transmissivity.
- resample_iMOD_data_2k.py: used to reproject the input data from WGS84 to Lambert.
- conversion_iff_to_csv.py: used to convert the output from iMOD (*.iff) to a conventional file type (*.csv).
- traveltimes_as.py: used to derive modern groundwater volumes from flow path simulations per continent, in this case Asia.
- compute_values_per_ID.py: used to calculate average values of porosity and transmissivity for the thirty aquifers analysed in section 3.2.

Appendix D

Table D1-4: Comprehensive overview of the aquifer area data considered during the sensitivity analysis for each of the four resolutions, including F_{modern}, which is the percentage of total groundwater younger than fifty years, and the modern groundwater volume in km^3.

Resolution = 10K						
Aquifer area	Data Points	Area (km²)	Points per km²	GW volume (km³)	F_{modern} (%)	Modern GW volume (km³)
Najd Aquifer	7952	5850	1.36	1841.16	0.00	0.00
Batinah Coastal Plain	12	700	0.02	35.88	0.00	0.00
Paleogene Limestone Aquifer	517	5300	0.10	83.66	0.00	0.00
Aquitaine Basin	1571	20100	0.08	699.24	0.00	0.00
Molasse Basin	15	400	0.04	30.18	0.00	0.00
Southern Turkey Coastal Aquifer	12	900	0.01	61.39	0.00	0.00
Paris Basin	7	700	0.01	21.34	0.00	0.00
Pisa Plain Coastal Aquifer	1	200	0.01	18.84	0.00	0.00
Average	1123	3839	0.19	310.82	0.000	0.00

Resolution = 5K						
Aquifer area	Data Points	Area (km²)	Points per km²	GW volume (km³)	F_{modern} (%)	Modern GW volume (km³)
Najd Aquifer	44052	58850	0.75	1848.16	1.11	22.60
Batinah Coastal Plain	1198	1500	0.80	54.24	24.79	40.28
Paleogene Limestone Aquifer	3155	5200	0.61	80.35	0.06	0.18
Aquitaine Basin	17024	19400	0.88	673.23	4.94	78.43
Molasse Basin	84	325	0.26	23.10	0.00	0.00
Southern Turkey Coastal Aquifer	1077	1500	0.72	45.03	23.86	24.18
Paris Basin	62	650	0.10	15.87	1.61	0.77
Pisa Plain Coastal Aquifer	199	325	0.61	31.27	10.55	4.33
Average	7451	9789	0.59	308.25	7.44	18.97

Resolution = 2K						
Aquifer area	Data Points	Area (km²)	Points per km²	GW volume (km³)	F_{modern} (%)	Modern GW volume (km³)
Najd Aquifer	384060	58860	6.52	1790.24	0.33	2.13
Batinah Coastal Plain	3846	1708	2.25	52.17	19.42	28.92
Paleogene Limestone Aquifer	41411	5228	7.92	76.86	0.00	0.00
Aquitaine Basin	21070	19272	1.09	652.09	4.73	115.21
Molasse Basin	730	312	2.34	21.69	0.14	0.04
Southern Turkey Coastal Aquifer	9542	1628	5.86	36.59	15.50	25.60
Paris Basin	1286	668	1.93	17.74	2.80	1.15
Pisa Plain Coastal Aquifer	1520	316	4.81	30.62	10.26	3.11
Average	51704	9812	4.29	297.83	5.91	19.57

Resolution = 1K						
Aquifer area	Data Points	Area (km²)	Points per km²	GW volume (km³)	F$_{modern}$ (%)	Modern GW volume (km³)
Najd Aquifer	2858063	58872	48.55	1794.97	0.14	5.14
Batinah Coastal Plain	102424	1575	65.03	54.92	1.94	9.74
Paleogene Limestone Aquifer	317488	5208	60.96	77.61	0.03	0.47
Aquitaine Basin	385599	18917	20.38	620.12	2.10	42.07
Molasse Basin	4878	308	15.84	21.24	3.57	1.05
Southern Turkey Coastal Aquifer	72680	1257	57.82	37.59	5.19	9.53
Paris Basin	9197	668	13.77	18.42	1.81	1.23
Pisa Plain Coastal Aquifer	1709	306	5.58	25.98	15.45	7.29
Average	418077	9714	35.80	294.82	3.36	8.50

Appendix E

Table E1-6: Comprehensive overview of the aquifer area data considered during the sensitivity analysis for all thirty aquifer areas at a resolution of 2K.

North America						
Aquifer area	Data Points	Area (km^2)	Points per km^2	GW volume (km^3)	F$_{modern}$ (%)	Modern GW volume (km^3)
California basin	47272	10384	4.55	1000.74	0.79	6.04
High Plains Aquifer	558965	80944	6.91	3506.90	0.02	8.74
Middle Rio Grande Basin	119473	10768	11.10	277.88	0.27	6.62
Mississippi embayment Memphis aquifer	1217	304	37.28	12.71	2.47	0.38
Pleistocene till plain	11332	1556	1.91	25.16	3.95	0.47
Villa de Reyes basin	2979	1120	126.03	22.52	6.58	2.50
Eastern Snake River Plain	141148	22848	0.76	312.00	1.79	15.41
Central Oklahoma alluvium and terrace aquifer	17328	4216	4.11	343.40	0.72	0.27

South America						
Aquifer area	Data Points	Area (km^2)	Points per km^2	GW volume (km^3)	F$_{modern}$ (%)	Modern GW volume (km^3)
Santiago Basin	12308	792	15.54	44.87	9.03	17.75

Europe						
Aquifer area	Data Points	Area (km^2)	Points per km^2	GW volume (km^3)	F$_{modern}$ (%)	Modern GW volume (km^3)
Najd Aquifer	384060	58860	6.52	1790.24	0.33	2.13
Batinah Coastal Plain	3846	1708	2.25	52.17	19.42	28.92
Paleogene Limestone Aquifer	41411	5228	7.92	76.86	0.00	0.00
Aquitaine Basin	21070	19272	1.09	652.09	4.73	115.21
Molasse Basin	730	312	2.34	21.69	0.14	0.04
Southern Turkey Coastal Aquifer	9542	1628	5.86	36.59	15.50	25.60
Paris Basin	1286	668	1.93	17.74	2.80	1.15
Pisa Plain Coastal Aquifer	1520	316	4.81	30.62	10.26	3.11

Africa						
Aquifer area	**Data Points**	**Area (km^2)**	**Points per km^2**	**GW volume (km^3)**	**F$_{modern}$ (%)**	**Modern GW volume (km^3)**
Kalahari Desert Ntane Sandstone	5157057	249472	20.67	36626.07	0.00	26.50
Tadla basin	105527	19720	5.35	484.75	2.34	21.13
Chad Sedimentary basin	351929	29964	11.75	3844.25	0.01	1.93
Karoo Aquifer	54156	6900	7.85	519.23	0.07	11.05
Sidi Bouzid plain	532	320	1.66	15.47	20.49	6.41

Asia						
Aquifer area	**Data Points**	**Area (km^2)**	**Points per km^2**	**GW volume (km^3)**	**F$_{modern}$ (%)**	**Modern GW volume (km^3)**
Bengal Basin	104696	65024	1.61	6516.05	1.77	123.51
Yinchuan plain	382753	10436	36.68	963.87	0.01	0.65
North China Plain	378762	46352	8.17	4644.15	1.89	266.71
Songnen Plain aquifer	6087437	47222	128.91	17234.97	0.01	5.65
Taoyuan Chungli Tableland	5743	888	6.47	23.83	23.54	6.71
KrishnaGodavari alluvial basin	13399	5184	2.58	580.58	13.27	154.68

Australia						
Aquifer area	**Data Points**	**Area (km^2)**	**Points per km^2**	**GW volume (km^3)**	**F$_{modern}$ (%)**	**Modern GW volume (km^3)**
Canterbury Plains	472	260	1.82	29.41	2.12	1.78
Perth Basin	4220	656	6.43	69.49	1.66	1.01
Ovens Catchment	342399	23664	14.47	2193.36	0.00	2.04

Appendix F

Table F1: Computed modern groundwater volume for each of the thirty considered aquifer areas, for all three methods.

Aquifer area	Tritium based modern GW volume (km³)	2-D modelled modern GW volume (km³)	3-D modelled modern GW volume (km³)
California basin	296.31	24.24	6.04
High Plains Aquifer	1013.99	152.22	8.74
Middle Rio Grande Basin	241.24	707.80	6.62
Mississippi embayment Memphis aquifer	1.05	1.17	0.38
Pleistocene till plain	4.13	154.11	0.47
Villa de Reyes basin	19.04	63.20	2.50
Eastern Snake River Plain	249.19	1211.30	15.41
Central Oklahoma alluvium and terrace aquifer	61.89	3.97	0.27
Santiago Basin	0.93	0.84	17.75
Najd Aquifer	1718.06	13.50	2.13
Batinah Coastal Plain	0.18	4.02	28.92
Paleogene Limestone Aquifer	32.44	8.02	0.00
Aquitaine Basin	208.55	460.67	115.21
Molasse Basin	5.90	1.82	0.04
Southern Turkey Coastal Aquifer	385.20	712.93	25.60
Paris Basin	1.89	50.58	1.15
Pisa Plain Coastal Aquifer	1.55	10.93	3.11
Kalahari Desert Ntane Sandstone	6003.38	219.96	26.50
Tadla basin	161.14	679.30	21.13
Chad Sedimentary basin	2674.36	28.86	1.93
Karoo Aquifer	40.55	2.65	11.05
Sidi Bouzid plain	4.69	1.22	6.41
Bengal Basin	664.23	36.85	123.51
Yinchuan plain	577.58	121.17	0.65
North China Plain	2760.35	88.84	266.71
Songnen Plain aquifer	3238.30	498.85	5.65
Taoyuan Chungli Tableland	9.64	0.08	6.71
KrishnaGodavari alluvial basin	38.57	4.07	154.68
Canterbury Plains	3.79	0.80	1.78
Perth Basin	5.27	0.03	1.01

Appendix G

Table G1: Overview of the continental data, and the global numbers as a result of the hydrogeological 3-D flow path simulations.

Continent	Data Points (10^6)	Area (10^6 km^2)	Points per km^2	GW volume (10^3 km^3)	F_{modern} (%)	Modern GW volume (10^3 km^3)
North America	153.14	21.33	7.18	352.69	1.56	8.18
South America	91.48	17.54	5.21	515.18	2.29	24.91
Europe	204.37	16.33	12.51	452.43	1.16	31.93
Africa	814.58	29.40	27.71	1345.98	0.35	33.59
Asia	377.52	34.99	11.11	917.78	0.84	37.99
Australia	147.42	10.38	14.20	642.88	0.27	10.38
World	**1788.52**	**128.97**	**13.87**	**4226.94**	**1.04**	**147.0**