Review of the methodologies used to derive groundwater characteristics for a specific area in The Netherlands

H.P. Ritzema[1*], G.B.M. Heuvelink[2], M. Heinen[3], P.W. Bogaart[4], F.J.E. van der Bolt[3], M.J.D. Hack-ten Broeke[3], T. Hoogland[3], M. Knotters[3], H.T.L. Massop[3], H.R.J. Vroon[†3], H. van den Bosch[2]

[1] Water Resources Management Group, Wageningen University & Research, Wageningen, The Netherlands

[2] ISRIC – World Soil Information, Wageningen, The Netherlands

[3] Wageningen Environmental Research, Wageningen University & Research, Wageningen, The Netherlands

[4] University of Utrecht, Utrecht, The Netherlands

* Corresponding author: H.P. Ritzema (henk.ritzema@wur.nl)

† Deceased 27 May 2016.

Highlights:

- The system approach to upscale groundwater characteristics in time and space has changed over time
- The selected combinations of upscaling methods affects the resulting groundwater characteristic
- A systematic approach that minimize or avoid the impact of errors in upscaling is not used
- Based on these findings, we cannot conclude whether drought stress is under- or overestimated

Abstract

In this paper, we analyze the methods that are used in The Netherlands to upscale in-situ groundwater measurements in time and in space, and how the selected combinations of upscaling methods affect the resulting groundwater characteristic. In The Netherlands, a three-step approach is used to obtain groundwater characteristics for a specific area: (1) in-situ monitoring of the water table depth; (2) temporal upscaling; and (3) spatial interpolation and aggregation. The three-step approach is, however, not standardized, but a combination of the following methods is used: (i) four methods to measure/monitor the phreatic water table; (ii) four methods for temporal aggregation; and (iii) four methods for spatial interpolation and/or aggregation. Over the past sixty years, several combinations of these methods have been used. Our review shows that the use of these different combinations in the approach to measure and interpret water table depths has resulted in significant systematic differences in the corresponding groundwater characteristics and that there are many sources of potential error. Error in the in-situ measurement of the water table depth can be as high as 1 meter. Errors in the temporal aggregation are in the range of 10 to 20 cm and for the spatial interpolation between 20 to 50 cm. We show that there has been no systematic assessment of how these errors influence the resulting groundwater characterization. Thus, we cannot answer the question of whether drought stress in The Netherlands is under- or overestimated. Based on these findings we give recommendations for a systematic approach to groundwater characterizations studies that can minimize the impact of errors.

Keywords: phreatic groundwater table; in-situ measurements; hydropedology, temporal aggregation; spatial interpolation

1 Introduction

In many parts of the world, (ground) water depths are intensively monitored. These water table depths vary in time and space and depend on the interactive pedological and hydrological processes and their properties in the (un)saturated zone (Lin, 2012). Because we cannot measure water table depths everywhere and all the time we use temporal and spatial interpolation and aggregation methods to characterize the fluctuating water table depths (e.g. Van Heesen 1970;

Finke *et al.* 2004; De Vos *et al*. 2010;). All these methods have their pros and cons, which one is best depends very much on the objective(s) of the research and the availability of data. In The Netherlands, water depth classes are based on the mean highest (MHW) and mean lowest (MLW) water tables. The MHW and MLW are calculated from time series of dip-well records which have been collected on a national scale since the beginning of the 1950's (Van der Sluijs and De Gruijter 1985). Using profile and field characteristics, the MHW and MLW of the observation points are extrapolated for larger areas. Relationships are derived making it possible to convert water table classes into duration classes of water table depth. The water table at the beginning of the growing season (MSW) can also be derived from the MHW and the MLW.

The spatial and temporal representation of the water table depth or its characteristics in an area depends on the accuracy of the basic data, i.e. the measured water table depth. Various monitoring methods are used, i.e. observation wells, piezometers, open boreholes, (De Ridder 2006), thus it is important to realize that the measured water level is not necessarily equal to the position of the phreatic surface (e.g. Brassington 1992; Chapuis 2005 & 2009; Elci *et al*. 2003; Paydar and Richardson 2002; Van Duijvenbooden 1981). Furthermore, soils and hydrological conditions are in general not homogeneous (Berg and Christensen 1992), which influences the accuracy of the temporal and spatial upscaling. Numerous studies have described the natural uncertainties (e.g. barometric pressure) and errors made in the interpretation of groundwater level data (e.g. Saines 1981; Curch and Granato 1996; Dalton *et al*. 2007) and in the temporal and spatial upscaling (e.g. Knotters 2001; Stein 1991), but how these measurement and methodological errors affect the final upscaling result is often lacking.

To analyze the effects and ultimate impact of these errors and uncertainties, we have used The Netherlands as a case study. However such errors are not unique to The Netherlands and we believe that the lessons learned from our analysis will be useful for many other countries or regions of the world. The Netherlands, a low-lying country in Western Europe (50^o - 54^o N and 3^o - 8^o E), consists of deltas and former flood plains of the rivers Rhine, Meuse and Schelde (Colenbrander, 1989; Overeem *et al.* 2001). The total territory, including inland lakes, estuaries and territorial waters, is 41,543 km2, of which 55% is agricultural, 12% is nature, 19% is open water and the remaining 14% is built-up area (CBS, 2014). The land consists mainly of alluvial deposits and about 25% of the country lies below mean sea level (MSL). The lowest point is some 7 m below MSL. In the absence of dunes and dikes, more than 65% of the country would be flooded at high sea and high river levels (Van de Ven 1996). Average rainfall (851 mm/year) is substantially higher than the potential evaporation (559 mm/year) (KNMI 2014), thus drainage is a fact of life as it is required to use the land: for the inhabitants, for agriculture and for nature.

After the Second World War, agriculture intensified and more intensive drainage was required, resulting in deeper water tables, increased drainage rates and more drought stress in dry periods. This process was further intensified by an increase in groundwater abstraction and land consolidation practices employed to reduce the problems of fragmentation of land holdings (Van den Noort 1987). These land consolidation activities were often combined with improvement of the water management and road infrastructure (Prak 2002; Stańczuk-Gałwiaczek *et al*., 2018). It is estimated that the resulting average drop of the water table in agricultural areas has been in the range of 20 to 40 cm (Kremers and Van Geer 2000; Van der Sluijs and Van Heesen 1989). The drops of the water table in agricultural areas also resulted in deeper water tables in the neighboring nature areas (Martens et al. 2013; Van Tol et al. 1998). The water table in many areas is now significantly lower than the target values set by the Ministry of Transport, Public Works and Water Management (1999), not only in the man-made polders but also in the higher sandy areas in the east and south of The Netherlands. To counteract the adverse effects that these deeper water tables have on the environment, in particular nature reserves, the government has initiated policies to reverse the trend. However, in order to make informed decisions, policy makers and practitioners need reliable information on groundwater levels (Lijzen *et al*. 2014; De Lange *et al*. 2014).

In this research we contribute to the emerging interdisciplinary science of hydropedology by presenting an integrated, iterative methods for improved understanding of methodologies to derive groundwater characteristics for multiple scales. Hydropedology is an intertwined branch of soil science and hydrology that encompasses multiscale basic and applied research of interactive pedological and hydrological processes and their properties in the unsaturated zone (Lin, *et al*., 2005). The objectives of our research were to analyze (i) the methods used to measure and upscale groundwater level information in both time and space and (ii) how the selected combinations of these methods affect the resulting groundwater characterization. For this analysis we reviewed all projects assessing the characteristics of seasonal fluctuation in groundwater behavior conducted by the Dutch Soil Survey Institute and its successors over the last 25 years. These projects have been documented in more than 170 reports and papers (Ritzema *et al*. 2012). One test/aspect of our results will be to see if we can answer the question of whether drought stress in The Netherlands (as determined by groundwater levels) is under- or over-estimated.

After a brief discussion of the various definitions used for groundwater and groundwater characteristics, this paper presents:

- How different measurement methods and measuring depths lead to discrepancies or errors in the measured or estimated water table depth;
- How different temporal aggregation techniques lead to discrepancies or errors in the estimated characteristics of the seasonal fluctuation of water table depths;
- How different spatial interpolation and aggregation techniques lead to discrepancies or errors in the estimated spatial characteristics of the seasonal fluctuation of water table depths;
- The extent or impact that errors in these methods and techniques may have on the accuracy of the steps and ultimate determination of groundwater characteristic.

We conclude with a recommendation for a systematic approach that can minimize the effects of uncertainties and interpretation errors to provide the most consistent and robust estimate of groundwater characteristics for a specific area.

2 Definitions of groundwater and groundwater characteristics

In The Netherlands, several definitions of the hydrological parameters to define the position of the water table are used simultaneously, i.e. water table depth, groundwater level, phreatic level, phreatic surface, etc. This data is stored in "*Aquo-standard*", the data base of standardized concepts and definitions for data storage, exchange and processing for the Dutch water sector (http://www.aquo.nl/aquo-standaard). Water table depth is relative to the ground surface, whereas the other parameters are relative to a reference level. There are also various methods to measure the groundwater level, e.g. a groundwater observation well, piezometer, borehole, etc.

All these definitions and methods are used interchangeably with the result that it is often not unequivocally determined whether the water table or a piezometric level at larger depths is measured.

On top of this, different definitions are used to characterize the groundwater level. The phreatic water table is the upper groundwater level measured from the soil surface, but sometimes this is a perched groundwater table above a poorly permeable layer on top of an unsaturated deeper aquifer. This is in contrast to the (geohydrological) groundwater level which is defined as the depth at which the pressure at the groundwater level is equal to zero irrespective of the presence of a perched water table. To avoid confusion, it is essential to explicitly indicate whether the measurements refer to the phreatic water table or the (hydrogeological) groundwater.

The use of these different definitions results in correspondingly different estimates of the average water table depth over certain time periods: spring, autumn, hydrological year, 30 years, etc. (Table 1). In this article, the parameters MHW, MLW and MSW are summarized as MxW.

+++ **Table 1**. Hydrological characterizations used in The Netherlands to define the water table depth over time.

3. In-situ measurement of water table depths

To measure water table depths in the field, four in-situ methods are used: groundwater observation wells, piezometers, open boreholes and field estimates based on expert knowledge. These methods are standardized: International (ISO, IEC), European (EN) and national (NEN) standards that are used in The Netherlands (Netherlands Normalization Institute/NEN, https://www.nen.nl). Next to the four methods, the actual recording is also done in different ways: (i) with a measuring tape; (ii) a measuring tape with a sounder; (iii) electronically; (iv) with a float, and; (v) with a pressure gauge/indicator (De Ridder 2006). In The Netherlands, groundwater is monitored by many private and (semi-) governmental organizations (Van Duijvenbooden 1981). Most measurements are stored in a national data base, the "*DINOloket*" (www.dinoloket.nl). The above mentioned standards, however, only specify a limited number of aspects like the type of well, installation and measuring methods and the data storage and processing, but don’t address factors like location, filter depth, filter length, frequency, hydrological conditions (i.e. is it a seepage area or an area with natural drainage) and the soil characteristics (heterogeneity, anisotropy, etc.), etc. (e.g. Bartholomeus 2009; Freeze and Cherry 1979; Fetter 1980; Nielsen and Nielsen 2007; Richards 1931; Vroon *et al*. 1988).

All these factors influence the nature and accuracy of the measurements (e.g. Brassington 1992; Chapuis 2005 & 2009; Elci *et al*. 2003; Paydar and Richardson 2002; Saines 1981; Van Duijvenbooden 1981). Thus, what has really been measured is often not known or clear. In a groundwater observation well you can measure the (geohydrological) phreatic groundwater level, but if there is a (often unknown) poorly permeable layer it can also be the perched water table. In a piezometer you don’t measure the phreatic groundwater table but the piezometric pressure. When you use an open borehole you normally know the soil profile, and therefore can be aware of the occurrence of poorly permeable layers. If these exist, several boreholes to different depths must be used. Clearly some improvement or better definition in this area is needed/would be useful.

We assessed the four in-situ measuring methods on the basis of a set of evaluation criteria (after Knotters *et al*. 2010) (Table 2). The use of groundwater observation wells and/or piezometers scored better than the other two methods, mainly because they can be better reproduced than the other two methods. Although previously mentioned, it should be emphasized that, while the water level in both groundwater observation wells and piezometers can be objectively identified, (i) in a piezometer it is *not* the phreatic water table that is measured but the piezometric pressure, and (ii) the relationship with the phreatic groundwater depends on the soil profile (and presence or absence of poorly permeable layers) and / or hydrological conditions (infiltration or seepage).

++ Table 2. Assessment of the four methods used to measure the water table depth (++ stands for "complies with criterion" and – for "does not comply at all")

4. **Temporal aggregation of observed water table depths**

The next step in the analysis is the upscaling of in-situ water table depths to characteristics that summarize the temporal variation. There are four temporal aggregation methods: (i) direct calculation from time-series; (ii) statistical models; (iii) process models and; (iv) expert knowledge.

Temporal aggregation of a time-series of water table depths to directly calculate groundwater characteristics (GWC) is straightforward when the time-series is sufficiently long and has a sufficiently high measurement frequency. A period of 30 years is considered to be sufficiently long to compensate for deviations caused by extreme dry and/or wet years (Heesen 1970; Knotters (2001). In the case of direct calculation of groundwater characteristics using a (too) short time-series, errors are introduced. If the time-series is not long enough, statistical models and/or process models can be used (e.g. Von Asmuth *et al*. 2002; Von Asmuth and Knotters 2004; Knotters and de Gooijer 1999; Knotters and van Walsum 1997; Bierkens *et al.* 1999). If there is no time-series at all, but only a few measurements at pre-selected locations, regression methods can be used to estimate the GWC for this location (Finke *et al*. 2004). The accuracy of the estimates depends on the number of measurements: few measurements results in a lower accuracy.

Time-series modelling based on transfer function-noise (TFN) models derives the water table depth at a certain time from a (usually linear) function of the groundwater at a previous step in time, often in combination with other relevant information such as the precipitation surplus in the considered time interval (Changnon *et al. 1988;* ,Yi and Lee 2004). Process models can also be used to generate time-series of water table depths (Cirkel *et al*. 2010). Physical-mechanistic models have the advantage that they are based on physical laws and can be used for extrapolation and/or scenario analysis. Disadvantages are that model construction and model calibration are potentially more laborious compared to time-series models and that more input data is required (Berendrecht *et al*. 2004; Bierkens *et al*. 2001; Knotters and Bierkens 2000; Webster and Heuvelink 2006; Young and Beven 1994). If the groundwater characteristic (MxW) is calculated from the deterministic component of the model, e.g. the TF part of a TFN model or a physical-mechanistic model, a problem can be that through calibration levelling-out of the resulting groundwater characteristic can occur. This can be solved by including the stochastic component in the simulations.

Temporal scaling can also be done by expert knowledge. An expert can, based on the soil profile and field characteristics, make an estimate of the GWC (Van Heesen 1970)). Based on existing data and maps, an expert will make additional boreholes to assess the soil profile characteristics (texture, moisture conditions, depth of the water table, etc.). A prerequisite for a good assessment of the temporal groundwater characteristic (MxW) is a good understanding of the region.

All four methods for temporal aggregation have their pros and cons. Which one is best depends very much on the objective(s) of the research and the availability of data. The same criteria used to assess the four in-situ measuring methods (Table 2) were used to assess the temporal aggregation methods (Table 3). In practice, a combination of the above mentioned temporal aggregation methods are often used, for example expert knowledge in combination with process models, or a combination of a process model with a statistical model.

++ Table 3. Assessment of the four temporal aggregation methods used to characterize the temporal groundwater level (++ stands for "complies with criterion" and – for "does not comply at all")

5. **Spatial interpolation and aggregation**

We distinguished four categories of methods used for spatial interpolation and aggregation of groundwater characteristics: (i) expert knowledge; (ii) methods based on random sampling theory; (iii) geostatistical models and; (iv) physical-mechanistic models. Each of these methods has advantages and disadvantages (Table 4).

++ Table 4. Assessment of the four spatial interpolation and aggregation methods used to characterize the spatial groundwater level (++ stands for "complies with criterion" and - for "does not comply at all")

Experts can, based on their knowledge of an area, make an estimate of the GWC at locations in the area and average these GWCs for the whole area or parts of it. An expert will not only use the GWC measurement locations in the area, but also from surrounding locations and additional information, such as topographic maps, soil maps, open water levels in drains, ponds, creeks, etc. A disadvantage of this approach might be its irreproducibility. The spatial estimates of GWCs depend on the insights of the expert and therefore estimates may vary among experts. A further disadvantage might be that the accuracy of the spatial estimates can only be assessed by an additional validation. An advantage of the approach might be that the resulting maps will show patterns that can be recognized in the field, which make the maps more acceptable for users.

Random sampling for spatial aggregation has the important advantages that it is model-free and quantifies the uncertainty about the aggregated GWC (Brus and De Gruijter 1993; De Gruijter *et al*. 2006). The locations are drawn using a prescribed drawing mechanism. Examples include simple random sampling and stratified simple random sampling (e.g. Cochran 1997; De Gruijter *et al.* 2006; Knotters and Brus 2012; Guo *et al*. 2011). A disadvantage might be that this approach results in aggregated information such as spatial averages rather than maps reflecting spatial patterns. Advantages are reproducibility, objectivity, and the approach not only provides a spatial estimate but also its accuracy.

Spatial aggregation can also be done using a geostatistical approach such as ordinary block-kriging (Isaaks and Srivastava 1989; Goovaerts 1997). In contrast to aggregation methods based on sampling theory, geostatistical methods do not require that locations are selected using probability sampling. Instead, a fairly uniform distribution of the sampling locations over the study area is advised, while also including short-distance comparisons to be able to estimate the spatial correlation structure (i.e. semi-variogram). A minimum number of locations at which the GWC is measured is required. For example, to estimate a semi-variogram at least 100 observations are needed (Webster and Oliver 2007). Another difference between this approach and methods based on random sampling is that the results of geostatistical methods only apply under certain model assumptions, e.g. linear relationships, constant variance of regression residuals, stationarity of the semi-variogram, etc. (Hengl *et al*. 2004).

Compared to the other three methods, physical-mechanistic models have the advantage that all kinds of (often non-linear) physical relationships can be included (e.g. Bierkens *et al*. 2000; Knotters *et al*. 2010; Burrough and McDonnel 1998; Webster and Oliver 2007; Stein 1991). Because these models are based on physical-mechanistic principles it is possible to extrapolate the results to other situations (in space and time) and to simulate proposed measures and/or scenarios. A disadvantage is that often some of the input data and model parameters are not known and are thus based on assumptions or used as calibration parameters. Quantification of the model errors is often difficult and therefore uncertainty analyses and validation studies are rarely done.

Similar to the methods used for temporal aggregation, the spatial aggregation and interpolation methods can be combined. For example, a random sample with good spatial coverage can be used for spatial aggregation based on sampling theory as well as for geostatistical interpolation.

6. Accuracy of the final groundwater characteristic.

To obtain groundwater characteristics for a specific area or region, a three-step process is used: (1) in-situ monitoring of the water table depth; (2) temporal upscaling, and (3) spatial interpolation and aggregation. As noted in the previous sections, for all three steps various methods or methodologies are used: (i) four methods to measure/monitor the phreatic water table ; (ii) four methods for temporal aggregation; and (iii) four methods for spatial interpolation and/or aggregation (Figure 1). Which combination of these method options is best for a particular study depends on the objectives of the study. We have analyzed the errors that can occur in each step to assess the accuracy of the final GWC.

++ Figure 1 Combinations of temporal and spatial interpolation and aggregation methods that can be used to derive groundwater characteristics for a specific area based on in-situ measurements.

6.1 Errors in in-situ monitoring

Errors made with the in-situ measuring are not only measurement errors but also interpretation errors. Often it is not clear what has been measured: Was it the phreatic groundwater level under steady or unsteady conditions? Or was it the depth to a phreatic water table? Or was it a perched water table? Or was it the piezometric head at a particular depth?

For steady-state conditions the error can be estimated for each measuring point. The hydrological conditions however can influence the measurements: for an observation well, in a relatively homogeneous soil layer, subject to infiltration, the measured groundwater level in a filter below the water table will be lower than the phreatic groundwater level, but in a seepage area it will be the other way around. In a layered soil, especially with low-permeability layers, the interpretation is much more complex (Saines 1981).

Errors are also introduced when, in course of time, the well has been replaced, moved, extended or when the length and/or the position of the screen has been changed. For manual recording, the time of recording and processing can introduce an error (U.S. Army Corps of Engineers 1993). It is not known, at the time of recording (manually or with a pressure transducer), whether the water level in the well is in equilibrium with the water level in the soil. The deviation can be caused by changes in the groundwater flux, changes in air pressures or in temperature, etc. (e.g. Brassington 1992; Chapuis 2005 & 2009; Elci *et al*. 2003; Paydar and Richardson 2002; Toll and Rasmussen 2007).

The measuring device itself can also introduce an error: (i) the calibrated references level can slowly change over time; (ii) the difference between day and night temperature can result is deviations of several centimeters; and (iii) sensors can malfunction or breakdown. Von Asmuth (2012), reported that in several monitoring networks up to 50% of the sensors had to be replaced.

The above mentioned errors can range from a few centimeters to tens of centimeters (Rasmussen and Crawford, 1997; Spane 2002). Little is known about the accumulation of these errors, thus it is hard to give a rule of thumb for the overall error made by in-situ measurements. In extreme cases, the total error (difference between the measured level and the actual phreatic groundwater level) can be as high as 1 meter (Van den Akker *et al*. 2010).

6.2 Errors in temporal aggregation

The next set of errors is introduced with the temporal aggregation of the in-situ measurements. The temporal variation of the water table depth at the location of an in-situ well is often characterized by a temporally aggregated groundwater characteristic (GWC), such as MHW, MLW and MSW. It is essential, of course, that a unique definition of such a GWC is used. Often, however, different definitions are used alongside each other.

Other errors in temporal aggregation result from the data that is used or missing. Sometimes, a GWC is estimated from an observed time series that is too short, i.e. that does not completely reflect the dynamics of the water table depth under the prevailing hydrological and climatic conditions. It can also happen that a time series contains data from multiple wells, because a well was moved or extended. Sometimes missing data are replaced by values predicted with a time series model or a physical-mechanistic model. It is often also unknown whether the hydrological conditions, for example a perched groundwater table, are considered when estimating the GWCs.

The ultimate effect of these errors is difficult to assess. On the one hand temporal aggregation can level out random errors made with the in-situ measurement, but on the other hand new errors related to the model assumptions made in the aggregation can be introduced. We conclude that the error in estimating the temporal groundwater characteristic that can occur with the different methods is in the range of (i) 10 to 20 cm for direct calculations of time-series (Heuvelink 1998; Knotters 2001); (ii) 10 to 20 cm for time-series models (e.g. Gupta *et al*. 2006; Knotters 2001; Knotters and Van Walsum 1994; McLeod and Hipel 1978; Refsgaard *et al*. 2006; Vrugt *et al*. 2008); (iii) 0.5 to 21 cm for physical-mechanistic models (e.g. Bierkens 1998; Knotters 2001; Knotters and De Gooijer 1999); and (iv) between 20 and 25 cm for direct field observations by experts (Finke *et al.* 2004). In general, the errors are larger at locations where only a short time-series are available.

6.3 Errors in spatial aggregation

For spatial aggregation studies, time-series from existing monitoring wells are usually used, if appropriate, in combination with newly collected data from open boreholes and expert knowledge obtained during field surveys. In general, these studies are not based on the optimum monitoring network configuration and monitoring frequency, but on existing monitoring networks. Geostatistical methods like kriging are widely used to assess this existing data to minimize redundant data and the cost of monitoring (Kim, et al. 2007; Narany *et al*. 2014; Stasch *et al*. 2014; Yang *et al*. 2008).

These spatial interpolation or aggregation methods introduce a next set of errors related to the structure, parameters and input data in the models and the demarcation of the area (Beven and Freer 2001). The resulting errors can be in the range 20 to 50 cm depending on the area and the method used (e.g. Brus 2000; Brus and Te Riele 2001; De Gruijter *et al*. 2006; Finke *et al*. 2004). On the other hand, spatial aggregation can also reduce the overall error because the random part of the error is largely averaged out. It is clear that each spatial aggregation method has its advantages and disadvantages, and that some give more accurate results than others, but it is still difficult to assess the overall effect for each combination of methods.

One reason is that in the studies we analyzed, identical methodologies have been used for different data sets and for different areas, thus an objective comparison is not possible. Another important aspect is that quantifying the validation accuracy assumes that sufficient independent observations, with negligible error, of the target variable (e.g. MxW) are available. In practice, validation observation sets are certainly not flawless and also often obtained through preferential sampling. Furthermore, the locations at which the validation data have been collected often do not correspond with the locations for which the groundwater characteristics were required, thus part of the differences are in fact caused by spatial variability.

Our review indicates that the accuracy of the final GWC of a particular area can be compromised due to various errors that can occur at each step of the process. Often these errors are due to lack of clarity or consistency regarding definitions, data sources, etc. Those working in this area would be wise to not just take information at face value, but to dig a little deeper to know as accurately as possible just what has been measured and how it has been upscale and aggregated.

7. Concluding remarks

Over the past sixty years the methodologies to measure and interpret water table depths to obtain groundwater characteristics for a specific area have changed. To obtain a reliable prediction of the groundwater characteristics in a specific area, a three-step approach is used: (1) in-situ monitoring of the water table depth; (2) temporal upscaling; and (3) spatial interpolation and aggregation. For all three steps, various methods or methodologies are used: (i) four methods to measure/monitor the phreatic water table are used; (ii) four methods for temporal aggregation; and (iii) four methods for spatial interpolation and/or aggregation. Various combinations of these different methodologies are used, usually based on the objective(s) of the particular study and the available data. The use of these different combinations of methods and techniques has certainly led to better understanding of the groundwater characteristics in the areas that were studied. However, the impact of errors in the observations or inaccuracies in the temporal and/or spatial interpolation and aggregation on the accuracy of the final result has never been systematically investigated.

We conducted an analysis of the various methods and techniques to try to determine the impact of the errors that can occur at each step. Our review shows that changes in the methods to measure and interpret water table depths has resulted in significant systematic differences in the groundwater characteristics. Although there are many sources of potential error in the in-situ measurement of the water table depth, the impact is hard to give. However in extreme cases it can be as high as 1 meter. Errors in the temporal aggregation are in the range of 10 to 20 cm and for the spatial interpolation between 20 to 50 cm. We could not, however, assess the accuracy of the final groundwater characteristic partly because of the remediating interactions between the errors and consequent steps in the process.

So we cannot answer the question of whether drought stress in The Netherlands is under- or overestimated. To make such an objective and process-independent assessment of the water table depth in a specific area it is a prerequisite to collect independent validation data at locations in the area that are randomly selected. Only in this way can a statement like "*the average MHW in this area has changed xx cm caused by this........ intervention*" be tested. Unfortunately, the existing monitoring wells in an area rarely satisfy this ‘randomly selected’ criterion. Thus at present it is not possible to accurately assess the overall error in the estimation of the aggregate groundwater characteristics.

What we have learned is that, based on the research question(s), an analysis should be made to select the most appropriate measuring locations (data sets) that can be used in the study and the most appropriate methods to aggregate the data sets both in space and time. It is also essential to make an analysis of the accuracy of the existing data sets in combination with the selected aggregation methods to assess whether the resulting accuracy is sufficient to answer the research question(s). This is an iterative process in which each step has to be repeated a number of times to check that the assumptions that are required to go through a particular step do not hamper the following steps.

Acknowledgments and Data

The data presented in this paper is the result of the research project "*Verdroging*" in which the authors and our colleagues from the Environmental Sciences Group (ESG) of Wageningen University and Research Centre participated. The project was initiated by Bram de Vos, our current Director, and jointly financed by the ESG research institute Alterra and the Knowledge Base Project "*Duurzame ontwikkeling van de groenblauwe ruimte*" of the Dutch ministry of Economic Affairs. This paper could not have been written without the data and support provided by these colleagues.

References

Bartholomeus, R.P., 2009., Moisture matters. Climate-proof and process-based relationships between water, oxygen and vegetation. PhD Thesis, Thesis, Institute of Ecological Science, Vrije Universiteit, Amsterdam, The Netherlands.

Berendrecht, W.L., A.W. Heemink, F.C. van Geer, Gehrels, J.C., 2004. A non-linear state space approach to model groundwater fluctuations. *Advances in Water Resources* 29: 959-973, doi:10.1016/j.advwatres.2005.08.009.

Beven, K., Freer, J., 2001. Equifinality, data assimilation, and uncertainty estimation in mechanistic modelling of complex environmental systems using the GLUE methodology. *Journal of Hydrology* 249: 11-29, doi: 10.1016/S0022-1694(01)00421-8.

Bierkens, M.F.P., Finke, P.A., de Willigen, P., 2000. Upscaling and downscaling methods for environmental research. Kluwer, Dordrecht.

Bierkens, M.F.P., Knotters, M., Hoogland, T., 2001. Space-time modeling of water table depth using a regionalized time series model and the Kalman filter. *Water Resources Research* 37: 1277-1290, doi:10.1029/2000WR900353.

Bierkens, M.F.P., Knotters, M., van Geer, F.C., 1999. Calibration of transfer function-noise models to sparsely or irregularly observed time series. *Water Resources Research* 35(6): 1741-1750, doi: 10.1029/1999WR900083

Bjerg, P.J., Christensen, Th.H., 1992. Spatial and temporal small-scale variation in groundwater quality of a shallow sandy aquifer. *Journal of Hydrology*, 131: 133-149, doi: 10.1016/0022-1694(92)90215-H.

Brassington, F. C., 1992. Measurements of head variations within observation boreholes and their implications for groundwater monitoring. *Journal of the Institution of Water and Environmental Management* 6(1): 91-100, doi:10.1111/j.1747-6593.1992.tb00742.x.

Brus, D.J., 2000. Using regression models in design-based estimation of spatial means of soil properties. *European Journal of Soil Science* 51, 159-172, doi:10.1002/env.3170040202.

Brus, D.J., de Gruijter, J.J., 1993. Design-based versus model-based estimates of spatial means. Theory and application in environmental soil science. *Environmetrics* 4: 123-152, doi:10.1002/env.3170040202.

Brus, D.J., te Riele, W.J.M., 2001. Design-based regression estimators for spatial means of soil properties: the use of two-phase sampling when the means of the auxiliary variables are unknown. *Geoderma* 104: 257-279, doi:10.1016/S0016-7061(01)00084-2.

Burrough, P.A., McDonnell R.A., 1998. Principles of Geographical Information Systems. Oxford University Press, Oxford.

Centraal Bureau voor de Statisitek (CBS), 2014. StatLine, electronic databank of Statistics Netherlands, http://statline.cbs.nl assessed: 2-07-2014.

Changnon, S.A., Huff, F.A., Chin-Fei Hsu, 1988. Relations between precipitation and shallow groundwater in Illinois. *Journal of Climate*, 1(12):1239–1250, doi: 10.1029/JZ069i004p00605.

Chapuis, R.P., 2005. Numerical modeling of rising-head permeability test in monitoring wells after lowering the water level down to the screen. *Canadian Geotechnical Journal* 42: 705-715, doi:10.1139/T05-003.

Chapuis, R.P., 2009. Numerical modeling of reservoirs or pipes in groundwater seepage. *Computers and Geotechnics* 36: 895-901, doi:10.1016/j.compgeo.2009.01.005.

Church, P.E., Granato, G.E., 1996. Bias in groundwater data caused by well-bore flow in long-screen wells. *Ground Water* 34(2): 262-273, doi: 10.1111/j.1745-6584.1996.tb01886.x

Cirkel, D.G., Witte, J.P.M, van der Zee, S.E.A.T.M., 2010. Estimating seepage intensities from groundwater level time series by inverse modelling: A sensitivity analysis on wet meadow scenarios. *Journal of Hydrology* 385: 132–142 133, doi: 10.1016/j.jhydrol.2010.02.009.

Cochran, W.G.. 1977. Sampling techniques. Wiley, New York.

Colenbrander, H.J. (Ed), 1989. Water in The Netherlands. Proceedings and Information/TNO Committee on Hydrological Research, The Hague, no. 37 – 96 pp.

Dalton, M.G., Huntsman, B.E., Bradbury, K., 2007, Acquisition and interpretation of water-level data. *In*: D.M. Nielsen, D.A. Nielsen, and G.L. Nielsen (eds.), The Essential Handbook of Ground-Water Sampling, chapter 7, pp. 173-201, CRC Press/Taylor & Francis.

De Gruijter, J.J., Brus, D.J., Bierkens, M.F.P., Knotters, M., 2006. Sampling for natural resource monitoring. Springer, Berlin.

De Lange, W.J., Prinsen, G.F., Hoogewoud, J.C., Veldhuizen, A.A., Verkaik, J., Oude Essink, G.H.P., van Walsum, P.E.V., Delsman, J.R., Hunink, J.C., Massop H.Th.L., Kroon, T., 2014). An operational, multi-scale, multi-model system for consensus-based, integrated water management and policy analysis: The Netherlands Hydrological Instrument. *Environmental Modelling & Software* 59: 98-108, doi: 10.1016/j.envsoft.2014.05.009.

De Ridder, N. A., 2006. Groundwater investigations. In: H. P. Ritzema (ed.), Drainage Principles and Applications, Alterra, Wageningen, pp. 33-75.

De Vos, J.A., van Bakel, P.J.T., Hoving, I.E., Smidt, R.A., 2010. Raising surface water levels in peat areas with dairy farming Upscaling hydrological, agronomical and economic effects from farm-scale to local scale. *Agricultural Water Management*, 97 (11): 1887–1897, doi: 10.1016/j.agwat.2010.06.017.

Elci, A., Flach, G.P., Molz, F.J., 2003. Detrimental effects of natural vertical head gradients on chemical and water level measurements in observation wells: identification and control. *Journal of Hydrology* 281: 70-81, doi:10.1016/S0022-1694(03)00201-4.

Fetter, C.W., 1980. Applied hydrogeology. Charles E. Merrill Publishing Co., Columbus, Ohio.

Finke, P.A., Brus, D.J., Bierkens, M.F.P., Hoogland, T., Knotters, M., de Vries, F., 2004. Mapping ground-water dynamics using multiple resources of exhaustive high resolution data. *Geoderma* 123: 23-39, doi:10.1016/j.geoderma.2004.01.025.

Freeze, R.A., Cherry, J.A., 1979. Groundwater. Prentice-Hall, Englewood Cliffs.

Goovaerts, P., 1997. Geostatistics for natural resources evaluation. New York, Oxford University Press.

Guoa, Y., Wanga, J.F., Yin, X.L., 2011. Optimizing the groundwater monitoring network using MSN theory. *Procedia Social and Behavioral Sciences* 21: 240–242, doi: 10.1016/j.sbspro.2011.07.033.

Gupta, H.V., Beven, K.J., Wagener, T., 2006. Model Calibration and Uncertainty Estimation, chapter 131 in: Encyclopedia of Hydrological Sciences, John Wiley & Sons.

Hengl, T., Heuvelink G.B.M, Stein, A., 2004. A generic framework for spatial prediction of soil variables based on regression-kriging. *Geoderma* 120: 75-93, doi:10.1016/j.geoderma.2003.08.018.

Heuvelink, G.B.M., 1998. Error Propagation in Environmental Modelling with GIS. London: Taylor & Francis, 127 pp.

Isaaks, E.H., Srivastava, R.H., 1989. Applied geostatistics. New York, Oxford University Press.

Kim, G.B., Lee, K.K., Lee J.Y., Yi, M.J., 2007. Case study for determination of a water level monitoring frequency for nationwide groundwater monitoring networks in Korea. *Journal of Hydrology* 342: 223– 237, doi: 10.1016/j.jhydrol.2007.05.034.

Knotters, M., 2001. Regionalised time series models for water table depths. PhD-thesis, Wageningen University, Wageningen.

Knotters, M., Brus, D.J., 2013. Purposive versus random sampling for map validation: a case study on ecotope maps of floodplains in The Netherlands. *Ecohydrology* 6, 425–434, doi: 10.1002/eco.1289.

Knotters, M., de Gooijer, J.G., 1999. TARSO modeling of water table depths. *Water Resources Research* 35: 695-705, doi: 10.1029/1998WR900049.

Knotters, M., van Walsum, P.E.V., 1997. Estimating fluctuation quantities from time series of water-table depths using models with a stochastic component. *Journal of Hydrology* 197: 25-46, doi:10.1007/s00477-002-0090-8.

Knotters, M., Bierkens, M.F.P., 2000. Physical basis of time series models for water table depths. *Water Resources Research* 36: 181-188, doi:10.1029/1999WR900288.

Knotters, M., Bierkens, M.F.P., 2001. Predicting water table depths in space and time using a regionalised time series model. *Geoderma* 103: 51-77, doi:10.1016/S0016-7061(01)00069-6.

Knotters, M., Bierkens, M.F.P., 2002. Accuracy of spatio-temporal RARX model predictions of water table depths. *Stochastic Environmental Research and Risk Assessment* 16: 112-126, doi:10.1007/s00477-002-0090-8.

Knotters, M., Heuvelink, G.B.M., Hoogland, T., Walvoort D.J.J., 2010. A disposition of interpolation techniques. Wageningen, Wettelijke Onderzoekstaken Natuur en Milieu (WOt), werkdocument 190.

Koninklijk Nederlands Meteorologisch Institute (KNMI), 2014. Climatology, daily data from the weather in The Netherlands. Available from: http://www.knmi.nl/kd/daggegevens/selectie.cgi

Kremers, A.H.M., van Geer, F.C., 2000. Trendontwikkeling Grondwater 2000. Analyseperiode 1955-2000 (in Dutch). TNO-rapport: NITG 00-184-B, TNO, Delft.

Lijzen, J.P.A., Otte, P., van Dreumel, M., 2014. Towards sustainable management of groundwater: Policy developments in The Netherlands. *Science of the Total Environment* 485–486: 804–809, doi: 10.1016/j.scitotenv.2014.02.081.

Lin, H. 2012. "Hydropedology: Summary and Outlook." In: Lin, H. (Ed) Hydropedology, Chapter 24: 759-781. Elsevier.

Lin, H., J. Bouma, L. P. Wilding, J. L. Richardson, M. Kutílek, and D. R. Nielsen. 2005. Advances in Hydropedology. Advances in Agronomy 85: 1-89, doi: 10.1016/S0065-2113(04)85001-6.

Martens, K., an Campa, M., Van Damme, D., Walraevens, K., 2013. Groundwater dynamics converted to a groundwater classification as a tool for nature development programs in the dunes. *Journal of Hydrology* 499: 236–246, doi: 10.1016/j.jhydrol.2013.06.045.

McLeod, A.I, Hipel, K.M., 1978. Simulation procedures for Box-Jenkins models. *Water Resources Research* 14(3): 969-975, doi: 10.1029/WR014i005p00969.

Ministry of Transport, Public Works and Water Management, 1999. Fourth National Policy Document on Water Management, Government proposals (Dutch with English Summary). 's-Gravenhage, 1999.

Narany, T.S., Ramli, M.F., Aris, A.Z., Sulaiman, W.N.A., Fakharian, K., 2014. Spatial Assessment of Groundwater Quality Monitoring Wells Using Indicator Kriging and Risk Mapping, Amol-Babol Plain, Iran. *Water* 2014, 6, 68-85, doi: 10.3390/w6010068.

Nielsen, D.M., Nielsen, G.L., 2007. The essential handbook of ground-water sampling. CRC-press, Boca Raton.

Overeem, I., Weltje, G.J., Bishop-Kay, C., Kroonenberg, S.B., 2001. The Late Cenozoic Eridanos delta system in the Southern North Sea Basin: a climate signal in sediment supply? *Basin Research*, 13: 293–312, doi: 10.1046/j.1365-2117.2001.00151.x.

Paydar, Z., Richardson, D.P., 2002. Water table depth and piezometers. Soil physical measurement and interpretation for land evaluation, Australian Soil and Land Survey

Handbook Series, K. C. N. McKezie and H. Cersswell, ed., CSIRO Publishing, Collingwood, Australia, 177-189.

Prak, H., 2002. Waternood: Working on Integrated Water Management in Rural Areas. ICID, Proc. of the 18th Congress. Paper Q51 – R2.08. Montreal, Canada. 16 pp.

Rasmussen, T.C., Crawford, L.A., 1997. Identifying and removing barometric pressure effects in confined and unconfined aquifers. *Ground Water* 35(3): 502-511, doi:10.1111/j.1745-6584.1997.tb00111.x.

Refsgaard J.C., van der Sluijs, J.P., Brown, J., van der Keur, P., 2006. A framework for dealing with uncertainty due to model structure error. *Advances in Water Resources* 29: 1586-1597, doi: 10.1016/j.advwatres.2005.11.013.

Richards, L.A., 1931. Capillary conduction of liquids through porous mediums. *Physics* 1: 318-333.

Ritzema H.P., Heuvelink, G.B.M. Heinen, M., Bogaart, P.W., Hack-ten Broeke, M.J.D, Hoogland, T., Knotters, M., Massop, H.T.L., van der Bolt, F.J.E., Vroon, H.R.J., 2012. *Meten en interpreteren van grondwaterstanden. Analyse van methodieken en nauwkeurigheid* (in Dutch). Alterra-report 2345, Alterra, Wageningen, 122 pp.

Saines, M., 1981. Errors in interpretation of ground-water level data. *Ground Water Monitoring and remediation*, 1(1): 56-61, doi:10.1111/j.1745-6592.1981.tb00798.x .

Spane, F. A. , 2002. Considering barometric pressure in groundwater flow investigations. *Water Resources Research*, 38(6): 14-18, doi:10.1029/2001WR000701.

Stańczuk-Gałwiaczek, M., K. Sobolewska-Mikulska, H. Ritzema, and J. M. van Loon-Steensma. 2018. Integration of water management and land consolidation in rural areas to adapt to climate change: experiences from Poland and the Netherlands. Land Use Policy 77: 498-511. doi: 10.1016/j.landusepol.2018.06.005.

Stasch, C., Scheider, S., Pebesma, E., Kuhn, W., 2014. Meaningful spatial prediction and aggregation. *Environmental Modelling & Software* 51: 149-165, doi: 10.1016/j.envsoft.2013.09.006.

Stein, A., 1991. Spatial interpolation. PhD thesis, Wageningen University, 222 p.

Toll, N.J., Rasmussen, T.C., 2007. Removal of barometric pressure effects and earth tides from observed water levels. *Groundwater* 45(1): 101-105, doi:10.1111/j.1745-6584.2006.00254.x.

U.S. Army Corps of Engineers - Wetlands Research Program, 1993. Installing monitoring wells/piezometeric in wetlands. WRP Technical Note HY-IA-3.1, ed., Arlington.

Van de Ven, G.P. (Ed.), 1996. Man-made Lowlands. History of water management and land reclamation in The Netherlands. Stichting Matrijs, Utrecht, The Netherlands, 293 pp.

Van den Akker, J.J.H., de Groot, W.J.M., Vroon, H.R.J., van der Bolt ,F.J.E., van Kekem, A.J., 2010. Stijghoogteverschillen en verdichting: een eerste Twentse verkenning in de praktijk (in Dutch). Alterra-rapport 1735, Alterra, Wageningen.

Van den Noort, P.C., 1987. Land consolidation in The Netherlands. *Land Use Policy*, Vol.4(1), pp.11-13, doi: 10.1016/0264-8377(87)90004-4.

Van der Sluijs, P., De Gruijter, J.J., 1985. Water table classes: A method to describe seasonal fluctuation and duration of water tables on Dutch soil maps. Agricultural Water Management, Vol.10(2), pp.109-125, doi: 10.1016/0378-3774(85)90001-0.

Van der Sluijs, P., van Heesen, H.C., 1989. Veranderingen in de berekening van de GHG en de GLG (in Dutch). *Landinrichting*, 29 (1): 18-21.

Van Duijvenbooden, W., 1981. Groundwater quality in The Netherlands - collection and interpretation of data. *The Science of the Total Environment*, 21: 221-232, doi: 10.1016/0048-9697(81)90154-6.

Van Heesen, H.C., 1970. Presentation of the seasonal fluctuation of the water table on soil maps. *Geoderma* 4: 257-278, doi: 10.1016/0016-7061(70)90006-6.

Van Tol, G., van Dobben, H.F., Schmidt, P., Klap, J.M., 1998. Biodiversity of Dutch forest ecosystems as affected by receding groundwater levels and atmospheric deposition. *Biodiversity and Conservation* 7, 221-228, doi: 10.1023/A:1008888519478.

Von Asmuth, J., 2012. Groundwater System Identification through Time Series Analysis. PhD-thesis, Delft University of Technology.

Von Asmuth, J.R., Bierkens, M.F.P, Maas, K., 2002. Transfer function-noise modeling in continuous time using predefined impulse response functions. *Water Resources Research* 38: 23-1-23-12, doi:10.1029/2001WR001136.

Von Asmuth, J. R., Knotters, M., 2004. Characterising groundwater dynamics based on a system identification approach, *Journal of Hydrology* 296: 118–134, doi: 10.1016/j.jhydrol.2004.03.015.

Vroon, H.R.J., Dekker, L.W., Hendricks, J.M.H., 1988. A method for measuring hydraulic properties of brittle soil horizons. *Soil Sci. Soc. Am*. J. 52:292-294.

Vrugt, J.A., Ter Braak, C.J.F., Clark, M.P., Hyman, J.M., Robinson, B.A., 2008. Treatment of input uncertainty in hydrologic modeling: doing hydrology backward with Markov chain Monte Carlo simulation. *Water Resources Research* 44: W00B09, doi: 10.1029/2007WR006720.

Webster, R., Heuvelink, G.B.M., 2006). The Kalman filter for the pedologist's tool kit. European *Journal of Soil Science* 57: 758-773, doi:10.1111/j.1365-2389.2006.00879.x.

Webster, R., Oliver, M.A., 2007. Geostatistics for Environmental Scientists. Second Edition. Wiley, Chicester.

Yang, F.G., Cao, S.Y., Liu, X.N., Yang , K.J., 2008. Design of groundwater level monitoring network with ordinary kriging. *Journal of Hydrodynamics* 20(3):339-346, doi: 10.1016/S1001-6058(08)60066-9.

Yi, M.J., Lee, K.K., 2004. Transfer function-noise modelling of irregularly observed groundwater heads using precipitation data. *Journal of Hydrology* 288: 272–287, doi: 10.1016/j.jhydrol.2003.10.020.

Young, P.C., Beven, K.J., 1994. Data-based mechanistic modelling and the rainfall-flow non-linearity. *Environmetrics* 5: 335-363, doi: 10.1002/env.3170050311.

Table 1. Hydrological characterizations used in The Netherlands to define the water table depth over time.

Hydrological parameter	Description
HG3	**Highest Annual Watertable**: Average of the depths to the three highest water tables in a hydrological year (1 April / 31 March) at a measuring frequency of twice a month (the 14th and 28th).
MHW	**Mean Highest Watertable**: average of the HW3 over a period of 30 years under the given climatic and hydrological conditions.
MLW	**Mean Lowest Watertable**: Average LW3 over a period of 30 years under the given climatic and hydrological conditions.
MSW	**Mean Spring Water table**: Average SW3 over a period of 30 years under the given climatic and hydrological conditions.
LG3	**Lowest Annual Watertable**: Average of the depths to the three lowest water tables in a hydrological year (1 April / 31 March) at a measuring frequency of twice a month (the 14th and 28th).
SG3	**Spring Watertable:** Average water table depths on March 14, March 28 and April 14 in a given calendar year.

Table 2. Assessment of the four methods used to measure the water table depth (++ stands for "complies with criterion" and – for "does not comply at all")

	Assessment of the measuring method			
Criterion	Groundwater observation well	Piezometer	Open bore hole	Expert knowledge
Level of detail	+	++	+	-
Can be used to calculate scenarios	n.a.[a]	n.a.	n.a.	n.a.
Frequency of measurements	+	+	+/-	-
Accuracy	+/-	++	+/- to +[b]	-
Need for additional information	+/-	+/-	+	+
Can be used to calculate fluxes	-	++	-	-
Can be used to extrapolate	n.a.	n.a.	n.a.	n.a.
Can be reproduced	++	++	+/-	-
Objectiveness	+	+	+	-
User friendliness	+	+	+	+
Quantification of (un)accuracy	+/-	+	+/-	-
Limited number of measurements needed	n.a.	n.a.	n.a.	n.a.
Interpolation needed to obtain results	n.a.	n.a.	n.a.	n.a.

[a] n.a. = not available; [b] With a piezometers or open bore hole, you know better what you measure than with a groundwater observation well, provided that the open bore hole is drilled up to a poorly permeable layer on which a rise in the piezometeric pressure or perched groundwater table occurs or that several open bore holes are drilled up to different depths to determine the occurrence of a perched watertable.

Table 3. Assessment of the four temporal aggregation methods used to characterize the temporal groundwater level (++ stands for "complies with criterion" and – for "does not comply at all")

	Assessment of the temporal aggregation method			
Criterion	Time-series	Time-series models	Physical-mechanistic models	Expert knowledge
Level of detail	n.a.	n.a.	n.a.	n.a.
Can be used to calculate scenarios	-	-	+	+/-
Frequency of measurements	n.a.	n.a.	n.a.	n.a.
Accuracy	++	+/-	+/-	- /+*
Need for additional information	--	+/-	+	+
Can be used to calculate fluxes	n.a.	n.a.	n.a.	n.a.
Can be used to extrapolate	n.a.	n.a.	n.a.	n.a.
Can be reproduced	++	++	++	-/+
Objectiveness	++	+	+	-
User friendliness	++	+/-	-	+
Quantification of (un)accuracy	+	++	+	+/-
Limited number of measurements needed	-	+/–	-	+
Interpolation needed to obtain results	n.a.	n.a.	n.a.	n.a.

* depending of knowledge of the location/area

Table 4. Assessment of the four spatial interpolation and aggregation methods used to characterize the spatial groundwater level (++ stands for "complies with criterion" and - for "does not comply at all")

	Assessment of the spatial aggregation method			
Criterion	Expert knowledge	Random sampling	Statistical models	Process models
Level of detail	+/-	-	+	+
Can be used to calculate scenarios	-	-	-	+
Frequency of measurements	+	+/-	+/-	-
Accuracy	-/+ *	+/-	+/-	+/-
Need for additional information	+	+	+	++
Can be used to calculate fluxes	+/-	-	-	++
Can be used to extrapolate	+	-	-	++
Can be reproduced	+/-	++	++	+
Objectiveness	-	++	-	–
User friendliness	+	+	-	-
Quantification of uncertainty	-	++	+	+/-
Limited number of measurements needed	+	-	-	-
Interpolation needed to obtain results	+	-	++	++

* depending of knowledge of the location/area

Figure 1 Combinations of temporal and spatial interpolation and aggregation methods that can be used to derive groundwater characteristics for a specific area based on in-situ measurements.