Sruthi Sathyadevan

Performance Assessment of Global Optimization Algorithms in Automatic Calibration of Groundwater Models

TU Delft

Performance Assessment of Global Optimization Algorithms in Automatic Calibration of Groundwater Models

By

Sruthi Sathyadevan

in partial fulfilment of the requirements for the degree of

Master of Science
in Civil Engineering

at the Delft University of Technology,
to be defended publicly on Friday October 12, 2018 at 04:00 PM.

Supervisor:	Prof. dr. ir. Ronald van Nooyen,	TU Delft
Thesis committee:	Prof. dr. ir. Nick van de Giesen,	TU Delft
	Dr. ir. Willem-Jan Zaadnoordijk,	TU Delft/TNO
	Ir. Bernard J. Meulenbreuk,	TU Delft

Preface

This report represents the culmination of my master thesis research that was undertaken as the final stage of my study Water Management at Delft University of Technology. I acquired a great deal of knowledge in groundwater flow modelling and on ways to efficiently do a scientific research. This could not have been made possible without the help and support of many.

First, I would like to thank my university supervisor Ronald van Nooyen for his continued involvement at all stages of my thesis and ensuring his support when needed. A special thanks to my company supervisor Koen van der Hauw at Sweco Nederland for his immense patience and abundant knowledge I feel honoured to have been able to learn from. Further I thank my graduation committee for their feedback and guidance during my thesis. Heartfelt thanks go to Willem-Jan Zaadnoordijk, Bernard J. Meulenbreuk and Nick van de Giesen for steering my research in the right direction at all times. Without their support and advice, this report could not have been completed. Further, I want to thank Heleen Westerhof-Graafstal at the Waterschap De Dommel for her continued support during my research by passing on her knowledge. Last but not the least, I would like to thank my family and friends for their presence and support in my life without whom this thesis could not have been completed.

Sruthi Sathyadevan
Delft, October 2018

Contents

Abstract

Calibration of groundwater models can be done manually by comparing the measured and computed groundwater heads or automatically using algorithms which do the work for you. Modules for this purpose are included in or linked externally to traditional groundwater modelling software and used. One such module PEST (Parameter ESTimation) uses the Levenburg Marquardt algorithm which is a combination of Gauss Newton and Gradient Descent methods to find the minima of a function. However, PEST is capable of finding only a local minimum for the objective function. A possibility exists that there is a global minimum that better fits our function and could give even better results for the optimization problem in, in our case, parameter estimation in groundwater models. A performance assessment has been done with a Genetic algorithm on a groundwater modelling problem in this study.

List of Figures

List of Tables

1. Introduction

1.1. Groundwater modelling

Groundwater modeling is a way to represent a groundwater system in another, simpler form to investigate its responses under different conditions. This allows for the prediction of future behavior of such a system as well. Groundwater flow models are typically used to calculate the rate and direction of movement of groundwater through aquifers and confining geological units in the subsurface by means of calculations known as simulations. The outputs of the model simulations are the hydraulic heads and flow rates which are in equilibrium with the hydrogeological conditions (aquifer boundaries, initial and transient conditions and sources or sinks) defined for the modelled area (Muyinda .N, et. Al. 2014). Groundwater models are essentially simplified versions of reality and as such are therefore imperfect. Be that as it may, models are very useful in hydrogeology.

Groundwater models can be classified into three broad categories: physical, analogue and mathematical. Physical models are scaled-down physical replicas of simple groundwater flow systems made in the laboratory. Analogue models are those models that illustrate physical processes that are governed by equations that are similar to equations of groundwater flow. Some such examples are heat transfer and electrodynamics. A mathematical groundwater model can be defined in several ways, one [Anderson & Woessner, 1992] of which is:

> "A mathematical model simulates groundwater flow indirectly by means of a governing equation thought to represent the physical processes that occur in the system, together with equations that describe heads or flows along the boundaries of the model."

In this study we consider a mathematical model and as such the focus will be only on that in the rest of the report.

Solution of a mathematical groundwater model can be either analytical or numerical. Analytical models can only be applied to simple problems which does not necessitate modifying the structure of the solution. In reality, the system is much more complicated than what analytical models are capable of representing. In such a case, numerical solutions come in handy and are widely used. Due to their ability to handle complex representations of the modelled system, together with rapid technological advancement in computer processors, numerical modelling has become the go-to method for groundwater modelling.

Numerical solutions to groundwater model problems used to be assessed on the basis of a pre-determined degree of accuracy to an analytical solution (Craig et. Al. 2010). Increase in computational power over the years has seen a blurring of the distinction between what used to be analytical and numerical approaches. This has led to advances and further easing of the boundary and understanding between the various types of solutions. Thereby exists semi-analytical methods which resort to mathematical tricks of analytical solution derivations such as superposition, integral transforms, etc. to help support numerical approximation of a solution.

With further scientific development, finite difference and finite element approaches with their flexibility have made it possible to model complex groundwater systems more than what was capable of using analytical approaches. Finite volume methods are also used similarly by incorporating finite-sized sub domains called control volumes. Hydraulic characteristics of an aquifer can also be estimated using semi-analytical pumping tests in wells by comparing the drawdowns at known distances to be used in formulation of a well-flow equation of a theoretical groundwater model.

Groundwater models can be a simple 1D analytical model or a complicated 3D one. It is always recommended to start with a simple model and build up the complexity in stages (Hill, 2006), as was done in this study as well. The choice of type of solution used in groundwater modelling is heavily dependent on the modeler's choice and the available tools at their disposal.

Mathematical groundwater models can be further classified based on various characteristics, such as design/schematization of the problem, hydrological processes considered and the intended application of the model.

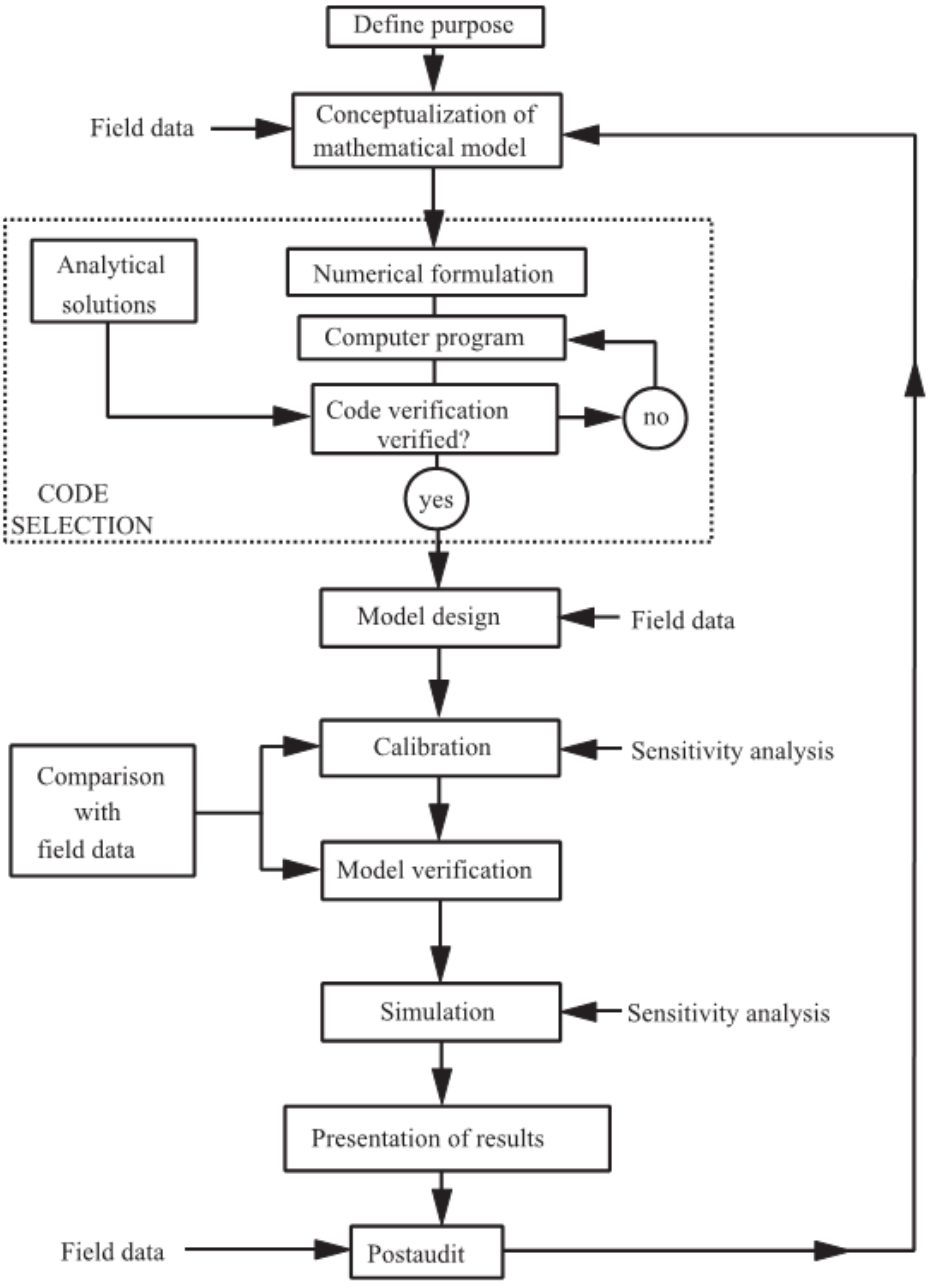


Fig. 1.1. Modelling methodology (G. H. P. Oude Essink, 2000)

It is essential to identify the objectives of creating a model before diving in. It needs to be decided whether a model is necessary to solve the hydrological problem at hand and if it is indeed, then the next step is to decide on the type of model based on the problem at hand.

Depending on the objectives decided and agreed upon, a selection of model approach is made. This also influences the data requirement of the model for the chosen area of study. If the area of study is small, then a fine-gridded model with high intensity data should be used, whereas for a model objective covering a larger area, a coarse grid could be satisfactory.

The identification of the purpose of the modelling effort should lead the modeler to deciding the type of model to be made: analytical or numerical, transient or steady-state, etc. A selection of a model approach is what leads to the next step in modelling methodology: conceptualization of the mathematical model (Fig. 1.1).

1.1.1. Conceptualisation of a mathematical model

After identification of objectives, the next step in numerical groundwater modelling is to define a conceptualize the mathematical model in the making. A schematization of the study area is done by defining the layer structure, steady state or transient flow, boundaries of the study area such as location and type of boundary conditions, etc. Based on the schematization prepared, a concept of the mathematical model is built. This step is essential to simplify the schematization suitable for numerical modelling. Building a concept defines system characteristics, processes and interactions among various hydrologic components in the groundwater system. Developing a good concept is crucial to attaining accurate predictions with the numerical model.

Identifying the boundary conditions for a study area is an important step in creating a concept for the mathematical groundwater model. Solving mathematical groundwater models which is a system of partial differential equations requires identification of boundary conditions to give a unique solution. If incorrect boundary conditions are provided then the behavior of the resulting model will not match the behavior of the physical system. The groundwater model would then represent a different groundwater system that is influenced by a different set of boundary conditions. Boundary conditions can be classified into three main types (G.H.P Oude Essink, 2000) :

- **Specified head (also called Dirichlet boundary).** It can be expressed in a mathematical form as: h (x,y,z,t)=constant. It describes head boundaries for which a head is given. E.g. Rivers which are in hydraulic connection with the aquifer being modelled. Lines ABC and EFG in Fig. 1.2. are examples of this type of boundaries, where a portion of the aquifer occurs underneath a reservoir.

- **Specified flux (also called a Neumann boundary).** In a mathematical form it is: ∇h (x,y,z,t)=constant. This describes boundaries for which a flow (the derivative of head) is given across the boundary. E.g. Recharge across the water table in a phreatic aquifer. Line AD & HI in Fig. 1.2. is an example of this type. No-flow boundaries which represent impermeable media is a special case of specified flux boundary.

- **Head-dependent flux (also called a Cauchy boundary).** Its mathematical form is: ∇h (x,y,z,t)+a*h=constant (where “a” is a constant). This describes boundaries for which flux across a boundary is calculated, given a value of the boundary head. E.g. A semi-confined aquifer, where the water head depends on the flux through the semi-confining layer. ABC and EFG in Fig. 1.2 represent this type of boundary.

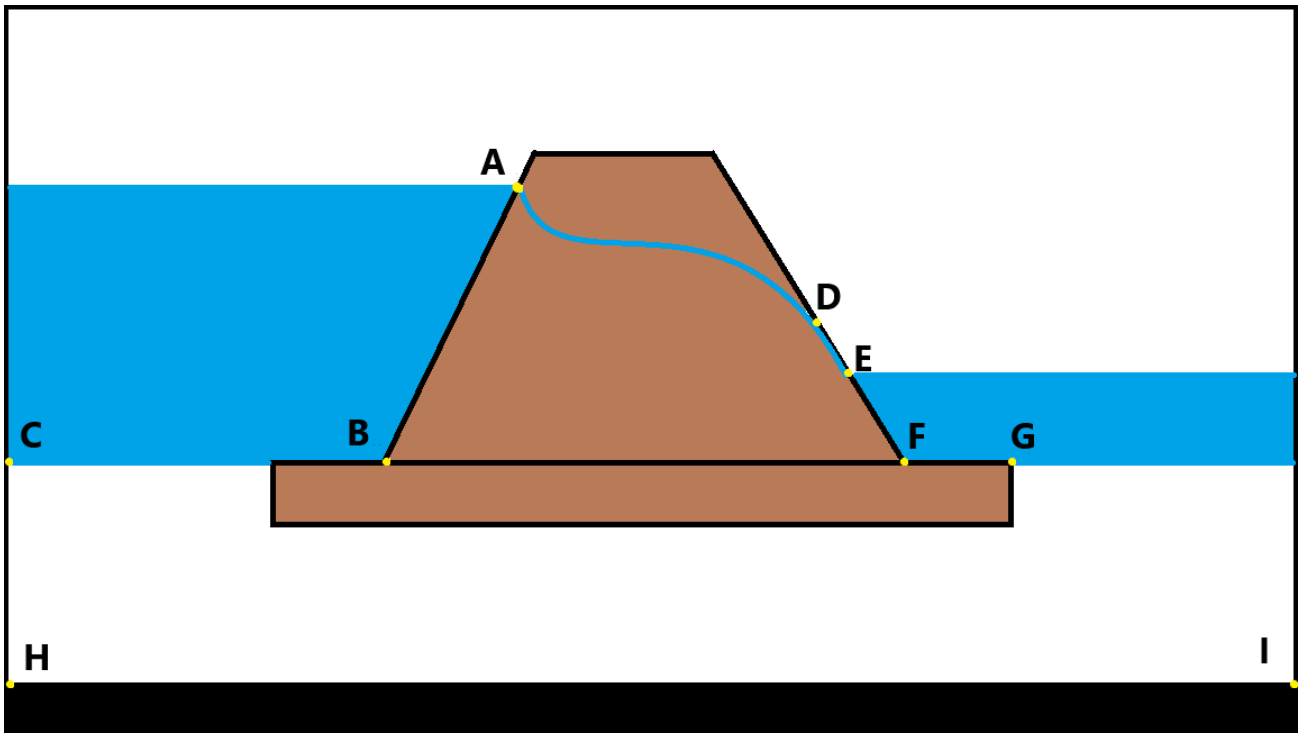


Fig. 1.2. Different types of boundaries. Line ABC & EFG represent Dirichlet & Cauchy Boundaries, Lines HI & AD, Neumann Boundaries

A further need to get accurate boundary conditions lies in the fact that they also represent sources and sinks (Reilly and Harbaugh 2004) using Neumann boundaries described earlier.

1.1.2. Mathematical description of hydrogeologic processes

Groundwater flow can be described by partial differential equations. The basic underlying principles of groundwater models are the equation of motion (Darcy's law) and the equation of continuity (mass balance equation). In this study we consider a steady state (time independent) groundwater flow system and as such, no changes in piezometric head as a function of time will be considered. Variations in fluid density due to changes in solute concentrations, pressure or temperature are also assumed to be negligible and thereby density of groundwater is considered constant in this study. According to Darcy's law, the flux is a function of the head loss, the distance in a certain direction x and the hydraulic conductivity (Watson & Burnett, 1995):

$$q_x = -K_{xx}\frac{\partial h}{\partial x}$$

Where k is assumed to be constant in space.

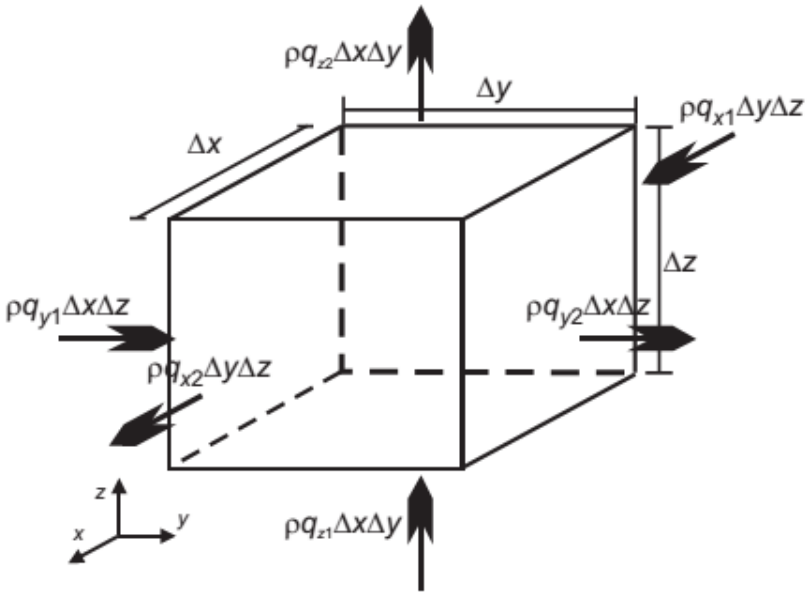


Fig. 1.3. Determination of steady state groundwater flow equation (G. H. P. Oude Essink, 2000)

Conservation of mass per unit of time (Fig. 1.3.) leads to:

$$\left(\frac{\rho q_{x2} - \rho q_{x1}}{\Delta x}\right)\Delta x\Delta y\Delta z + \left(\frac{\rho q_{y2} - \rho q_{y1}}{\Delta y}\right)\Delta x\Delta y\Delta z + \left(\frac{\rho q_{z2} - \rho q_{z1}}{\Delta z}\right)\Delta x\Delta y\Delta z = S_s$$

If the solution is time independent then $S_s = 0$ and therefore:

$$\frac{\partial \rho q_x}{\partial x} + \frac{\partial \rho q_y}{\partial y} + \frac{\partial \rho q_z}{\partial z} = 0$$

Combining this with Darcy gives:

$$\frac{\partial\left(K_{xx}\frac{\partial h}{\partial x}\right)}{\partial x} + \frac{\partial\left(K_{yy}\frac{\partial h}{\partial y}\right)}{\partial y} + \frac{\partial\left(K_{zz}\frac{\partial h}{\partial z}\right)}{\partial z} = 0$$

when ρ is constant in time and space.

If external sources or sinks are present in the system:

$$\frac{\partial\left(K_{xx}\frac{\partial h}{\partial x}\right)}{\partial x} + \frac{\partial\left(K_{yy}\frac{\partial h}{\partial y}\right)}{\partial y} + \frac{\partial\left(K_{zz}\frac{\partial h}{\partial z}\right)}{\partial z} + W_{ext} = 0$$

Where K_{xx},K_{yy},K_{zz} = hydraulic conductivity along x,y,z co-ordinate axes (m/day)

(assuming the principal axes of the hydraulic conductivity are aligned with the coordinate axes.)

h = potentiometric head (m)

W_{ext} is the external volume flux per unit area e.g. well pumpage (outflow), or groundwater recharge (inflow)

S_s=specific storage of porous material (per meter)

The storage term in the groundwater flow equation is zero for steady state groundwater modelling (Harbaugh et al., 2000). To model groundwater flow, the above equation is solved. The partial

differential equations can be solved using analytical or numerical methods. However, analytical solutions are available only for very simplified models of reality and rely heavily on assumptions of isotropy and homogeneity of an aquifer which are not valid in general. Analytical solutions therefore are not applied to complex groundwater models.

Numerical methods have been developed to cope with complexity in groundwater systems. There are different groundwater modelling techniques, based on the numerical method used, to solve the differential equations. Numerical models in general involve solving a set of algebraic equations at discrete head values at selected nodal points. As discussed before, the finite difference method is one of the methods that can be used to solve a numerical groundwater model and has been implemented in this study as well.

1.1.3. Finite difference method

In the finite difference method, the groundwater flow equation is numerically solved by dividing the system into a grid with cells. The grid cells usually have a rectangular shape and nodes are located within cells (block-centered nodes) or on the intersection of grid lines (mesh-centered nodes). For each node an equation for groundwater flow is defined (Wang and Anderson, 1995). The head is calculated for a finite number of nodes (center of each cell) in space and time. The space discretization consists of rectangular grids for which cell geometry has to be defined. For the time discretization time steps are used, for which the length and number of time steps is predefined. Both steady state and transient conditions can be modelled using this method.

The water balance continuity equation for one cell, when assuming a constant groundwater density, is expressed as (McDonald and Harbaugh, 1988):

$$\sum Q_i = SS\frac{\Delta h}{\Delta t}\Delta V$$

By combining this equation with Darcy's law, the flow from one cell to another in one direction (Fig.1.6.) is obtained (McDonald and Harbaugh, 1988):

$$q_{i,j-1/2,k} = KR_{i,j-1/2,k}\Delta c_i \Delta v_k \frac{h_{i,j-1,k} - h_{i,j,k}}{\Delta r_{j-1/2}}$$

Numerical modelling softwares extends this principle by calculating the flow from one cell to all adjacent cells in the x, y and z direction and doing this for all cells in the model's grid. In addition, external fluxes are taken into account. By indicating initial head values at t=0, the groundwater model can be solved and groundwater heads are calculated for each cell and for each time step in case of a dynamic (transient) model.

Figure 1.5. further illustrates spatial discretization of an aquifer system using the finite difference method with a grid of blocks called cells, the locations of which are described using rows, columns, and layers. While formulating the equations of the groundwater model, an assumption was made that layers would generally correspond to horizontal geohydrologic units. Within each cell there is a point called a "node" at which the head is to be calculated. In groundwatermodelling softwares like MODFLOW the nodes are at the center of the cells (Modflow-2005 Manual) .

The different types of boundary conditions on such a discretized aquifer can be seen in Fig. 1.4. The area bounded by the conditions are also designated values that denote whether the groundwater heads in them can vary or not.

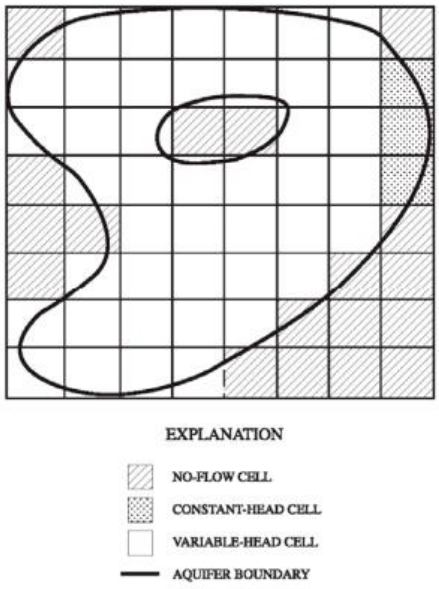


Fig. 1.4. Cell designations and boundary conditions (Modflow-2005 maual)

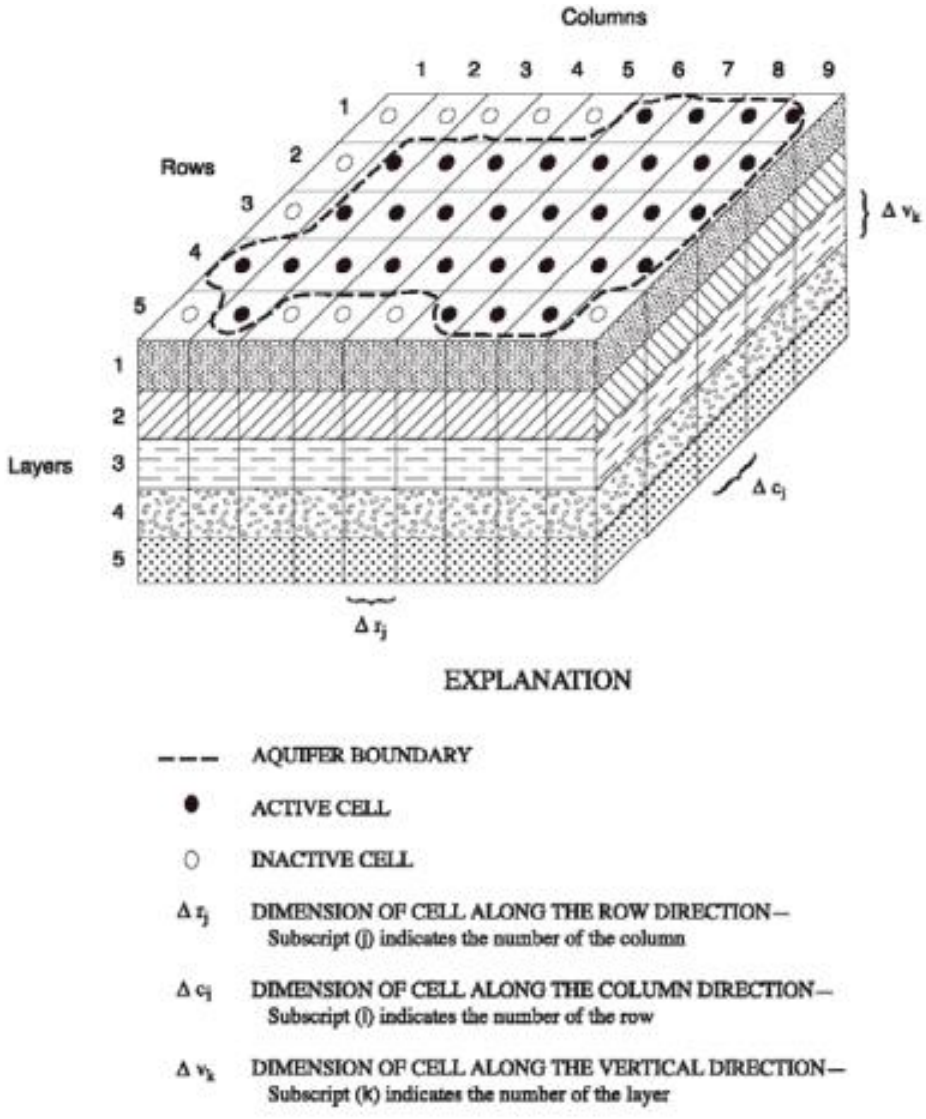


Fig. 1.5. A discretized hypothetical aquifer system (Modflow-2005 maual)

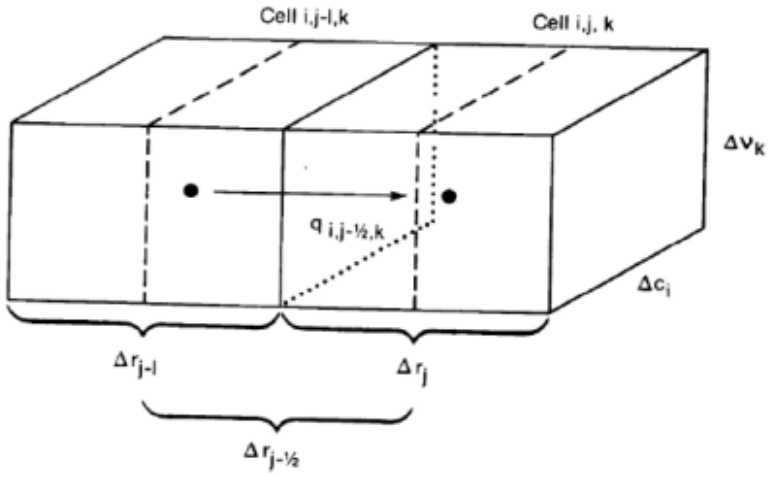


Fig. 1.6. Flow between two cells in finite difference modelling (MODFLOW-2005, Manual)

1.1.4. iMOD

iMOD (interactive MODeling) is a user interface for the groundwater modelling software package developed by Deltares and based on the concepts of Finite Difference groundwater modelling software MODFLOW which is developed by the United States Geological Survey (USGS). iMOD contains an accelerated version of MODFLOW called iMODFLOW as its flow simulation program which facilitates large, high resolution modeling and faster simulations. An improvement of iMODFLOW over earlier versions of MODFLOW is that for spatial data unequal resolution files can be used as well as sub-models at different resolutions can be easily used without recreating sub-sets of original model data.

Groundwater flow is modelled in iMOD following the same principles as MODFLOW. The influence of abstraction or pumping wells, rivers, geological faults on the groundwater flow is simulated using different packages in iMOD. iMOD will then call the different packages in order to simulate the groundwater flow.

1.2. Need for parameter estimation

After a first successful model run, the results may be different from field measurements. This can be attributed to approximations and assumptions of reality considered while setting up the model, as well as inevitable computational errors.
Model calibration is aimed at fine-tuning the model results to match field observations, by modifying the model parameters. In the case of groundwater modelling, the calculated head is forced to match the head at measured points. As a result of calibration, the model parameters (e.g. hydraulic conductivity) needs to be changed to obtain the best match between the calculated and measured head values. Calibration involves an optimization process to minimize the error criterion that is set for a particular groundwater model.

Nevertheless for a model to be considered well-calibrated, there needs to be ample quantitative data on model sources/sinks and parameters (Baalousha, 2011). However, in most real-world systems being modelled, data required for groundwater modeling purposes is not readily available through direct measurements or *a priori*. In such cases, we are forced to deploy parameter estimation methods through inverse modelling or calibration problems. Observed heads at measurement points are usually the calibration targets which are compared against the computed output variables.

In trial-and-error methods of calibration, an initial set of values for the model parameters being estimated (e.g. hydraulic conductivity) are assigned to each node in the groundwater model grid and the piezometric heads are calculated for the model. This is then compared against field measurement

values and the model parameters are adjusted accordingly by a certain factor decided by the modeler, so as to minimize their difference further. At the end of the calibration process, the minimized objective function value should demonstrate an adequate fit.

The parameter estimation problem can be defined in several ways using the minimization function. The parameters to be minimized depend on the specific groundwater model in consideration. Parameter estimation is conducted by iterative methods which aim to minimize a pre-defined objective function used to evaluate the fit between observed and calculated heads at each iteration. The parameter values are subsequently adjusted at the end of each iteration, until the objective function in minimized.

There are many factors that combine to make selection of an algorithm for parameter estimation a complex problem. Complex models can be very computationally demanding and this can be a major constraint in doing sufficient runs to get accurate results, this means algorithms must make do with relatively few evaluations of the objective function (Blasone et al. 2007). Strong correlations between two parameters can also make it difficult to estimate them separately as well as introduces more uncertainty into the solution by being non-identifiable. As such, strongly correlated parameters (e.g. thickness and conductivity) are often combined as one and estimated (e.g. transmissivity).

1.2.1. Use of algorithms in parameter estimation

Parameter estimation methods typically use some optimization algorithm to minimize the objective function defined for the problem at hand. While earlier this used to be done using a 'trial-and-error' approach, now most parameter estimation methods make use of optimization algorithms to do the work for them.

1.2.1.1. Previous research

The use of optimization algorithms in model calibration has been explored in the past. However, most of these studies focus on optimization methods that are capable of finding only the local minimum of the model error function (objective function). While most of the traditional tools for groundwater model calibration consider the minimization function as a single extremum function, inevitably leading to accept a local minimum that such tools are capable of finding, there is no proof that exists that suggests the problem of groundwater model calibration is not a multi-extremum function.

Arsenault et. Al. (2013) compared the performance of several optimization algorithms for a hydrological model and recorded the dependency on dimensionality of the model as a major influential factor. With higher dimensional models, it is impossible to calibrate other than by using automatic algorithms (Moradkhani and Sorooshian 2009; Tolson and Shoemaker 2007).

The purpose of such heuristic methods is to search the parameter space which is bound by a set of lower and upper parameter values, and find the global minimum of the error function. However, hydrological problems are highly non-linear and non-convex. As such there is a high probability of presence of local minima. While a global optimization algorithm may find the best solution it is impossible to prove with 100% accuracy that a minimum is global in a non-convex problem. In theory there are probably algorithms that could find a global maximum as long as the function to be minimized is locally smooth, but these involve splitting up the search space in "sufficiently small" components which in practice results in run times exponential in the number of dimensions. This is compounded by the need to run the groundwater model for many, many points. The best methods will be the ones who converge to the best quality minimum, even if it is the best local optimum, with least use of computation time.

Over the past few decades, several algorithms have been used to perform minimization in groundwater models, some of which are the Gauss-Newton algorithm ,the Levenberg-Marquardt algorithm and the Broyden-Fletcher-Goldfarb-Shanno (BFGS) algorithm (Yeh, 2015), GLUE and DREAM.

1.2.1.2. Need for this study

Most of the algorithms used in the optimization process of groundwater models assume the problem to a single-extreme minimization problem. However, in the present study, we make no such assumptions on the existence of just a single minimum for the error function; the optimization problem is posed as a global optimization problem, with possibility of existence of local *and* global minima. Traditional optimization tools like PEST or iPEST (as it is used in iMOD) are incapable of consistently finding global optima (Vermeulen et. al. 2017). Existence and subsequent discovery of a global optimum could lead to better estimation of aquifer characteristics for a groundwater model. Genetic algorithms (GAs; Goldberg 1989) have been used before in groundwater management problems (Madsen & Perry, 2010). Heuristic search techniques such as Simulated Annealing and Tabu Search have been used in parameter structure identification (Zheng and Wang 1996) as have been other global search algorithms such as “ant colony”, “particle swarm”, and “honey bees mating”. The key advantage of using these algorithms is that they can handle non-differentiable and non-convex functions, but at the expense of computation time. Solomatine et. al. (1999) compared the performance of several global optimization techniques on a Triwaco Flairs model and compared it against manual calibration done on measurement data. However, the identification of the parameters selected for the estimation purpose was largely subjective and thus may not be the best selection method. Furthermore, several parameters were left out as they had already been set by the manual calibration.

In this study, we have tried to make a standalone and independent assessment of the optimization process. The basic form of objective function as defined in Section 2 has been used for the optimization problem in this study. This choice has been tested and proved to have multiple local minima during the course of this study. iPEST inbuilt in iMOD groundwater modelling software, has been used over the last few years in parameter estimation. However iPEST, like PEST, is capable of only finding a local optimum (Vermeulen et. Al., 2017). The performance of such a local and other global optimization algorithms in discovering the global minimum for the chosen objective function have been put to test.

1.3. Objective and research questions

In this study, a global optimization algorithm is selected and its performance compared against that of the inbuilt iPEST optimization tool on an iMOD groundwater model in a selected study area. The aim of this study is to assess whether a global optimization algorithm can give more reliable estimates of parameters or the best possible values during calibration for a groundwater model than a local search algorithm that is being traditionally used. The performance of the global optimization algorithm will be measured on factors like computation time, success rate and plausibility of results. This brings forth our research questions:

1) Can we have a better calibrated groundwater model using a global optimization algorithm than the traditional ones in use?
2) Are the parameter values estimated by the global optimization algorithm physically plausible?
3) Is there a trade-off between computation time and results in the performance of both algorithms?

We have considered steady state modelling and thus will not be considering the time dependency of the fluxes in and out of the system.

2. Choice of Algorithms

2.1. iPEST

In our present study, we have used iPEST (Vermeulen et. Al. 2017) in-built calibration tool of iMOD groundwater modelling software, as the base optimization algorithm we are assessing the performance of. iPEST uses Levenberg-Marquardt algorithm (LMA) which is a numerical method used for non-linear functions. The LMA uses a combination of Gauss-Newton Algorithm (GNA) and the method of gradient descent. iPEST has been developed based on the concepts of LMA and PEST by Doherty (2010).

PEST (Doherty, 2016), one of the most popular tools for automatic calibration of groundwater models, is a non-linear parameter estimation software that is widely used to estimate model parameters. The user needs to inform PEST of the model parameters in the input model files that it must adjust. The corresponding output files containing the values that must be matched to field measurements (observation well heads) should also be informed to PEST console.

Before starting the optimization, PEST writes a set of model input files containing parameter values which PEST would like the model to use on that run. PEST then runs the groundwater model several times adjusting the model parameters every time. PEST begins each iteration with the calculation of the Jacobian matrix which contains the partial derivatives of each observation with respect to each parameter. Calculation of the Jacobian matrix thus takes as many model runs as there are adjustable parameters. Once a Jacobian matrix has been calculated, it is now used to calculate an improved set of parameters within the parameters upper and lower bounds. After creating a new parameter upgrade vector, a model run is initiated and objective function value is calculated and tested if it has been lowered. Otherwise steps are taken to create new parameter adjustments.

For the purpose of parameter estimation, PEST minimizes a least-squares objective function. This is the sum of squared weighted differences between measurements and corresponding model outputs. The difference between a measurement and a model output it corresponds to is referred to as a residual, as designated below:

$$r_i = h_{calc} - h_{obs}$$

Where r_i = i^{th} residual
h_{calc} = calculated piezometric head
h_{obs} = observed piezometric head
The objective function, Φ, is calculated using the following equation:

$$\varphi = \min \sum r_i^2$$

The observation values are assigned a default weight of 1.0. Weights are assigned to measurements which are more trustworthy or to give equal weight to observations with different units such as Heads vs. Fluxes.

The objective function would then be defined as:

$$\varphi = \min \sum (w_i r_i)^2$$

At the end of the PEST run, provided the adjusted parameters are reasonable, the results are saved which can be used as inputs for a new model run. PEST stores the Jacobian matrix corresponding to the best set of parameters achieved up to any stage of the parameter estimation process in a special

binary file which is updated as the parameter estimation process progresses. The basic principle of optimization used in iPEST is similar to PEST.

The Levenberg-Marquardt algorithm is applied to minimize the objective function value by multiplying the individual values for the parameter vector with a factor which is adjusted as the calibration process progresses, so as to obtain a minimal value for the objective function. The Gradient Descent part comes into play by adjusting each parameter according to the gradient of the objective surface), thus leading the algorithm to faster convergence. Each parameter is assigned an upper and lower bound for its multiplication factor to vary. If at any point during iteration, the paramteres exceed their upper or lower bounds, this can influence the optimization process and other parameters. This is avoided by temporarily freezing such parameters at their upper or lower bounds. They will be included in the optimization process later if the parameter vector finds itself calculating values that move these frozen parameters back within their bounds.

2.2. Genetic Algorithms

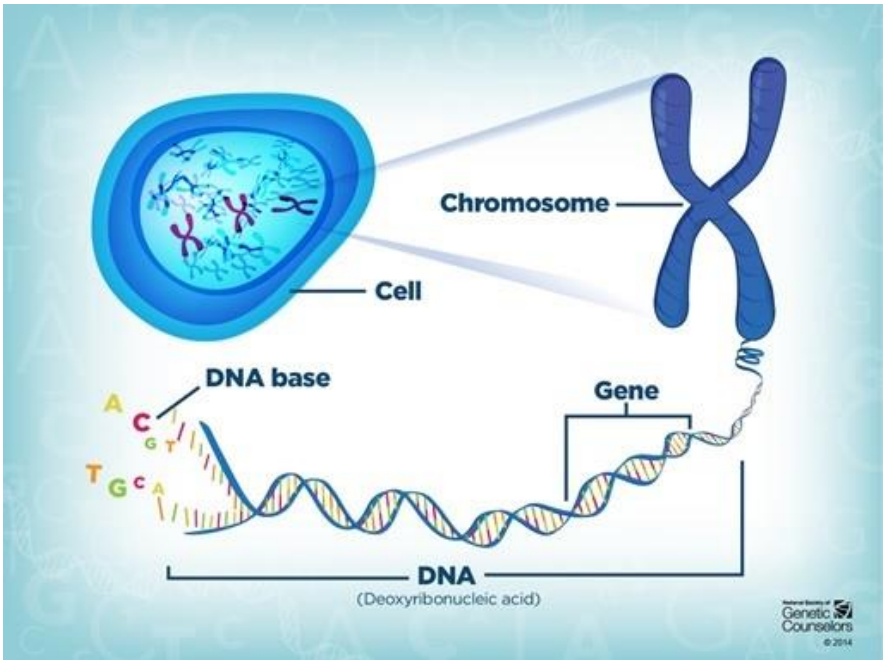
Genetic algorithms (GA) have been used for optimization problems across several fields and disciplines. GAs are known to be flexible and can be applied to any real world problem, provided it is structured numerically (Goldberg, 1989). Genetic algorithms use the mechanics of natural evolution to solve mathematical problems and can be modified to represent functions that mimic biological processes such as mutation, genetic selection. Genetic Algorithms can be easily coded using several programming languages such as Python and Visual Studio C#. Genetic algorithms are also advantageous in the fact that they can work with non-differentiable and discontinuity of the objective function. (Yeh, 2015). Genetic algorithm was chosen as the choice algorithm in this study due to their intrinsically parallel ability to explore different directions of the solutions space at the same time, unlike linear search methods. They can directly sample only small regions of the fitness landscape and successfully finds optimal or very good results at the very least in a short period of time (Bajpai et. al., 2010). While GAs are rather time-consuming, it is possible to reduce this by choosing the right settings for the algorithm. GAs are capable of escaping being stuck in local minima and possible ruggedness of the solution space. Its ability to have multiple starting points can cover more areas of the solution space and lead to confidence in the results obtained.

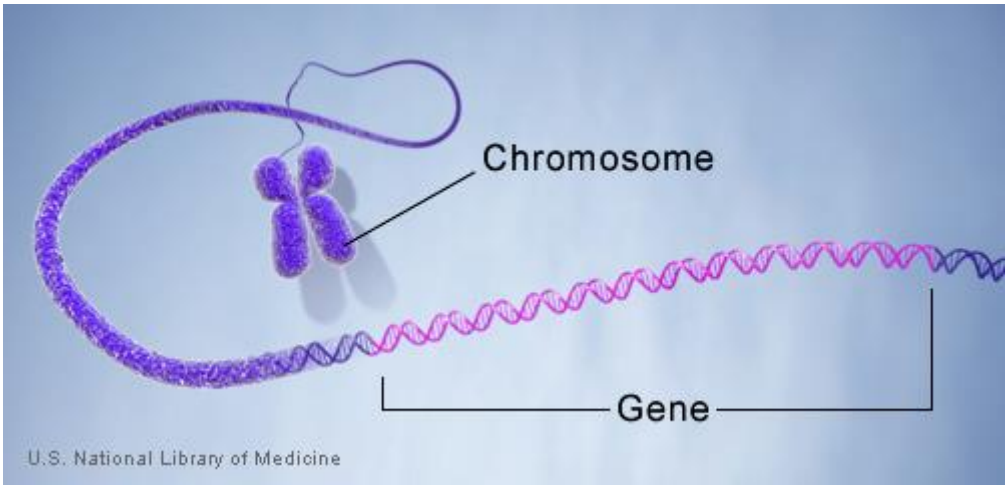
2.2.1. Basic terminology

Genetic algorithms are search algorithms which bases its working on the mechanism of natural selection in biological evolution. Similar to evolution, genetic algorithm works in a random manner, however, it allows the user to control the level of such randomness. GAs have been considered far more efficient than other random search algorithms due to their ability to handle discontinuity and non-linearity in optimization problems (Carr, 2014).

Genetic algorithms are inspired by Darwin's theory of natural evolution and reflects the process of natural selection where the fittest individuals are selected for reproduction in order to produce offsprings for the next generation. These offsprings inherit the characteristics of the parents and will be added to the next generation. If parents have better fitness, their offspring will be better than parents and have a better chance at surviving. This process keeps on repeating over the generations and at the end, a generation with the fittest individuals of all past generations will be left in the end. In the same way genetic algorithms try to mimic human evolution. This notion can be applied for a search and optimization problem. In the most simplified manner, we consider a set of solutions for a problem and select the set of best ones out of them.

Due to its basis in biology, most of the terminology used by Genetic Algorithms are synonymous to relevant evolutionary processes among biological species. A few of the terminology used by genetic algorithms can be listed as below:





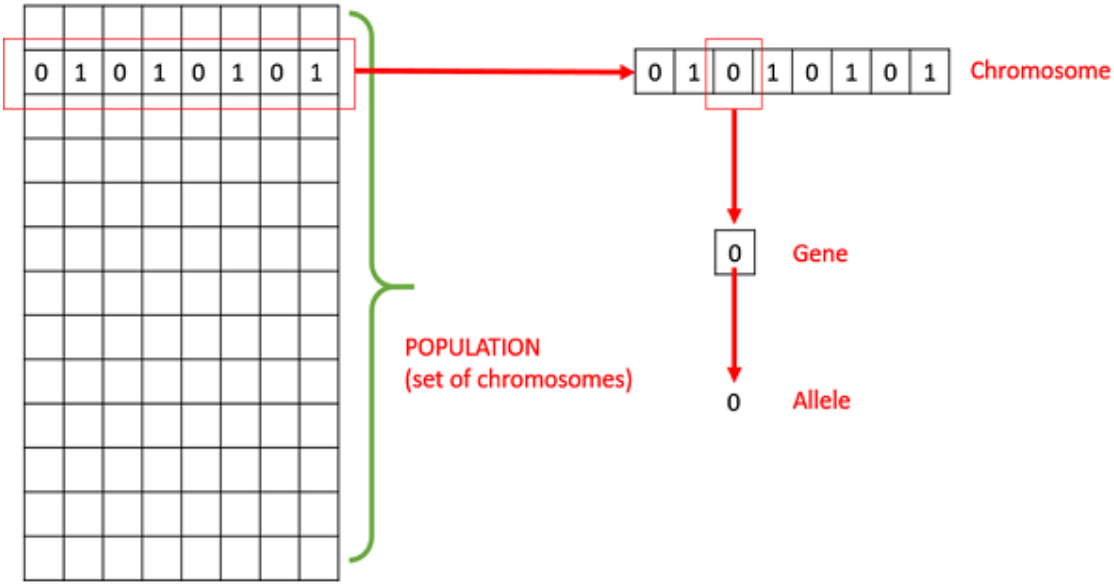


Fig. 2.1. Representation of GA terminology (Ref. U.S. National Library of Medicine, Tutorials Point)

- **Population** − It is a subset of all the possible solutions to a given problem. The population for a GA is analogous to the population of human beings which are capable of reproducing and creating new fitter generations of humans. In the case of genetic algorithms, except that instead of human beings in the context of reproduction, we have candidate solutions for a particular problem at hand.

- **Chromosomes** − A chromosome is one such solution to the given problem. A set of chromosomes constitute a population of solutions (Fig. 2.1). In biology, our body is made up of billions of cells which contains genetic information in the form of DNA. Inside a cell, DNA is packaged into groups of structures called chromosomes (Fig. 2.1).

- **Gene** − A gene in Genetic Algorithm denotes the position of one element in a chromosome. A combination of genes (elements) thus constitute a chromosome, similar to a human body (Fig. 2.1).

- **Fitness Function** − A fitness function is any function which takes the solution as input and produces the suitability of the solution as the output. It is a positive number assigned to a solution as a measure of its goodness/suitability/fitness. The value of fitness function evaluates how close a particular solution is to achieving a preset aim for the problem at hand.

- **Genetic Operators** − These alter the genetic composition of the offspring so as to introduce certain level of variety into the population. This is accomplished by using operators such as Crossover, Mutation, Elitism, etc. (Section 2.2.3.)

- **Generations** – These are successfully created offspring populations as a result of the actions of genetic operators of the parent population.

  The process workflow in a basic Genetic Algorithm is shown in Fig. 2.2. We start off with an initial population (which may be generated at random or chosen from prior initial knowledge about the problem in hand), and select parents from this population for mating. The genetic operations (crossover, mutation, etc.) are applied on the parents to generate new off-springs. As a next step these off-springs replace the existing individuals in the population and the process repeats until a particular termination criteria is achieved.

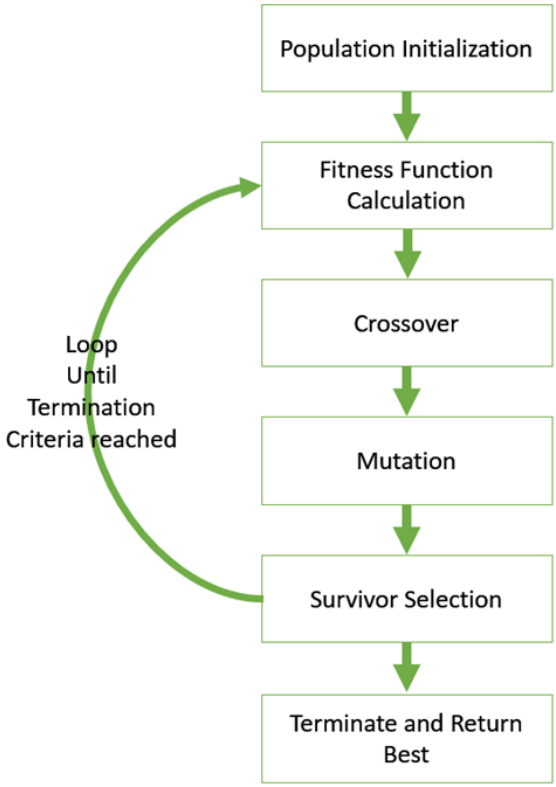


Fig 2.2. Genetic Algorithm process (Ref. Tutorials point)

2.2.2. Representation

Genetic Algorithms work with coded versions of candidate solutions. It is necessary to choose the proper representation/gene coding to get the best performance. Representation is highly problem specific. For example, problems can be represented using binary values, integers or real values. In this study, real valued representation has been adopted.

0.5	0.2	0.6	0.8	0.7	0.4	0.3	0.2	0.1	0.9

Fig. 2.3. Sample real-valued representation of a chromosome.

2.2.3. Genetic Operators

Several genetic operators are applied to a chromosome to create a new and varied population. Some of them are mentioned below:

1) Crossover:

The crossover operator is analogous to reproduction in biology. In this operation, two or more parents are selected and one or more offsprings are produced by combining the genetic material of the parents. A high crossover probability rate is assigned to GA in order to apply this. A few types of crossover used in GA is mentioned below:

- One Point / Single Point Crossover

In this type, two chromosomes are chosen and then a random crossover point is selected and the portions (genes) of its two parents (chromosomes) are swapped to get new offsprings.

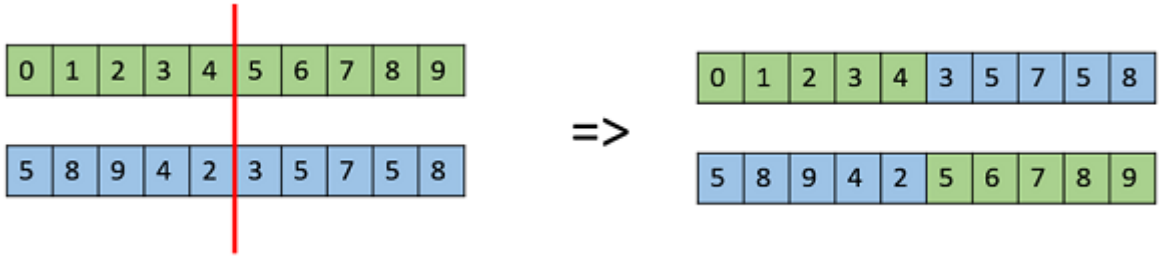


Fig. 2.4. One point crossover

- Multi Point Crossover

Multi point crossover is the generalized form of the one-point crossover wherein alternating segments (genes) are swapped among two parents (chromosomes) to get new offsprings. One commonly used type is Double point crossover as shown in Fig. 2.5.

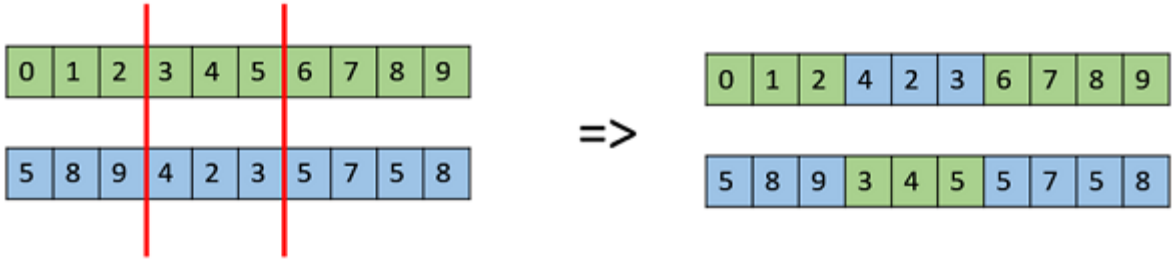


Fig. 2.5. Double point crossover

- Uniform Crossover

In this type of crossover between chromosomes, each gene is treated separately for recombination. A probability is assigned to each gene to decide whether it will be included in the creation of an offspring.

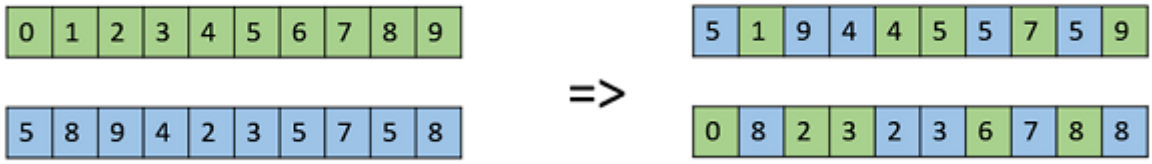


Fig. 2.6. Uniform Crossover

2) Mutation:

Mutation can be defined as a small change in the gene structure in the chromosome, to get a new solution. It is used to introduce diversity in the genetic population and is usually applied using with a low probability value.

Mutation is a very important part of the GA as it is related to the exploration of the search space.

3) Elitism:

In a fitness based selection as used in this study, the children tend to replace the least fit individuals in the population. Elitism is the operation by which the current fittest member of a population is always carried forward to the next generation. It ensures such a good solution does not get replaced during a generation. A certain percentage of the fittest solutions are ensured to survive each generation using the Elitism operator.

2.2.4. Termination Criteria

The termination criteria of a Genetic Algorithm is essential in determining when a GA run will be finished. As GA is a computationally expensive algorithm, it is important to choose the termination criteria wisely so as to not to keep on running the algorithm even after the best solution has been found.

One of the following termination criterion can be used −

- When there has been no improvement in the population for a certain number of iterations.
- When we reach a specific number of generations.
- When the objective function value has reached a certain pre-defined acceptable value.

3. Study area

The area chosen for this study is the stream valley Oude Strijper Aa, in the province of North Brabant.

The river Strijper Aa consists of two parallel streams, Oude Strijper Aa and Nieuwe Strijper Aa. The upper course of the Strijper Aa belongs to the European network of Natura 2000 protected areas ‘Leenderbos and Groote Heide’ & ‘De Plateaux’. During the land consolidation in 1973 a new stream was dug on the east side (WDD, 2016). On the west side, closer to the Leenderbos, is the Oude Strijper Aa. The Nieuwe Strijper Aa starts at the Dutch-Belgian border and then flows past the Soerendonks Goor nature reserve and to the east of Leenderstrijp to Leende , where it joins the river Buulder Aa to form the Groote Aa. (Fig.3.2., 3.3, 3.4.)

The study area falls under the control and jurisdiction of Waterboard De Dommel, Boxtel in the southern part of the Netherlands marked in Fig. 3.1.

Fig. 3.1. Service area of Waterschap De Dommel

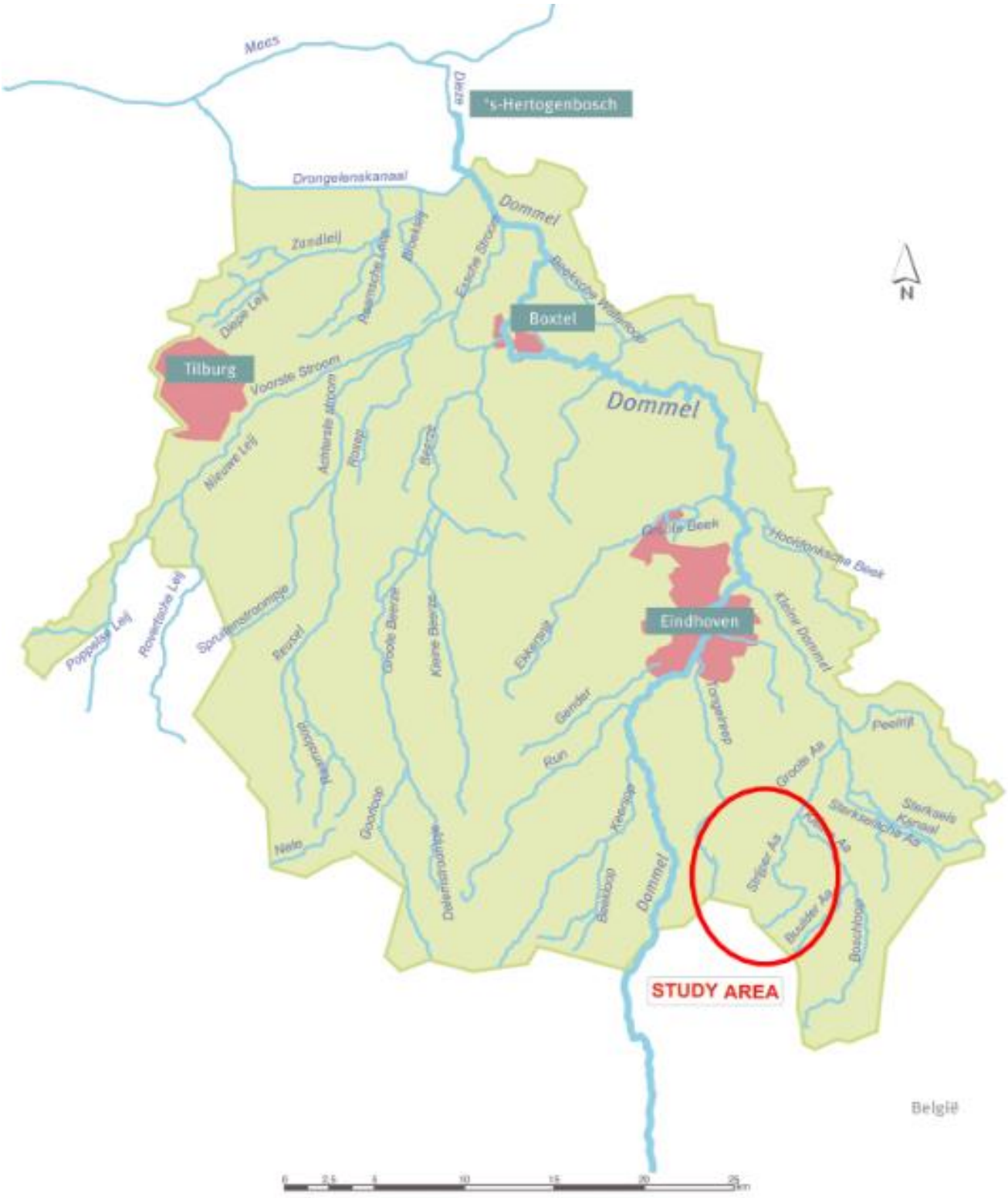


Fig.3.2. Service Area of Waterboard De Dommel (Ref. Waterboard De Dommel website)

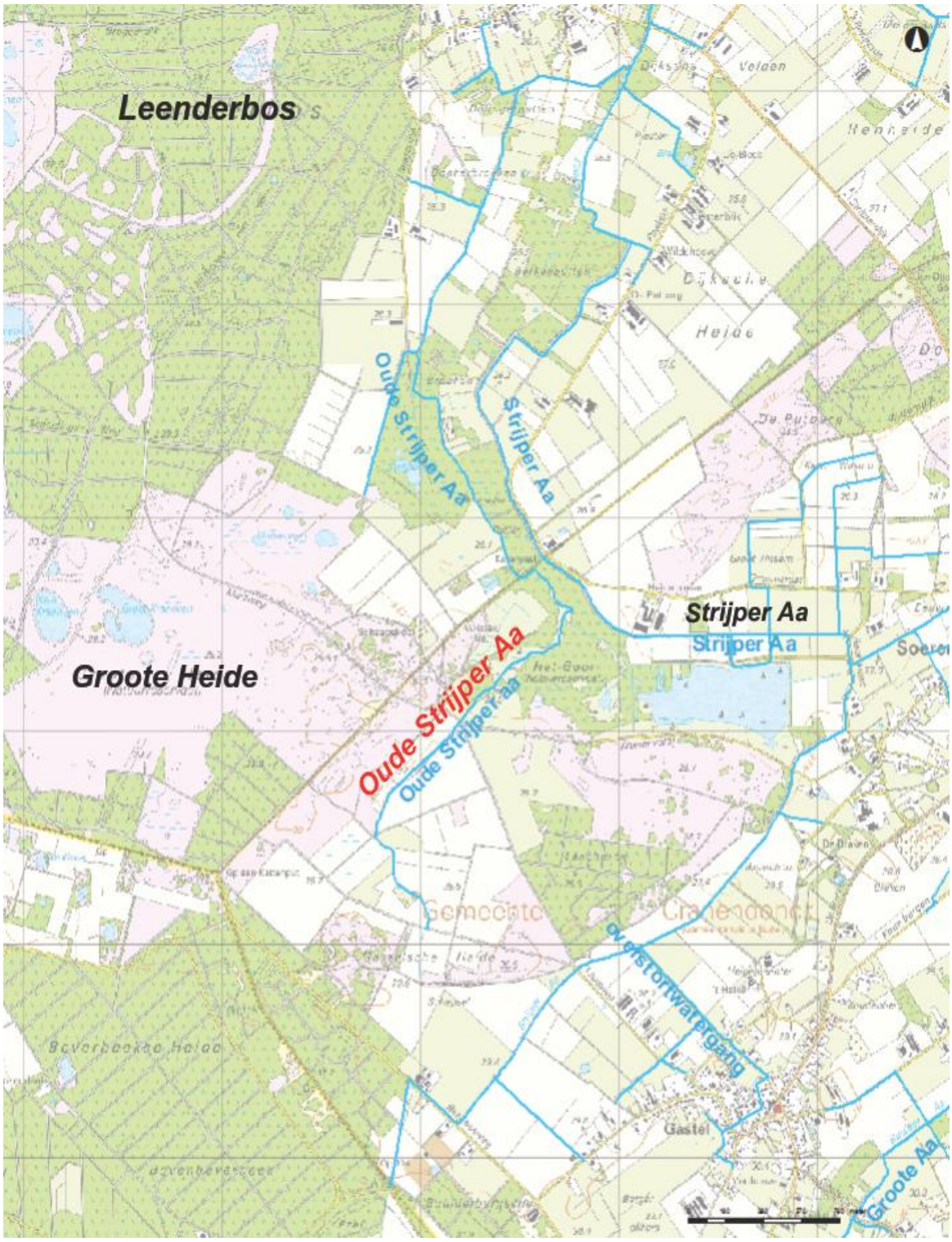


Fig. 3.3. Study Area & surroundings (Ref. Waterschap De Dommel)

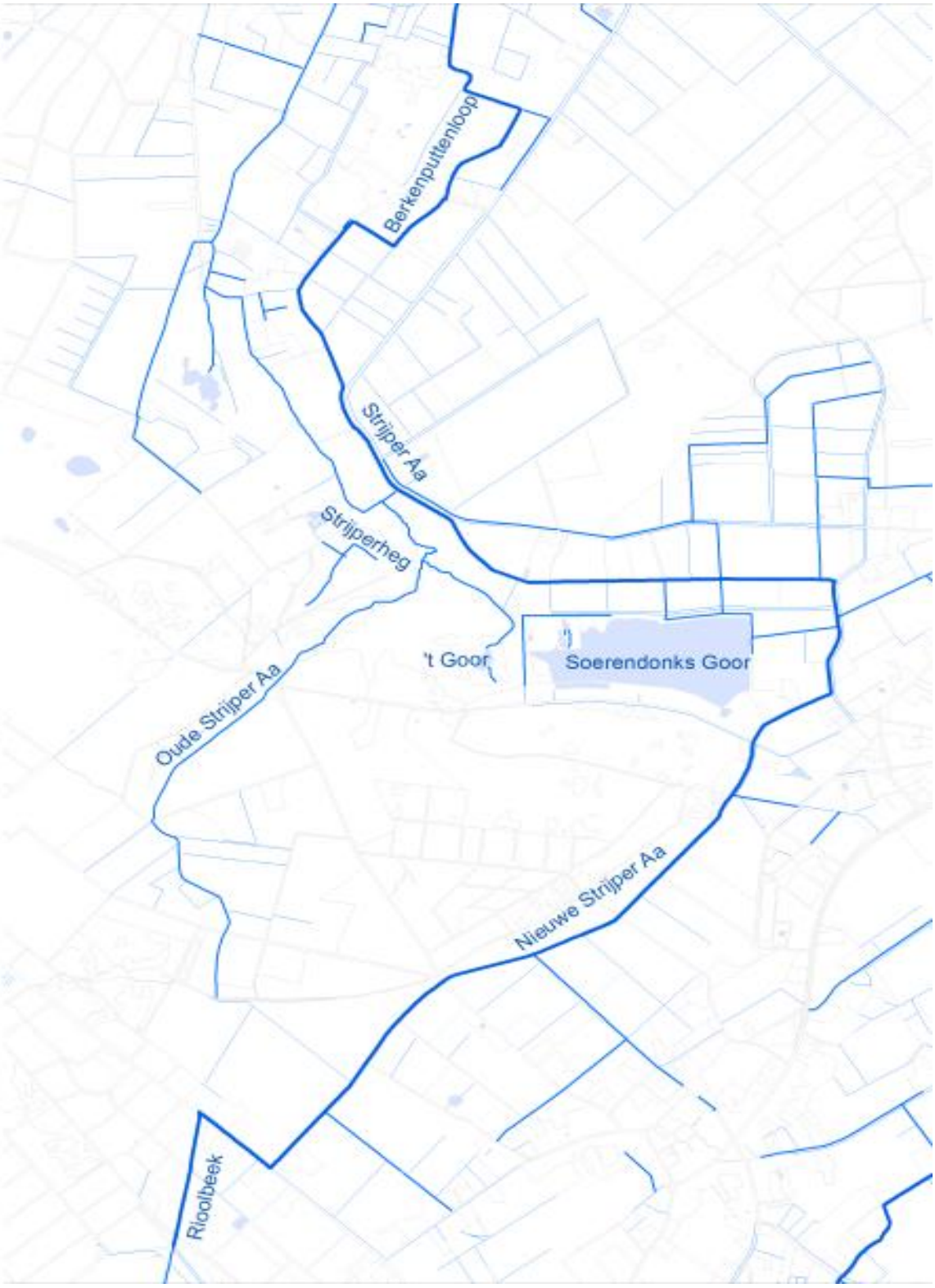


Fig. 3.4. Oude Strijper Aa & connected streams (Ref. Waterschap De Dommel)

3.1. Why this area?

At present, the Oude Strijper Aa has become deeper in the landscape and has little room for a natural course. The water now flows out of the area too quickly, causing the surrounding fens and forest areas to suffer dehydration. Morphological processes such as erosion, sedimentation and meandering hardly take place anymore. The brook is also deeply cut so that it has a draining character.

Fig.3.5. Oude Strijper Aa – field conditions (Ref. Waterschap De Dommel)

Due to this, fish can no longer migrate upstream. There are also several factors in the catchment area of the Oude Strijper Aa that leave a mark on water quality. For example, quality is strongly influenced by nutrients as a result of agriculture and Waste Water Treatment Plant (WWTP) overflows from Belgium.

De Oude Strijper Aa does not sufficiently meet the objectives of the various policy documents set for the Natura 2000 areas in the current situation. Among them, a partial function of water nature has been established from the origins of the Oude Strijper Aa to the inflow into the canalized Strijper Aa.

The Waterschap de Dommel aims to increase the flow rate in the stream and bring back a better stream morphology in the old Strijper Aa, which will cause erosion, sedimentation and meandering to take place again. This gives the drain a more natural flow character, reduces the dehydration in the surrounding areas and ultimately complies with the objectives as laid down in the Water Framework Directive. The groundwater level may not sink too far either to ensure natural flow in the stream. Research needs to be carried out to better understand the soil structure and groundwater flow in this area. By creating a groundwater model and comparing it with field conditions, the Waterschap De Dommel aims to be capable of accurately predicting the effect of planned measures in restoring the Oude Strijper Aa. As such, this area was chosen as the project area for this study.

3.2. Model Area

The model edge for the groundwater model for the Oude Strijper Aa is more than around the project area chosen. This includes the adjacent stream valleys within the model edge. Fig. 3.6. shows the location of the model for the Oude Strijper Aa within the regional model of the Dommel water Board again. The project area is displayed with a red boundary and the modelled area with a green boundary.

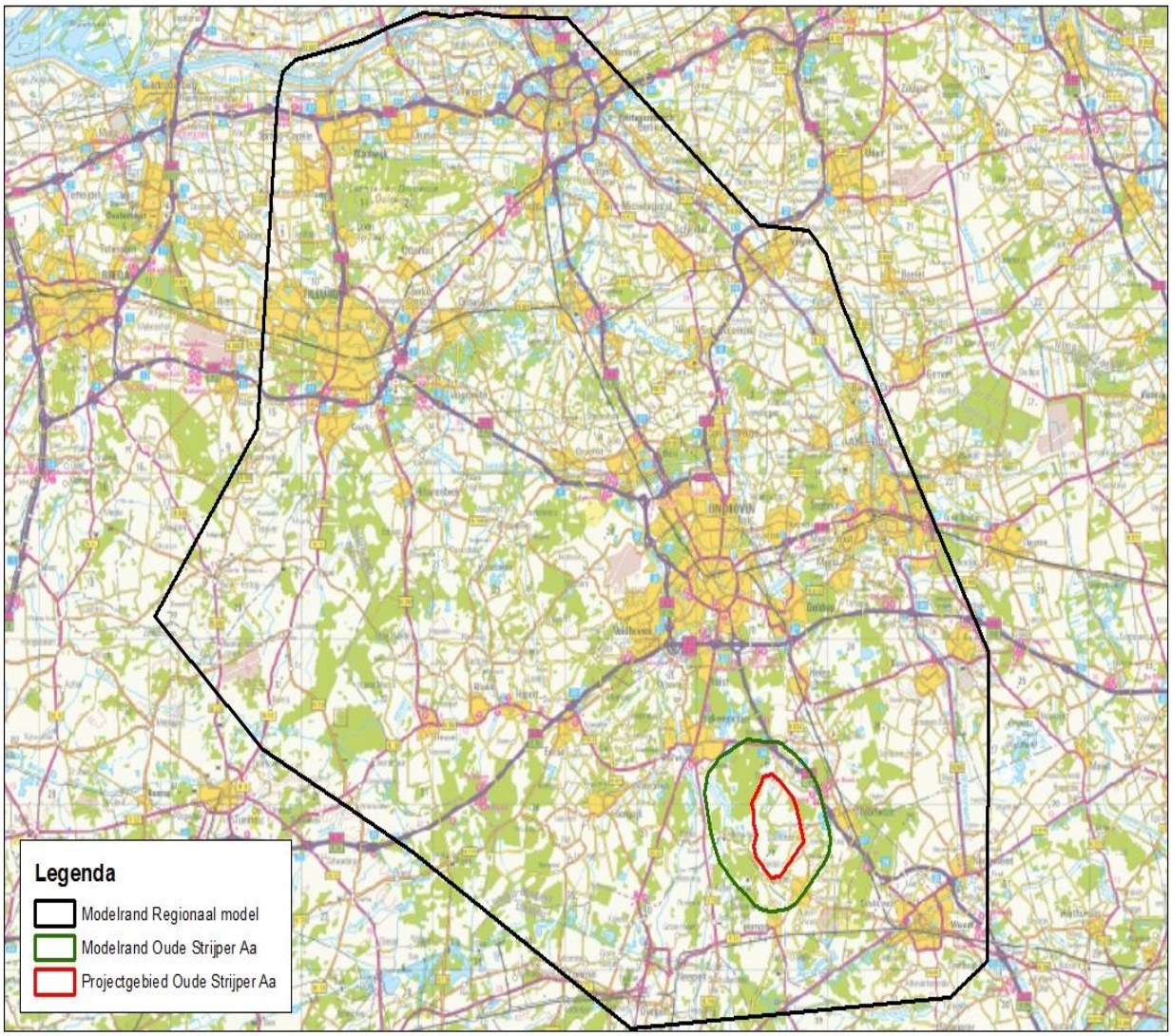


Fig. 3.6. Area modelled by Regional and local Triwaco Flairs models (Ref.Waterschap De Dommel)

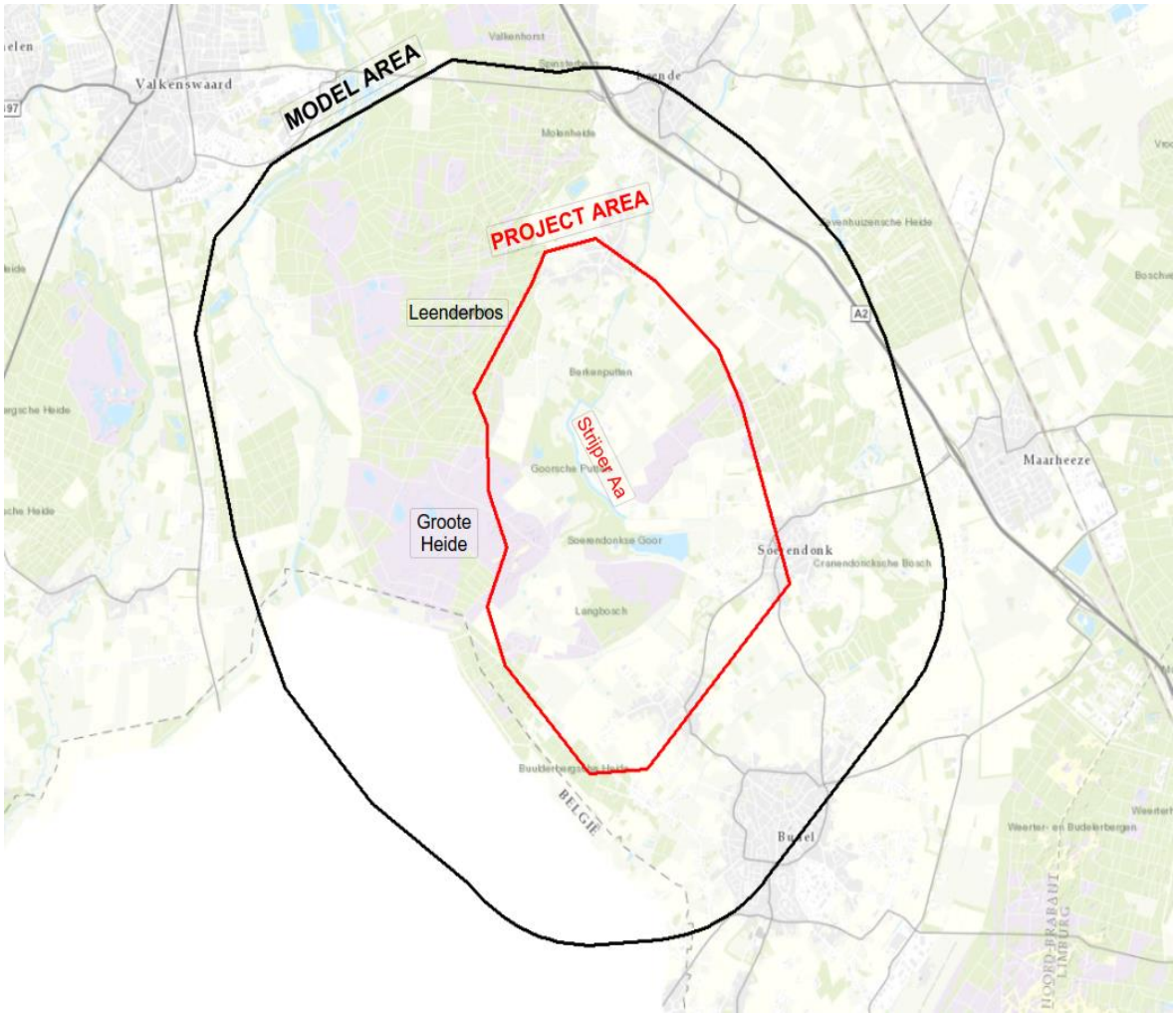


Fig. 3.7. Model Area & Project Area

3.3. Properties of the area

Fig. 3.8. shows the main topographical features of the model area in this study. The legend denotes the variation in elevation across the study area. The red dotted lines indicate the locations of the fault lines and the blue lines the locations of the rivers that flow through the area.

3.3.1. Elevation

From Fig. 3.8., a decline in elevation from south to north direction in the study area can be observed. As such, the groundwater flow in this area is expected to follow the trend by flowing from a higher head potential in the south to a lower head potential in the north. The river valleys in the model area are deeply incised with much lower elevations than that of the surroundings which is expected to provide a sink for groundwater from the surrounding elevated areas to flow into.

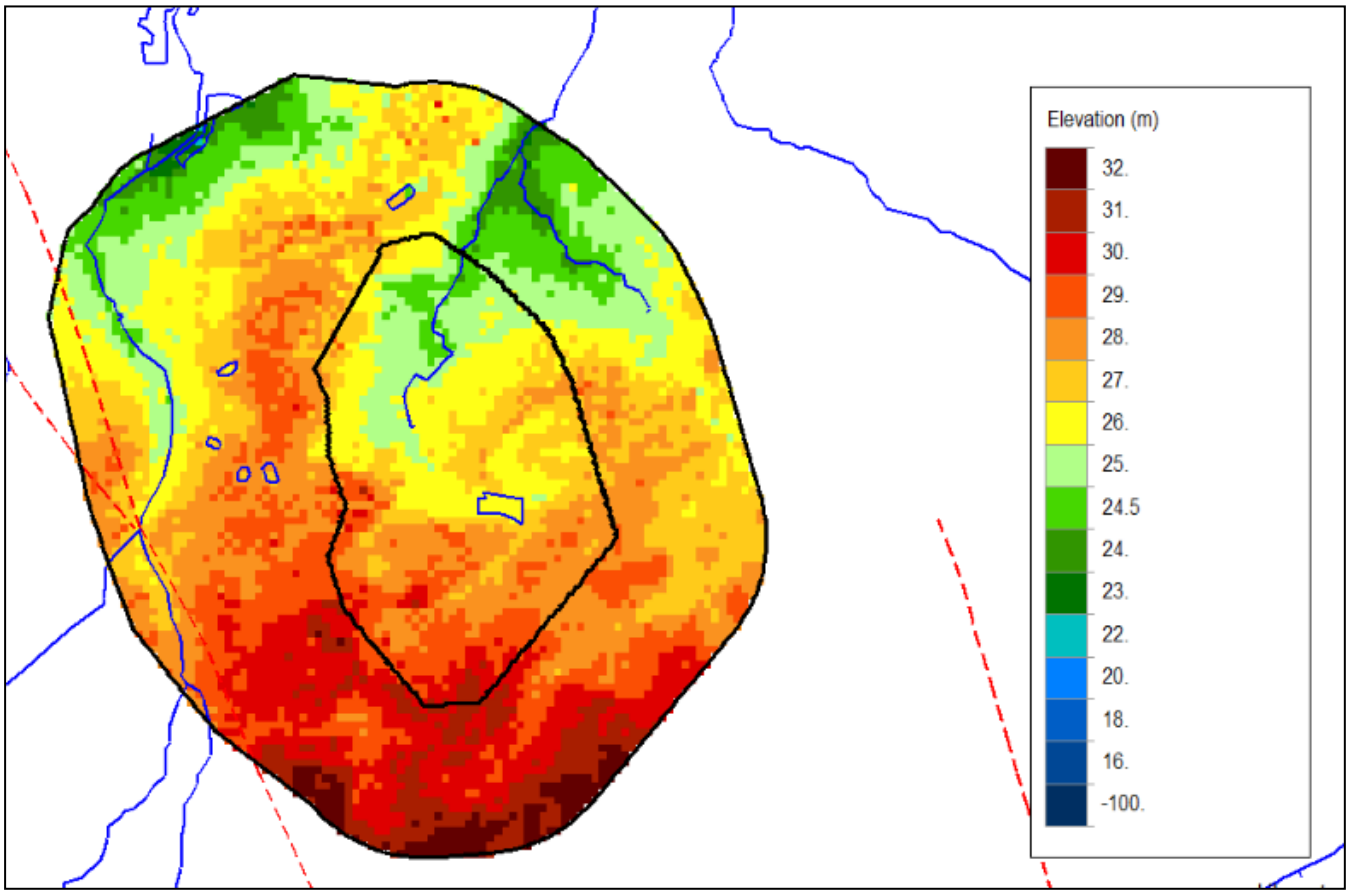


Fig. 3.8. Elevation map of Model Area and Project Area. Fault lines (dotted red) and major river systems (blue) are displayed.

3.3.2. Fault lines

Within the management area of water board De Dommel and the immediate environment, there are a number of fault systems. The fractions that are assumed to have the most influence on the groundwater flow are the Peelrand fracture and the Feldbiss. The zone between the Peelrand and the Feldbiss forms the Centrale Slenk. The Feldbiss fault line traverses the modelled area in this study (Fig. 3.8.).

The fractures in the subsurface could have an important influence on the behavior of the groundwater at certain locations. Rupture between geological fractions can ensure that geological layers that lie at a relatively short horizontal distance can get vertically offset with respect to each other. Depending on the geology in an area, this can have a significant influence on the flow of the groundwater. For example, An offset of a clay layer can provide an extra resistance to flow between two layers of sand that had not existed without breakage. For certain regions in the management area of the Waterschap de Dommel, there are indications that fracture effect has a significant influence on the groundwater levels. This is derived from measured groundwater levels that differ significantly at relatively short horizontal distances. This may indicate that groundwater flow is not possible or is heavily hindered due to the fractures.

3.3.3. Rivers

Two main rivers of Strijper Aa (Dommel) and Tongelreep intersects with the aquifer modelled in this study. It is important to understand their influence on the groundwater in the project area since the aim is to establish natural flow back in to the stream valley of Oude Strijper Aa and with that, a hydraulic connection. (Fig. 3.8.)

3.4. Expected outcomes

There are a few aims for this project in the chosen study area. In addition to understanding and rejuvenating the Oude Strijper Aa as mentioned in Section 3.1, the Waterschap De Dommel also has made it a priority to make a push into the territory of automated data processing when it comes to groundwater models. There is keen interest in learning more about optimization methods that are used and the results an optimization algorithm can come up with. Lastly, the waterboard is interested in testing which are the most influential hydrological factors in this region, by receiving the best possible calibration results. This can aid them in obtaining insight into the effect of proposed measures for reinstating a natural flow for the project area.

4. Methodology

4.1. Overview

To achieve the defined objectives, this study was developed in 3 phases.

Phase 1. Building an iMOD groundwater model for study area

Phase 2. Calibration of iMOD model using iPEST

Phase 3. Calibration of iMOD model using Genetic Algorithm

4.2. iMOD Groundwater Model

Data was collected from an existing groundwater model that was built using Triwaco Flairs model code by the Waterschap De Dommel. Details of the source data collected are described below. Further details of conversion of the source data to an iMOD compatible format and basic differences between the Triwaco Flairs and iMOD in their implementation can be found in Annex.

4.2.1. Data collection

The Waterschap De Dommel already has a regional groundwater model for the Brabant province which was set up and calibrated in 2014 and further based on this, a local groundwater model for the current focus area Oude Strijper Aa. To build a groundwater model in iMOD, data corresponding to the required model input files needed to be collected from the Triwaco Flairs model. The data collected for the aforementioned purpose is listed and described below.

- **Model extent & resolution**

The best fit between the grid of the source Triwaco Flairs data files and proposed iMOD model grid was matched after several trials and the model area boundary limits were decided as:

XLL: 160000.0

YLL: 363100.0

XUL: 170500.0

YUL: 374100.0

Units are given in meters (m) on a geographic projection of Amersfoort /RD new.

The model resolution was chosen as 100x100m, similar to the source/basis data for the Triwaco Flairs model.

- **Layer model data**

The original Triwaco Flairs model for the study area was modelled in 19 layers. The layer structure is based on REGIS 2.1 Geohydrological Model (Netherlands portion) and DOV Model (Flanders portion) for the study area. Each model layer out of the 19 total consists of an aquifer and an aquitard. Fig. 4.1. shows the types and distribution of geologic formations that are present along the schematization of the study area in this study. A combination of such geologic formations are often present in each model layer. The layer boundaries are not depicted.

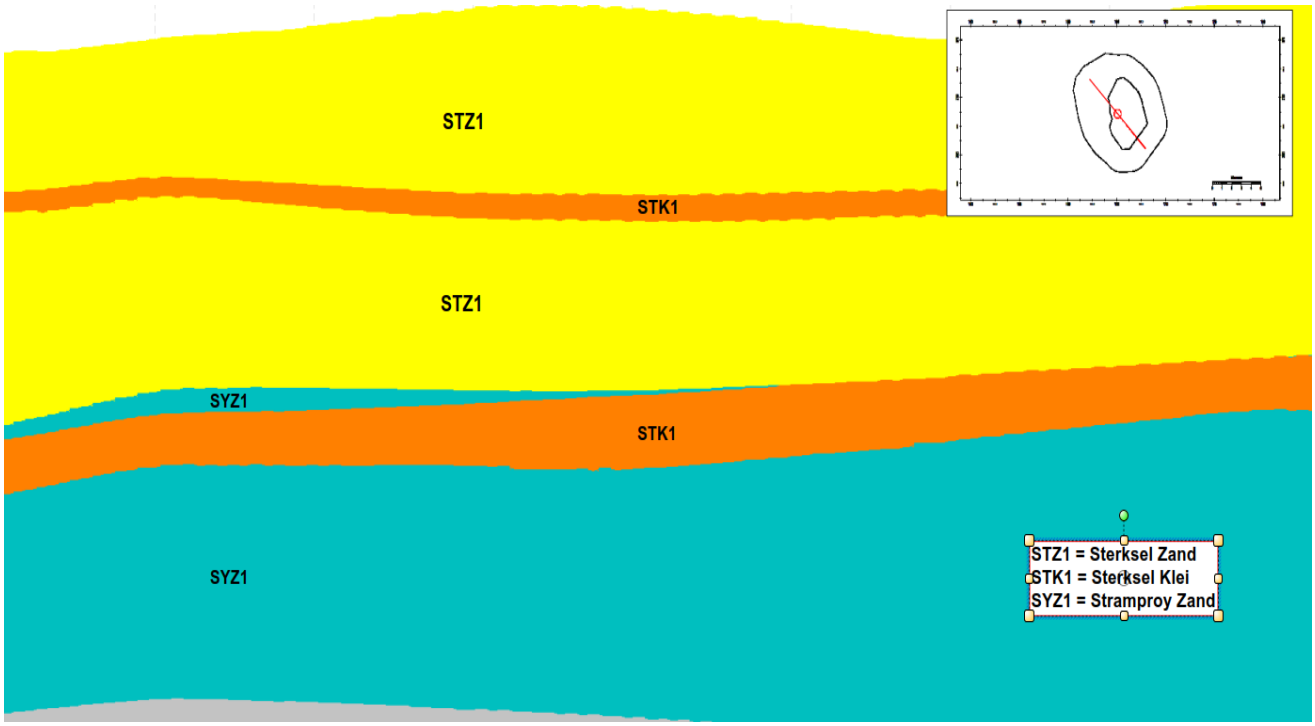


Fig 4.1. Cross-section view of the model structure with geological formations from REGIS 2.1.

This layer structure is the basis used to describe other hydrogeological parameters used in the groundwater model, such as hydraulic conductivity and layer resistance but also the distribution of abstraction wells among the different layers.

- **Boundary conditions**

  The cells at the model boundary are set as having a constant head boundary, similar to the Triwaco Flairs model. The cells outside are considered to have zero thickness and hence assigned a null value. The cells within the model boundary with an aquifer thickness are assigned to have varying heads with the model simulation.

- **SOBEK models**

  Two major contributing rivers (Dommel/Strijper Aa and Tongelreep) (Section 3.3.3) and their corresponding SOBEK river models are implemented in the iMOD model.

- **Drain data**

  Drainage pipes and ditches in the model area are implemented similar to Triwaco Flairs, using the Dataset for Secondary Watercourses developed by Deltares. These represent pathways by which water is removed from the iMOD model when the calculated head in a model layer exceeds the elevation of the drainage system. From the legend, it can be observed that the stage levels (m) of water in the drains follow the topographic elevation trend with high stage level in the south and decreasing to lower at the north, with the lowest being in the river valleys (Section 3.3.1.).

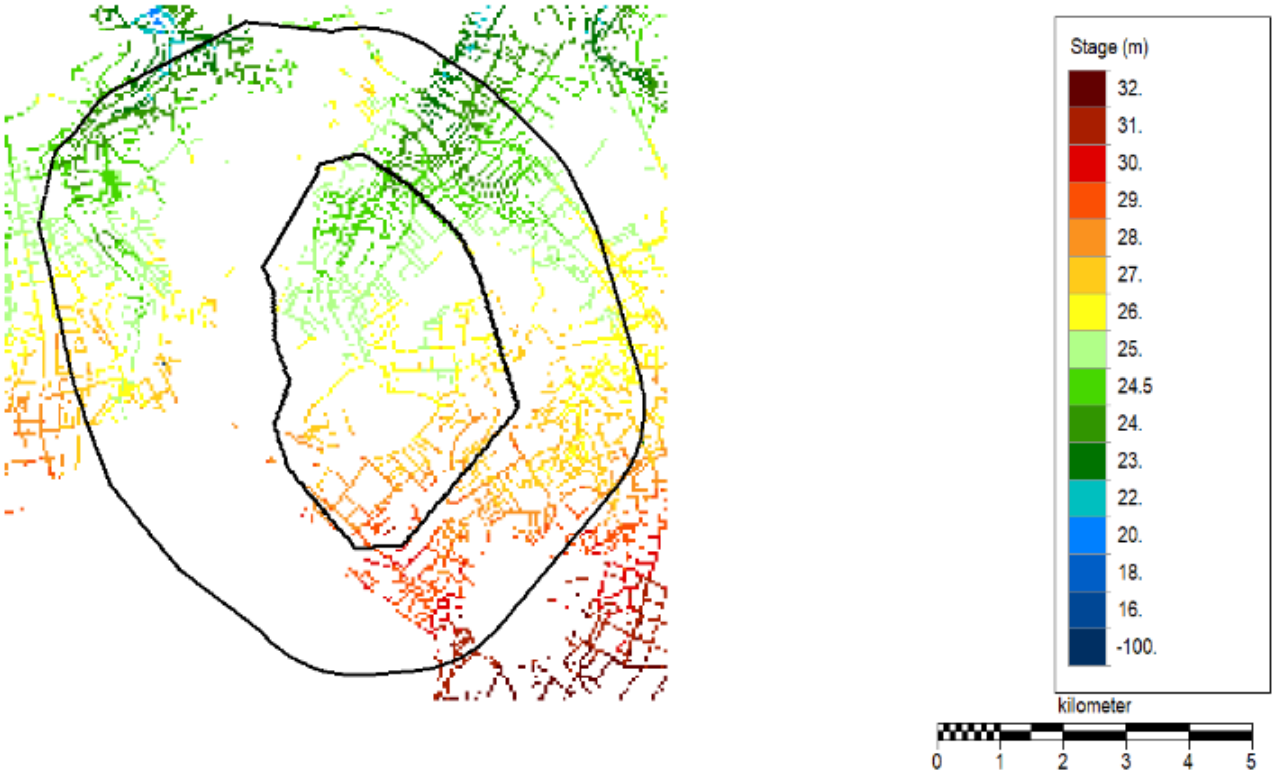


Fig. 4.2. Location of drains in the model area

- **Fault lines**

  The locations and conductance factors of fault lines (Section 3.3.2) used in the Triwaco Flairs are imported to iMOD for conversion to resistance factors to groundwater flow.

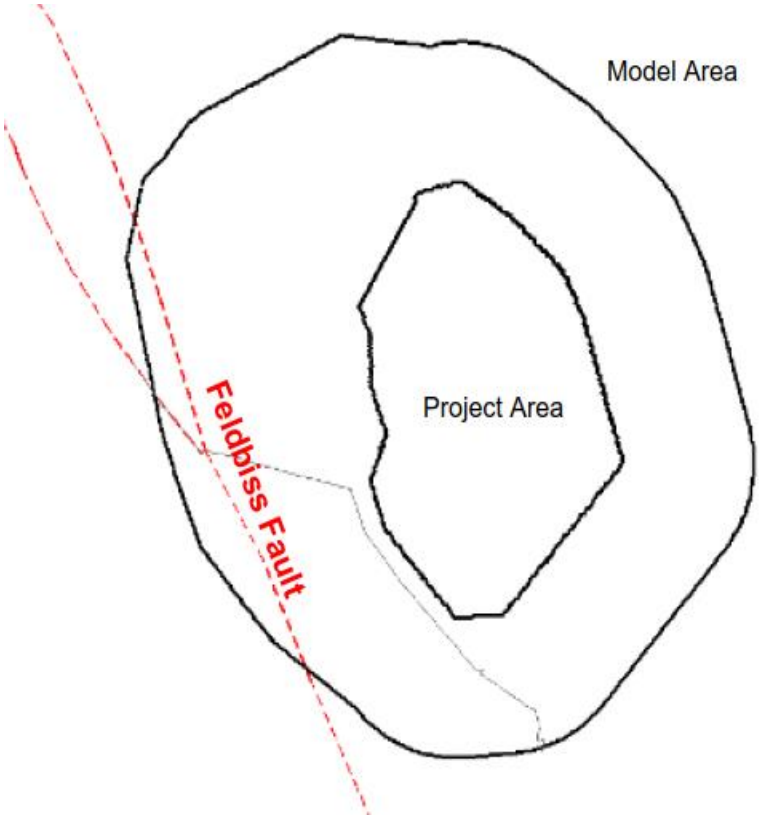


Fig. 4.3. Location of fault lines in the model area

- **Abstraction wells**

  Information such as location and rates of abstraction and pumping wells in the model area were collected from the Triwaco Flairs and incorporated into the iMOD model.

- **Starting head values**

  The initial set of head values for the start of iteration needs to be provided to iMOD, same as that in Triwaco. Starting head values are thus taken from the calculated heads of Regional Model of Brabant (2014), similar to Triwaco local model for the study area.

- **Observation well data**

  Information regarding location, filter and measured head of observation wells distributed across the model layers are carried over from the source used for Triwaco Flairs. A total of 149 observation well data are available for this study area, distributed across 19 model layers (Fig. 4.4, where N denotes number of wells in a particular layer).

Table 4.1. Distribution of observation wells among the model layers.

Layer	N
1	3
2	8
3	91
4	26
5	7
6	2
7	1
8	3
9	1
10	1
11	1
12	1
14	1
16	2
17	1
Total	**149**

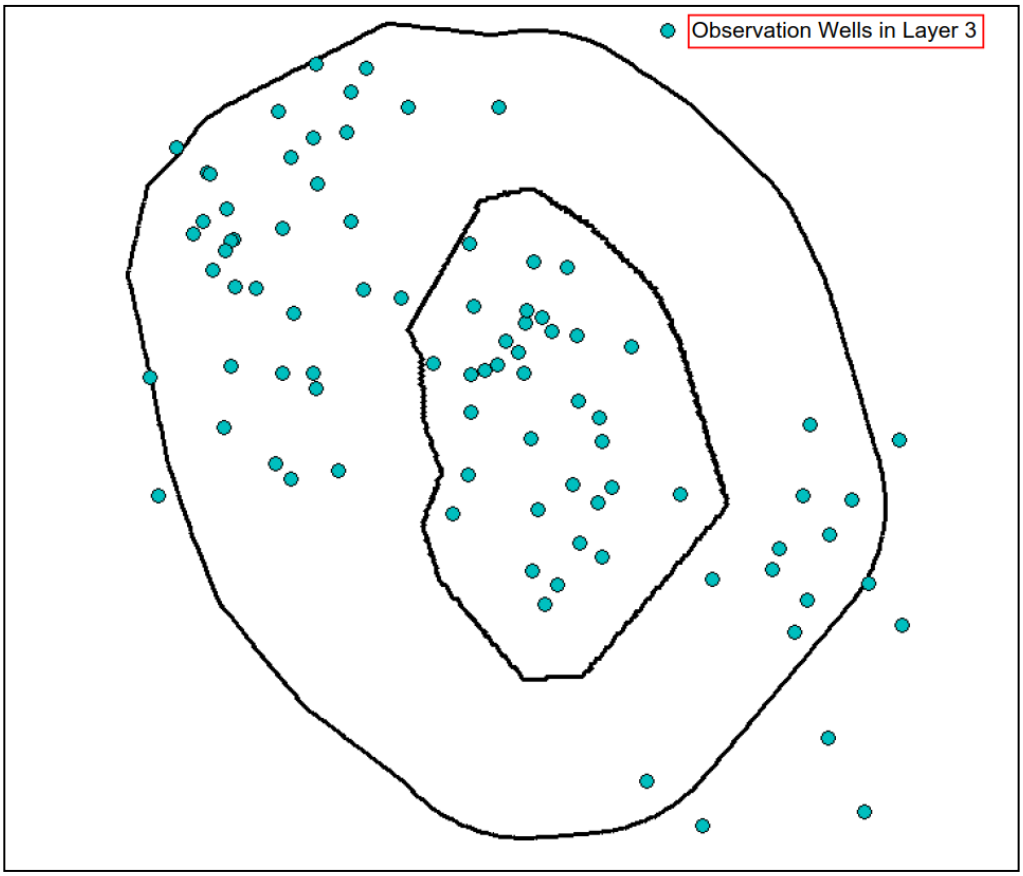


Fig. 4.4. Spatial distribution of observation wells in Model Layer 3

- **Surface level**

  The surface level for the model area was taken as the same as Triwaco Flairs, from AHN Database.

4.3. Calibration concepts

4.3.1. Zones

For our study, the concepts of zonation used to grouping of parameters has been implemented. Zonation can be based on any criteria set by the user and depends on the type of parameter being optimized. For example, the zone files for optimization of hydraulic conductivity values are based on geological distributions with basis in REGIS 2.1 geohydrological model. These geological distributions are spread across multiple layers so it is important to extend the zoning capability of iPEST to calculate fractions within each model layer. In this way, it is possible to optimize only a part of the permeability of a model layer, in accordance with a particular geological formation. The main advantage of this zoning approach is that this maintains the relationship between the geological formations and their contributions to the model layers intact. Similarly, zones can also defined for conductance of water courses or fracture resistances along fault lines.

In this study, we have calibrated the following parameters:

1) horizontal (KH) and vertical hydraulic conductivities (KV)
2) vertical anisotropy (VA) of layers
3) river conductance (RC) of two main rivers
4) drain conductance (DC)
5) fault line resistances (HF)

wherein KH, KV, VA, RC, DC, HF are the parameter codes that are recognized by iMOD for calibration.

Once zoning is done, parameters with similar defining characteristics are grouped together to make sure the optimization of these parameters happens in a linked way to generate similar changes in values during the calibration process.

In iPEST, each parameter needs to be associated with a zone file that defines its spatial distribution. If a parameter KH which is the hydraulic conductivity of Boxtel Zand (BXZ) exists in layer 8, then its spatial distribution needs to be represented by a zone file which has a unique number for that parameter alone. This zone number will have two parts. Each cell in that zone file will have this unique zone number as the integer part of its value and the decimal part will be the fraction of Boxtel Zand in that cell, say,

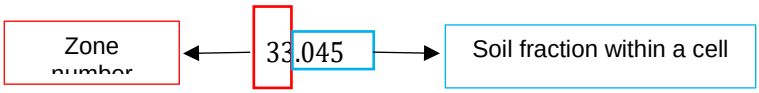


Fig. 4.5. Zonation in iPEST explained.

A Zone tool was created in C# to automatically create Zone files for our model area.

4.3.2. Objective function

Taking over the concepts of its underlying source algorithm of PEST, the objective function Φ_m (p) is defined in iPEST as depending on a parameter vector p with elements $p_i \rightarrow i = 1, N_p$ where N_p denotes the number of unknowns to be optimized (i.e. the amount of parameters). In general the objective function Φ_m (p) is the sum of squares sum of the individual errors notated as:

$$\Phi_m\,(p) \;=\; (y \;-\; \varphi(p))^T\; Q\; (y \;-\; \varphi(p))$$

where y are the measurements with elements $y_i \rightarrow i = 1,2,\ldots..N_h$;

N_h denotes the number of observations;

φ(p) are the computed head (obtained as the result of a model run by iPEST during calibration iteration) for the parameters defined in p, and

Q is the diagonal weight matrix with individual weight values along the diagonal, assigned to the observations (varying between 0 and 1 with 1 being the highest)

4.4. Calibration with iPEST

4.4.1. Sensitivity Analysis

A sensitivity analysis is the last step before an iPEST calibration. In highly parameterized hydrological problems, some parameters will inevitably be more sensitive than others. This relative difference in sensitivity can affect the progress of calibration when an optimization algorithm tries repeatedly to vary the insensitive parameter to no avail, resulting in wastage of computation time and resources. The relative sensitivity for a parameter is computed in iPEST by:

$$s_i = \frac{m^{-1} \sum_{j=1}^{m} w_j J_{ij}}{\sum s_i} . 100\%$$

where s_i is the sensitivity of the i^{th} parameter and is the product of the observational weight w_j times the Jacobian value (J) for that particular observation j in relation to the parameter i, divided by the total number of observation wells m.

The results give an indication of which parameter changes can have significant impact on the model results (i.e. sensitive parameters) and rule out the ones which do not have significant influence. The result of the iPEST sensitivity analysis is then used as a pre-selection method to include only the sensitive parameters in the first calibration run.

4.4.2. Calibration

After conducting the sensitivity analysis prior to calibration run, the sensitive parameters are now chosen to be calibrated with 20 iterations via iPEST within iMOD.

For an iPEST calibration run, a RUN-file needs to be generated containing a PST-block which has details on the initial multiplication factor for each parameter being calibrated, upper and lower bounds for the multiplication factor, step size as well as associated zone files for each parameter as shown in Fig 3-6.

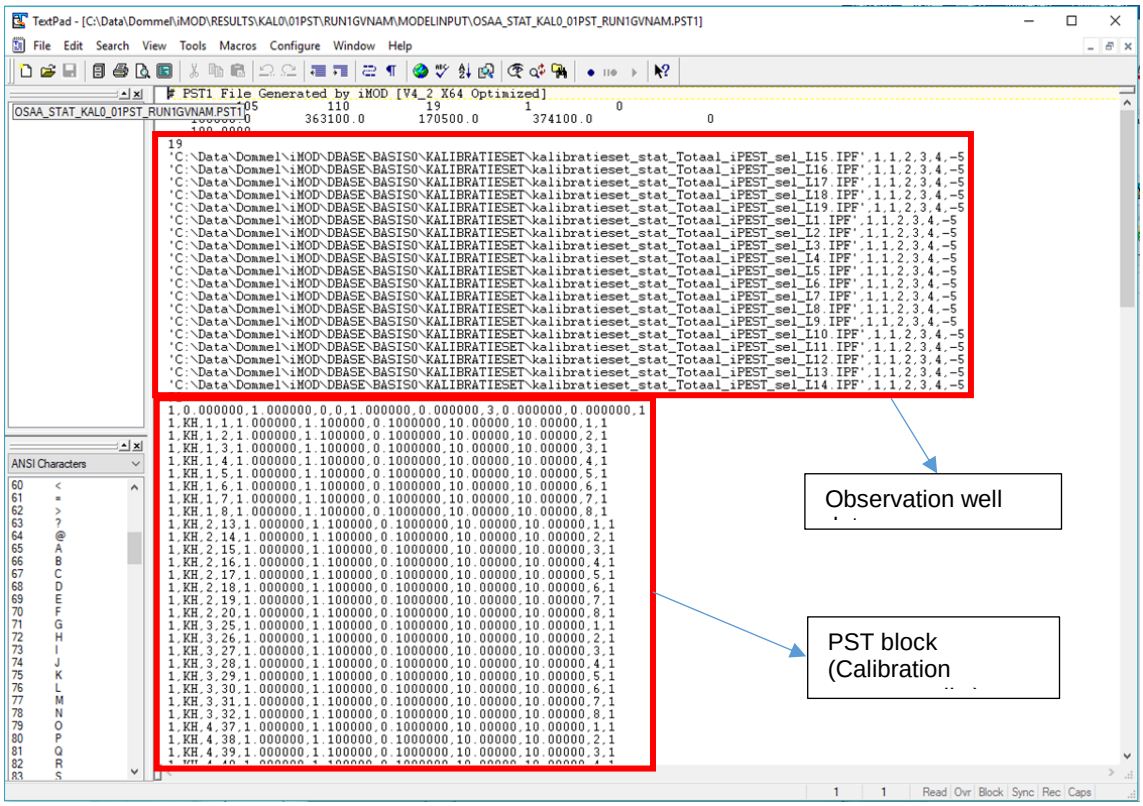


Fig 4.6. iPEST RUN-file with PST block

Given below is a representational line from the PST block above, showing a sample calibration parameter (in this case, the horizontal hydraulic conductivity KH).

```
1,KH,1,1,1.000000,1.100000,0.1000000,10.00000,10.00000,1,1
```

Each value in that line corresponds to a keyword that has already been defined by iMOD for running iPEST, such as :

PACT, PPARAM,PILS, PIZONE, PINI, PDELTA,PMIN,PMAX,PINCREASE,PIGROUP,PLOG

Where,

PACT → A number denoting if the parameter is to be changed in the current iPEST run (value of 0 or 1)

PPARAM → Type of parameter (KH, KV, etc.)

PILS → Layer number for the parameter

PIZONE → Zone number (integer part of the zone file) for the parameter

PINI → initial multiplication factor for the parameter

PDELTA → step size to be used (>1.0)

PMIN → Lower limit for the multiplication factor

PMAX → Upper limit for the multiplication factor

PINCREASE → Maximum increase in the value of the parameter

PIGROUP → Group number assigned to the parameter

PLOG → A number denoting if the parameter needs to be log-transformed (default is 1.0 implying log transformation)

During calibration, iPEST multiples the calibration parameters with an initial value equal to PINI initially. The value of PINI is then allowed to vary between the ranges of PMIN and PMAX as calibration proceeds. During each iteration, the multiplication factor used per parameter as well as the resulting value of the calibration parameter per layer are stored by iPEST. At the end of calibration run, this gives us a chance to see quickly access and check the resulting new values for the model parameters for interpretation.

4.5. Calibration with Genetic Algorithm

In this study, a specific number of 100 generations has been set as the termination criteria for GA to ensure the best possible solution was found during testing of GA.

In this study, a tool in Visual Studio was developed to use the concepts of GA in optimizing the iMOD groundwater model built for the study area. The same model input files used for running iPEST have been used to run the GA tool.

4.5.1. Adapted Genetic Algorithm

- **Concepts**

  In this study, the calibration parameters as shown in Fig. 4.6. are coded using a Real-valued representation in the GA tool developed. Each **column** (PINI, PMIN, PMAX, etc.) of values in the PST block denote a **Gene**.

  One set of the calibration parameters (as shown above, a whole PST block) represents a **Chromosome**.

  A collection of several such chromosomes will then represent a **Population**.

- **Genetic Operators (crossover, mutation, elitism)**

  - **Crossover**

  Single pointed crossover and Uniform Crossover have been tested in this study. A crossover probability of 0.85 has been used.

  - **Mutation**

  A fixed mutation rate of 0.02 has been used.

  - **Elitism**

  In order to make sure good solutions are not replaced, a certain percentage of the fittest chromosomes have been ensured to be carried forward during each generation. This has been implemented using an elitism percentage of 20% for each population size that has been tested in this study.

- **Fitness**

  - **Fitness function**

  The fitness function is a function that takes a candidate solution from a population of solutions and checks how ‘good’ or ‘fit’ the solution is for the problem at hand. In this study, the fitness function is chosen the same as the objective function for our iPEST calibration (Section 4.3.2).

- **Workflow**

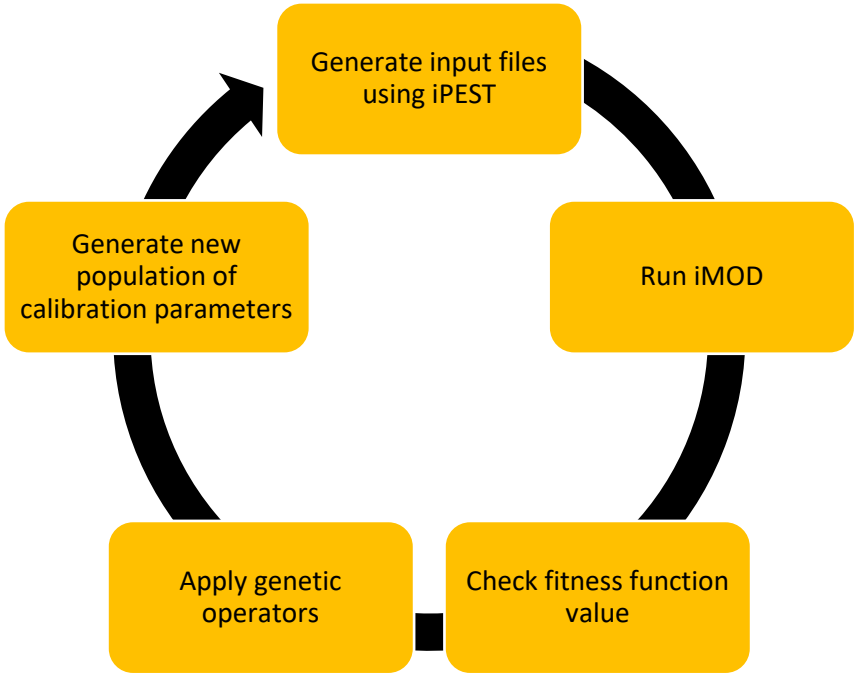


Fig. 4.7. Workflow diagram of GA tool.

The GA tool takes as input the same input files that are used to run iPEST. For each generation, iPEST package of iMOD is called as GA by itself cannot at this time in this study create the initial set of iMOD files for the calibration parameters. (Note: These are the files that contain updated values of the calibration parameters at the start, end and each stage of calibration and at this point in the research study, GA needs the support of iPEST to generate them) Now the GA is ready to calculate the fitness of the first population of candidate solutions and reproduce with the help of genetic operators that have been implemented in the tool. After each generation of offspring, the top 10% of the fittest solutions are chosen and from them, the fitness of the best individual (fittest chromosome) so far is calculated and carried forward to next generation as per the elitism percentage defined by the user.

This process is repeated until a defined termination criteria is reached, which in this study is creation of 100 generations. Once the process is terminated, a sensitivity analysis is triggered on the best chromosome among the 100 generations.

4.5.2. Sensitivity Analysis

Since a calibration run involves an initial sensitivity analysis to weed out the insensitive parameters, iPEST has the advantage of being able to focus on optimizing the more important parameters. However, the method used to calculate the sensitivity in iPEST is a rather internal and iterative one as mentioned in Section 4.4.1. For this study, the GA tool runs the sensitivity analysis at the end of the simulation. The sensitivity of each parameter is calculated as:

$$S_{P_i} = \frac{F_b - F_p}{F_b}$$

Where S is the sensitivity of the parameter P_i

F_b is the (basic) fitness of the best chromosome of the last generation

F_p is the fitness of the same chromosome when only the PINI value (genes) for the parameter P_i have been set to its basic value of 1.0

An extra step to calculate a comparable (to iPEST) sensitivity is also included as:

$$\log S = \frac{|\log(F_b) - \log(F_p)|}{|\log(1.0) - \log(P_i)|}$$

The user also has the flexibility to mention a cut-off sensitivity percentage so as to provide flexibility in choosing the parameters that are most significant to the model calibration.

4.5.3. Experiments with GA

- **Population size**

Population size is an important notion in Genetic algorithms. While traditional search methods go from one step to the other based on some transitional value, GAs rely on a population of candidate solutions. The size of the population is an important factor in getting good solutions using GA (Sastry, et.al 2005) in allowing to cover as much of the search space as possible. While a smaller population size can lead to premature convergence due to finding local optima, a larger size can lead to unnecessary wastage of expensive computational runtime. Thus the optimum population size for a particular problem needs to be tried and tested so as to choose an acceptable value. Given that Genetic Algorithms allow the user to have control over the population size is one of the advantages that can lead to better solutions than other random search methods.

In this study, population sizes of 10, 20 and 50 have been tested and their results analyzed and interpreted.

- **Crossover type**

While several types of cross-over methods already exist (Goldberg 1989), in this study the single-point crossover was chosen as the base method to be used. An improvement was also tested using UX (Uniform Crossover) on the best population size chosen in the previous step.

- **Crossover and Mutation rates**

Several combinations of crossover (p_c) and mutation (p_m) rates were also tested. An educated guess from Lin et. al. 2003 and Stanhope et.al. 2005 was used to assign the ranges of crossover and mutation rates to be used in this study. Based on this, the following combinations of crossover and mutation rates for the problem at hand were considered for running trials to find the combination that gives the lowest objective function value:

Combination Type	Crossover rate (p_c)	Mutation rate (p_m)
Original Combination (default)	0.85	0.2
Combi1	0.85	0.02
Combi2	0.5	0.2
Combi3	0.5	0.02

Table 4.2. Combinations of Crossover and Mutation rates used

4.6. Comparison strategy

For comparing the performance of the two algorithms considered in this study, namely iPEST and GA, a few performance factors have been decided upon as follows.

4.6.1. Success rate

For each of the population sizes (10, 20, 50) that was tested, the simulation was run for 10 times so as to assess the number of times GA found a better solution than iPEST, with a lower value of the objective function.

4.6.2. Computation time

The computation time for one simulation of GA was compared against that of iPEST.

4.6.3. GW head residuals

The resulting best parameters given by iPEST and GA were input into iMOD and run as two iMOD models so as to compare the calculated head output with the validated model of Triwaco. The residuals (difference between observed and calculated heads at measurement points) are then plotted along the model area and compared as well.

5. Results

5.1. iMOD Modeling

5.1.1. Model creation

For creating the iMOD model, it was decided to add the collected data in stages while starting with a layer model as the basis iMOD model.

Several iMOD models were made in stages with increasing complexity due to addition of more data per step. The iMOD models that were made and their construction is listed in Table 5.1.

Data / Model name	Layers	GW Recharge	River 1 (Dommel)	Abstraction/Pumping wells	Fault lines	River 2 (Tongelreep)	Drains
BASIS0	X	X					
BASIS1	X	X	X				
BASIS2	X	X	X	X			
BASIS3	X	X	X	X	X		
BASIS4	X	X	X	X	X	X	
BASIS5	X	X	X	X	X	X	X

Table 5.1. iMOD models and data addition

1. BASIS0
   The first model that was built had the layer structure from REGIS 2.1 model, together with precipitation data used by Triwaco Flairs as the groundwater recharge.

2. BASIS1
   The next model included addition of the Dommel river SOBEK model to BASIS0 iMOD model.

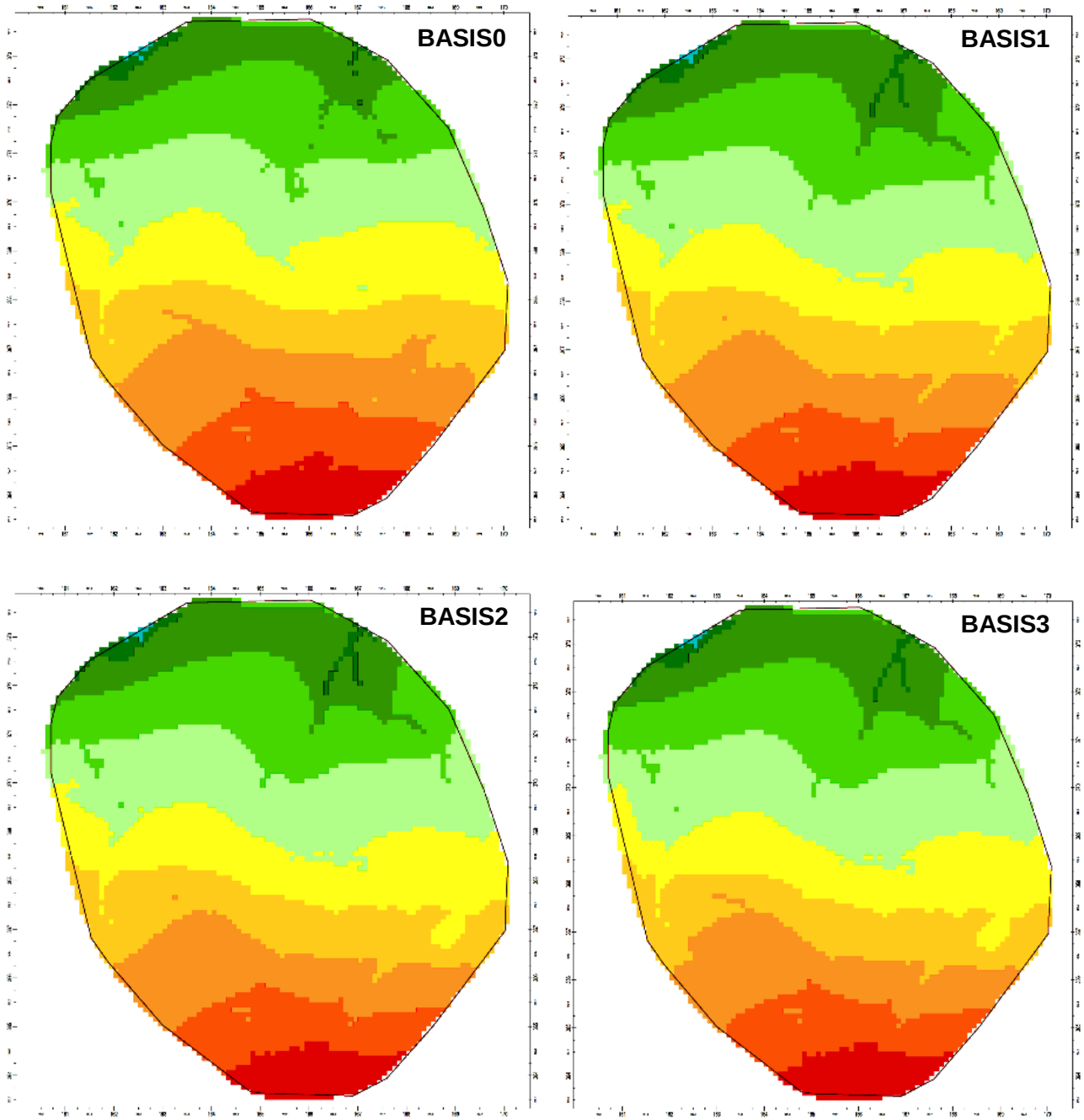
3. BASIS2
   Location and abstraction or pumping rates of wells in the study area was added to BASIS1 iMOD model.

4. BASIS3
   Geological faults and their resistances to groundwater flow were added to BASIS2. A proportional conversion of the fault conductance used in Triwaco was implemented in iMOD.

5. BASIS4
   A second river of Tongelreep was added to BASIS3 using a SOBEK model.

6. BASIS5

Location, Conductance and stage data of drains in the study area was added to BASIS4 using the Deltares dataset for secondary watercourses. This was the final iMOD model made in this study which produced comparable results to Triwaco Flairs model.



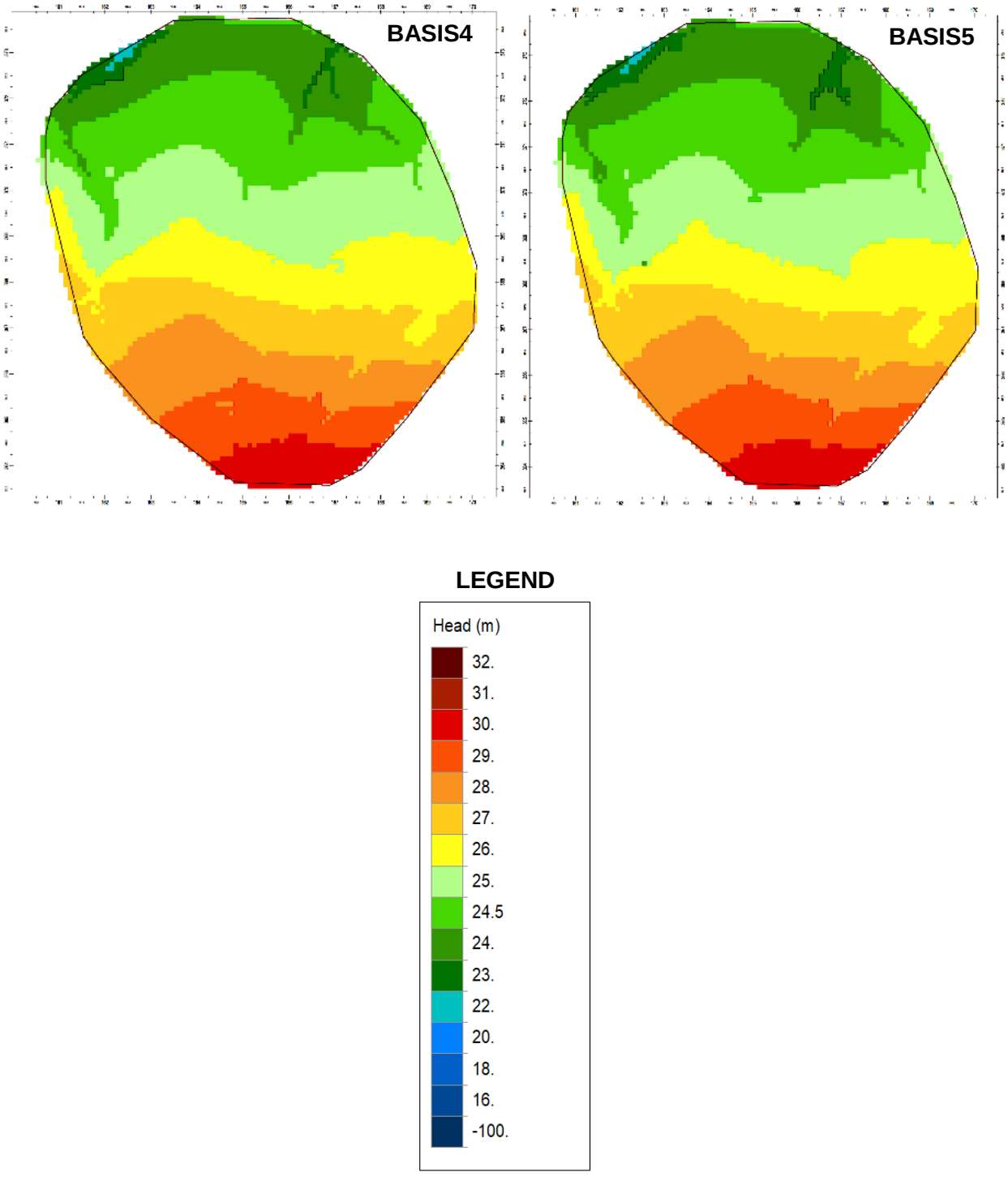
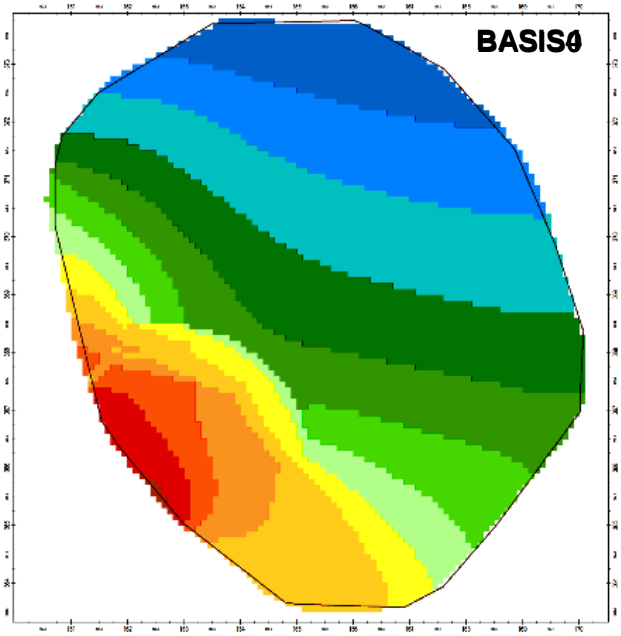
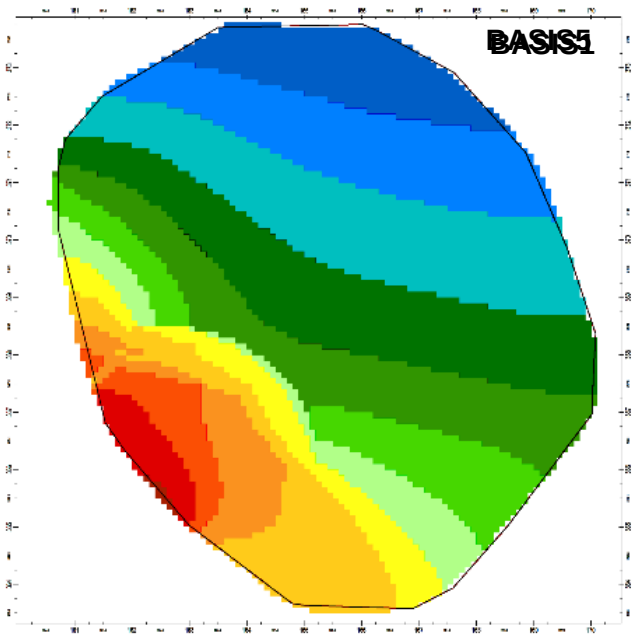
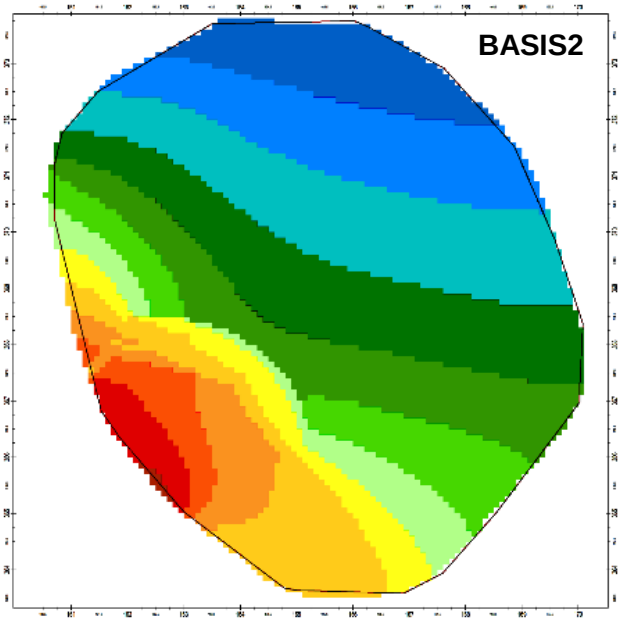
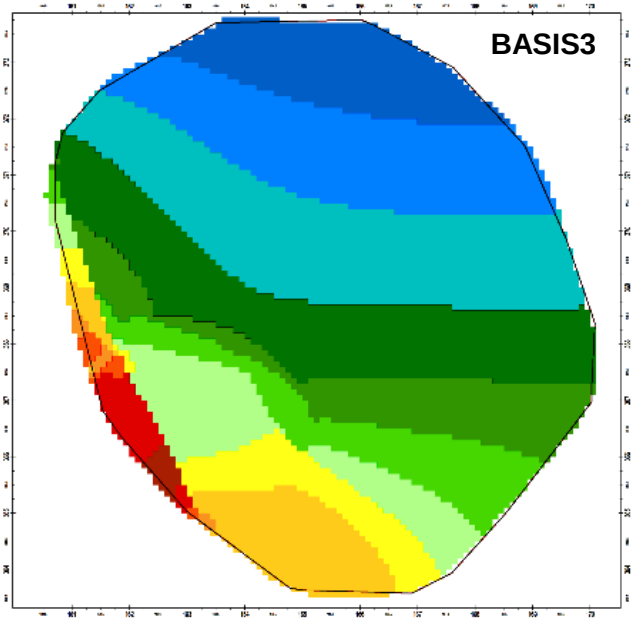


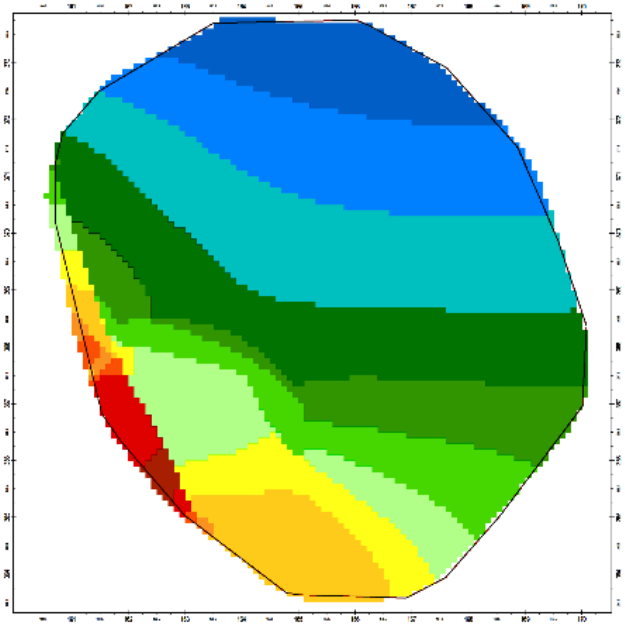
Fig. 5.1. Calculated heads in Layer 1 for BASIS0, BASIS1, BASIS2, BASIS3, BASIS4, BASIS5 models from left to right and top to bottom

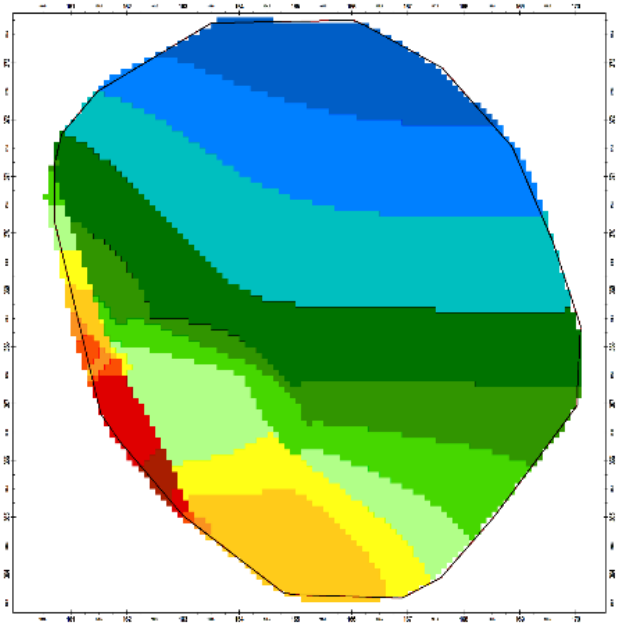

BASIS0


BASIS1


BASIS2


BASIS3





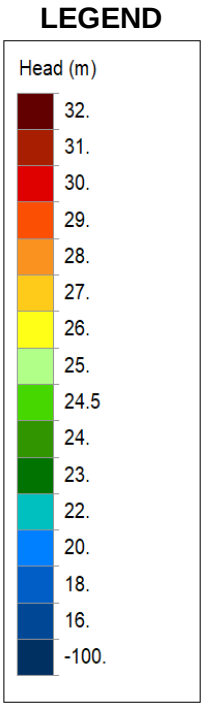


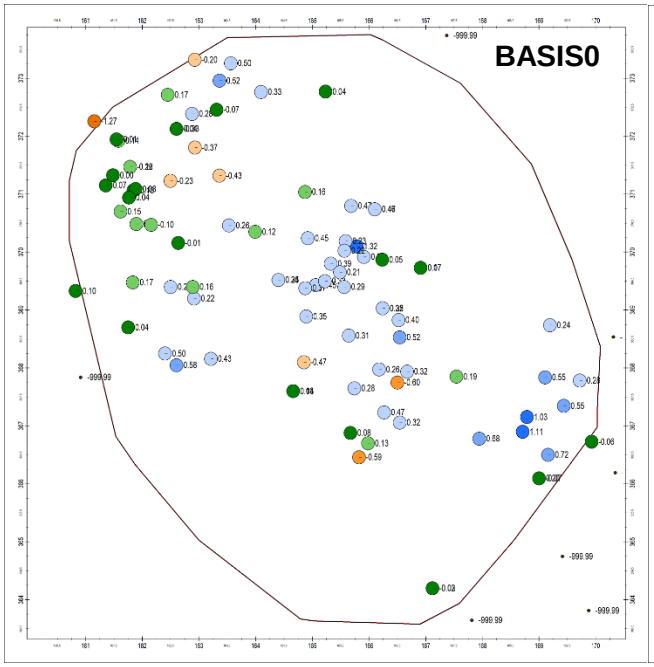
Fig. 5.2. Calculated heads in Layer 15 for BASIS0, BASIS1, BASIS2, BASIS3, BASIS4, BASIS5 models from left to right and top to bottom

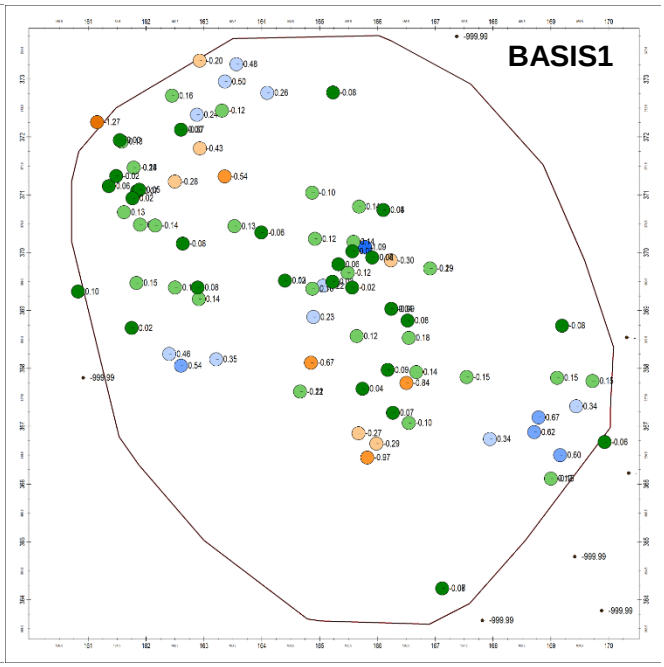
5.1.2. Comparison of iMOD models

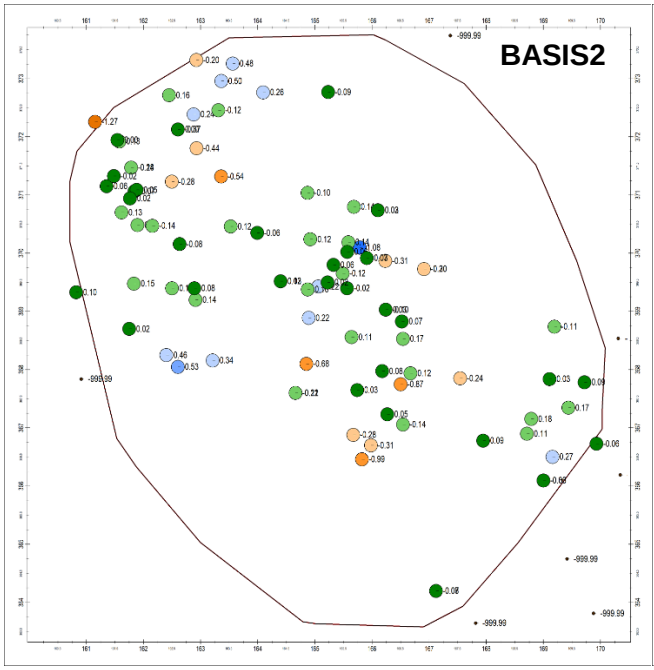
A comparison among all the iMOD models created in stages and listed in Table 5.1. was done. This was done using calculated heads (Fig. 5.1, 5.2) as well as with groundwater residuals between calculated heads and observed well measurements (Fig. 5.3 and Table 5.2.) at measurement point locations on the field. A representative layer no. 3 which has the most number of observation wells have been chosen to be analyzed.

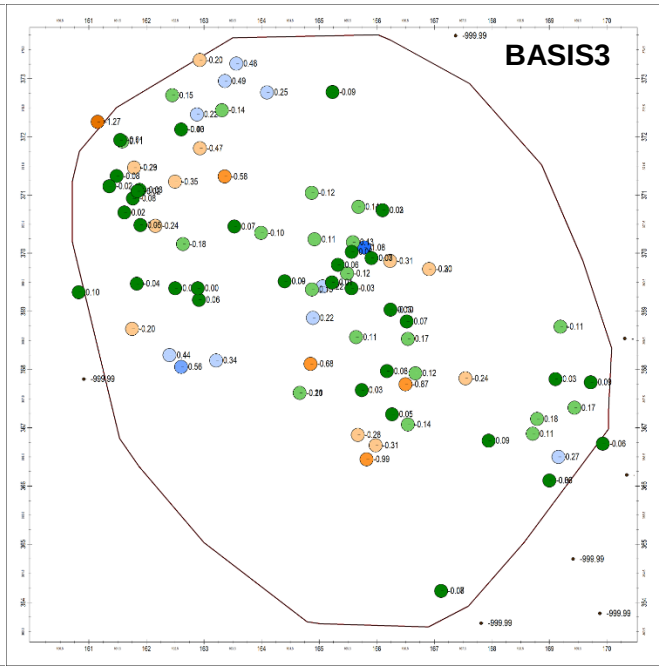
Figures 5.1. and 5.2. show the variation in calculated heads with addition of more data in two representative layers (topmost and 15th layers) of the iMOD model used in this study. From the fig 5.1., the groundwater head along the river valleys can be seen to be lowering (draining into the river) once river data is incorporated into the iMOD model. Fig 5.2. shows a very significant change in groundwater head at the location of fault lines once incorporated into the model, by restricting groundwater flow through it. The change in groundwater heads with addition of SOBEK river models illustrate the draining characteristic of the project area as is the field condition. The latter iMOD models further correspond to the assumption made in Section 3.3.2 on the impact of fault lines on groundwater flow.

As can be observed in Table 5.2, the total value (calculated as the mean of residuals in all layers) have reduced with further addition of data in stages in iMOD from an initial mean residual error value of 0.25 for BASIS0 model to a final of -0.16 for the BASIS5 model. The mean error (ME) is used to evaluate the overall state of the groundwater model from ‘too wet’ to ‘too dry’. Our final model BASIS5 with the maximum data added in comparison to the original Triwaco Flairs model for the study area is a relatively dry model as is visible in Fig. 5-3 and Table. 5-2.


BASIS0


BASIS1


BASIS2


BASIS3

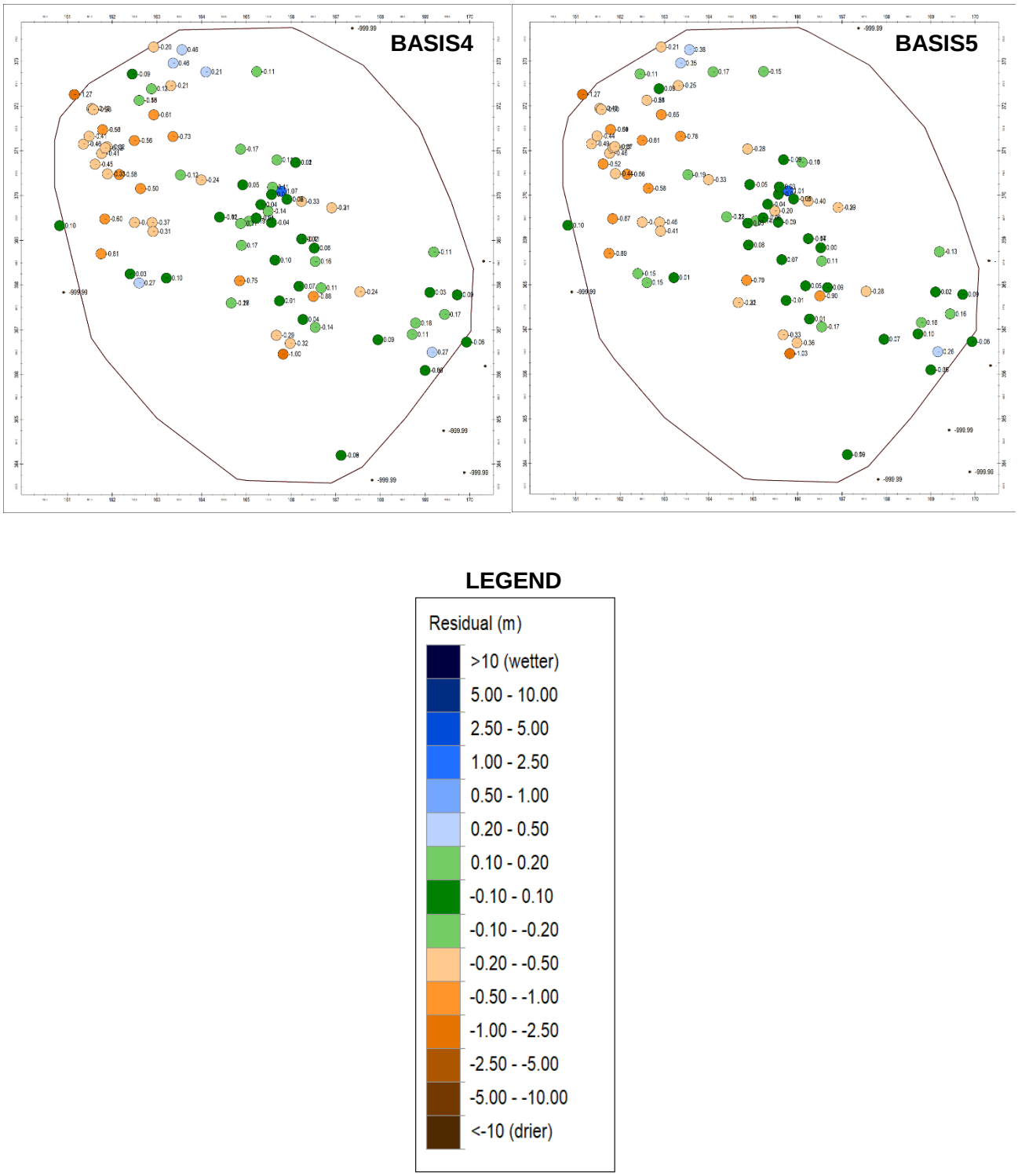


Fig. 5.3. Groundwater residuals in Layer 3 for iMOD models BASIS0, BASIS1, BASIS2, BASIS3, BASIS4, BASIS5 from left to right and top to bottom

			ME					
Set	Layer	N	BASIS0_B/	BASIS1_B/	BASIS2_B/	BASIS3_B/	BASIS4_B/	BASIS5_BAS
KALIBRATIESET	1	3	0.01	-0.20	-0.21	-0.22	-0.39	-0.41
	2	8	0.36	0.25	0.25	0.19	-0.13	-0.20
	3	91	0.19	0.02	-0.01	-0.04	-0.15	-0.21
	4	26	0.39	0.19	0.11	0.10	0.04	-0.01
	5	7	0.21	0.11	-0.06	-0.10	-0.24	-0.27
	6	2	0.20	0.02	-0.08	-0.09	-0.10	-0.13
	7	1	-0.01	-0.10	-0.10	-0.16	-0.28	-0.33
	8	3	0.42	0.27	0.20	0.17	0.12	0.09
	9	1	0.40	0.15	0.11	0.09	0.06	0.03
	10	1	0.79	0.54	0.50	0.48	0.45	0.42
	11	1	0.03	-0.22	-0.26	-0.28	-0.31	-0.34
	12	1	0.80	0.55	0.51	0.49	0.46	0.43
	14	1	0.06	-0.19	-0.22	-0.24	-0.27	-0.31
	16	2	0.56	0.56	0.56	0.15	0.15	0.15
	17	1	-0.06	-0.06	-0.06	-0.13	-0.13	-0.13
	Total	**149**	**0.25**	**0.08**	**0.03**	**0.00**	**-0.11**	**-0.16**

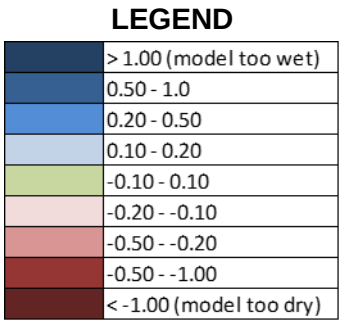


Table 5.2. Groundwater residuals (ME = Mean Error) of iMOD models

5.1.3. Issues encountered

- Elevation

It was observed that several dummy layers existed in the layer structure in the source REGIS model with thickness of 10 cm. However, the assignment of the layer thickness among these dummy layers were mixed up in the Flanders (Belgium) region. For example, in the layer no. 6 in the Belgian portion of the study area, the bottom elevation was assigned to be above the top elevation of the same layer.

This was not a problem in the original Triwaco Flairs model as Flairs does not use the elevation data for any calculations of hydraulic conductivity. In contrast, the elevation data and subsequent thickness of layers is an important parameter in iMOD due to its usage in calculation of hydraulic conductivity values.

To bypass this error in the layer structure, a re-assignment of the elevation of these dummy layers was done with the same earlier thickness of 10cm each, for the layers 6-14. All iMOD models were run based on this corrected elevation data.

- iPEST Zones

The REGIS 2.1 geohydrological model data was used to create zones based on the distribution of geological formations across the model layers. It was observed that the same geological formation was split along the depth of several layers and sometimes sufficient geological information was not available to find all the formations that contributed to the thickness of one model layer. As a result, the soil fractions in the zone files for a layer, when summed up, did not always add up to the value of 1.0, as expected.

To bypass this error, an iPESTzone tool was developed in Visual Studio C# which intakes the REGIS formation data and iMOD model layer elevations and notes down the missing soil thickness fractions per layer, should it exist, and considers them as separate zones. These are also assigned to parameters during calibration so that their sensitivity can be assessed and inferences drawn whether the discretization of the model for the study area is missing an aquifer or aquitard which contributes significantly to the model results.

5.2. Final iMOD model v/s Triwaco Flairs model

The iMOD model BASIS5 was input the maximum amount of data available and from a perusal of the results in Sections 5.1.3 was agreed upon by the Waterschap De Dommel as being closest to the original Triwaco Flairs model in terms of amount of hydrological data input. This was then confirmed by a comparison between the outputs of the iMOD BASIS5 model and the Triwaco Flairs model, with a comparison between the calculated heads as well as residuals as shown below in Fig. 5.4. to 5.6. BASIS5 was chosen as the iMOD model to proceed with calibration as the residuals were low in total in the model area as well as the project area of Oude Strijper Aa within the modelled area shown had residuals in an acceptable range of -0.20m to 0.20m. It also very closely matched the head plots of Triwaco Flairs model.

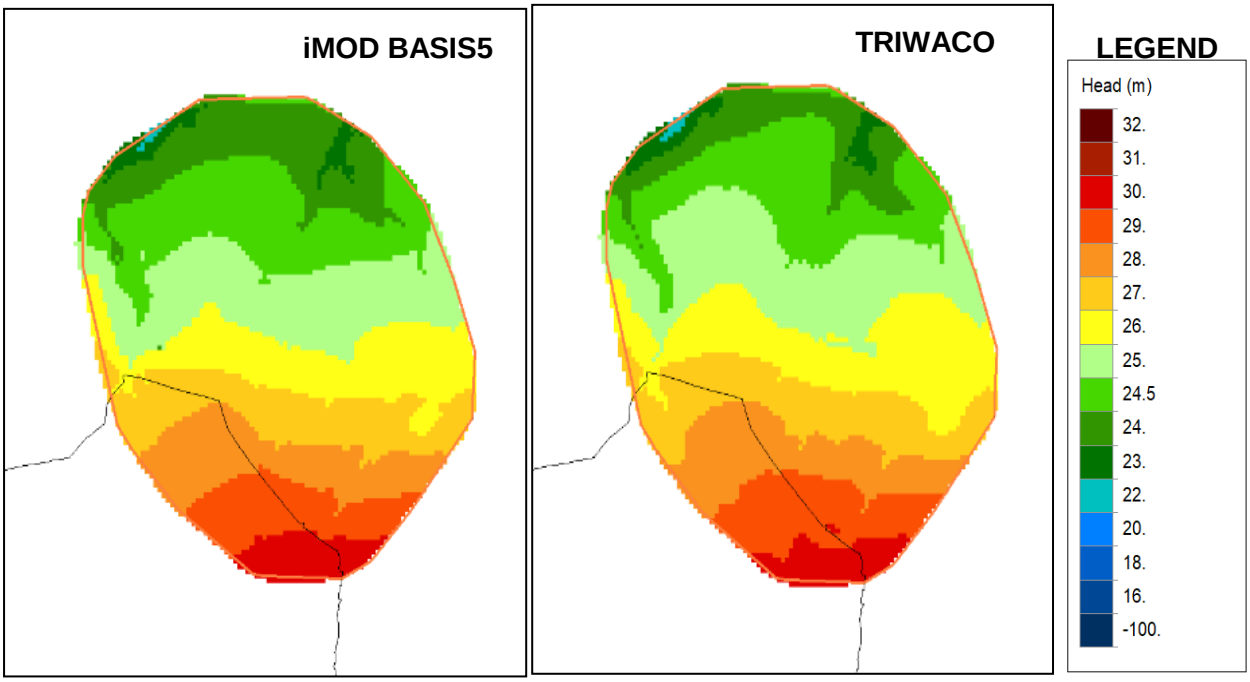


Fig. 5.4. Comparison of calculated heads in Layer 1 (iMOD BASIS5 v/s Triwaco Flairs)

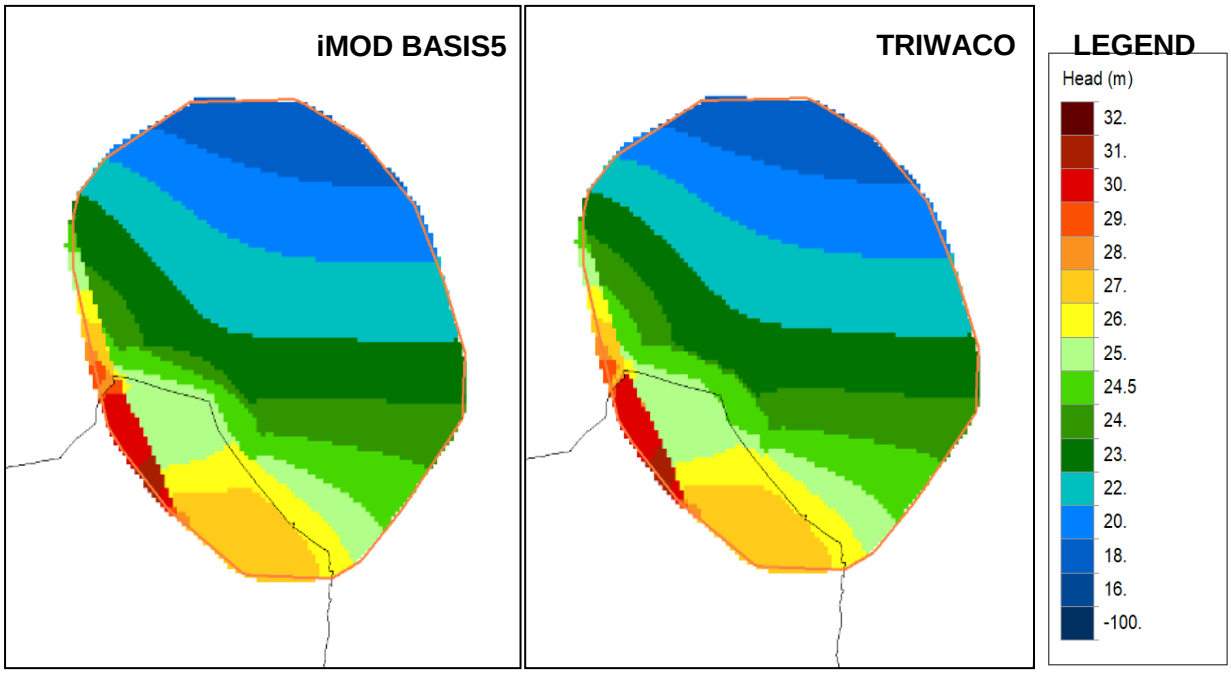


Fig. 5.5. Comparison of calculated heads in Layer 15 (iMOD BASIS5 v/s Triwaco Flairs)

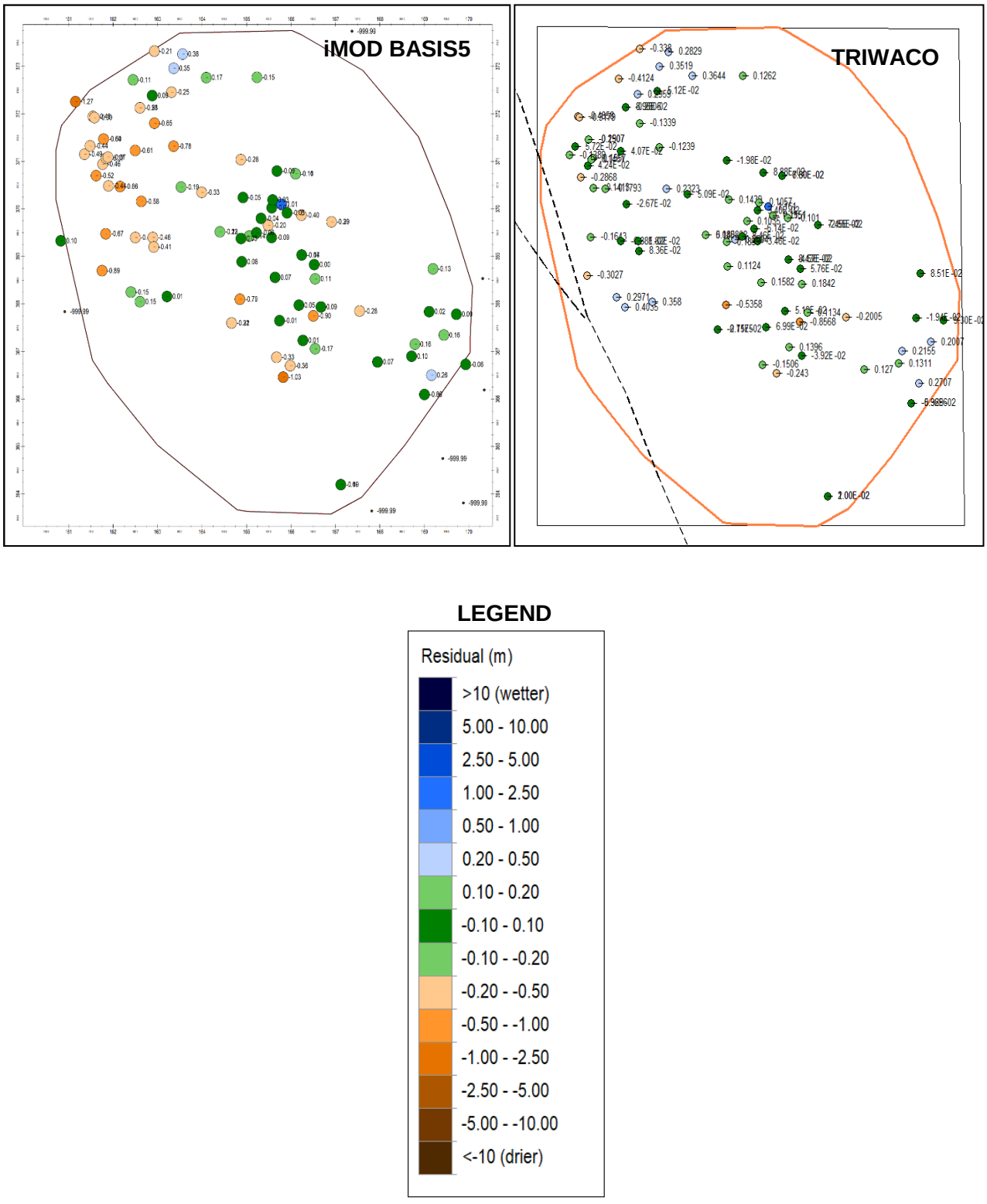


Fig. 5.6. Comparison of groundwater residuals in Layer 3 (iMOD BASIS5 v/s Triwaco Flairs)

5.3. Calibration using iPEST

5.3.1. Zones per parameter type

A total of 401 parameters (Table 5.3.) were chosen for calibration estimated using combinations of zones based on geological formations and their presence per layer. Similar geological formations were grouped together since their values need to change by the same multiplication factor during the calibration. This is done so as to ensure consistent optimization in value for the same formation in different layers.

Parameter Notation	**Parameter type**	**No. of zones**	**No. of groups**	**Source**
KH	Horizontal conductivity	124	36	Geological formations
KV	Vertical conductivity	83	30	Geological formations
VA	Vertical anisotropy	124	36	Geological formations
RC	River conductance	2	2	2 rivers grouped separately
DC	Drain conductance	14	14	Bodemkaart van Nederland
HF	Fault resistance	1	3	Fault lines Feldbiss
	Total	**401**	**121**	

Table 5.3. Calibration Parameters and iPEST zones

Data	**Source**
Geological Formations	REGIS II hydrogeological model, DINOloket.nl
Bodemkaart van Nederland	Geoplaza.nl (Vrije Universiteit, Amsterdam)
Surface level	Actueel Hoogtebestand Nederland (AHN.nl)
Geological Faults	Waterschap de Dommel, Boxtel

Table 5.4. Data source used for zonation

This was further reduced as a result of sensitivity analysis prior to the calibration run. Several models were made in stages and sensitivity analysis for each chose different number of independent freely adjustable parameters due to addition of more data per model (Table 5.5.).

Model name	**KAL0**	**KAL1**	**KAL2**

No. of sensitive parameters	17	16	17

Table 5.5. Number of independent calibration parameters

The nomenclature for the parameters have been set by iPEST as explained in an example shown in Table 5.6. and will be used in this report to be consistent.

Calibration Parameter name	Parameter Type	Layer No.	Zone No.	Group No.
KV003081-021	**KV**	**003**	**081**	**021**

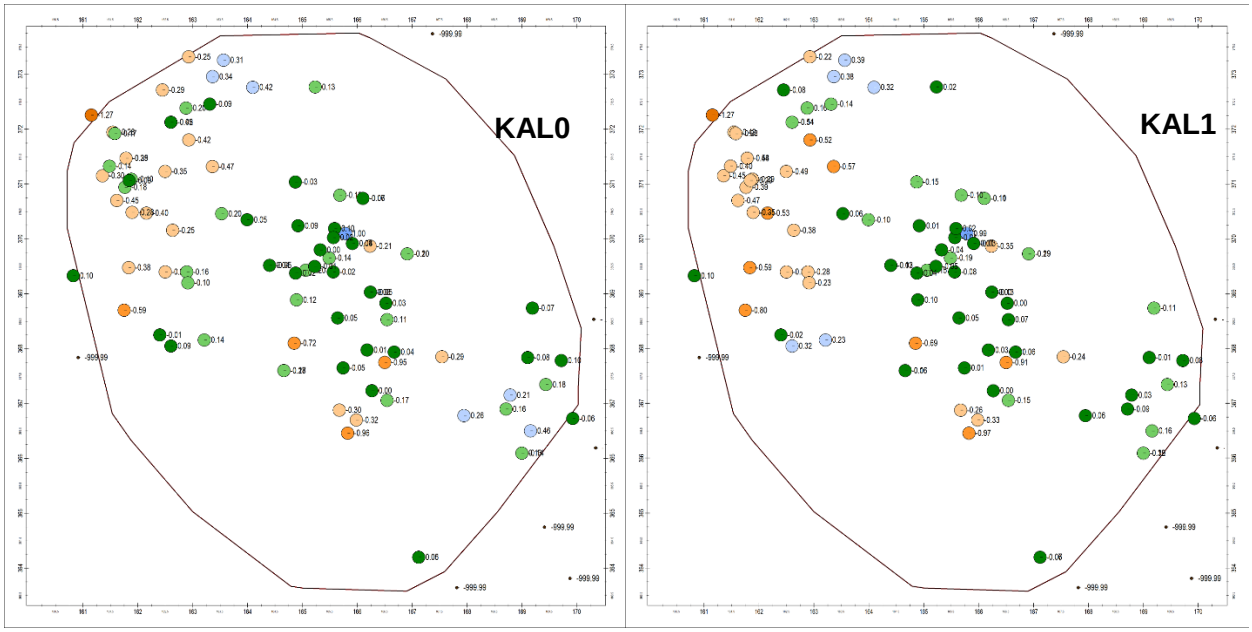
Table 5.6. Example of nomenclature for iPEST calibration parameters

5.3.2. iPEST calibration models

Calibration using iPEST was done on the final iMOD Model BASIS5 (decided as per Section 5.2.) Modifications were done in stages for the calibration model as well as listed below in Table 5.7.

Model name	Modifications
KAL0	No weights to measurements, calibration of KH, KV (Table 5-3)
KAL1	KAL0 with weighted measurements in project area Oude Strijper Aa.
KAL2	KAL1 model with calibration parameters VA, RC, DC, HFB added (Table 5-3)

Table 5.7. Calibration models used in iMOD



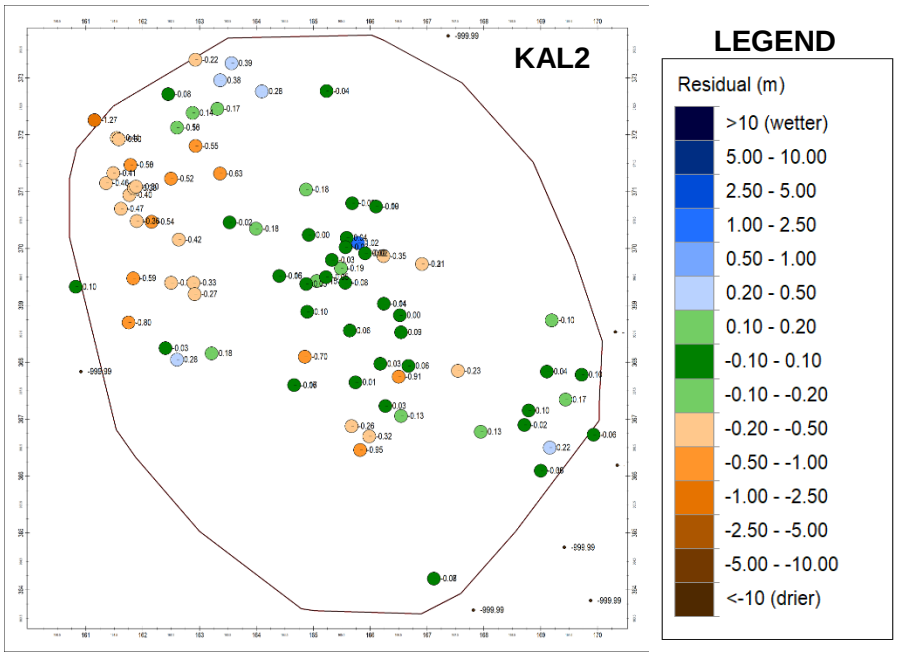


Fig. 5.7. Groundwater residuals in Layer 3 for KAL0, KAL1, KAL2 iMOD models

			WME			
Set	Layer	N	BASIS5_BAS	KAL0_01PST_I	KAL1_01PST_I	KAL2_01PST_I
Totaal	1	3	-0.41	-0.33	-0.39	-0.39
	2	8	-0.20	-0.09	-0.17	-0.17
	3	91	-0.21	-0.09	-0.16	-0.16
	4	26	-0.01	0.05	0.01	0.02
	5	7	-0.27	-0.17	-0.28	-0.26
	6	2	-0.13	-0.08	-0.09	-0.08
	7	1	-0.33	-0.20	-0.21	-0.26
	8	3	0.09	0.16	0.15	0.14
	9	1	0.03	0.05	0.11	0.10
	10	1	0.42	0.43	0.49	0.48
	11	1	-0.34	-0.33	-0.26	-0.28
	12	1	0.43	0.44	0.51	0.49
	14	1	-0.31	-0.30	-0.23	-0.24
	16	2	0.15	0.15	0.15	0.15
	17	1	-0.13	-0.13	-0.13	-0.13
	Total	**149**	**-0.16**	**-0.06**	**-0.12**	**-0.12**

LEGEND
> 1.00 (model too wet)
0.50 - 1.0
0.20 - 0.50
0.10 - 0.20
-0.10 - 0.10
-0.20 - -0.10
-0.50 - -0.20
-0.50 - -1.00
< -1.00 (model too dry)

Table. 5.8. Groundwater residuals for KAL0, KAL1, KAL2 iMOD models

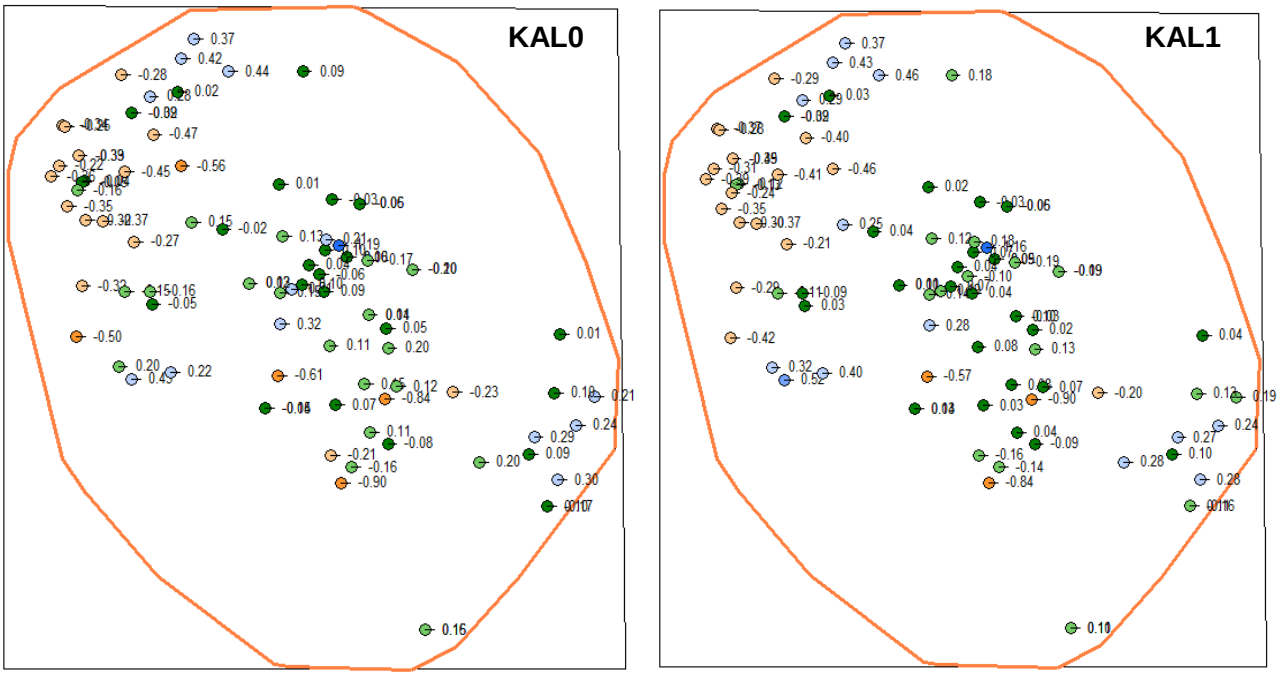
From observing the residuals in Fig. 5.7. and Table 5.8. for the iPEST calibrated models, a large reduction in residuals was seen in most layers for the first unweighted calibration model KAL0. But by giving weights to the observation well measurements in the project area of Oude Strijper Aa (Fig. 3.7.), the residuals were again analyzed for the interest area of the main project. The model (KAL1) now becomes drier when the mean residual error is considered. Observation wells in Layers 10-14 have consistently high values of residuals. This could be due to the availability of only single wells that fall in the weighted project area in each of those layers, thereby leading to them being given undue significance in the calibration process. Further modifications (KAL2) by including more calibration parameters does not change the mean residual of the model significantly.

5.4. Calibration using GA

Genetic algorithm was run for KAL0, KAL1, KAL2 iMOD models initially created for iPEST calibration. The same input files used for iPEST calibration were fed into the Genetic Algorithm tool to ensure consistency in starting point location in the search space. Genetic Algorithm then creates subsequent candidate solutions spread across the search space by implementing genetic operations such as crossover, mutation, etc. For the calibration runs in this study using the GA tool, the following default values were used to define the GA:

GA Tool Setting	Value
Population size	20
Maximum number of generations	100
Crossover probability	0.85
Mutation probability	0.2
Elitism %	20

Table 5.9 Genetic Algorithm default settings used in this study.



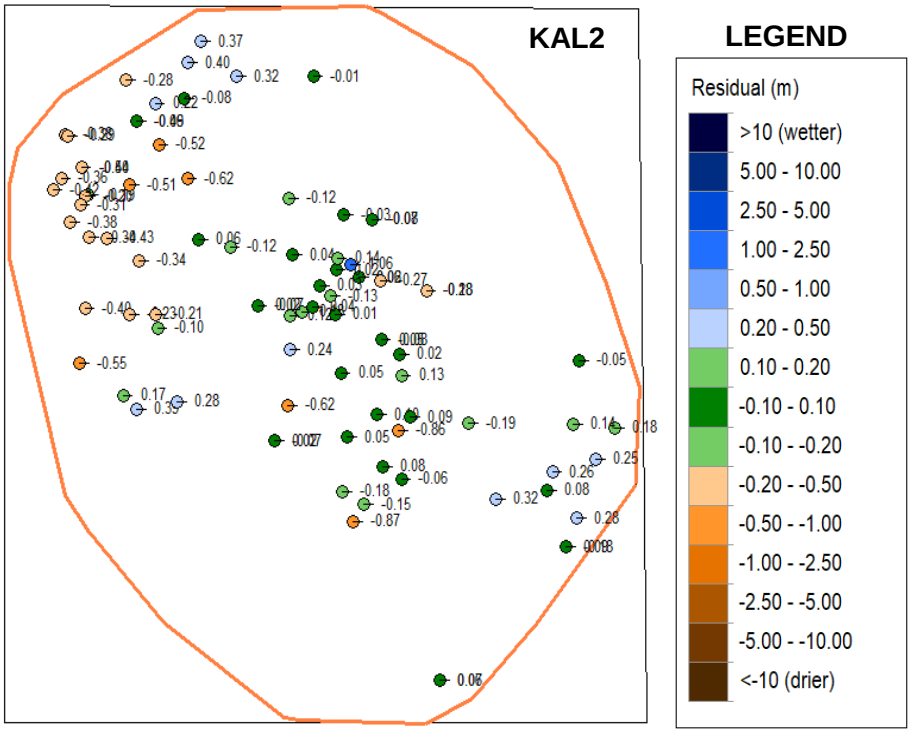


Fig. 5-8. Groundwater residuals for KAL0, KAL1, KAL2 in Layer 3 by GA

			ME								
Set	Layer	N	Generation1	Generation2	Generation3	Generation4	Generation5	Generation10	Generation20	Generation50	Generation100
Totaal_iPEST	1	3	-0.40	-0.40	-0.40	-0.35	-0.36	-0.31	-0.31	-0.34	-0.34
	2	8	-0.15	-0.15	-0.15	-0.14	-0.11	-0.07	-0.07	-0.01	-0.01
	3	87	-0.15	-0.15	-0.15	-0.13	-0.14	-0.09	-0.09	-0.02	-0.02
	4	25	0.02	0.02	0.02	0.06	0.03	0.09	0.09	0.10	0.10
	5	7	-0.26	-0.26	-0.26	-0.17	-0.25	-0.15	-0.15	-0.17	-0.17
	6	2	-0.05	-0.05	-0.05	0.02	-0.04	0.05	0.05	-0.10	-0.10
	7	1	-0.23	-0.23	-0.23	-0.22	-0.19	-0.13	-0.13	-0.13	-0.13
	8	3	0.16	0.16	0.16	0.24	0.18	0.23	0.23	0.16	0.16
	9	1	0.09	0.09	0.09	0.23	0.13	0.27	0.27	-0.06	-0.06
	10	1	0.52	0.52	0.52	0.64	0.55	0.68	0.68	0.38	0.38
	11	1	-0.28	-0.28	-0.28	-0.15	-0.24	-0.11	-0.11	-0.42	-0.42
	12	1	0.54	0.54	0.54	0.67	0.58	0.71	0.71	0.40	0.40
	14	1	-0.20	-0.20	-0.20	-0.06	-0.16	-0.03	-0.03	-0.35	-0.35
	16	2	0.16	0.16	0.16	0.16	0.15	0.19	0.19	0.16	0.16
	17	1	-0.06	-0.02	-0.02	-0.02	-0.07	-0.11	-0.11	-0.02	-0.02
	Total	**144**	**-0.11**	**-0.11**	**-0.11**	**-0.08**	**-0.09**	**-0.04**	**-0.04**	**-0.01**	**-0.01**

LEGEND	
	> 1.00 (model too wet)
	0.50 - 1.0
	0.20 - 0.50
	0.10 - 0.20
	-0.10 - 0.10
	-0.20 - -0.10
	-0.50 - -0.20
	-0.50 - -1.00
	< -1.00 (model too dry)

Table 5.10. Groundwater residuals progression across Generations for KAL2 by GA

From Fig. 5.8. it can be seen that all of the residuals in models KAL0, KAL1 and KAL2 are more or less similar to residuals of iPEST calibration shown in Fig. 5.7. A closer comparison of the residuals in the project area of Oude Strijper Aa (Fig. 5.8) will show that the residuals at more wells now fall in the acceptable range of -0.20m to 0.20m. Further scrutiny of the residuals change during calibration for calibration model KAL2 (Table 5.10.) will show that the residuals have decreased drastically with each generation by using Genetic Algorithm with a final mean residual of -0.01m at the end of calibration which is negligibly small.

5.5. Comparison GA vs iPEST calibration output

5.5.1. Computation time

Algorithm	Computation Time
iPEST	20 mins
GA	8 hours

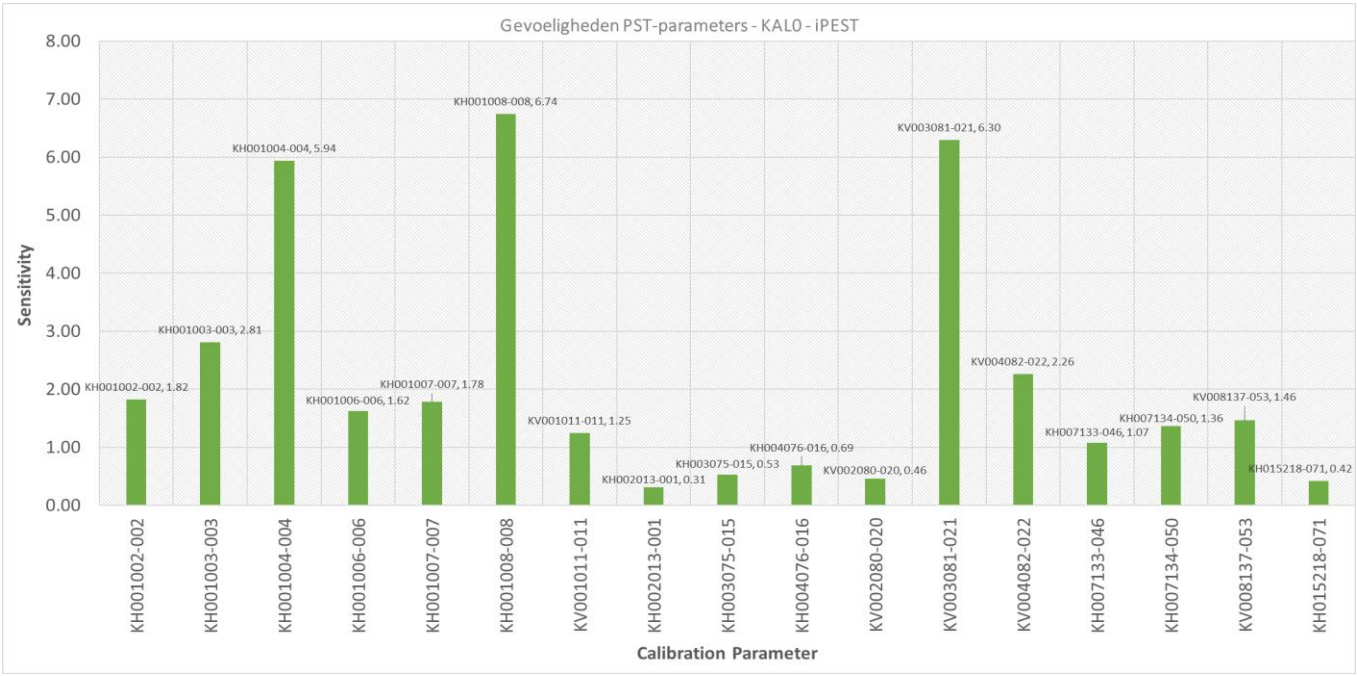
Table. 5.11. Comparison of computation time iPEST vs GA

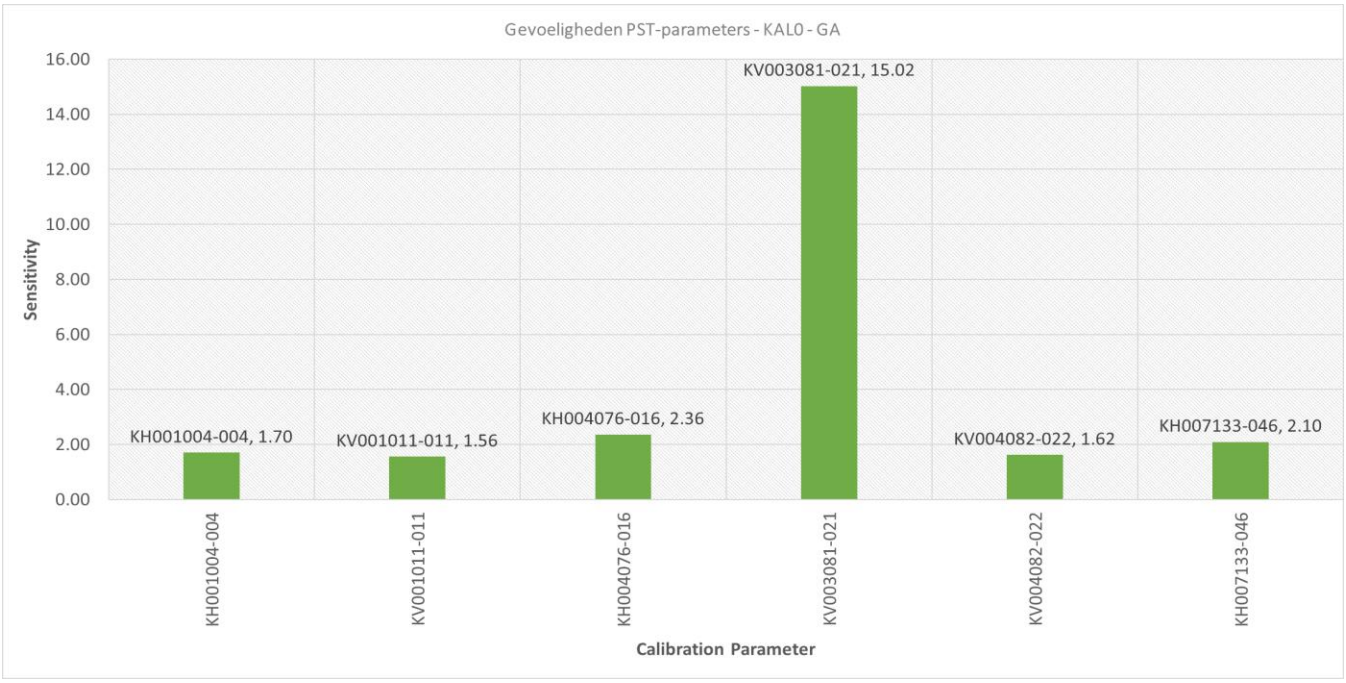
For the model KAL2, Genetic Algorithm tool used a computation time of 8 hours while iPEST took 20 mins to finish calibration (Table. 5.11.). Model KAL2 has been chosen as the final calibration model in this study.

5.5.2. Sensitivity Analysis

Prior to calibration using iPEST, the sensitivity analysis inbuilt in iPEST module was run. This weeded out the parameters that to which the value of the objective function was relatively insensitive according to iPEST. This was proceeded by deactivating those insensitive parameters in the PST-block of the RUN-file for iPEST and calibration was done using only the active (sensitive) calibration parameters. On the other hand, GA calibration is done by keeping all the parameters active. Since GA is a random-search algorithm, it was expected to show its strength by finding similar sensitive parameters as iPEST without a similar initial sensitivity analysis. After the GA calibration module was run, a sensitivity analysis was done on the parameter calibration changes during the run. It was then observed that GA chose the same parameters already chosen by iPEST. The log-senstivity test (Section 4.5.2) also chose the same parameters as chosen by the GA sensitivity calculation. The Triwaco Flairs model used by the Waterschap De Dommel contains a correction factor for a missing layer during calibration of their model. The ability of GA in choosing the right sensitive parameters is shown by the fact that it has consistently selected the parameter corresponding to this missing layer (KV003081-003 or Missing Aquitard Fraction in Layer 3) as an important parameter.

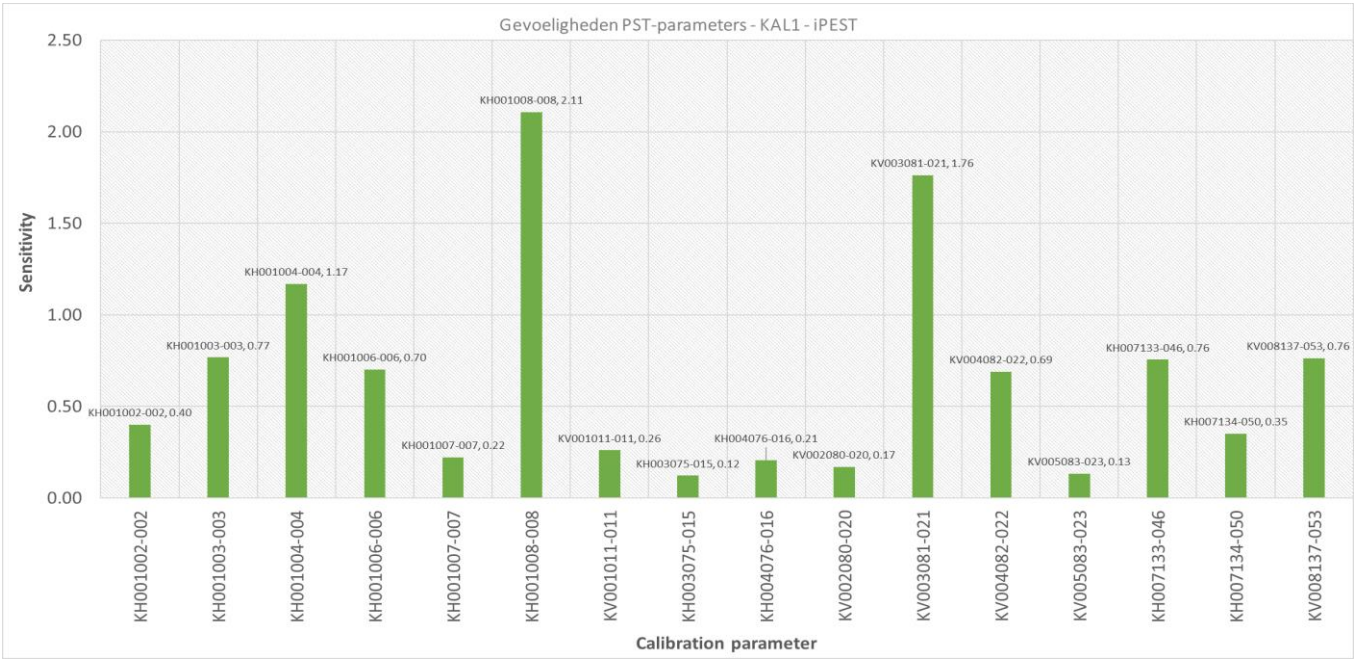
The only difference was in KAL2 model when GA found Drain Conductance in Layer 1 to be sensitive whereas iPEST did not. However, this difference can be attributed to the fact that GA uses a user-defined sensitivity measurement and cut-off percentage that defines which parameters are sensitive, whereas iPEST has its own measurement methodology and cut-off limits which is not accessible to the user for modifications. As the fundamental method used for measuring sensitivity in both the calibration methods in this study are different, the key take-away from Fig. 5.9 to 5.11 is the consistency in peaks/trends of the same parameters.

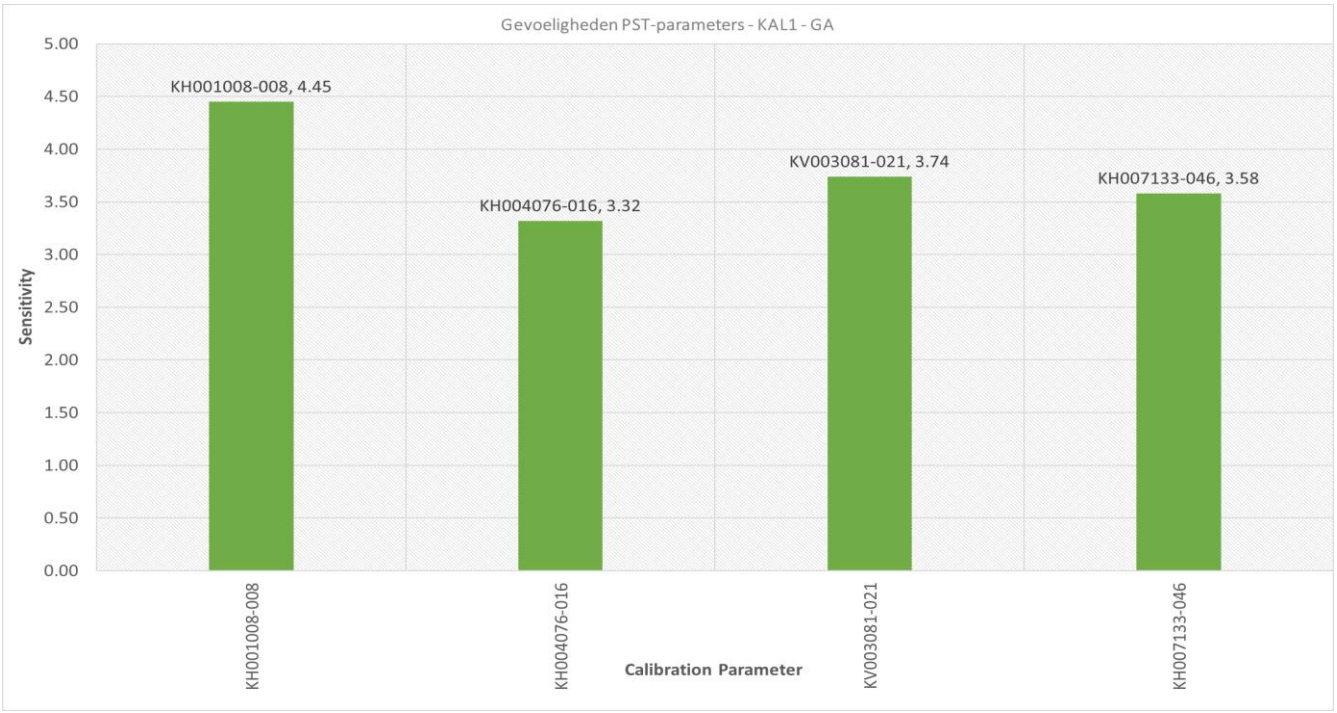




Parameter code	Zone ID	Formation
KH001002-002	ZandOv	Zand Overig
KH001003-003	ZMF	Zand Matig Fijn
KH001004-004	ZMG	Zand Matig Grof
KH001006-006	ZUG	Zand Uiterst Grof
KH001007-007	ZZF	Zand Zeer Fijn
KH001008-008	ZZG	Zand Zeer Grof
KV001011-011	Leem	Leem
KH002013-001	Grind	Grind
KH003075-015	MissingAquifer	
KH004076-016	MissingAquifer	
KV002080-020	MissingAquitard	
KV003081-021	MissingAquitard	
KV004082-022	MissingAquitard	
KH007133-046	syz3	Stramproy Zanden
KH007134-050	syz4	Stramproy Zanden
KV008137-053	kik1	Reuverklei/Venloklei
KH015218-071	kiz3	Zanden van Pey

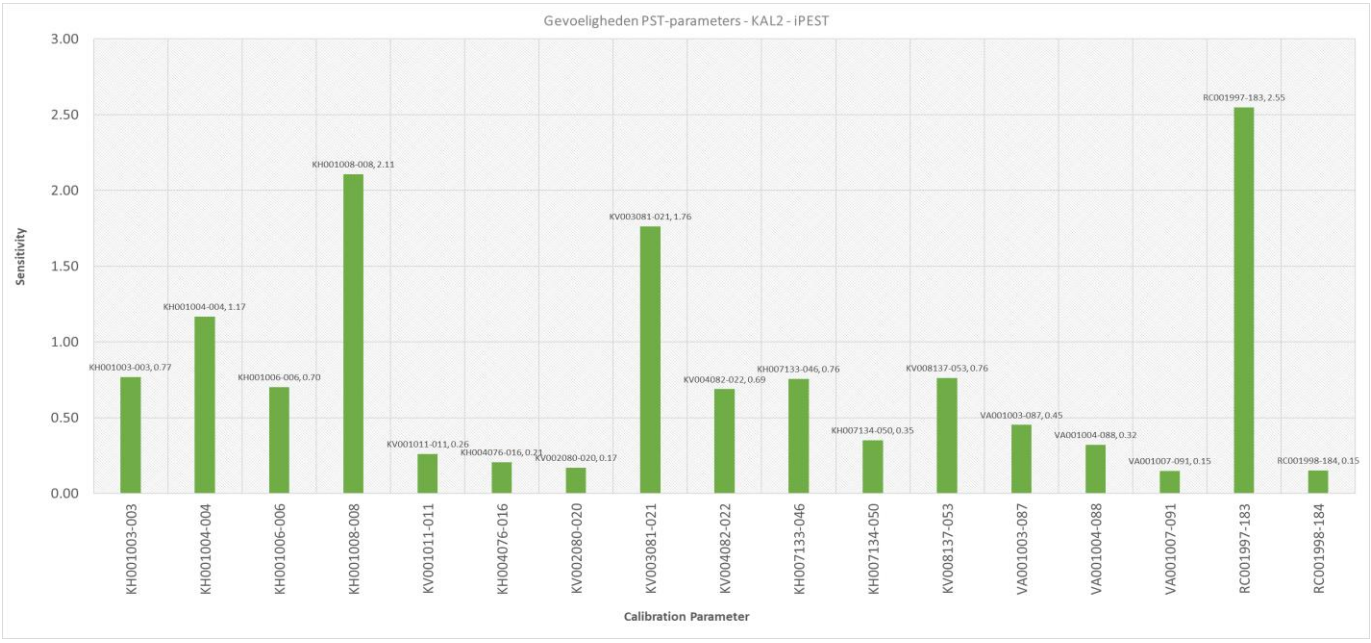
Fig. 5.9. Sensitivity plots KAL0 : (iPEST (top) and GA (middle), legend (bottom)

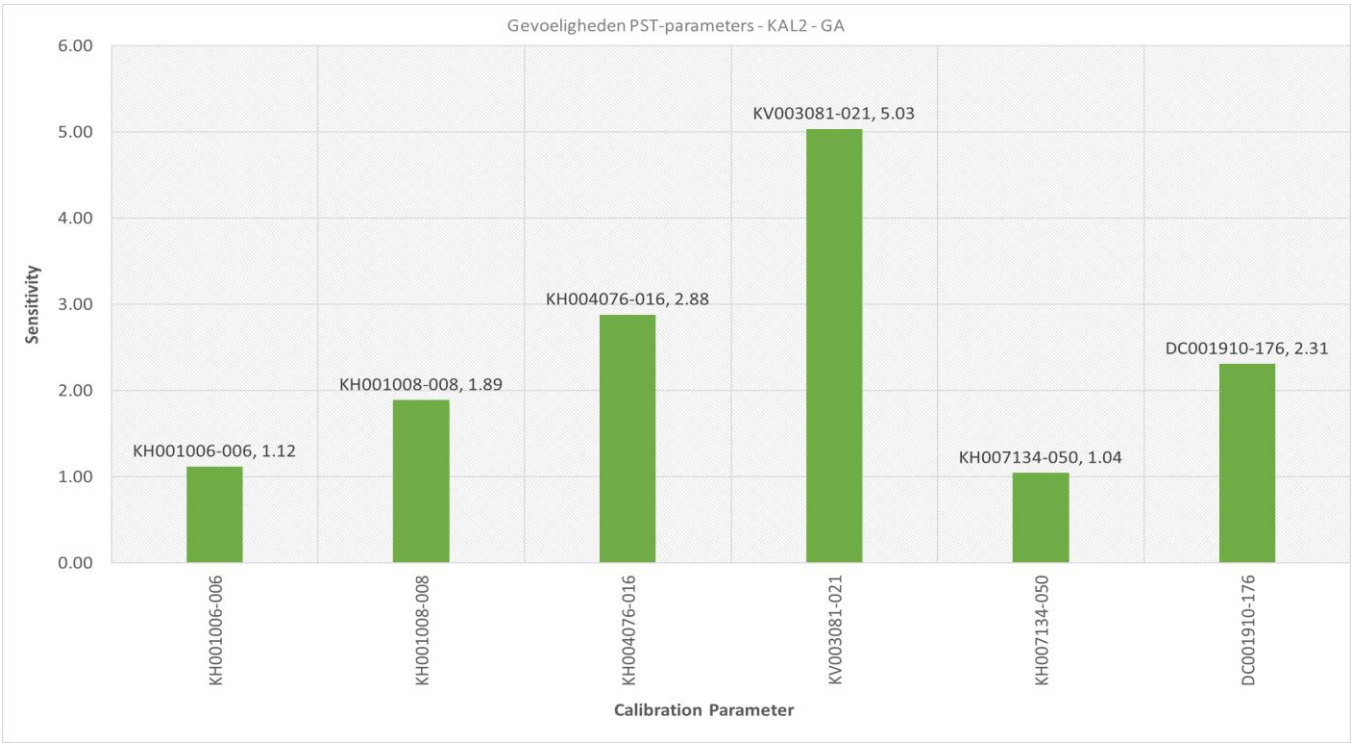




Parameter code	Zone ID	Formation
KH001002-002	ZandOv	Zand Overig
KH001003-003	ZMF	Zand Matig Fijn
KH001004-004	ZMG	Zand Matig Grof
KH001006-006	ZUG	Zand Uiterst Grof
KH001007-007	ZZF	Zand Zeer Fijn
KH001008-008	ZZG	Zand Zeer Grof
KV001011-011	Leem	Leem
KH003075-015	MissingAquifer	
KH004076-016	MissingAquifer	
KV002080-020	MissingAquitard	
KV003081-021	MissingAquitard	
KV004082-022	MissingAquitard	
KV005083-023	MissingAquitard	
KH007133-046	syz3	Stramproy Zanden
KH007134-050	syz4	Stramproy Zanden
KV008137-053	kik1	Reuverklei/Venloklei

Fig. 5.10. Sensitivity plots KAL1 : (iPEST (top) and GA (middle), legend (bottom))





Parameter code	Zone ID	Formation
KH001003-003	ZMF	Zand Matig Fijn
KH001004-004	ZMG	Zand Matig Grof
KH001006-006	ZUG	Zand Uiterst Grof
KH001008-008	ZZG	Zand Zeer Grof
KV001011-011	Leem	Leem
KH004076-016	MissingAquifer	
KV002080-020	MissingAquitard	
KV003081-021	MissingAquitard	
KV004082-022	MissingAquitard	
KH007133-046	syz3	Stramproy Zanden
KH007134-050	syz4	Stramproy Zanden
KV008137-053	kik1	Reuverklei/Venloklei
VA001003-087	ZMF	Zand Matig Fijn
VA001004-088	ZMG	Zand Matig Grof
VA001007-091	ZZF	Zand Zeer Fijn
RC001997-183	RIVZONES - DOMMEL-997	River 1
RC001998-184	RIVZONES - TONGELREEP-998	River 2

Fig. 5.11. Sensitivity plots KAL2 : (iPEST (top) and GA (middle), legend (bottom))

5.5.3. Parameter history

During calibration using iPEST and GA, the parameters chosen for calibration are multiplied using a calibration factor. As calibration progresses, the calibration factor is changed within its upper and lower bounds as per the value of the objection function. This is recorded by iPEST as well as GA. Plots of the changes in the calibration factor for the parameters was made for each of the models KAL0, KAL1 and KAL2 for iPEST. From the plots in Fig.5.12 to 5.14, you can see that iPEST changes the calibration factors progressively while assessing the value of the objective function. GA was observed to follow a similar progress in updation of calibration factors and thus, is not visualized here again.

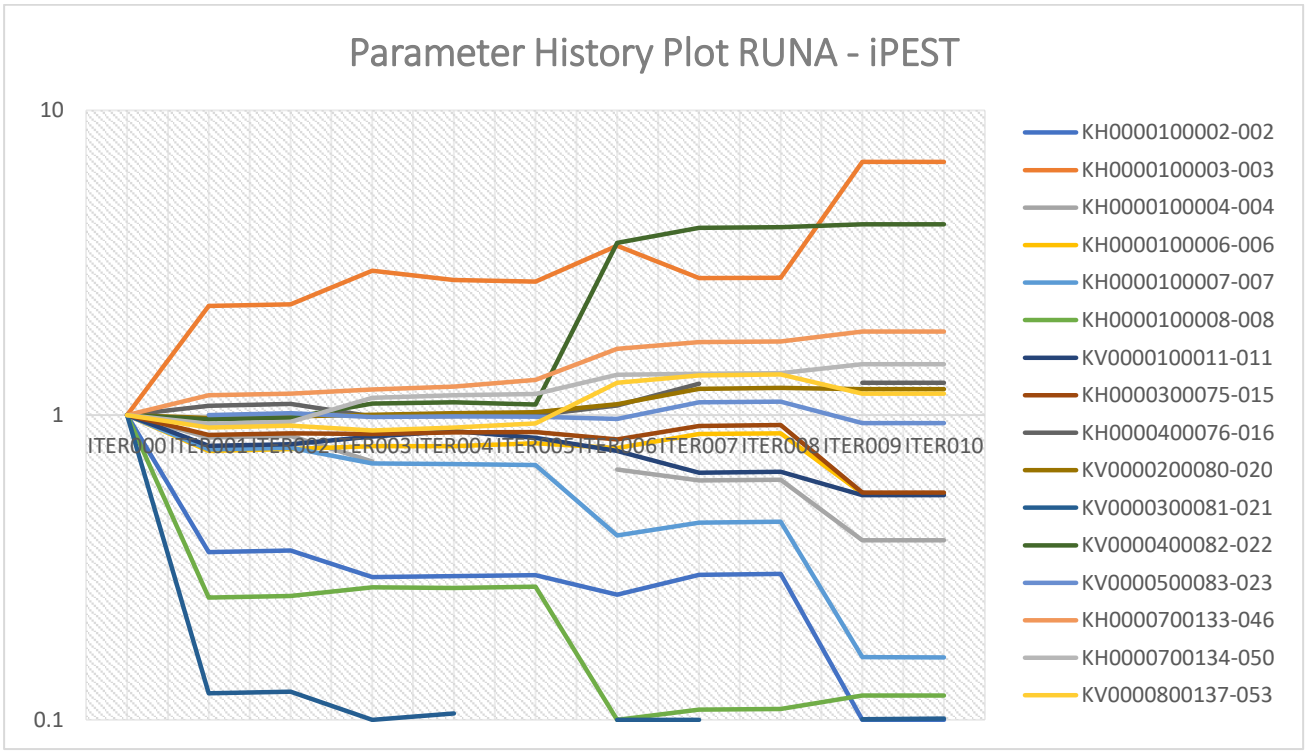


Fig. 5.12. Parameter History plots KAL0 : iPEST

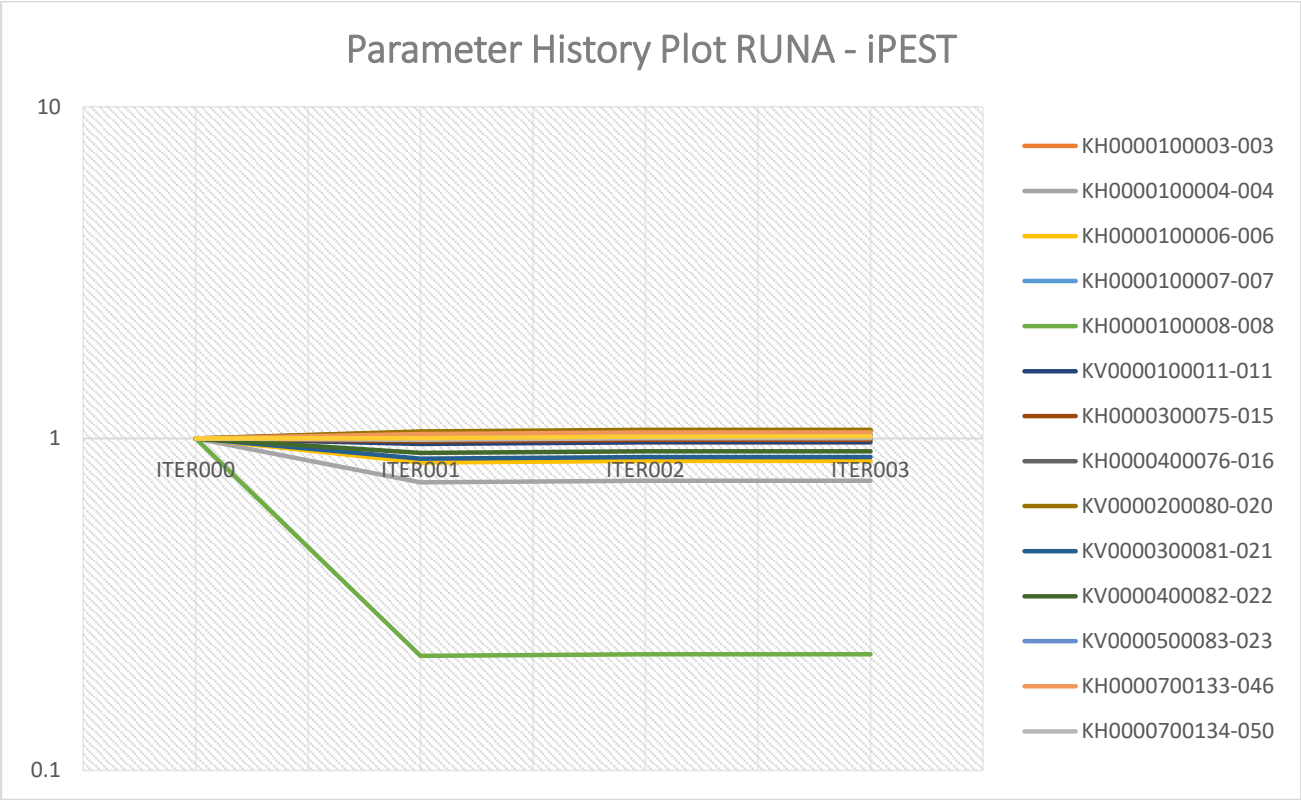


Fig. 5.13. Parameter History plots KAL1 : iPEST

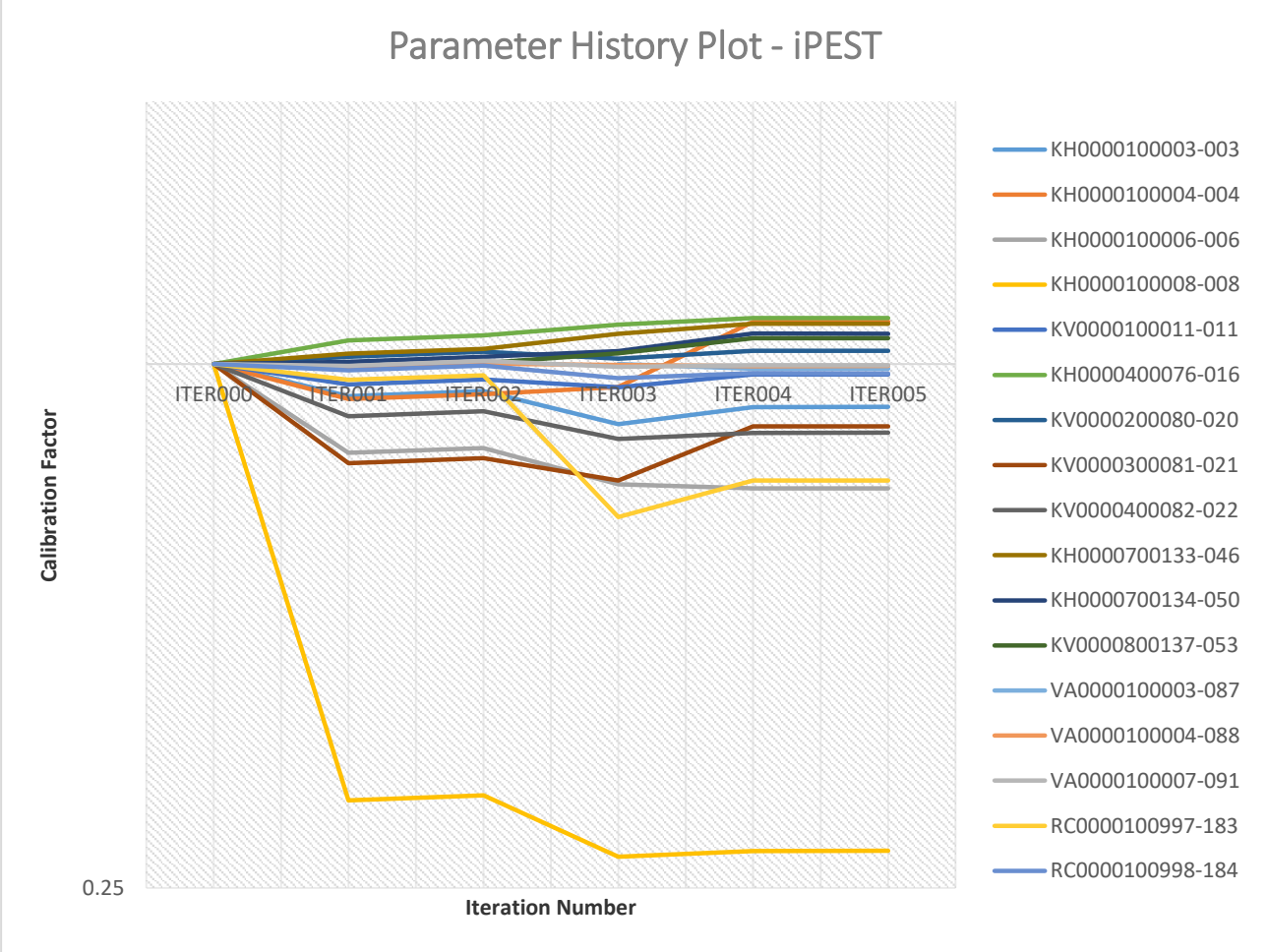


Fig. 5.14. Parameter History plots KAL2 : iPEST

5.5.4. Objective function value

The values of the objective function during both the process of calibration by iPEST and GA were recorded and compared. From Figures 5.15 to 5.17 it can be observed that the GA consistently reaches a lower value of objective function earlier than iPEST during calibration. In Fig 5.15, iPEST suddenly increases after reaching its lowest value. This can be attributed to keeping a very low tolerance value for convergence criteria in iPEST, thereby causing iPEST to proceed to the next iteration and obtaining a higher value which is not as low as the last one. The iteration was terminated then due to iPEST getting stuck in that local minima and not being able to find its way back to its earlier minima.

A tabular form of the lowest value (SSE) achieved by the objective function using iPEST and GA is also presented in Table. 5.12.

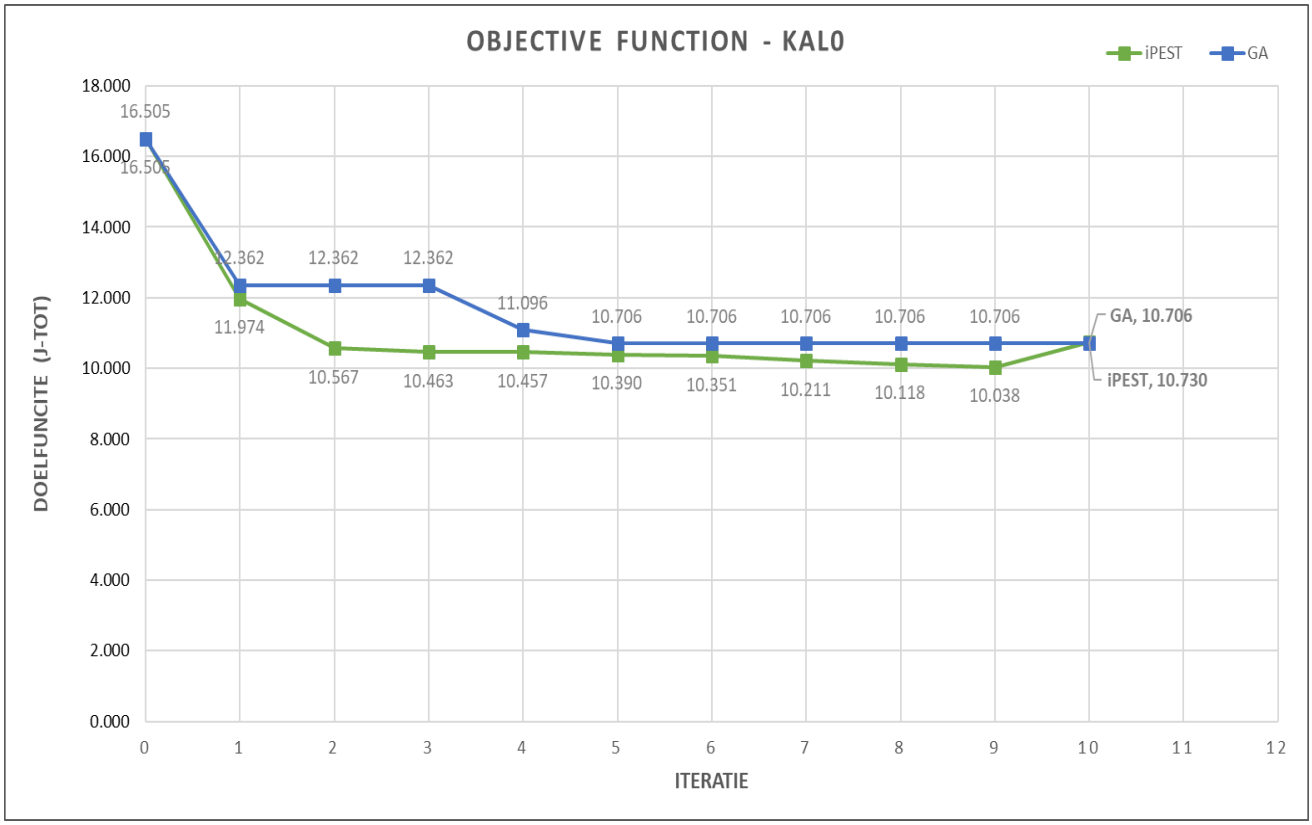


Fig. 5.15. Objective function plot KAL0 : iPEST vs GA

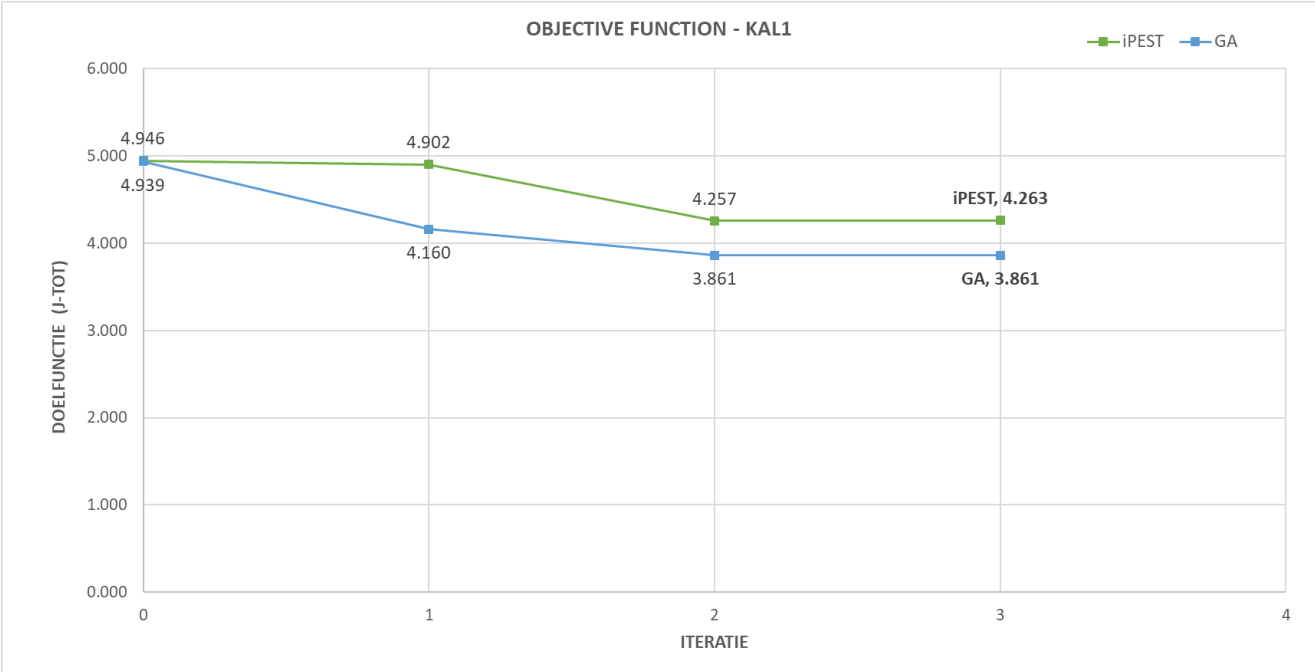


Fig. 5.16. Objective function plot KAL1 : iPEST vs GA

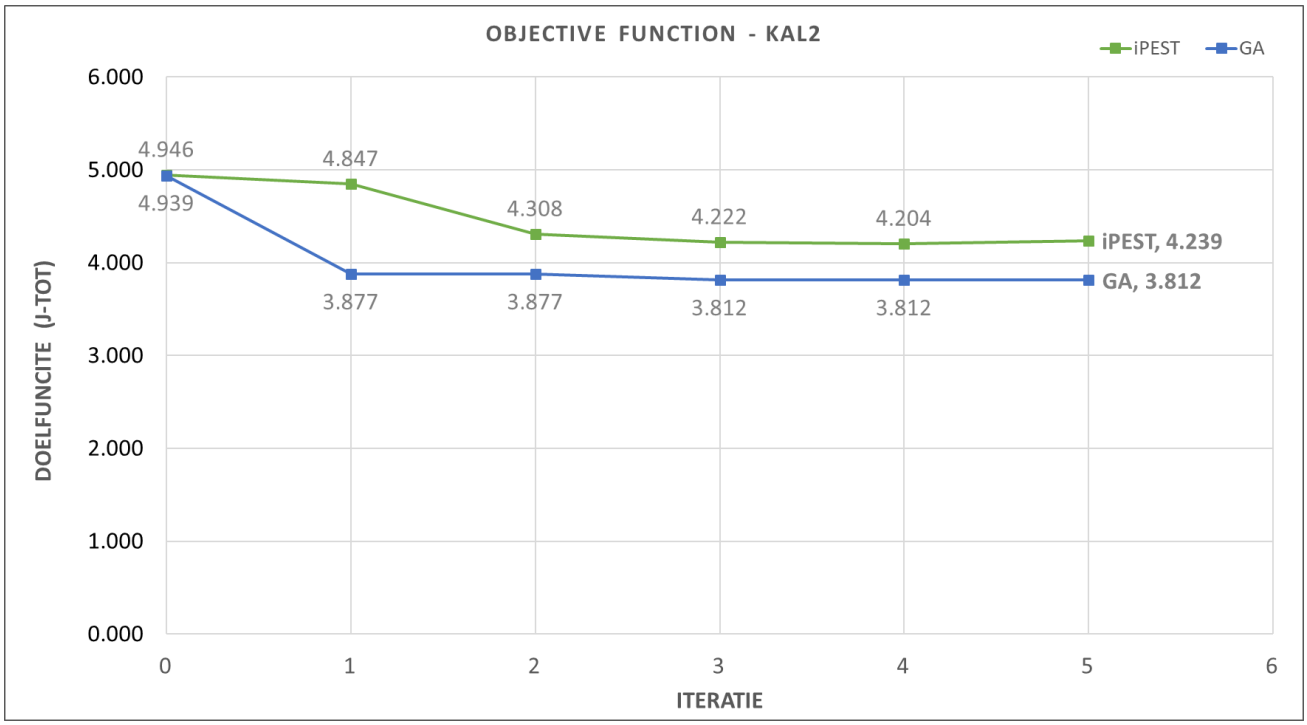


Fig. 5.17. Objective function plot KAL2 : iPEST vs GA

Final SSE / Algorithm	KAL0	KAL1	KAL2
iPEST	10.729	4.263	4.239
GA	10.705	3.860	3.812

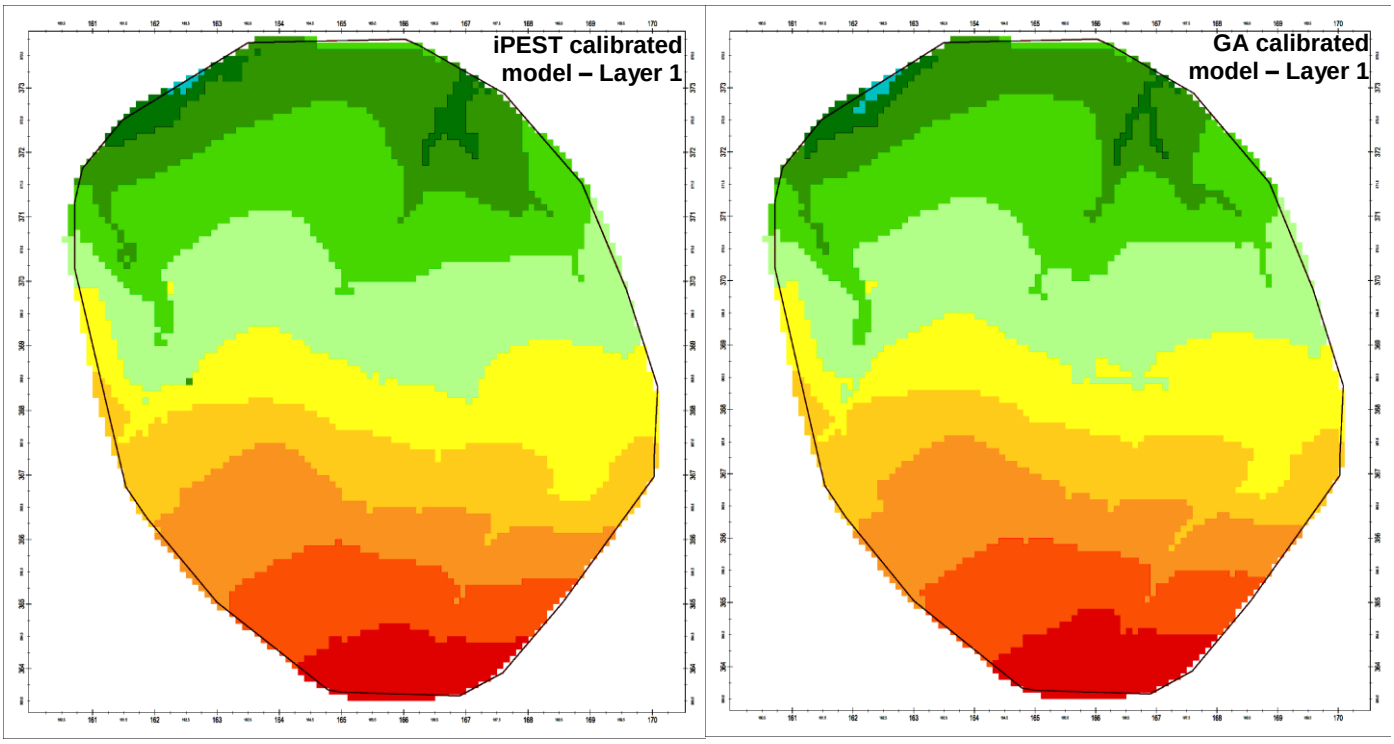
Table 5.12. Comparison of final error value: all iMOD MODELS

5.6. Final iMOD models

The calibrated parameters obtained as the result of the calibration model KAL0 were used to revise the original iMOD model input files and run two final iMOD models, one for each algorithm used in this study. The resulting heads and residuals have been analyzed below.

5.6.1. Comparison

Figs. 5.18 to 5.19 shows the resulting calculated groundwater heads in layers 1 and 15 for the study area in comparison to those of the original Triwaco Flairs model. It can be observed in comparison to Triwaco Flairs model for the study area, that both iPEST and GA calibration gives very similar heads and flow patterns to the original.
On comparing the residuals of the two final models in comparion to Triwaco Flairs model (Fig. 5.20), it can be observed that the GA-calibrated model better follows the trend in residuals in the original Triwaco Flairs model.



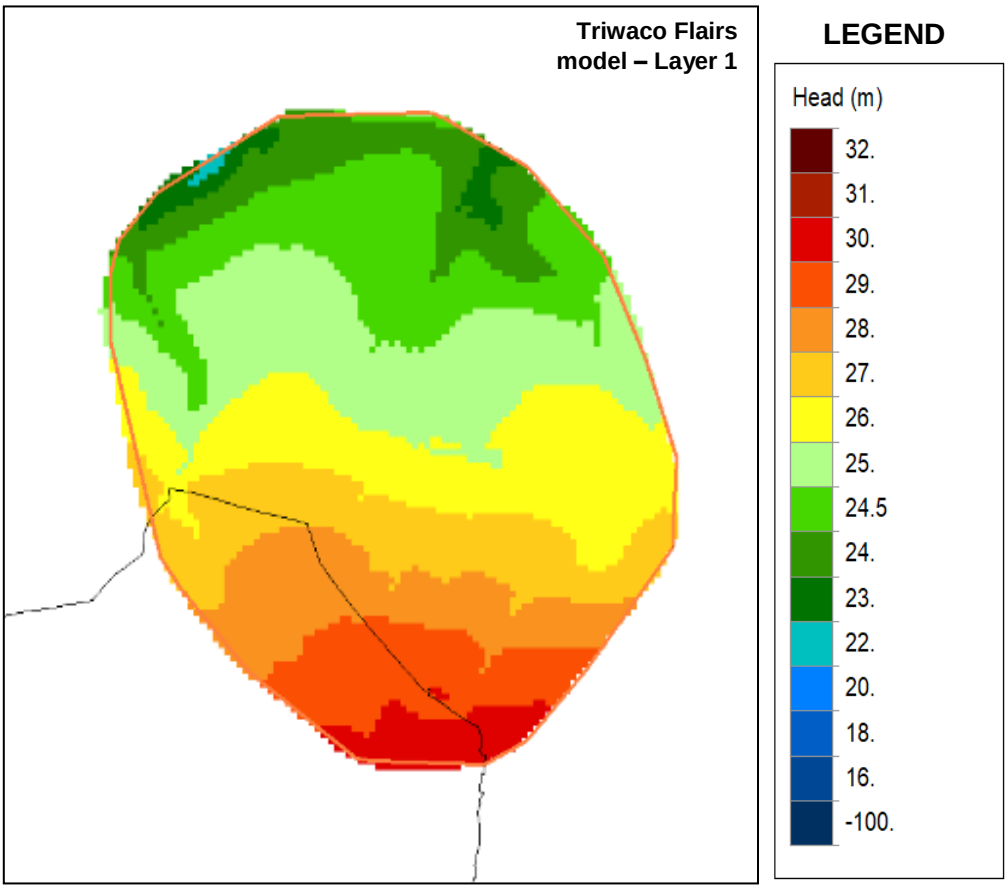
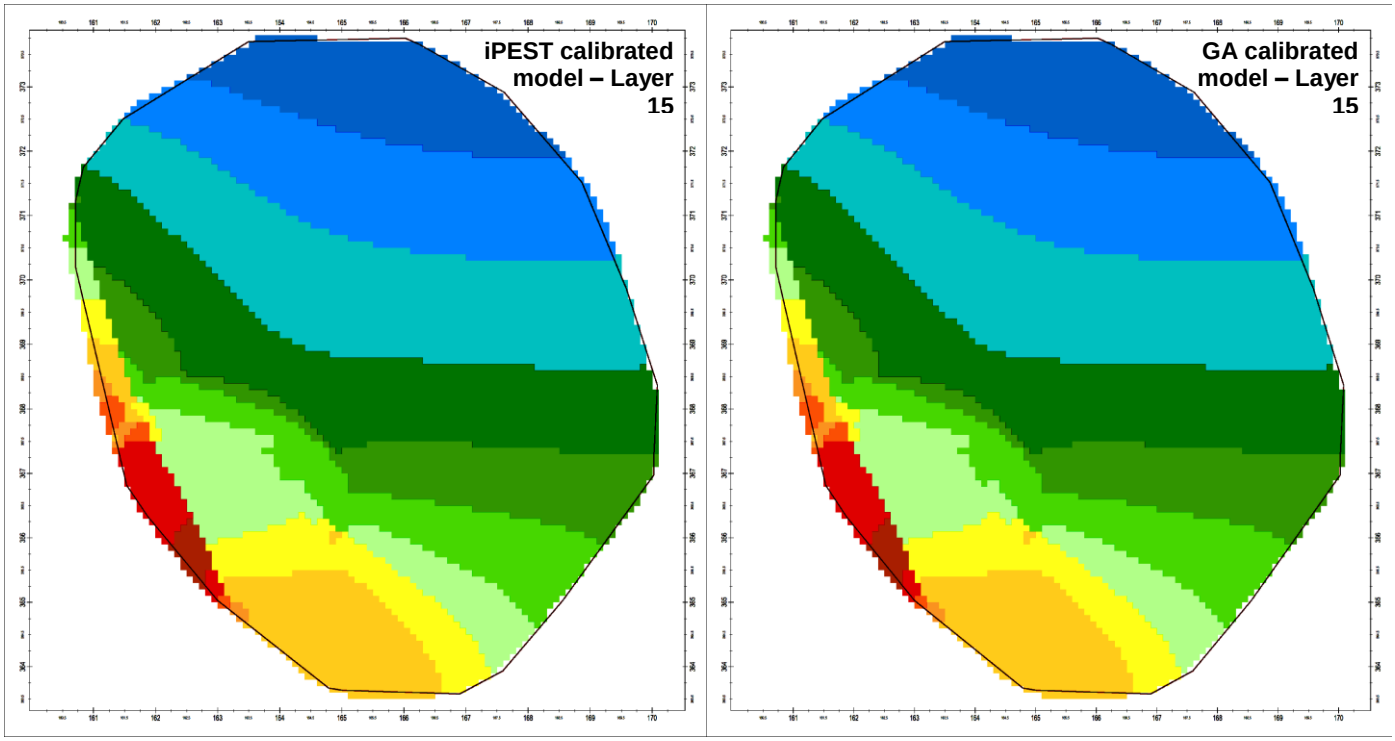


Fig. 5.18. Calculated heads in Layer 1 for results of iPEST, GA calibrated models and original Triwaco Flairs model from left to right and top to the bottom



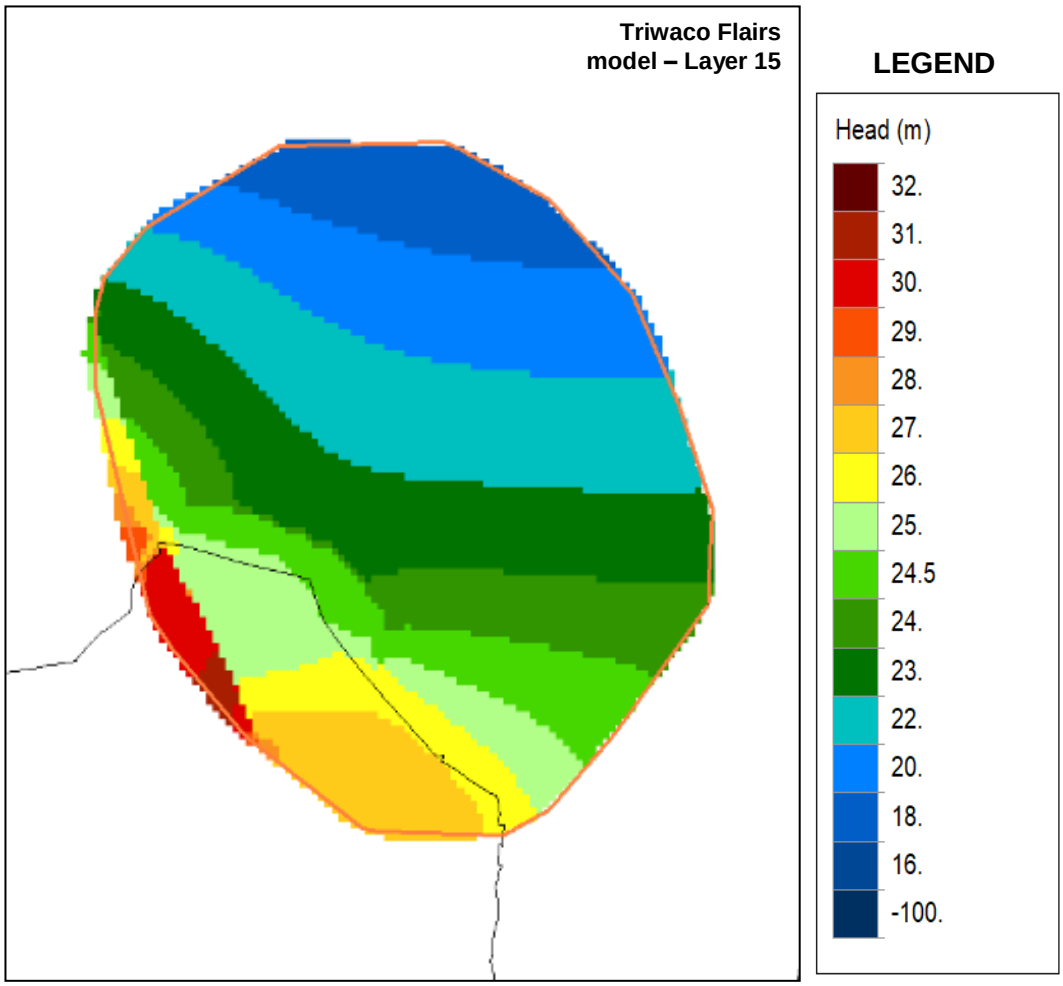
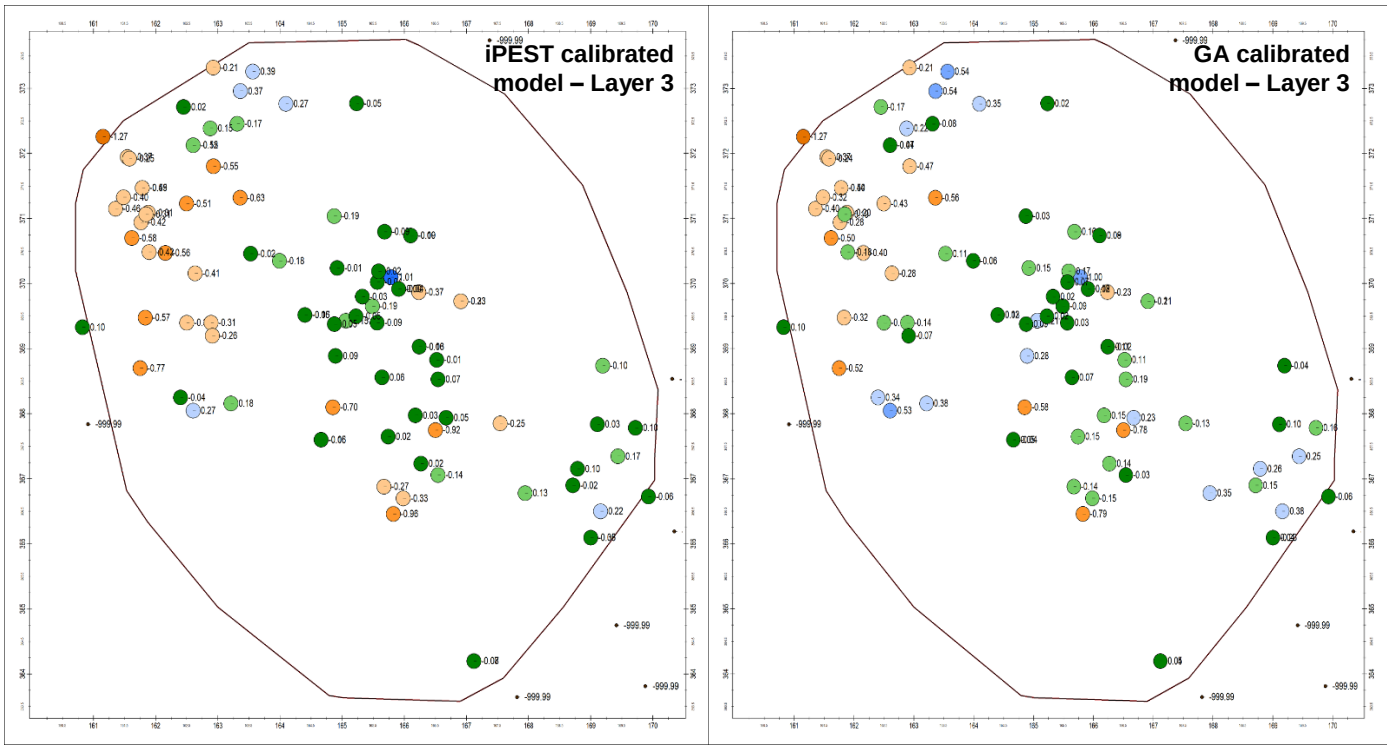


Fig. 5.19. Calculated heads in Layer 15 for results of iPEST, GA calibrated models and original Triwaco Flairs model from left to right and top to the bottom



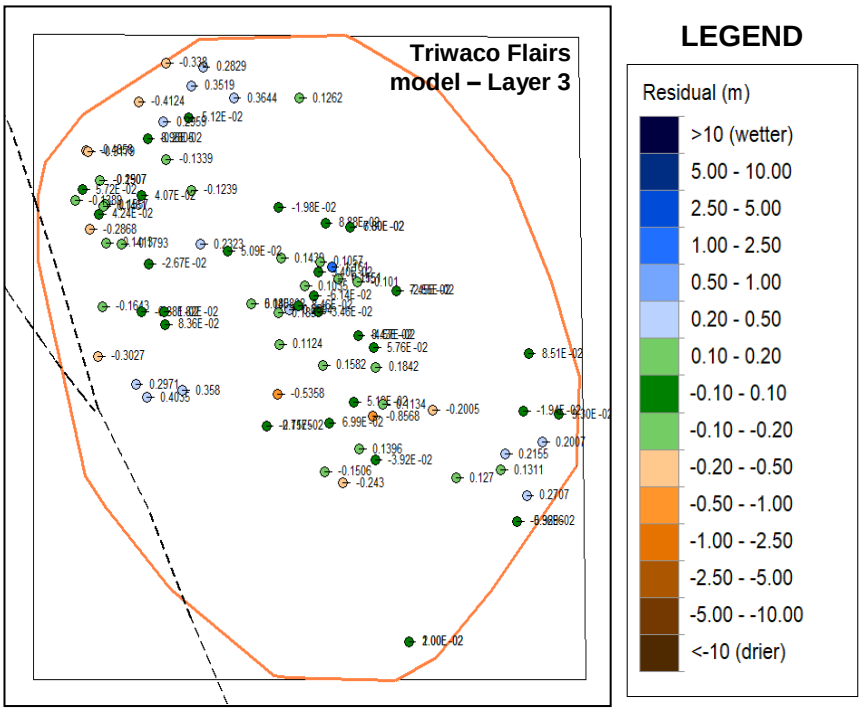
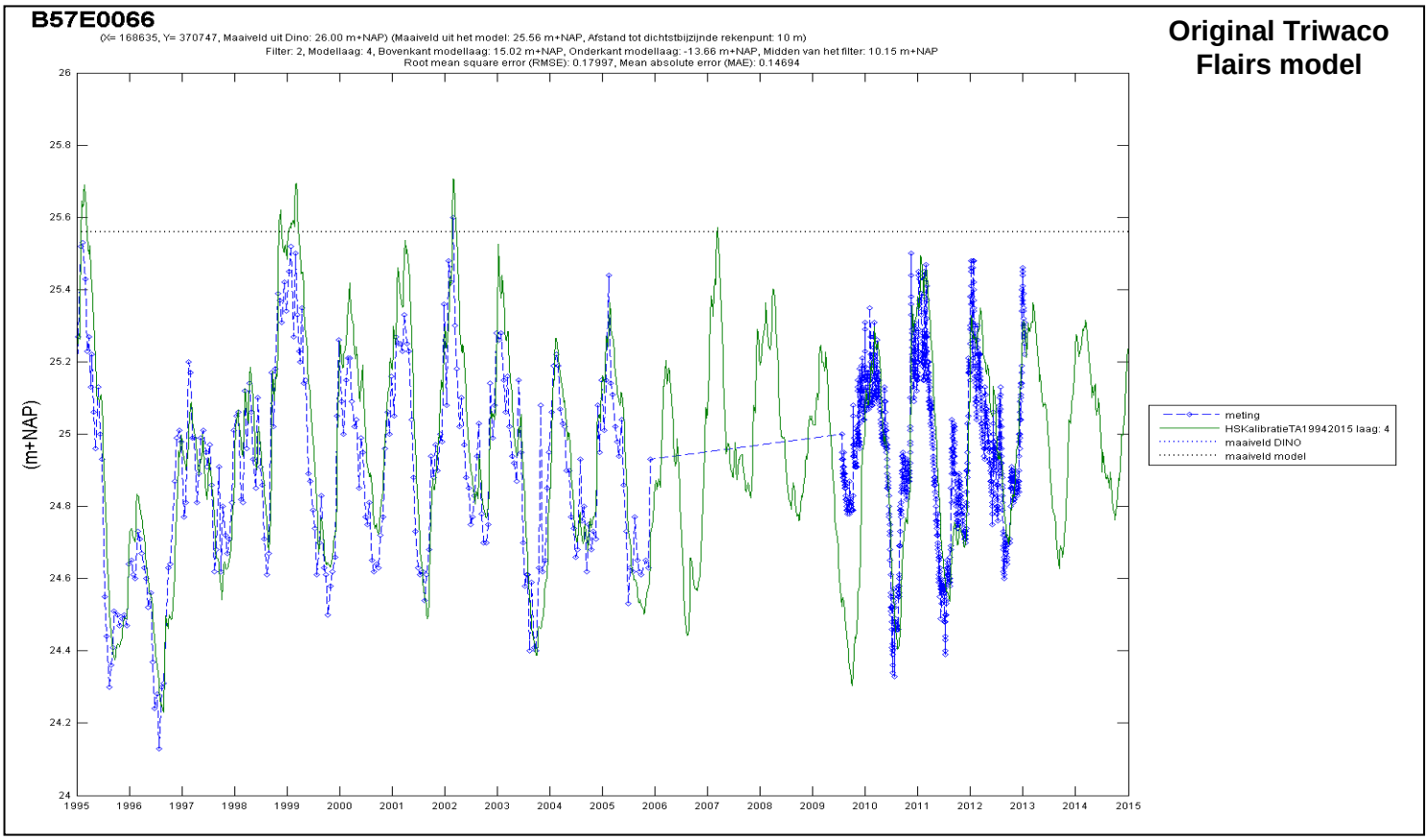


Fig. 5.20. Groundwater residuals in Layer 3 for results of iPEST, GA calibrated models and original Triwaco Flairs model from left to right and top to the bottom

5.7. Validation

A validation of the GA calibration results was done by feeding the GA-calibrated parameters of KAL2 model to a transient model run of the original Triwaco Flairs model for the study area and comparing it to previous results of the Triwaco Flairs model.

Fig. 5.21. shows the result of a transient run of Triwaco Flairs model at one observation well B57E066 in Layer 3 of the groundwater model during the selected time period of 1994 to 2015. From the figures it is clear that the calculated heads (green) of the GA-calibrated model gives much better fit along the peaks and crests of the observations (blue) when compared to the original Triwaco Flairs model. Similar trend was seen in other observation wells as well, more of which can be found in Annex.



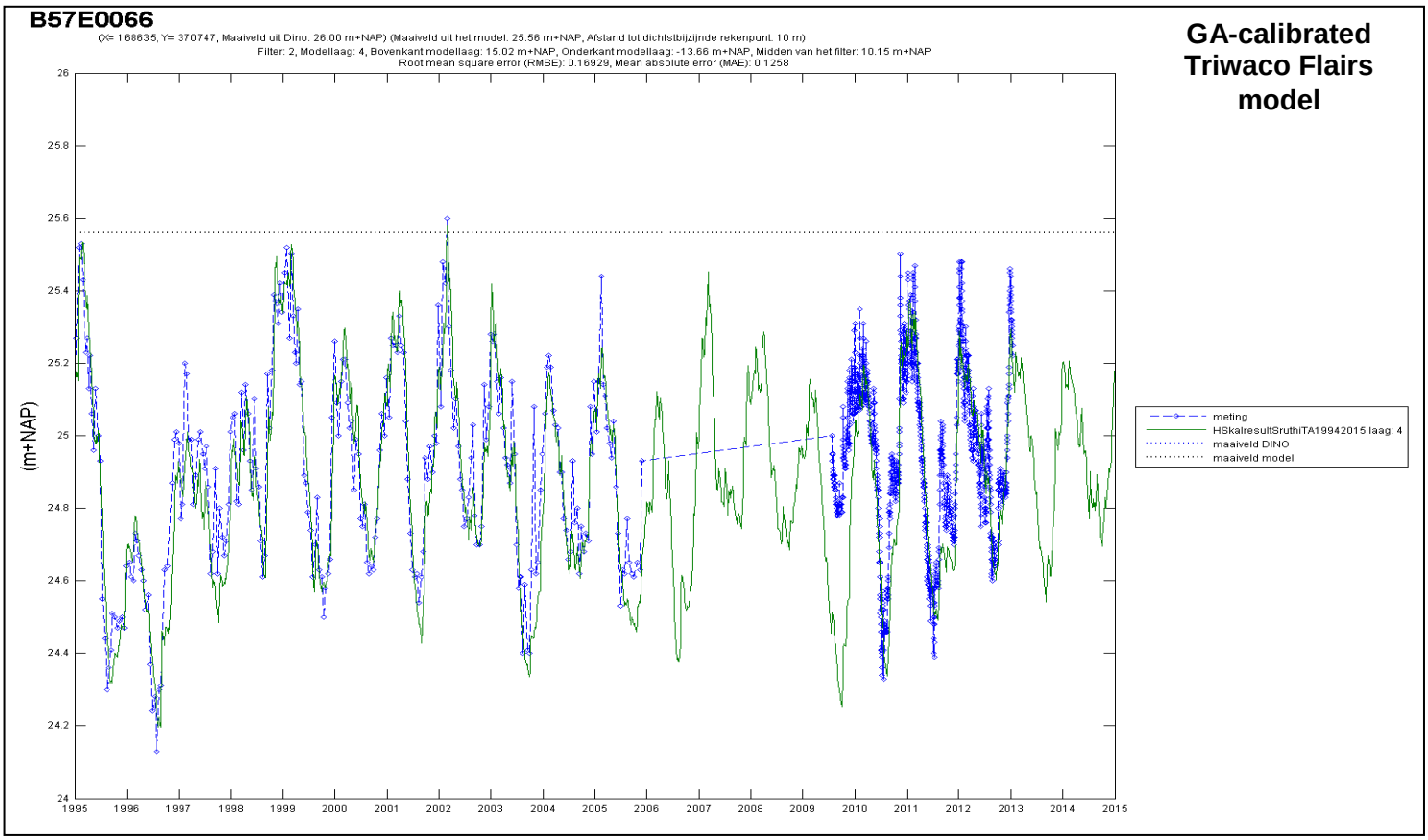


Fig. 5.21. Comparison of groundwater head variation in observation well ID: B57E0066 during the years 1994-2015 using Triwaco Flairs (Blue dotted lines represent observed head and Green lines represent modelled head)

5.8. GA modifications

It was decided to vary the settings of the GA and test for improvements in performance. A few tests were thus done by varying the population size as well as the functioning of the crossover operator.

5.8.1. Experiment: Population size

Population size of the set of solutions for running GA was varied among 10, 20 and 50 and the calibration was done. While a smaller population size of 10 took less time of 4 hours, bigger population sizes of 20 and 50 took 8 and 15 hours respectively. A linear relationship can be inferred between population size and computation time (Fig. 5.23). However it was also observed that, calibration run using a population size of 50 was able to find a much lower error function value within 10 generations (out of the maximum 100 generations) as is visible in Fig. 5.22, while it took other population sizes up to 50 generations to reach an error function value which was still higher than that of population size 50. Consistency was tested with 10 trial runs for each population size. While GA with a population size 10 was not giving consistently better results than iPEST, those with sizes 20 and 50 have shown consistently better results in all the 10 trial runs.

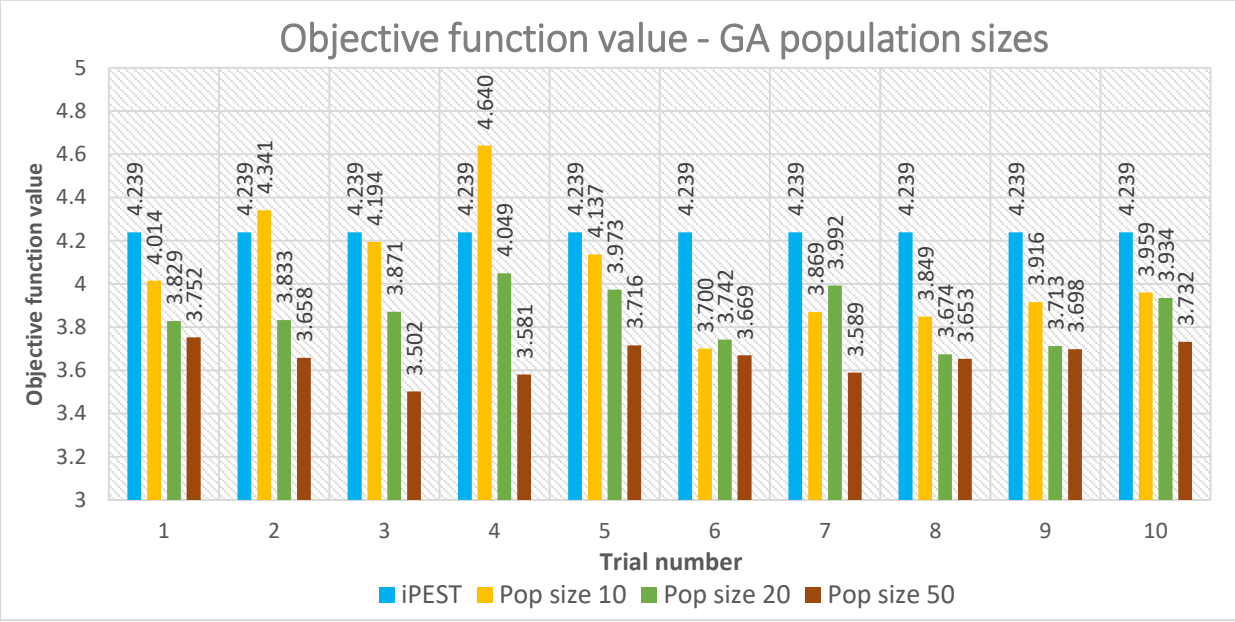


Fig. 5.22. GA experiments in Population sizes 10, 20, 50 vs iPEST

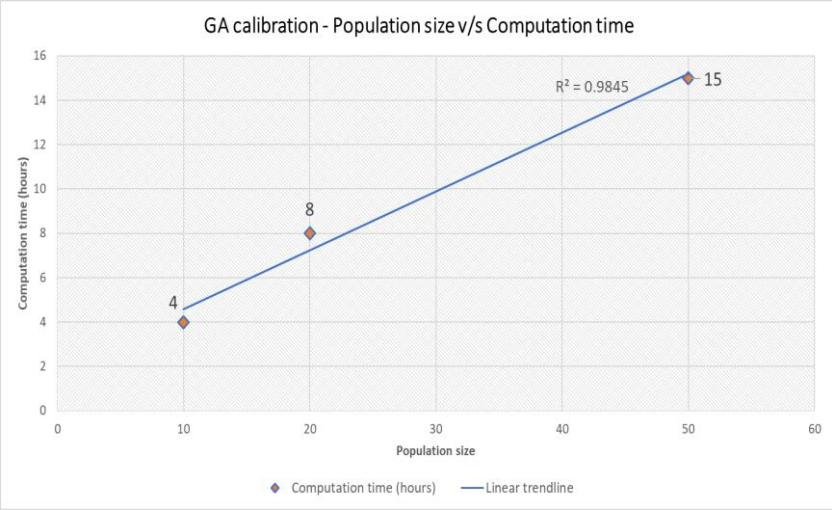


Fig. 5.23 Relationship between computation time and population size of GA

5.8.2. Experiment: Uniform Crossover (UX)

Uniform Crossover (Section 2.2.3) was implemented in KAL2 model with population size 50 and the results were observed. Fig. 5.24. shows a comparison in objective function value with progress in one GA calibration run. The default type of crossover operator used in this study is the Single Point Crossover (Section 2.2.3). This was compared against Uniform Crossover operator. From Fig. 5.24. it can be observed that using a uniform crossover operator achieves lower objective function value from the beginning of the calibration run and consistently maintains this trend until ending with a lower final value.

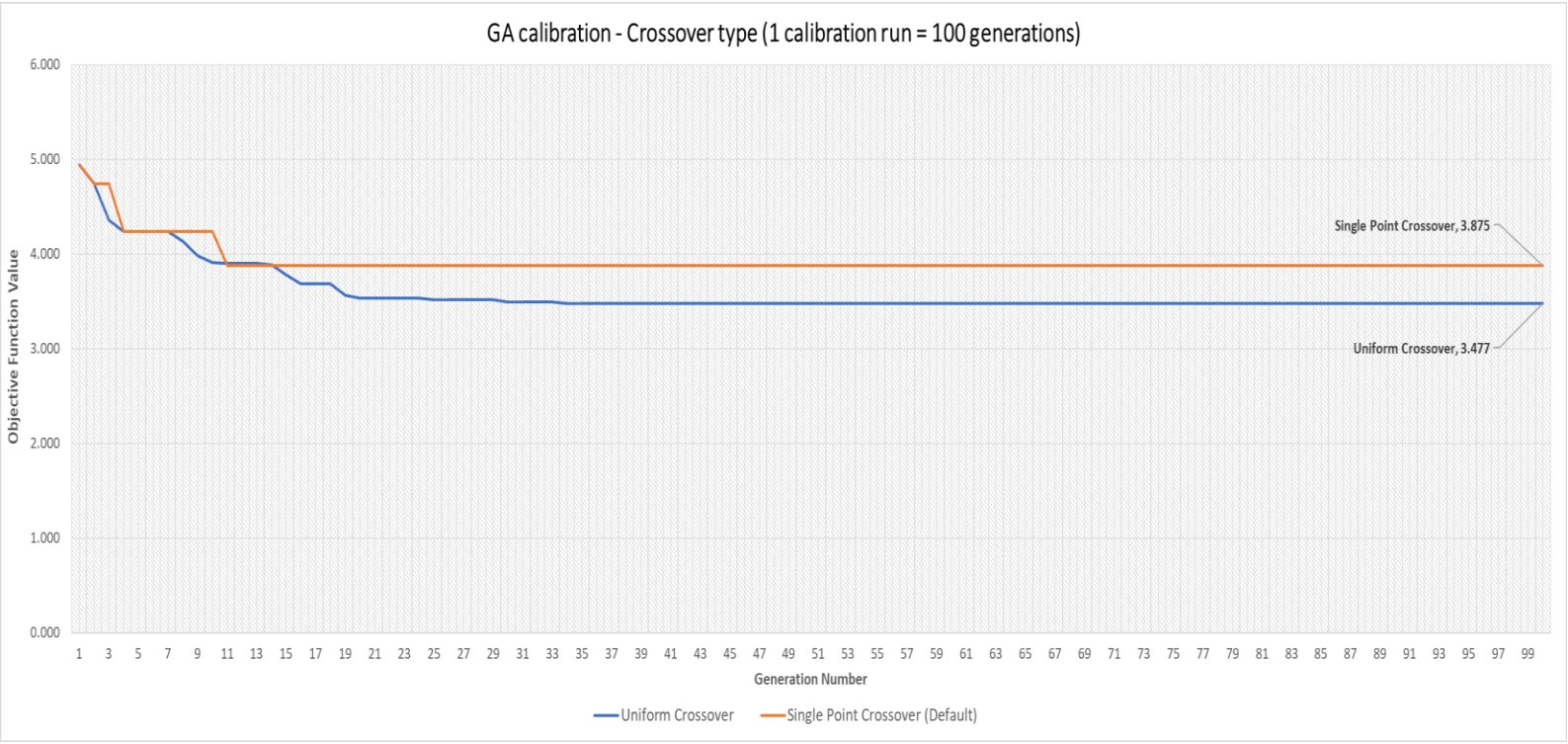


Fig. 5.24. Variation of objective function value using different types of Crossover.

5.8.3. Experiment: Crossover and Mutation rates

Population size of 50 on KAL2 calibration model was chosen for this experiment. Fig. 5.25 shows the progression of objective function value during one GA calibration run under the various combinations mentioned in Table 4.5.3. It can be observed that Combi2 and Combi3 takes more number of generations to reach the same objective function value as the Original combination and Combi1. Combi1 seems to give much lower objective function value than the other combinations. This implies that a high cross over rate with a low mutation rate should be used in the optimization problem at hand: parameter estimation in groundwater modelling.

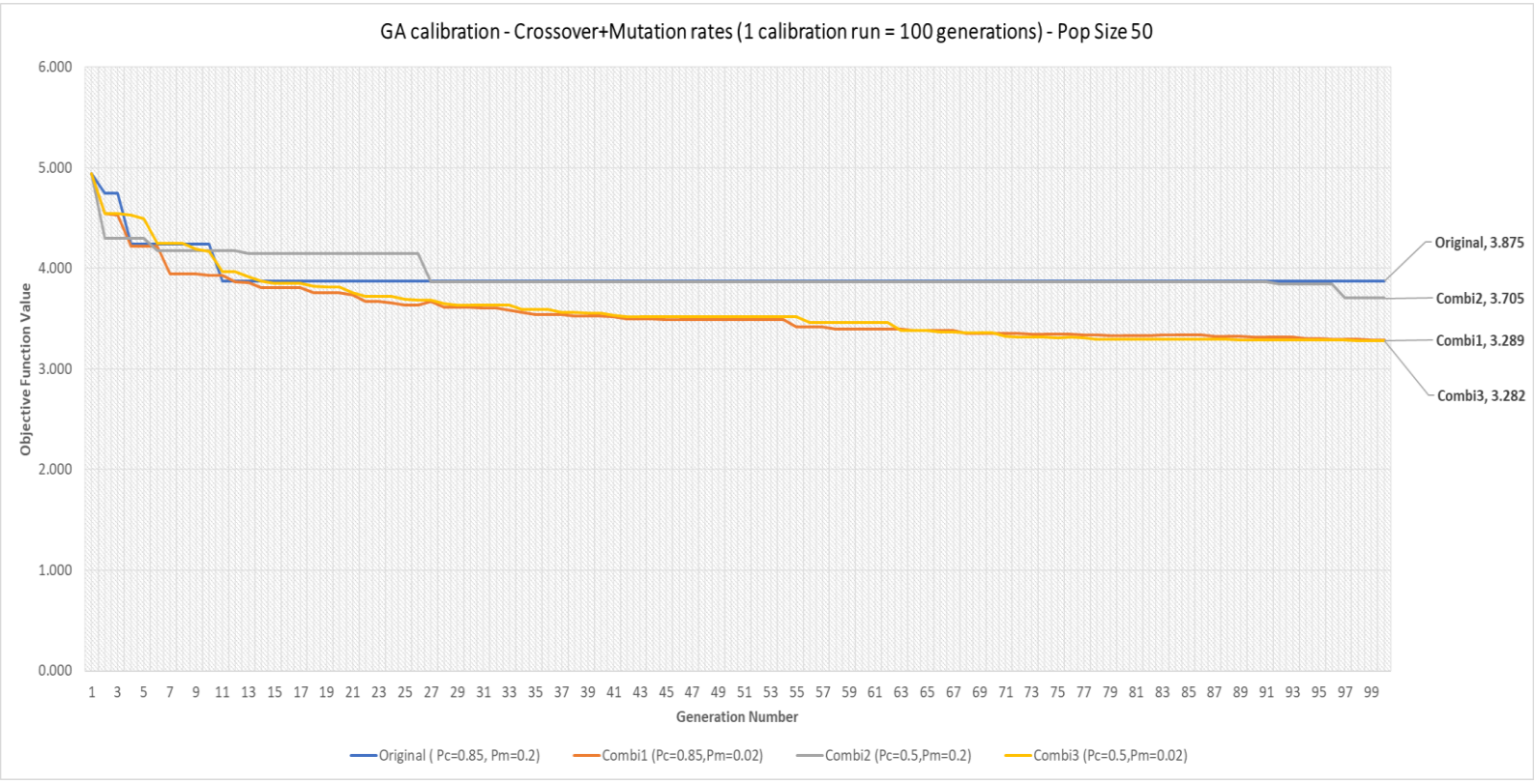


Fig. 5.25. Variation of objective function value (population size 50) using combinations of Crossover & Mutation rates.

6. Conclusions

6.1. Comparison of GA vs iPEST

The performance of the global optimization algorithm GA was compared against iPEST algorithm in terms of performance factors already defined. The fact that iPEST consistently ended up in the same final objective function value while GA was able to find several lower final objective function values indicate the presence of several local minima, the tendency of iPEST to get stuck in the same local minima and the ability of GA to not do the same.

6.2. Performance factors

The comparison between iPEST and GA was done using the results of their calibration of the same groundwater model KAL2. While computationally GA took more time than iPEST, GA has consistently produced much lower error function values. The computation time is a function of the model evaluations/generations the user inputs. From the fitness function plots (Fig. 5.15 to Fig. 5.17), it is visible that GA does not find a better solution after the 50^{th} iteration. As such, this cuts down the function evaluations by half and subsequently the computation time. A trial population size of 50 has reached a much better solution of all sizes in less than 10 generations. A study of the linear relationship between population sizes (of 10, 20 and 50) and computation time (Fig. 5.23) further supports the conclusion that by choosing the right population size, the computation time can be brought down.

On studying the performance of the GA tool under various combinations of its basic settings such as population size, crossover and mutation rates, it can be concluded that a population size of 50 is able to find the best solution for the parameter optimization problem in this study. A lower number of generations with a higher population size ensures lower computation time for an otherwise computationally expensive optimization method. A comparison of the various types and combinations of crossover and mutation operators further leads to the conclusion that a population size of 50 with a Combi1 (p_c = 0.85, p_m = 0.02) setting gives the lowest minima for the optimization problem used in this study.

6.2.1. Groundwater heads and residuals

From observing the groundwater heads and residuals calculated by two iMOD models using the results of iPEST and GA calibration (Section 5.6.1), it can be seen that iPEST and GA both give similar outputs with GA having an upper hand in matching the residuals of the original Triwaco Flairs model.

6.3. Physical plausibility of the results

The physical plausibility of the calibration results have to be checked and validated to ensure the final values of the calibration parameters stay within a reasonable range of values.

1) Horizontal and Vertical hydraulic conductivity (KH, KV)

During a comparison of the hydraulic conductivity parameters that were chosen for calibration by both iPEST and GA, both algorithms have created final values close to plausible range of values. In fact, GA has kept the vales in a lower range compared to iPEST. The values such as that in Layer 1 in Table 6.1. that are exceeding out of the plausible ranges are in very small regions which can be attributed to errors in interpolation methods during calculation of soil thickness. The missing aquifer in Layer 4 and missing Aquitard in Layer 3 could be layers of ZMG (Zand Matig Grof) sand and clay (Klei), estimating from the final parameter value obtained after GA calibration.

Parameter	Formation	Type	Layer	Calibrated value (mean)	Plausible value*
KH	ZUG	Aquifer	1	250	50-200
KH	ZZG	Aquifer	1	400	80
KH	Missing Aquifer	Aquifer	4	36.7	ZMG?
KH	Stramproy Zand (ZMF-ZMG)	Aquifer	7	15.3	1-50
KV	Missing Aquitard	Aquitard	3	$2.5 * 10^{-4}$	Klei?

Table 6.1. GA results for hydraulic conductivities KH, KV: A comparison (Ref. Bram Bot (2011), Grondwaterzakboekje*)

2) Field conditions

From studying the results of the validation using a transient run using Triwaco Flairs for the study area, the observed and calculated heads show a good fit. This is important for the study area and in turn for the project area to understand the hydrology closely. If the calculated head during a particular year is not following the peaks or troughs (Fig.5.21.) by a significant extent, it then becomes a case for further detailed study of the hydrology in that location. Subsequent steps would involve deeper analysis of the localized region for mismatches in surface elevation or geology of the area to understand changes in the system. This is an essential step to understanding the flow of groundwater in the project area and as such, will aid the Waterschap De Dommel in moving forward with the rejuvenation of the Oude Strijper Aa. On studying Fig. 5.21 and head variations in the rest of the wells (Annex) the GA results were seen to follow the peaks and troughs to a satisfactory level and most of the times better than the results of existing Triwaco Flairs model.

7. Recommendations

The previous section has concluded that using the global optimization algorithm Genetic Algorithm gives better results over local optimization algorithm iPEST used in parameter estimation of groundwater modelling. Recommendation for further research on the topic has been mentioned in this section.

First of all, it is recommended to further tune usage of Genetic Algorithm to ensure attainment of a global minimum by exploring as much of the search space as possible. Due to the large number of parameters that are initially available, this could end up being an over-parameterized optimization problem due to relatively smaller number of observations available for the study. While iPEST brings down the number of parameters after an initial sensitivity analysis, the GA tool at this stage of study is unable to do the same. The status quo can be changed by incorporating a pre-calibration sensitivity analysis similar to iPEST after which only the most sensitive parameters are chosen for calibration using the Genetic Algorithm. A reduction in the number of parameters being optimized allows a better exploration of the search space. This, together with larger population size, high crossover rate and a low mutation rate as chosen in this study should be tested to give the best performance of Genetic algorithm for the groundwater model optimization problem.

The zonation of the calibration parameters also plays a very important role in calibration. Choosing the right property for zonation depends on the purpose of the model as well. In this study, each of the two rivers flowing through the study area have been grouped as a whole and their conductance calibrated. A recommendation for a more detailed and accurate zoning would be to use the geomorphological map for the area to better group the river conductance. Another factor to be considered for zoning of the rivers can be flow velocities which will vary along the stream length and give better understanding of their influence on groundwater flow. Parameters can also be grouped based on prior knowledge, thereby combining several zones into one and reducing the total number of parameters being estimate and developing a well-defined optimization problem.

A good fitting cannot guarantee unique parameter estimation, due to correlations among the parameters. Correlation between multiple parameters also needs to be studied. Higher order interrelationships among parameters, wherein more than two parameters are correlated together but not in pairwise, cannot be detected by either iPEST or GA at this moment and needs to studied to ensure parameter uniqueness. This in turn is important to estimate confidence limits for parameters being optimized and in turn provide lower values of the objective function.

Due to correlations, certain regions in the parameter space may correspond to good model predictions. This would mean that the residual value remains low even if the parameters vary in certain regions. A tracking of the fitness function value among the different areas needs to be studied during the progression of a calibration run. Based on the results, the weights of the observations can be modified accordingly.

Since GA is a random search algorithm, it lacks the ability to evaluate the correlation in fitness values among different areas within a project area. A tracking of the fitness function value change needs to be done to better understand correlation and to ensure good solutions for specific areas of interest are not discarded by the algorithm.

8. Bibliography

1. Arsenault, R., Poulin, A., Côté, P., & Brissette, F. (2014). Comparison of Stochastic Optimization Algorithms in Hydrological Model Calibration. *Journal of Hydrologic Engineering*, *19*(7), 1374–1384. http://doi.org/10.1061/(ASCE)HE.1943-5584.0000938
2. Bajpai, P., & Kumar, M. (2010). Genetic algorithm–an approach to solve global optimization problems. *Indian Journal of Computer Science and Engineering*, *1*(3), 199–206.
3. Blasone, R. S., Madsen, H., & Rosbjerg, D. (2005). Comparison of parameter estimation algorithms in hydrological modelling, *304*(June 2005), 67–72.
4. Brikowski, T. (2010). GEOS 5311 Lecture Notes: Boundary Conditions. *Vers*, *114*.
5. Carr, J. (2014). An Introduction to Genetic Algorithms. *Whitman College Mathematics Department*, 1–40. http://doi.org/10.1016/S0898-1221(96)90227-8
6. Carrera, J., Alcolea, A., Medina, A., Hidalgo, J., & Slooten, L. J. (2005). Inverse problem in hydrogeology. *Hydrogeology Journal*, *13*(1), 206–222. http://doi.org/10.1007/s10040-004-0404-7
7. Craig, J. R., & Read, W. W. (2010). the Future of Analytical Solution Methods for Groundwater Flow and Transport Simulation, 1–8.
8. Datta, B., Chakrabarty, D., & Dhar, A. (2009). Simultaneous identification of unknown groundwater pollution sources and estimation of aquifer parameters. *Journal of Hydrology*, *376*(1–2), 48–57. http://doi.org/10.1016/j.jhydrol.2009.07.014
9. El Harrouni, K., Ouazar, D., Walters, G. A., & Cheng, A. H.-D. (1996). Groundwater optimization and parameter estimation by genetic algorithm and dual reciprocity boundary element method. *Engineering Analysis with Boundary Elements*, *18*(4), 287–296. http://doi.org/10.1016/S0955-7997(96)00037-9
10. Harbaugh., A. W. (2005). Design of the Ground-Water Flow Process. *MODFLOW–2005, The U.S. Geological Survey Modular Ground-Water Model*, 1–12.
11. Franklın, W., & Zhang, H. (2003). Fundamentals of groundwater, (July), 583.
12. GHP Oude Essink (2000), Groundwater Modelling, Utrecht University
13. Bram Bot (2011), Grondwaterzakboekje
14. Haddad, O. B., Tabari, M. M. R., Fallah-Mehdipour, E., & Mariño, M. A. (2013). Groundwater Model Calibration by Meta-Heuristic Algorithms. *Water Resources Management*, *27*(7), 2515–2529. http://doi.org/10.1007/s11269-013-0300-9
15. Harbaugh, A. W. (2013). Chapter 2: Derivation of the Finite-Difference Equation. *Chapter 2*, 1–17.
16. Koen van der Hauw (1996), Evaluating and Improving Steady State Evolutionary Algorithms on Constraint Satisfaction Problems, Leiden University
17. Lin, W.-Y., Lee, W.-Y., & Hong, T.-P. (2003). Adapting Crossover and Mutation Rates in Genetic Algorithms. J. Inf. Sci. Eng., 19, 889–903.
18. Madsen, K. M., & Perry, A. E. (2010). Using Genetic Algorithms on Groundwater Modeling Problems in a Consulting Setting. *Proceeding of the Annual International Conference on Soils, Sediments, Water and Energy*, *15*(June), 103–114.
19. Mallawaarachchi, V. (2017). Introduction to Genetic Algorithms — Including Example Code. *Towards Data Science*, 1–12. Retrieved from https://towardsdatascience.com/introduction-to-genetic-algorithms-including-example-code-e396e98d8bf3
20. Mathworks India. (2018). Least-Squares (Model Fitting) Algorithms. *MATLAB Help*, 1–6. Retrieved from https://in.mathworks.com/help/optim/ug/least-squares-model-fitting-algorithms.html
21. Moradkhani, H., & Sorooshian, S. (2008). General Review of Rainfall-Runoff Modeling: Model Calibration, Data Assimilation, and Uncertainty Analysis. In S. Sorooshian, K.-L. Hsu, E. Coppola, B. Tomassetti, M. Verdecchia, & G. Visconti (Eds.), *Hydrological Modelling and the Water Cycle:*

Coupling the Atmospheric and Hydrological Models (pp. 1–24). Berlin, Heidelberg: Springer Berlin Heidelberg. http://doi.org/10.1007/978-3-540-77843-1_1
22. Reilly, B. T. E. (n.d.). System and Boundary Conceptualization in Ground-Water Flow Simulation.
23. Rijksdienst voor Ondernemend Nederland (2017), Natura 2000-beheerplan: Leenderbos, Groote Heide & De Plateaux
24. Sastry, K., Goldberg, D., & Kendall, G. (2005). Genetic Algorithms. *Compute*, 97–125. http://doi.org/10.1007/978-1-60761-842-3_19
25. Solomatine, D. P. (1998). Genetic and other global optimization algorithms - comparison and use in calibration problems. *3rd International Conference on Hydroinformatics*, 1021–1028. http://doi.org/10.1.1.2.2629
26. SOLOMATINE, D. P. , DIBIKE, Y. B., & KUKURIC, N. (1999) Automatic calibration of groundwater models using global optimization techniques, Hydrological Sciences Journal, 44:6, 879-894, DOI: 10.1080/02626669909492287
27. Stanhope, S. A., & Daida, J. M. (1998). Optimal mutation and crossover rates for a genetic algorithm operating in a dynamic environment. In V. W. Porto, N. Saravanan, D. Waagen, & A. E. Eiben (Eds.), Evolutionary Programming VII (pp. 693–702). Berlin, Heidelberg: Springer Berlin Heidelberg.
28. Thomas E. Reilly and Arlen W. Harbaugh (2004), Guidelines for Evaluating Ground-Water Flow Models
29. Tolson, B. A., & Shoemaker, C. A. (2007). Dynamically dimensioned search algorithm for computationally efficient watershed model calibration. *Water Resources Research*, *43*(1), 1–16. http://doi.org/10.1029/2005WR004723
30. Update databank en grondwatermodel (2014), Kwaliteitsrapportage, Royal HaskoningDHV
31. Vermeulen, P.T.M, L.M.T. Burgering, F.J. Roelofsen and B. Minnema, J. Verkaik, 2017. iMOD user manual. version 4.2, December 21, 2017. Deltares, The Netherlands. (http://oss.deltares.nl/web/iMOD).
32. Wang, H. H. F., & Anderson, M. M. P. (1995). Introduction to groundwater modeling: finite difference and finite element methods. http://doi.org/10.1016/0022-1694(83)90201-9
33. Wu, J., & Zeng, X. (2013). Review of the uncertainty analysis of groundwater numerical simulation. *Chinese Science Bulletin*, *58*(25), 3044–3052. http://doi.org/10.1007/s11434-013-5950-8
34. Yeh, W. W.-G. (2015). Review: Optimization methods for groundwater modeling and management. *Hydrogeology Journal*, *23*(6), 1051–1065. http://doi.org/10.1007/s10040-015-1260-3

9. Annex

9.1. Triwaco

Triwaco is a numerical flow and transport simulation package developed by the consultancy Royal Haskoning. In the Triwaco Flairs modeling environment, a combination of finite elements and finite differences is used by the inbuilt FLAIRS model code to compute groundwater heads/flow by iteration for a specific area and several possible scenario outputs can be generated. In the current study, the existing geological and hydrological data is in the form of a Triwaco Flairs groundwater model developed by the Waterboard De Dommel, Boxtel.

9.2. Triwaco to iMOD

9.2.1. Data conversion

- File Format
  While Triwaco Flairsling environment used .ADO file formats, iMOD uses .IDF grid file format for its model input files, which is easy to convert to from ASCII or Raster grid formats using GIS softwares like ArcGIS or QGIS. During all file conversions from Triwaco to iMOD, the final output file was ensured to be in .IDF format.

- KH, KV calculation
  - Hydraulic Conductivity (based on borehole and REGIS data)

  For the shallow model layers (the top 40 meters), borehole data (from DINOloket) was used to estimate the hydraulic conductivity of the upper 40 meters of the model layers.

  For the deeper layers as well as shallow layers where there is no borehole data available, transmissivity data from REGIS 2.1 model was converted to hydraulic conductivity data.

  Mathematical expressions based on soil fractions (from borehole data) and geological classifications used in the Triwaco Flairs were used to calculate the hydraulic conductivity per layer grid in the iMOD model (Appendix). An interpolation was first done from soil borehole data using ArcGIS 10.5 and then the expressions used in Triwaco Flairs were used to calculate hydraulic conductivity.

- RIV model implementation
  SOBEK river models for the Dommel and Tongelreep rivers were converted to ISG (iMOD Segment) files to feed into iMOD model. While there exists an option to directly import a SOBEK model to iMOD, it was decided to not do so to eliminate possible unmanageable errors during conversion. Instead, the SOBEK models were converted to ISG file format and then to IDF which is a grid file format recognized by iMOD.

  Similar conversion was done to the Drain package data files. The Drain dataset for secondary watercourses developed by Deltares was downloaded and used in the iMOD model.

9.2.2. Differences

- Discretization
  Due to the discretization of Triwaco Flairs following the finite element method with triangular elements, this allowed for a lot of flexibility in setting the model boundaries. It also allows for finer element sizes at sources/sinks or hydraulic barriers such as fault lines. On the other hand, iMOD uses a finite difference method for grid definition, without flexibility in grid sizes. The advantage of iMOD lies in the

fact that numerous sub-models of smaller or large resolutions than the original model can be made very easily without taking up memory space.

- KH,KV vs TX calculation

  While Triwaco directly calculates the Transmissivity values for layers from the borehole and REGIS geological data, iMOD requires the use of layer thickness to calculate the hydraulic conductivity values as the end result (model input). The layer thickness in Triwaco exists only for the discretization and viewing purpose whereas iMOD actively uses the layer thickness for linking hydraulic conductivity.

- Interpolation method (IDW vs TIN)

  While calculating KH, KV using mathematical expressions based on soil fractions and geology, Triwaco uses TIN interpolation method for estimating soil fraction thickness that is already built inside its modelling environment. On the other hand, IDW interpolation method is used via ArcGIS to generate the soil fraction thickness as iMOD does not have an inbuilt interpolation method.

- Fault lines representation

  Fault lines are represented as line elements similar to rivers with denser node density around them in Triwaco discretization grid. In the Triwaco Flairs for the study area, the faults are explicitly included as a line in the model nodes and, with a special allocator internally available in Triwaco, the effect of its presence in the permeability is be adjusted locally. This means extra nodes are included in the model grid along the fault lines to create a much finer discretization locally. The conductance along the fault lines are represented by a factor separately input per layer.

  However, iMOD does not have an internal allocator to do the same. iMOD on the other hand uses a package that simulates the presence of thin vertical low-permeability geologic features that impede the horizontal flow of groundwater. The resistance of these barriers can be fed to iMOD as a factor per layer.

- Calibration

  Triwaco uses manual (?) calibration wherein the calculated heads from a Triwaco Flairs model run are compared against a set of observations and difference maps can be displayed as a result, whereas iMOD uses its inbuilt calibration module to do the work.

9.3. RESULTS

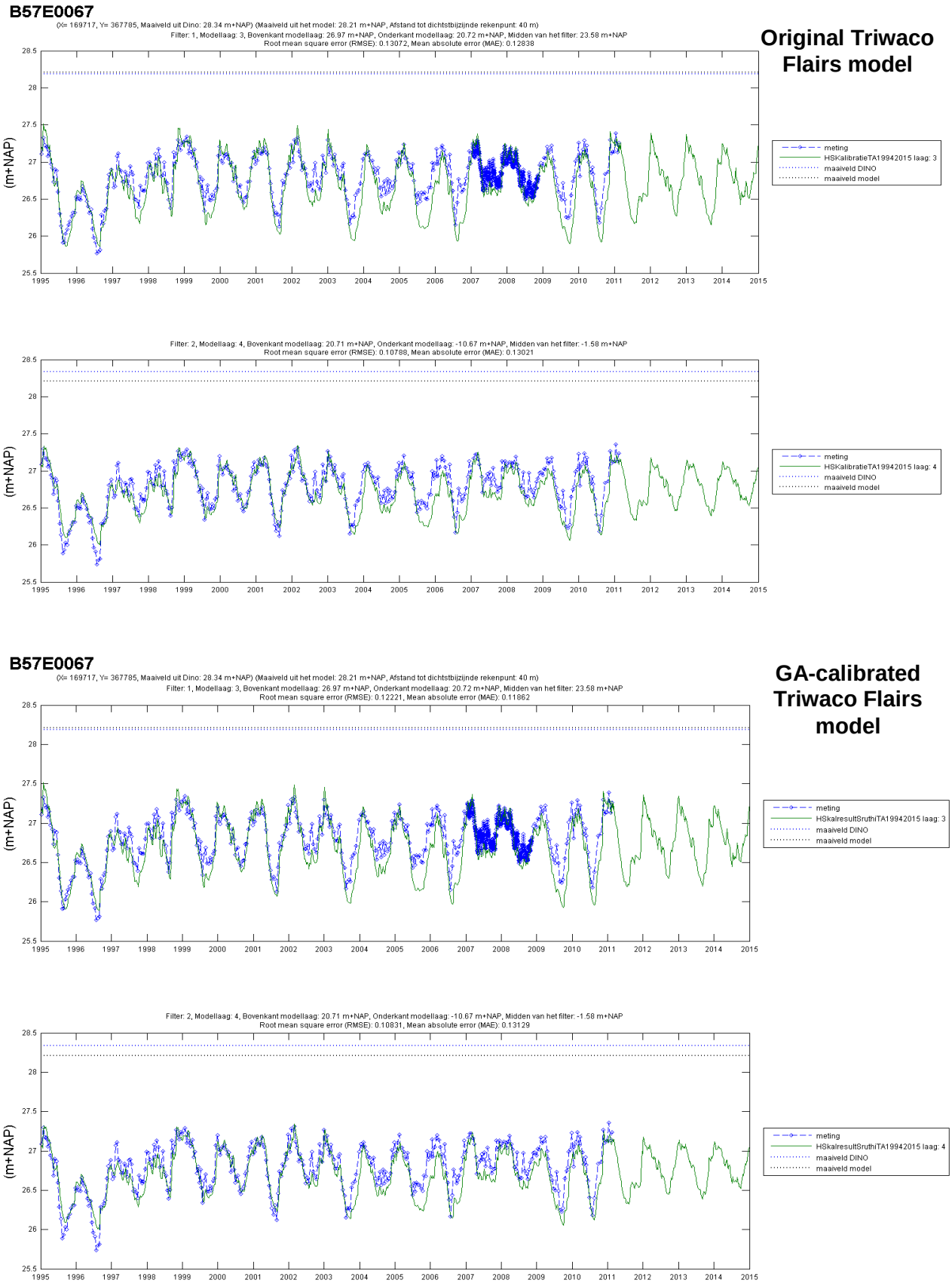
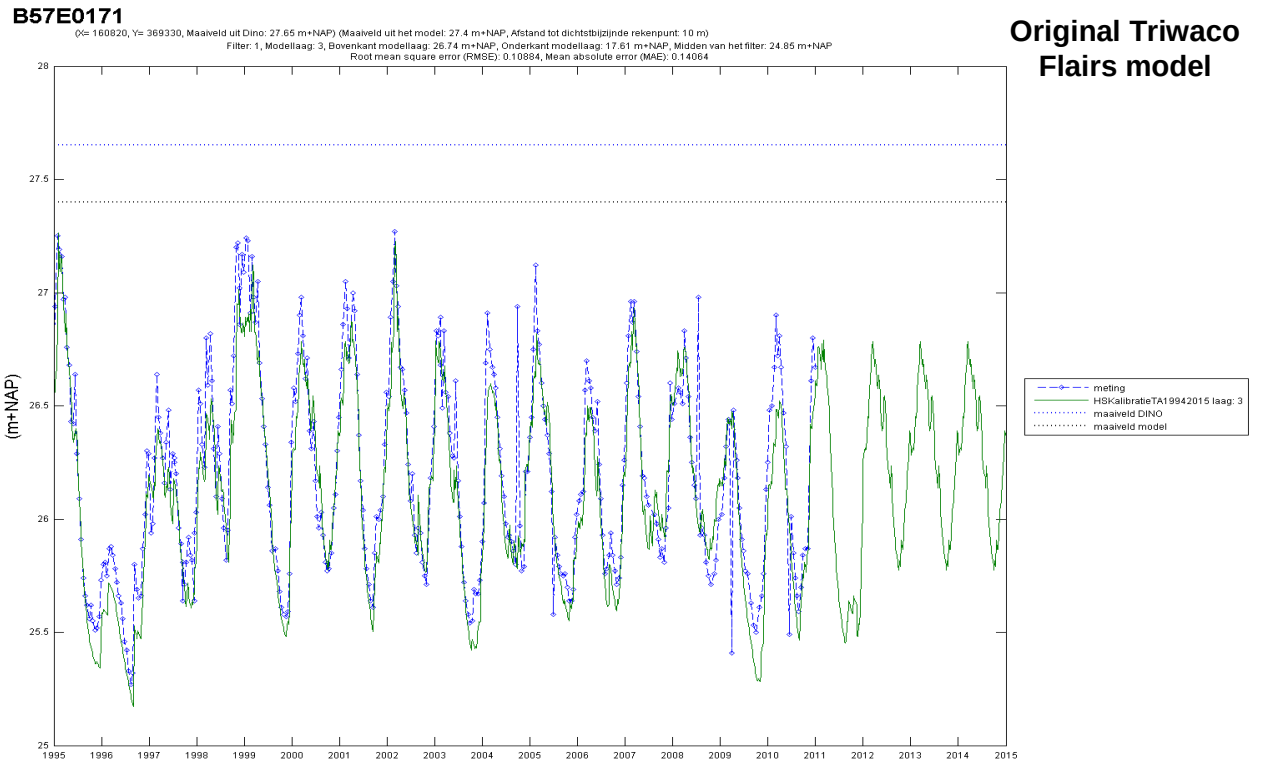


Fig. 9.1. Comparison of groundwater head variation in observation well ID: B57E0067 during the years 1994-2015 using Triwaco Flairs (Blue dotted lines represent observed head and Green lines represent modelled head)



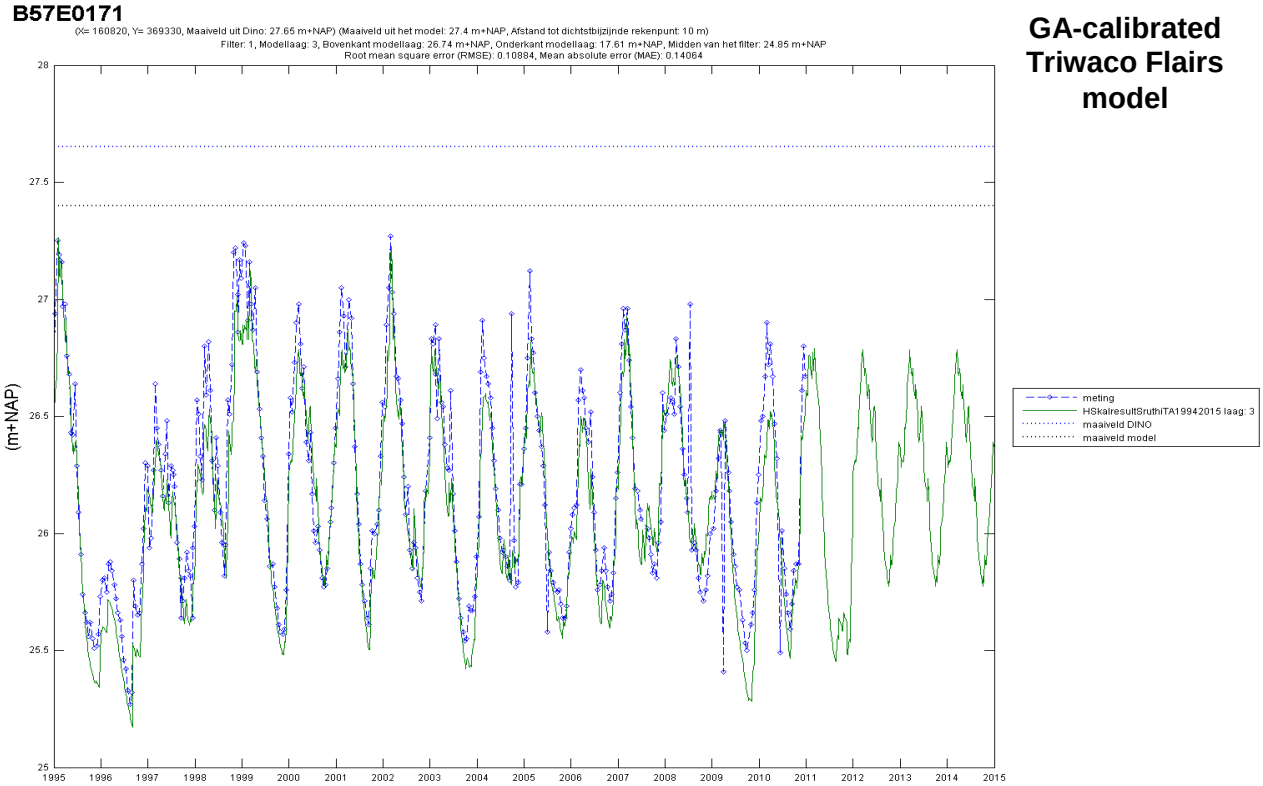


Fig. 9.2. Comparison of groundwater head variation in observation well ID: B57E0171 during the years 1994-2015 using Triwaco Flairs (Blue dotted lines represent observed head and Green lines represent modelled head)

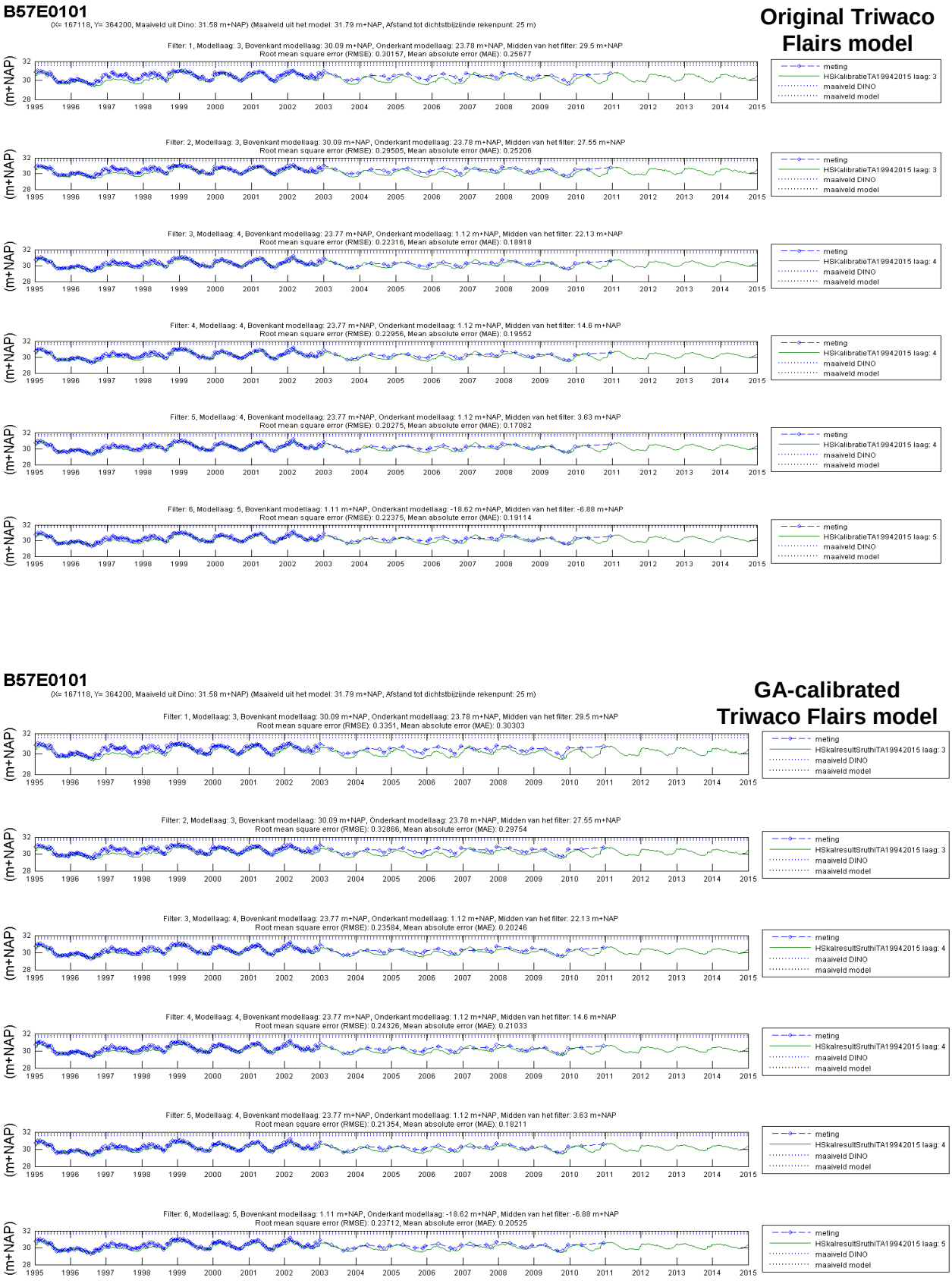


Fig. 9.3. Comparison of groundwater head variation in observation well ID: B57E0101 during the years 1994-2015 using Triwaco Flairs (Blue dotted lines represent observed head and Green lines represent modelled head)

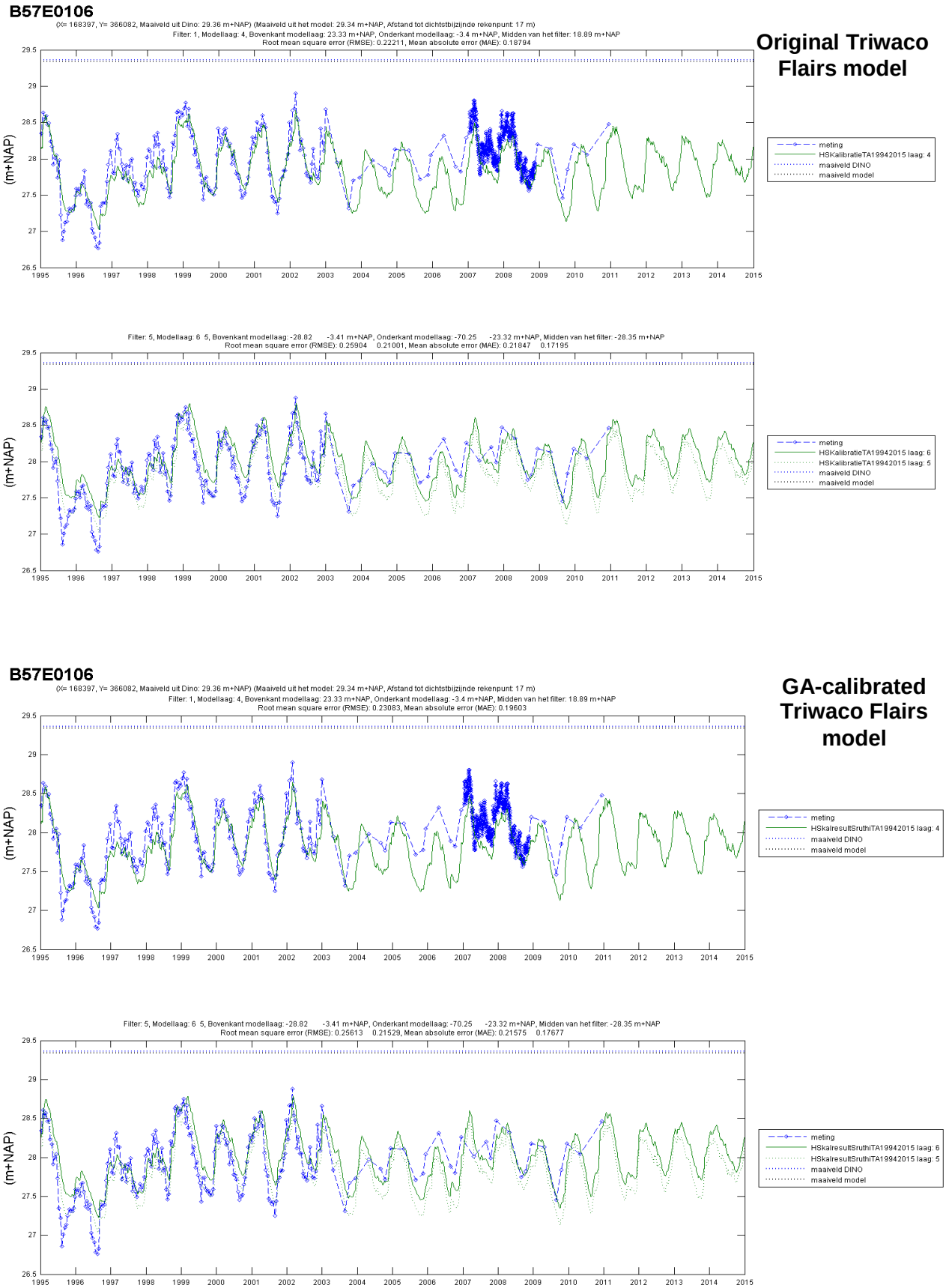


Fig. 9.4. Comparison of groundwater head variation in observation well ID: B57E0106 during the years 1994-2015 using Triwaco Flairs (Blue dotted lines represent observed head and Green lines represent modelled head)