HEATSTORE

Model validation for subsurface dynamics

Prepared by: François Diaz-Maurin (ed.), UPC
Maarten W. Saaltink (ed.), UPC

With contributions from:
Martin Bloemendal, KWR
Stijn Beernink, KWR
Niels Hartog, KWR
Mariëlle Koenen, TNO
Dorien Dinkelman, TNO
Luca Guglielmetti, University of Geneva
Alexandros Daniilidis, University of Geneva
Peter Alt-Epping, University of Bern
Sigrún Tómasdóttir, OR
Gunnar Gunnarsson, OR
Charlotte Rey, Storengy
Charles Maragna, BRGM
Isabella Nardini, Fraunhofer IEG
Florian Hahn, Fraunhofer IEG
Torsten Seidel, Delta-h
Geoffroy Gauthier, PlanEnergi

Checked by: Sebastia Olivella, UPC

Approved by: Marc Perreaux, Storengy, WP5 lead
Holger Cremer, TNO, HEATSTORE coordinator

Please cite this report as: Diaz-Maurin, F. & Saaltink, M.W. (eds.) 2021: Model validation for subsurface dynamics, GEOTHERMICA – ERA NET Cofund Geothermal. 110 pp.

This report represents HEATSTORE project deliverable number D5.3

HEATSTORE (170153-4401) is one of nine projects under the GEOTHERMICA – ERA NET Cofund aimed at accelerating the uptake of geothermal energy by 1) advancing and integrating different types of underground thermal energy storage (UTES) in the energy system, 2) providing a means to maximise geothermal heat production and optimise the business case of geothermal heat production doublets, 3) addressing technical, economic, environmental, regulatory and policy aspects that are necessary to support efficient and cost-effective deployment of UTES technologies in Europe.

This project has been subsidized through the ERANET cofund GEOTHERMICA (Project n. 731117), from the European Commission, RVO (the Netherlands), DETEC (Switzerland), FZJ-PtJ (Germany), ADEME (France), EUDP (Denmark), Rannis (Iceland), VEA (Belgium), FRCT (Portugal), and MINECO (Spain).

About HEATSTORE

High Temperature Underground Thermal Energy Storage

The heating and cooling sector is vitally important for the transition to a low-carbon and sustainable energy system. Heating and cooling is responsible for half of all consumed final energy in Europe. The vast majority – 85% - of the demand is fulfilled by fossil fuels, most notably natural gas. Low carbon heat sources (e.g. geothermal, biomass, solar and waste-heat) need to be deployed and heat storage plays a pivotal role in this development. Storage provides the flexibility to manage the variations in supply and demand of heat at different scales, but especially the seasonal dips and peaks in heat demand. Underground Thermal Energy Storage (UTES) technologies need to be further developed and need to become an integral component in the future energy system infrastructure to meet variations in both the availability and demand of energy.

The main objectives of the HEATSTORE project are to lower the cost, reduce risks, improve the performance of high temperature (~25°C to ~90°C) underground thermal energy storage (HT-UTES) technologies and to optimize heat network demand side management (DSM). This is primarily achieved by 6 new demonstration pilots and 8 case studies of existing systems with distinct configurations of heat sources, heat storage and heat utilization. This will advance the commercial viability of HT-UTES technologies and, through an optimized balance between supply, transport, storage and demand, enable that geothermal energy production can reach its maximum deployment potential in the European energy transition.

Furthermore, HEATSTORE also learns from existing UTES facilities and geothermal pilot sites from which the design, operating and monitoring information will be made available to the project by consortium partners.

HEATSTORE is one of nine projects under the GEOTHERMICA – ERA NET Cofund and has the objective of accelerating the uptake of geothermal energy by 1) advancing and integrating different types of underground thermal energy storage (UTES) in the energy system, 2) providing a means to maximize geothermal heat production and optimize the business case of geothermal heat production doublets, 3) addressing technical, economic, environmental, regulatory and policy aspects that are necessary to support efficient and cost-effective deployment of UTES technologies in Europe. The three-year project will stimulate a fast-track market uptake in Europe, promoting development from demonstration phase to commercial deployment within 2 to 5 years, and provide an outlook for utilization potential towards 2030 and 2050.

The 23 contributing partners from 9 countries in HEATSTORE have complementary expertise and roles. The consortium is composed of a mix of scientific research institutes and private companies. The industrial participation is considered a very strong and relevant advantage which is instrumental for success. The combination of leading European research institutes together with small, medium and large industrial enterprises, will ensure that the tested technologies can be brought to market and valorised by the relevant stakeholders.

Document Change Record

This section shows the historical versions, with a short description of the updates.

Version	Short description of change
1	First version, initiated by F. Diaz-Maurin (UPC) and M. W. Saaltink (UPC)
2	Filled by M. Bloemendal (KWR), M. Koenen (TNO), D. Dinkelman (TNO), L. Guglielmetti (U. of Geneva), A. Danillidis (U. Geneva), P. Alt-Epping (U. of Bern), S Tómasdóttir (OR), G. Gunnarsson (OR), C. Rey (Storengy), C. Maragna (BRGM), I. Nardini (IEG), F. Hahn (IEG), T. Seidel (Delta-h), and G. Gauthier (PlanEnergi)
3	Checked by S. Olivella (UPC)
4	Approved by M. Perreaux (Storengy, WP5 lead)
5	Final edited version (H. Cremer, TNO, HEATSTORE coordinator))

Table of Content

1 Introduction

Complex time-consuming numerical subsurface models are required for the design of Underground Thermal Energy Storage (UTES) systems. Such models imply large computation time and cost due to the large number of numerical evaluations required for each model. In the framework of the HEATSTORE project, UTES systems require to develop both (1) detailed numerical models of the subsurface dynamics requiring significant simulation time (WP2 – Modelization: Tools and processes to model underground flows), and (2) surrogate/proxy (simplified) models that can provide quick answers essential for the design optimization (WP3 – UTES Integration and optimization of the network) and monitoring (WP5 – Monitoring/Validation of the models for the system efficiency) of UTES systems.

Important modeling efforts have been performed as part of the HEATSTORE project are for example:

1. Development and benchmarking of detailed (3D) numerical models of subsurface heat storage dynamics fine characterization of UTES systems (with explicit account of heterogeneities) (Peter Alt-Epping and Mindel, 2020); and
2. Development and calibration of simplified (proxy) models that are analytical approximations of UTES systems for the design optimization and uncertainty analysis (with implicit account of heterogeneity) (Rohmer et al., 2020).

This report focuses on the validation of the detailed numerical models of the subsurface dynamics. Section 0 provides definitions used in the literature about model validation, presents validation methods, and lists examples of validation studies relevant to UTES applications. Section 0 then presents validation tests and experiments conducted on numerical models used for the different HEATSTORE case studies. Finally, section 4 summarizes and discusses the validation case studies.

2 Terminology, process and examples

2.1 Terminology

The validation of numerical models of subsurface dynamics has been the subject of much scientific discussions (e.g., Konikow, 1992; Konikow and Bredehoeft, 1992; Oreskes et al., 1994; Rykiel, 1994; Sterman, 1994; Tsang, 1991; Younker and Boak, 1994). Yet, as observed by Hassan (2004), most of the controversy over validation comes from discrepancies about the meaning and purposes of the term. In discussions about model validation, different terms are commonly used such as validation, verification, calibration, and confirmation. Yet, because the terms verify, validate, confirm, substantiate, etc. are essentially synonymous in ordinary language, they are often used interchangeably (Hassan, 2004). This situation thus creates confusion about whether a "validated model" can actually make reliable predictions for decision making. For this reason, it is important to make the distinction between calibration, verification, and validation to support the discussion of model validation.

- Calibration

Model calibration (also called *benchmarking*) is the "process of tuning the model to identify the independent input parameters by fitting the model results to field or experimental data" (Hassan, 2004). As such, calibration is part of the characterization-conceptualization-calibration-prediction loop of the modeling process which includes site characterization activities to determine model parameters (Tsang, 1992). Calibration does not seek to assess the validity of a model used to make predictions because it is only a limited demonstration of the reasonableness and reliability of the model for a set of experimental conditions. Calibrated models, alone, thus do not have predictive value for decision-making because they require to go through verification and validation procedures. Yet, a calibrated model can still be used to make predictions, test the model and *invalidate* it, which will lead to an improved model and understanding (Konikow, 1992), although this does not constitute a model validation process.

- Verification

There is a clear distinction between *code verification* and *model verification* (Hassan, 2004). Code verification and testing is limited to checking the correctness of the computer code with respect to the criteria and requirements for which it has been designed. Code verification thus consists of comparison of a numerical solution with one or more analytical solution(s) or with other numerical solutions (Anderson and Woessner, 1992). Code verification ensures that the computer program accurately solves the equations that constitute the mathematical model. The verification of governing equations can be performed through a series of verification tests. In numerical models, however, only their mathematical components may be verifiable, just as a code (or algorithm) within a computer program may be verifiable (Oreskes et al., 1994). That is, mathematical components may be subject to verification because they are part of closed systems. However, the models that use these components are never closed systems, thus they cannot be verified in the same way a code does.

Model verification is a completely different process than code verification (Hassan, 2004). Model verification can be defined as a process aimed at establishing a greater confidence in the model by using a set of calibrated parameter values and stresses to reproduce a second set of field data (Anderson and Woessner, 1992). That is, a model is said to be verified if it is demonstrated that its accuracy and predictive capability lie within acceptable limits or error by tests independent of the calibration data (Konikow, 1978). In this sense, model verification can be understood as a form of *strong validation* (see next section). In practice, however, it is impossible to verify a model because only one set of field data is available which is already needed for the calibration step. Therefore, a calibrated-unverified model can still make predictions but this requires that sensitivity analyses of both the calibrated model and the predictive model are performed and assessed (Anderson and Woessner, 1992). In this case, the predictions can serve to demonstrate the model's accuracy and predictive capability to mimic *past* behaviour, using historical data. As the model verification process is part of the development stage of the model, modelers can thus modify the model conceptualization if the calibrated model fails to reproduce the verification data set. Such *ad hoc* modifications of a calibrated model to fit a new data set, however, are not possible in a formal process of model validation.

- Validation

Whereas model calibration and verification can demonstrate the model's accuracy through historical data validation, the *model validation* process determines whether the model can predict future behaviour, that is,

whether it has some predictive capability. Validation, however, is probably the most ambiguous of the three terms discussed here. Validation can either suggest that a model is an accurate representation of physical reality. Validation can also mean verification when model predictions are consistent with observational data. Finally, validation can be used to determine that the model is adequate for its intended use. Depending on the context, validation can thus refer to the establishment of truth, accuracy, or legitimacy, respectively. Yet, in ordinary language, if a model is said to be validated, the immediate logical inference is that the model can make reliable predictions (Konikow, 1992). In his review, Hassan (2004) distinguishes between the various interpretations and perceptions of the meaning of the term validation—from the inherently unattainable "proof of truth" (scientific and philosophical views of validation) to more pragmatic approaches (operation and confidence-building views of validation) with a subjective assessment of whether a model is good enough to support decision making.

Attempting to summarize these different definitions, Hassan (2004) proposed the qualifiers *weak* and *strong* to refer to model and model validation. Figure 2.1.1 provides a visual representation of the relationship between model development and model validation applying the weak/strong distinction. In this view, *weak model* refers to a model that is an early stage of development. A weak model includes mathematical equations and simplifying assumptions, but little or no input data. A weak model can be used in an analysis mode to test hypotheses, explore designs, or increase understanding. But because of its simplifications, a weak model can only pretend to a weak form of validation, such as code verification. In contrast, a *strong model* includes the previously mentioned model components as well as the parameter values, boundary conditions and system geometry. A strong model is thus intended to simulate the full system with all its known properties and expected governing equations. A strong model can then go through the next steps of the modeling process, that are the model calibration, history matching, benchmarking, and testing. As mentioned earlier, a calibrated model can be used to make predictions to test the model's accuracy in predicting historical data, but they do not have predictive value for decision-making. Calibrated models can therefore only pretend to a weak form validation, like model verification tests.

Once a model successfully passes all calibration, benchmarking and verification tests, it can then be used to predict system behavior under modified conditions. Although good calibration does not necessarily imply equally good prediction (Hassan, 2004), a calibrated and verified model may be considered sufficiently accurate so that it can be used to make predictions and go through validation procedures (see section 2.2). Whereas "absolute validation" neither is theoretically possible nor is a regulatory requirement (Hassan, 2004), *strong validation* can be achieved by using new data sets to iteratively test the validity of a model. Given the high cost of acquiring new data for testing, however, the strong validation process requires an equally strong commitment to testing by the regulators, decision-makers, or users. Such an iterative approach to validation can build confidence in the modeling process and in the reliability of a model, which is a critical aspect of the decision-making and regulatory processes.

Finally, a strong model that followed validation procedures will contribute to increasing the scientific knowledge and understanding, independently of whether the model successfully passed or failed the validation tests. Eventually, any model is ultimately "invalidated", either because new scientific knowledge (theories) become available or because the model cannot predict or explain new field data acquired during operation and monitoring. In the latter situation, one shall not continue to work on the model until it achieves a fit (Oreskes et al., 1994). Instead, the whole modeling process must start again with a new characterization-conceptualization-calibration-prediction loop accompanied with verification and validation procedures. Scientific knowledge refers to the fundamental understanding of the physical and chemical processes happening in nature that are formulated by laws and theories. In the scientific process, like in the modelling process, theories are always tested so they are considered to be "true" until they are "falsified" by new experimental data (Popper, 1959). For this reason, scientific knowledge at any moment in time can be considered being strongly valid, whereas theories are weak representations (models) of reality because of their high level of abstraction and generalization.

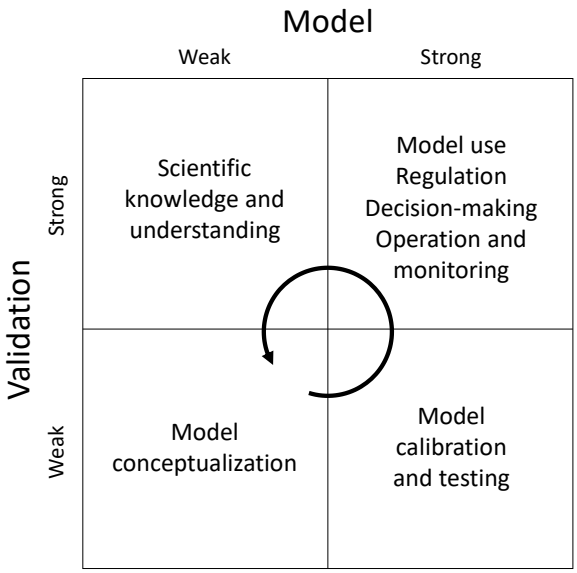


Figure 2.1.1: Relationship between model development and model validation. Source: F. Diaz-Maurin, after definitions in (Anderson and Woessner, 1992; Hassan, 2004; Tsang, 1991).

2.2 Validation process and methods

Although confidence is not synonymous with validation (Konikow, 1992), users, decision-makers and regulators still need to be confident that the model results used to support decision making are considered sufficiently reliable for the set of conditions they have been applied. The assessment of the reliability of model predictions is thus a basis for decision-making because it builds confidence in the models. Yet, just as confidence building is a long-term, iterative process, *model validation* is to be understood as a process which is an integral part of the model development process (Hassan, 2004; Tsang, 1991). Therefore, an important aspect of supporting and guiding a model-based decision-making process is to develop a validation process that allows to continuously test and refine the model. Yet, as observed by Hassan (2004), most of the debate in the literature over validation has focused on the terminology used rather than on defining a process of validation (which applies to Hassan's paper itself). Few have attempted to propose validation processes for models. For instance, Tsang (1991) proposed to consider model validation as an integral part of the modelling process. In his approach, validation needs to be carried out at every step of the modelling process. To achieve validation, he recommends that a multiple assessment group approach is adopted (Tsang, 1991), doubled with peer review and open discussions at every step of the modelling process (Tsang, 1992). However, in practice, even when adopting a rigorous set of validation procedures on a highly detailed and reasonably accurate characterization of the subsurface parameters, achieving validation may still be difficult. First, in presence of significant heterogeneity and limited data, it is not possible to make an objective judgement about the accuracy of model predictions and thus whether there is sufficient confidence in the model (Hassan, 2004). Moreover, even when there is enough data available, selecting the quantitative criteria on which to base the comparison between predicted and measured values (i.e., the quantitative measure of validation) requires consent or agreement that may be difficult to attain. Last, when validation procedures indicate discrepancies between predictions and measurements, it may be difficult to attribute these discrepancies to inadequacies in either the conceptual model, mathematical formulation, or input data.

These difficulties have led most definitions in the literature about validation of models used in geosciences to acknowledge the need to demonstrate that a model is a good or sufficient representation of reality, and that this demonstration requires to make subjective decisions of the validation criteria. However, they have not sufficiently recognized the need to measure the accuracy of model calculations (Hassan, 2004). The most significant effort to provide a quantitative assessment of model validation comes from the field of systems engineering and operations research (or decision science). Most notably, Roy and Oberkampf (2011) offers a practical approach to model validation through the assessment of model accuracy by way of quantitative comparison of simulation results with experimental data.

Because models cannot be validated in the absolute sense without qualifications (Tsang, 1992), validation thus requires further qualifier (e.g., a process, a site-specific system) as to the conditions for which the model has been validated and those for which it should not be used (Anderson and Woessner, 1992). When an assessment of the performance of a model is carried out, it is necessary to state explicitly what performance measures has been used as one of the qualifiers of the validation.

In his review, Hassan (2004) presents four main strategies to model validation proposed in the literature in groundwater flow and solute transport simulations. We add a fifth validation strategy coming from the literature in systems engineering that may be relevant to geothermal energy storage applications.

Strategy #1: Tailored 10-step validation experiment

First, Davis and Goodrich (1990) proposed a validation procedure that seeks to demonstrate model accuracy in simulating the real system's behaviour in 10 steps, given pertinent regulatory criteria have been selected:

1. Define a validation issue.
2. Develop a conceptual model or models.
3. Develop a mathematical model.
4. Identify and/or design an experiment that addresses the validation issue.
5. Define performance measures to be used for model comparisons.
6. Quantify the uncertainty associated with the input data and the data available for comparison with the model output.
7. Define the acceptance criteria or acceptable model error based on regulatory requirements and data uncertainty.
8. Simulate the experiment.
9. Perform the experiment in the laboratory or field.
10. Evaluate model results based on the acceptance criteria.

Although very practical, these validation procedures are not intended to an iterative process of predictive model validation (strong validation) unless they are run several times for different validation experiments using independent data sets.

Strategy #2: Statistical testing

Another validation method was proposed by Flavelle (1992) that focuses on the quantitative evaluation of model accuracy when calibrating and validating a model. This method consists in:

1. Performing a regression analysis of predicted values and measured data with the regression coefficient of the regression line interpreted as an empirical indicator of model bias and the standard error interpreted as the uncertainty in the validation.
2. Then, statistically testing the assertion that the slope of the regression line is unity and that the intercept of the line is 0. Hypothesis testing can be used for this purpose with the null hypothesis for the slope.

This approach compares the validation and calibration statistics so that it is possible to assess whether simulated conditions have changed and, thus, whether the model adequately accounts for all significant processes or not. The method considers a normal distribution of the regression residuals which can be considered as a more reasonable assumption than considering a normal distribution of the deviations between predicted values and observed values. As pointed by Hassan (2004), in addition to the linear regression analysis, other statistical tests, hypothesis tests, and qualitative tests are all necessary tools for the assessment of the model input, structure, and output. In fact, an important aspect of a validation process must be that it ensures that a diversity of statistical tests are being carried out and a diversity of model features are being tested.

More recently, a relatively similar approach of model validation has been proposed that is based on Bayesian updates and prediction related rejection criteria (Babuška et al., 2008). In this approach the prior density is related to the current candidate model and the posterior density is obtained by conditioning on the validation experiments. Although this approach comes from systems engineering, Bayesian approaches specific to uncertainty quantification of subsurface systems have also been developed (Scheidt et al., 2018) that allow to assess the value of acquiring new data about the system to support the decision-making process (Eidsvik et al., 2016).

Strategy #3: Stochastic approach by error decomposition
Luis and McLaughlin (1992) proposed a stochastic approach to model validation that relies on decomposing the differences between the predicted and measured values of the variable of interest into three components or error sources: (1) measurement errors, (2) spatial heterogeneity, and (3) model error. By expressing measurement residuals in terms of these three components, the authors use a perturbation analysis and derive the relationship between the measurement residual variance, actual variable variance, and measurement error variance that is only related to the measuring device (Hassan, 2004). However, this relationship assumes that model errors are negligible, and once developed, the relationship can be used to develop statistical tests that check the hypothesis that the model error is indeed negligible (see strategy #2).

Another approach to error decomposition was used by Mummert (1996) that consists in a point validation method where accuracy for point predictions is assessed by calculating the coefficient of determination, the relative error, and the standard error.

Strategy #4: Statistical distribution of model predictions
In another approach, Mummert (1996) used Monte Carlo simulations to obtain the statistical distributions of model predictions. Once the distributions are obtained, the hypothesis that the field data represent reasonable samples from the distribution of model predictions is tested by checking if observed values lie within the 5% and the 95% quantiles of the distribution.

Strategy #5: Nondeterministic predictive uncertainty quantification
Another approach to error (uncertainty) decomposition was proposed in the field of systems engineering by Roy and Oberkampf (2011). This approach consists of a series of procedures of verification, validation and uncertainty quantification (VV&UQ) for estimating the predictive uncertainty of numerical simulations treating both types of uncertainty (aleatory and epistemic). In this approach, aleatory (random) uncertainties in model inputs are treated as random variables, whereas epistemic (lack of knowledge) uncertainties are treated as intervals with no assumed probability distributions (like in strategy #4). First, numerical approximation errors (due to discretization, iteration, and computer round off) are estimated using different verification techniques. Then, model form uncertainty is quantified using model validation procedures (e.g., statistical comparisons of model predictions to available experimental data) and through the extrapolation of this uncertainty structure to points in the application domain where experimental data are not available. This approach to validation implies a paradigm shift from a deterministic to a nondeterministic approach to numerical simulations where a single solution to the mathematical model is no longer sufficient and where, rather, a set of calculations must be performed to map the uncertain input space onto the uncertain output space.

The advantage of the VV&UQ framework is that it allows to show clearly and distinguish how the different sources of uncertainties contribute to the uncertainties in predicted values of interest, thus facilitating the decision-making process. A new, open source tool called EasyVVUQ has been development that facilitates the verification, validation and uncertainty quantification of numerical simulations (Richardson et al., 2020).

2.3 Relevant examples

In subsurface systems applications, the long-term validation process typically seeks to ensure that the predicted consequences are not underestimated. In this case, confidence means that model-based decisions will not result in unacceptable risks to present or future populations (e.g., radioactive waste geological disposal) or degradation of the natural environment (e.g., ground water management). But not all subsurface dynamics models need to undergo a validation process, nor all models requiring validation need the same level of rigor in the process (Hassan, 2004). The determinant factor often is the amount of risk (financial, health and environmental) associated with possibly making a wrong decision based on results of an “unvalidated” model against the cost associated with going through a rigorous validation process. In UTES applications, not all subsurface dynamics models will require the same level of validation depending on the temperature range (e.g., high- vs. low-temperature), design (e.g., aquifer-, borehole-, pit-, or mine-based designs) and development stage (e.g., experimental, demonstration, commercial) that may imply different risk levels. Generally, in UTES applications, the risks entailed may be considered as relatively low. However, the reliability of the models still needs to be assessed to ensure that the estimated output results of relevance to decision-makers and investors (e.g., transient pressure, temperature profile) are not overestimated. Ultimately, the level of model validation—as well as the choice of criteria to demonstrate such validation—shall be decided by the future users, decision-makers, and regulators.

The HEATSTORE project focuses on high-temperature (HT) UTES systems that are at different development stages (Kallesøe and Vangkilde-Pedersen, 2019). Therefore, the validation requirements may differ between the systems (see section 0). In his review, Hassan (2004) discusses several subsurface model validation studies (mostly in the fields of nuclear waste management and hydrogeological sciences) that used one or several of the validation methods presented in the previous section. Other validation studies not cited by Hassan (2004) but of potential relevance to UTES applications include: Robinet et al. (1999), Olivella and Gens (2005), Gens et al. (2009), and Nishimura et al. (2009). Below, we provide an overview of these examples and characterize them according to the terminology of section 2.1 and validation methods and strategies of section 2.2.

2.3.1 Robinet et al. (1999)

Robinet et al. (1999) present a set of validation tests of numerical models used to simulate the thermo-hydro-mechanical (THM) behavior of radioactive waste geologic repositories as part of the EU project CATSIUS CLAY. The validations tests were organized in the form of benchmarks based on analytical laboratory and in-situ experimental results. First, it verified algorithms for solving equations of mass transfer (exercise 1-1) and heat transfer (exercise 1-2) by comparing them to analytical solutions. Second, it validated rheological models and THM calculation codes at laboratory scale by comparing them to the results of oedometer tests (exercise 2-1) and a reduced-scale model (exercise 2-2). Finally, the third step consisted of a comparison of HM (exercise 3-1) and THM (exercise 3-2) calculations with the full scale FEBEX (Full-scale Engineered Barriers EXperiment) experiment in crystalline host rock. System quantities for the model comparisons included the injected water through the confinement cell (in kg), the total thermal power provided by the two heaters (in W), as well as temperature (in °C), relative humidity (in %) and radial stress (in MPa).

The code comparison for the decoupled THM problems (exercises 1-1 and 1-2) showed that the iterative algorithms used (Newton Raphson, Picard, and Picard associated with the line search technique) correctly address problems of mass transfer and heat transfer affected by strong non-linearity over long time scales (over several thousands of years) with numerical solutions systematically converging with little deviations between the various algorithms. This apparent robustness of the numerical solutions was possible because of the slow evolution of the non-linear problems over time that allowed to use iterative algorithms. The second step (exercises 2-1 and 2-2) allowed to compare two approaches for the elasto-plastic rheological models where the hardening is produced either by volumetric deformations or suction. The model comparison on compacted unsaturated swelling clays showed that the two approaches gave very similar results and were able to simulate experimental data. Finally, the coupled HM and THM problems (exercises 3-1 and 3-2) showed that full-scale simulations resulted in significant deviations between numerical results and experimental data. Moreover, it was shown that these deviations of numerical results grew larger with the increasing number of parameters used when increasing the coupling and that these discrepancies were enhanced by the treatment of transient problems over the long term.

Using the terminology of section 2.1, all validation tests presented by Robinet et al. (1999) correspond to *code verification* by the intercomparison of several numerical solutions in solving the same problems, as well as by the comparison of these numerical solutions to analytical solutions and to observations. The code comparison of decoupled THM problems (exercises 1-1 and 1-2) and of elasto-plastic rheological models (exercises 2-1 and 2-2) corresponds to a case of *weak validation* of *weak models*, whereas the comparison of HM and THM numerical results to the full-scale FEBEX experiment (exercises 3-1 and 3-2) represents a case of *weak validation* of a *strong model*. As per the validation process (section 2.2), no formal validation strategy and procedures were followed other than through the direct visual comparison of numerical results with laboratory and in-situ experimental data.

2.3.2 Olivella and Gens (2005)

Olivella and Gens (2005) present thermo-hydro-mechanical (THM) analyses, simulating the in situ heating test Drift Scale Test (DST), performed at the Yucca Mountain nuclear waste repository project in Nevada. THM coupling was achieved through a double structure modeling approach based on the superimposition of two meshes. This modeling strategy was used to account for the role of double porosity, hence, double permeability from the contribution of the matrix and the fractures in the hydrological problem. In fact, in tuff, fractures can desaturate at very low capillary pressures (few kPa) whereas the matrix requires higher capillary pressures to desaturate (hundreds of kPa) due to smaller pores. In order to test the accuracy of their modeling approach, Olivella and Gens (2005) considered different cases of intrinsic permeability variations and their influence on

the calculated temperatures, degree of saturations and gas permeabilities. The different cases of intrinsic permeability variations corresponded to different levels of coupling and included (1) a minimum coupling, where the intrinsic permeability is considered constant (i.e., independent deformation changes) and where the mechanical problem is solved using a standard linear thermo-elastic model (BASE CASE); (2) a moderate coupling, where variable intrinsic permeability variations are considered following a cubic law and deformations are calculated in the same way (CASE 1); and (3) a stronger coupling, where intrinsic permeability variations are introduced via a cubic law too, but deformations are calculated with a thermo-elastic model modified with a dilatancy term (CASE 2). Numerical calculations for the three cases were then compared to in situ measurements of temperatures and gas permeabilities collected over a period 4 years.

Results showed that, when intrinsic permeability variations were introduced (CASES 1 and 2), both temperatures and gas permeabilities were influenced by deformations, translating a coupling strategy from mechanical to hydraulic and then to thermal. The comparison of measurements of temperatures and gas permeabilities with calculated results in CASE 1 seemed to indicate that the reduction of intrinsic permeability was overestimated (shorter calculated temperature stabilization periods at 100°C and larger reduction of calculated gas permeabilities). However, CASE 2 showed that a modified version of elasticity that includes dilatancy significantly improved the simulation of overall system behavior and showed better agreement of model predictions of gas permeability with measurements, even in zones where hydrological effects are small. Overall, the model was able to reproduce the measured gas permeability variations, when shearing-induced dilatancy was included, which seemed to demonstrate that the double structure approach proposed by Olivella and Gens (2005) was appropriate to simulate the DST test.

The three analyses presented in Olivella and Gens (2005) were performed to test the relative appropriateness of three different modeling approaches to simulate THM behavior of a in situ heating test. As such, they do not constitute a model validation exercise but rather a model testing exercise as part of the model development process (see section 2.1). No benchmarking of the model parameters was performed as hydraulic, thermal and mechanical properties of tuff for the three different rock units were considered to be known from other studies. Also, no formal validation strategy was followed in this study (section 2.2).

2.3.3 Gens et al. (2009)

Gens et al. (2009) describe the performance, observations and interpretation of the FEBEX (full-scale engineered barrier experiment) in situ test, a full-scale heating test carried out at the Grimsel test site (GTS), an underground laboratory excavated in granite rock in the Swiss Alps. The FEBEX in situ heating test is intended to simulate a geological disposal concept for heat-emitting, high-level nuclear waste. The test focuses on the thermo-hydro-mechanical (THM) behavior of the near-field region constituted by the compacted bentonite barrier surrounding the heater and the immediately adjacent rock. The study discusses the thermal, hydraulic and mechanical observations in the bentonite barrier and in the host rock over the five-year period of the heating stage. Several parameters were measured, including temperatures, relative humidity (total suction), pore water pressures, total pressures and displacements.

Because in situ tests in underground laboratories are fully instrumented, they provide a large amount of independent data required to *calibrate* (section 2.1) the parameters of the coupled THM numerical model adequately. In situ tests are also used to validate, using field observations, coupled THM formulations and associated computer codes. This form of *model validation* (see section 2.1) consists in the comparison between predictions and observations. In the study, test observations were plotted alongside the predictions of the numerical simulation to assist in the interpretation, and to assess both the reliability of the test results and the performance of the model. The model validation was conducted *a posteriori*, that is, after the heating test was completed and dismantled, thus observations collected over the five-year period corresponded to historical data (called “post-mortem data” in the study) to which the model predictions were compared.

Gens et al. (2009) correctly observe that the two steps of model calibration and prediction have to use independent observational data sets (see section 2.1). In their study, model calibration used the best information available on initial and boundary conditions and on material parameters at the time, whereas the evaluation of the predictive capability of the model used observations collected over the five-year period of the heating test and the following dismantling. Because of these procedures, the model validation presented by Gens et al. (2009) could be seen as a form of *strong validation* (section 2.1). However, although the authors conclude that their study has proved the capability of the numerical formulation to provide adequate predictive capacity, it did not follow an iterative process of model validation using independent validation tests (see

section 2.2). Moreover, their study did not define what the acceptance criteria of an adequate model are, not it provided a quantitative assessment of model errors against observations.

2.3.4 Nishimura et al. (2009)

In another model development effort, Nishimura et al. (2009) present a fully coupled thermo-hydro-mechanical (THM) finite element (FE) formulation aimed at simulating freezing and thawing of pore fluid in water-saturated soils. The model formulation relies on combining ice pressure, liquid pressure and total stress as state variables to account for frozen and unfrozen behavior within a unified effective-stress-based framework. After introducing the governing equations of thermodynamic equilibrium, mass/heat transfer and mechanical equilibrium, the authors formulated the constitutive model through the combination of the two stresses (net stress representing external confinement and suction) along with the deviatoric stress. This model formulation was then tested for shear strengths and freeze-thaw cycles through comparison of the model simulations with previously reported experimental results.

The performance of the THM model and its numerical implementation were then evaluated with reference to published pipeline frost heave experiments. First, the computed liquid flux and the liquid flux, separately calculated from the segregation potential (a coefficient relating the liquid flux into the frozen fringe and temperature gradient across it) theory, were compared. It was shown that the THM model predicted the same order of liquid flux as the segregation potential method. Second, the simulated pipeline heave developments were compared with the field measurements. The analysis showed that the simulations generally predicted substantial suppression of heaving up to a certain point in time, but that they started to overpredict observations from this point onwards, leading to a significant final overprediction of heave. In another analysis, the measured water content profile at the pipe centreline was compared to the simulated water contents. Finally, the THM model's ability to predict the stress and strain states developed in the soil, and their variations, during freezing and thawing was tested. After careful examination of the analytical predictions and the field test data, the authors concluded that the THM model could simulate, with fair accuracy, the field patterns of pipeline heave, water migration and ice accumulation.

Using the terminology of section 2.1, the first set of analyses performed by Nishimura et al. (2009) correspond to *model verification* tests that seek to validate the accuracy of the model formulation in reproducing a set of field data. However, no formal validation strategy was adopted (section 2.2).

3 Case studies

3.1 Koppert-Cress, Monster, the Netherlands

Koppert Cress is horticulture company specialized in the production of cress located in the western part of the Netherlands. The ATES system of Koppert Cress obtained a pilot license in 2015 and started to inject heated groundwater with temperatures above the conventional maximal injection temperature of 25 °C (Bloemendal et al., 2020; Bloemendal et al., 2019). As this is one of the few HT-ATES locations in the Netherlands, extensive monitoring and analysis of the ATES system operation was performed. The greenhouses of Koppert Cress have a relatively large heat demand in winter, compared to their cooling demand in summer. Therefore, excess heat from several sources is to be stored in the warm wells in summer and used in winter. This comprises of multiple 'passive' heat sources from e.g. solar panels, aquathermal heat generation and waste-heat from a combined heat and power (CHP) plant. After the start of the transition from LT-ATES to HT-ATES in 2015, these heat sources were gradually added to the heating and cooling system (Bloemendal et al., 2020).

3.1.1 Site description

The wells of the ATES system of Koppert Cress are located around their main greenhouse (Figure 3.1.1). The first aquifer in the subsurface at the Koppert Cress location is not available for the ATES system because this aquifer is reserved for application of fresh water storage and recovery, a technology many greenhouse also use for their fresh water supply. The deeper aquifers consisting of the Oosterhout and Maassluis formations are less frequently used compared to the shallow aquifer, resulting in limited data and uncertainty on their characteristics. The ATES system utilizes 2 aquifers of 20m thickness with screens up to ±170m depth (Figure 3.1.2). With 4 warm and 4 cold wells (Figure 3.1.1), in total 16 well screens are used for the ATES system.

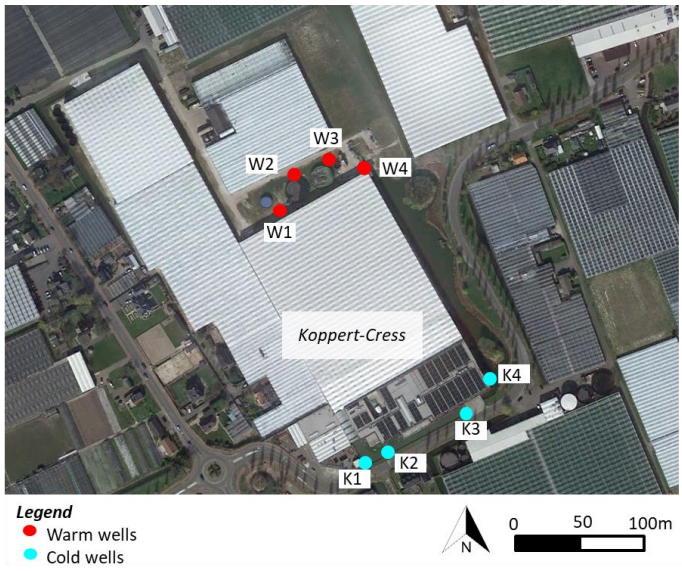


Figure 3.1.1: Overview of Koppert-Cress site with the warm and cold well locations. The individual warm and cold wells are 40 to 50m apart. The distance between the cold and warm wells is about 250m.

To obtain insights in the performance, heat spreading and water quality changes associated to the ATES with increased storage temperatures, the ATES operational data was analysed and the site was intensively monitored with Distributed Temperature Sensing (DTS) at 4 locations near warm well 1 (Figure 3.1.2). This allows to monitor the temperature profile and heat distribution around the warm well.

The goal of this study was to validate the SEAWATv4 groundwater model with the monitored operational and DTS data, and improve the insights in the operational and subsurface characteristics of the Koppert Cress ATES system.

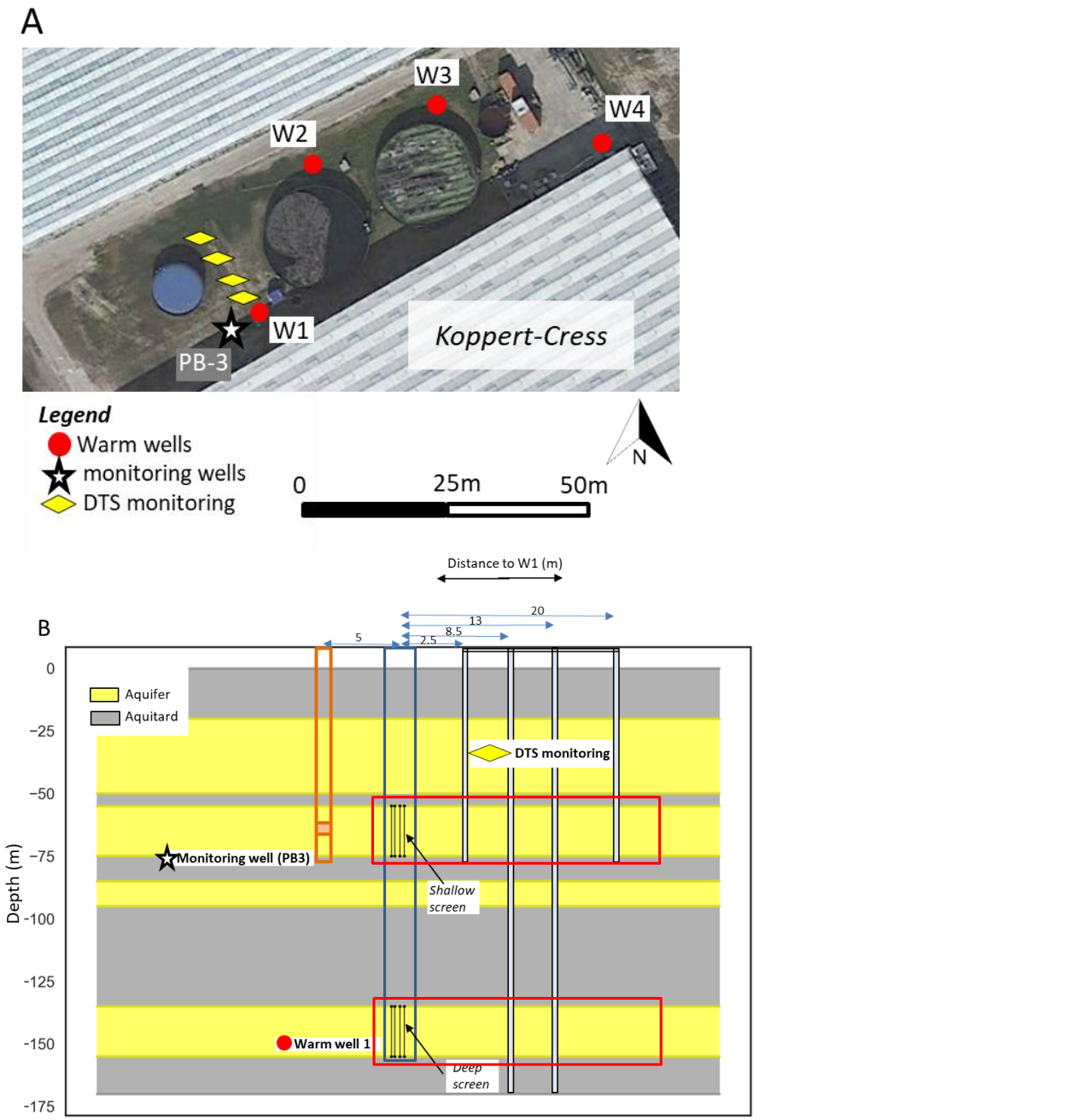


Figure 3.1.2: A) Location of groundwater monitoring wells and the DTS monitoring near warm well 1. B) Schematic cross-section at warm well 1 with the monitoring well and the 4 DTS locations. Red box indicates the monitored area with DTS: 4 shallow sections, 2 deep sections.

3.1.2 Models, model codes and data used

3.1.2.1 Simulator: SEAWATv4

The simulations for this study are carried out in SEAWATv4 (Langevin et al., 2008). SEAWAT is a model that couples the finite-difference code MODFLOW (Harbaugh et al., 2000) and MT3DMS (Zheng & Wang, 1999) which are flow and transport (Hecht-Mendez et al., 2010; Langevin et al., 2010) respectively. SEAWATv4 allows the inclusion of temperature dependent viscosity and density effects. SEAWATv4 uses the governing equations for groundwater flow and solute transport as well as the equations of state for fluid density and viscosity. Recently, the modelling package used in this study was successfully tested in a benchmark study for HT-ATES (Mindel et al., 2021). Furthermore the model is used in previous studies (Bloemendal & Hartog, 2018; van Lopik et al., 2015), indicating that governing processes are well captured in the model code. In this validation approach the goals is to identify model parameter and input set to match model output to measured data.
Van Lopik et al. (2016) calibrated an axisymmetric model of a high temperature (80 °C) ATES system against monitoring data, in which buoyancy flow was a dominating process. The initial model set-up and parameter values in this study follow their work.

Implementation of density and viscosity
SEAWATv4 utilises a linear relationship between groundwater temperature and density (Langevin et al., 2008; van Lopik et al., 2015; Thorne et al., 2006). However, for HT-ATES this leads to a relatively large difference with the actual non-linear relationship shown in Figure 3.1.3.

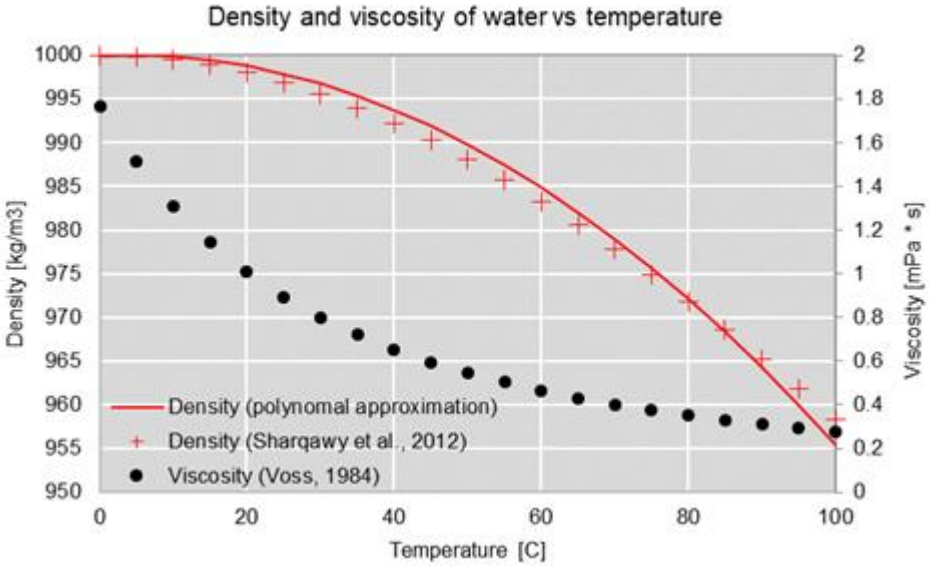


Figure 3.1.3: Non-linear relationship of viscosity and density with temperature.

To use this non-linear relationship, we altered the original SEAWAT executable and implemented the following relationship:

$$\rho(T) = 1000 - \frac{(T-4)^2}{207}$$

This is explained in more detail in the Heatstore WP2.2 report (Tomasdottir & Gunnarsson, 2019). The temperature also affects fluid viscosity (μ [kg/m/d]), to which the hydraulic conductivity is proportional (Fetter, 2001). The relation between viscosity and temperature may be approximated following Voss (1984);

$$\mu(T) = 2.394 \cdot 10^{-5} \cdot 10^{\frac{248.8}{T+133.2}}$$

3.1.2.2 Axisymmetric model setup

The injection volumes that are being distributed over the four warm wells in the ATES system are relatively small compared to the volumes that were anticipated with the well placement design. In effect, the resulting maximal thermal radii (<20m, Bloemendal et al. (2020)) for each individual warm well is much smaller than the distances between the warm wells (40-50m, Figure 3.1.1). Rather than a single thermal volume, the ATES system operates (sub-optimally) with individual thermal storage volumes for each well. Therefore, model validation was focused on monitoring the thermal impact at one of the warm wells (W1) where all DTS monitoring installations were located (Figure 3.1.2), in addition to the groundwater monitoring wells (Figure 3.1.2). An axisymmetric model was used to simulate this single well, as it is computationally less expensive compared to a 3D model, and thus allows for higher spatial and temporal discretization of the model. By assuming radial symmetry, it is however not possible to include lateral heterogeneities or ambient groundwater flow (Langevin, 2008; Louwyck et al., 2014).

Discretization and boundary conditions

To ensure accurate representation of buoyancy flow, a vertical discretization of 1m is used. This is small enough to appropriately take into account all relevant processes as further reduction of layer thickness did not result in improved results. In radial direction, a cell size of 0.1m is used for the first 50m around to the well. Further away, the cell size increases logarithmically up to a size of 25m, until the outer edge of the grid is reached at 2500m to prevent boundary conditions to affect the simulation results.

The horizontal (K_h) and vertical (K_v) hydraulic conductivity of both aquifers are estimated to be K_h=35 and K_v=7 m/d. The regional model REGIS II is used to determine this (TNO, 2019). A hydrogeological schematization consisting of 9 layers is set as input for the groundwater model, Table 3.1.1. The porosity is set at a constant value of 0.3 for all layers and the ambient groundwater temperature at 15 °C. Additional hydrogeological and thermal parameter settings used for the model are presented in Table 3.1.2.

Table 3.1.1: Hydrogeological layering used for the model, the location of the two screens are indicated in the first column.

Well screen	Layer	Top of layer (m)	Bottom of layer (m)	K horizontal (m/d)	K vertical (m/d)	Porosity (-)	Ambient groundwater temperature (°C)	Type
	1	0	-20	0.05	0.01	0.3	15	aquitard
	2	-20	-50	35	7	0.3	15	aquifer
	3	-50	-55	0.05	0.01	0.3	15	aquitard
	4	-55	-75	35	7	0.3	15	aquifer
	5	-75	-85	0.05	0.01	0.3	15	aquitard
	6	-85	-95	35	7	0.3	15	aquifer
	7	-95	-135	0.05	0.01	0.3	15	aquitard
	8	-135	-155	35	7	0.3	15	aquifer
	9	-155	-200	0.05	0.01	0.3	15	aquitard

Table 3.1.2: Input parameter values and corresponding package

Parameter	Value	Package
Solid heat capacity*	710 kJ/kg °C	RCT
Water reference density	1,000 kg/m^3	RCT
Solid density*	2,640 kg/m^3	RCT
Water thermal conductivity	0.58 W/m/°C	RCT
Solid thermal conductivity	2 W/m/°C	RCT
Thermal distribution coefficient#	1.7 ·10^{-4} m^3/kg	RCT
Thermal retardation+	2.21	RCT
Porosity	0.3	BTN
Specific storage aquifer	6 · 10^{-4} /m	LPF
Longitudinal dispersion	0.5 m	DSP
Transversal dispersion	0.05 m	DSP
Vertical dispersion	0.005 m	DSP
Effective molecular diffusion heat#	0.15 m^2/day	DSP
Effective molecular diffusion salt	8.64·10^{-6} m^2/day	DSP

3.1.2.3 Data and pre-processing

The ATES system of Koppert Cress is operational since 2012 and the total operational history of the system is available (production/injection temperature, volume, per 5 mins) up to mid-2021. Previous studies showed that the ATES system of Koppert Cress is highly imbalanced (Bloemendal et al., 2020). More energy is produced from the warm well than is stored during the summer. Therefore, the warm wells are depleted at the end of each production season. This means that the system starts at ambient groundwater temperature, stores heat, extracts all heat (20% more volume extracted for heating on average), and goes back to ambient conditions. This is confirmed by both the production temperature of the wells and the measured temperature with DTS in the subsurface. As a result, the subsurface temperature is the same (i.e. ambient temperature) each spring, when the first heat is stored. Due to this condition, previous years of operation have no considerable impact on subsurface temperature distribution, allowing us to use the period January 2020 to May 2021 (17 months) for our analyses, as this period consists of a complete unloading – loading – unloading cycle (Figure 3.1.6), for which DTS monitoring data is complete.

Two types of data were used to validate the numerical model (Table 3.1.3). Firstly, the 5-min injection and production raw logging data was used a) to set the model input (volume & temperature), and b) to validate the simulated production temperature with the measured production temperature. Secondly, the DTS-data is used to validate the modelled subsurface temperature with the measured subsurface temperatures at 4 different distances from warm well 1 (Figure 3.1.2).

Table 3.1.3: Overview of used data

Data	Used for
Injection data	Model input
Production data	Validate with production model output
DTS data: 6 sections	Validate with subsurface temperature model output

Operational data

The operational 5-min data for the chosen time-period (1-1-2020 to 26-5-2021) is shown in Figure 3.1.5, Figure 3.1.6 and Figure 3.1.6. The following observations are of importance:

- **Short cycle storage and recovery**

Heat storage is not only utilized seasonally, the ATES system is also frequently used for daily heat storage (Figure 3.1.5 and Figure 3.1.6). This is observed throughout the year, but is most visible in spring and autumn. During these periods, excess heat is often available during the day, stored, and used during the night.

- **Difference injection and extraction temperature**

The extraction temperature (unloading) results from the heat that was injected before and the subsurface interaction that occurred during storage. As the injection temperature is variable, the extraction temperature also varies. Subsequently, during storage, the heated groundwater tends to average out over the previously injected temperature. Because of these two processes, the extraction temperature is not as extreme (lower peaks and higher lows) compared to the injection temperature (Figure 3.1.5-B).

- **Injection temperature in warm well is sometimes smaller than ambient**

Ambient groundwater temperatures in The Netherlands, for aquifers at these depths, are typically at 11-12°C. For this ATES system however, determining the ambient temperature based on abstracted temperatures from the warm well was not straightforward due to the high temporal dynamics of the system and because the cold well stores heat at temperatures below ambient temperature (like conventional ATES systems). When cooling demand is low, it may occur that the temperature stored in the warm well is lower than T_{amb}. This may happen when, for example, groundwater is extracted from the cold well at 7 °C, heated to 10 °C by cooling of the greenhouse, and stored in the warm well at 10 °C (Figure 3.1.5). This is in contrast with the maximally observed injection temperatures of 40 °C during the hottest days in summer. This condition occurs frequently, and resulted in the injection temperature of the warm actually being lower (down to 7°C, Figure 3.1.5), when there was insufficient heat available in the building climate system and any of the connected environmental sources of heat. Therefore, the best available indication of the ambient groundwater temperature is based on the extraction temperatures from the wells in combination with the DTS measurements, which suggested a relatively high ambient groundwater temperature of 15°C (Figure 3.1.8).

Data processing

From the 5-min data, the aggregated hourly injection/extraction temperature are calculated. Following the rule of energy balance, the average injection or extraction temperature are calculated as:

$$T_{hourly} = T_{amb} + \frac{\Delta E_{hourly}}{(\Delta V_{hourly} \cdot C_w)}$$

$$\Delta E_{hourly} = \sum (E_{inj\_5\min} - E_{ext\_5\min})$$

$$E_{inj\_5\min} = (T_{inj\_5\min} - T_{amb}) \cdot V_{5\min} \cdot C_w$$

With T_{amb} being the ambient groundwater temperature at the location. The hourly data is used as input for the model, and to validate the model output to the measured data.

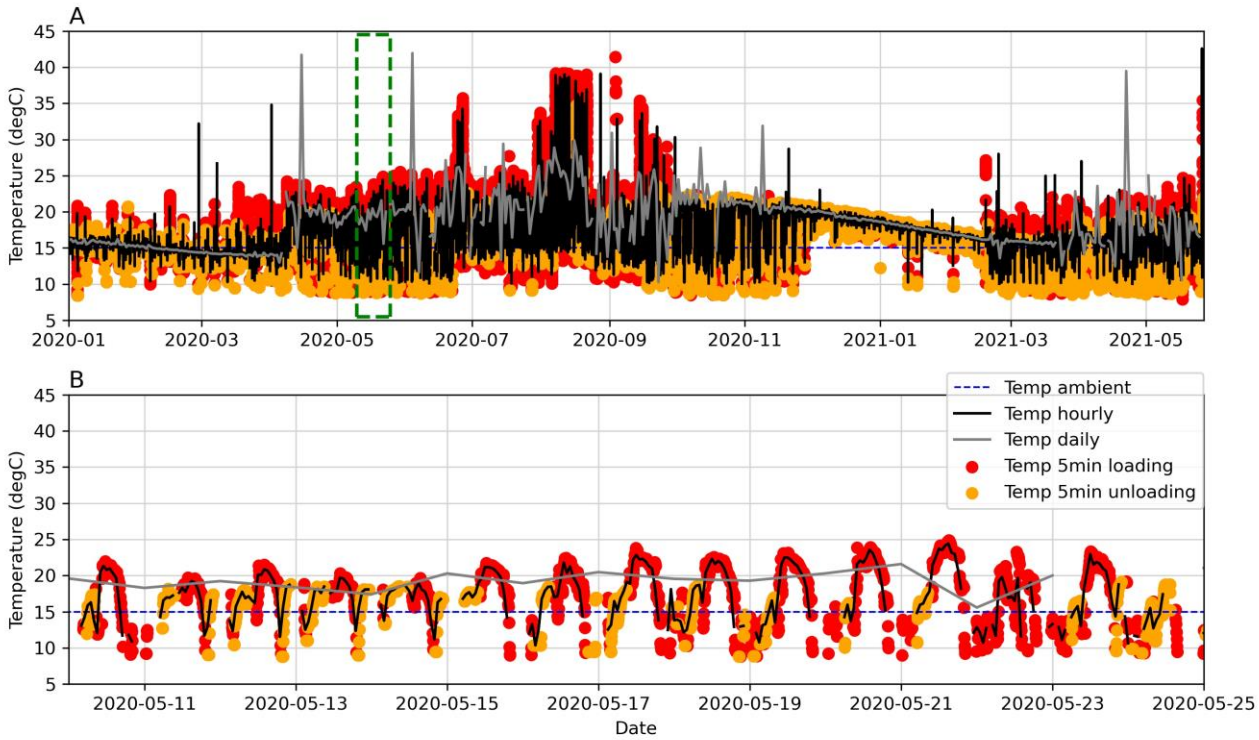


Figure 3.1.5: A. The 5-min temperature data that is extracted (unloading) and injected (loading) in the warm wells, and the aggregated hourly and daily temperature data. B. zoomed in on the period 2020-08-10 to 2020-08-25 (green frame in A).

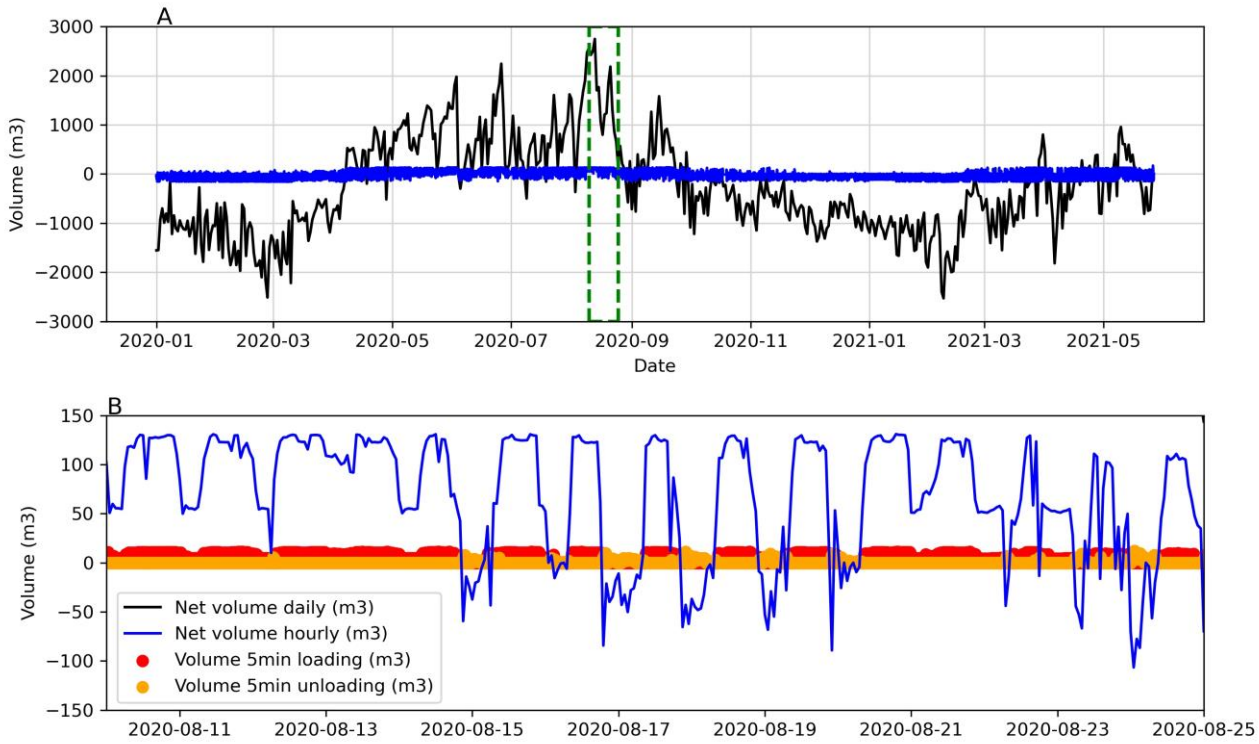


Figure 3.1.6: Aggregated hourly volume data (A) and the 5-min loading and unloading volumes for the period of 10-08-2020 to 25-08-2020 (B) (green frame in A).

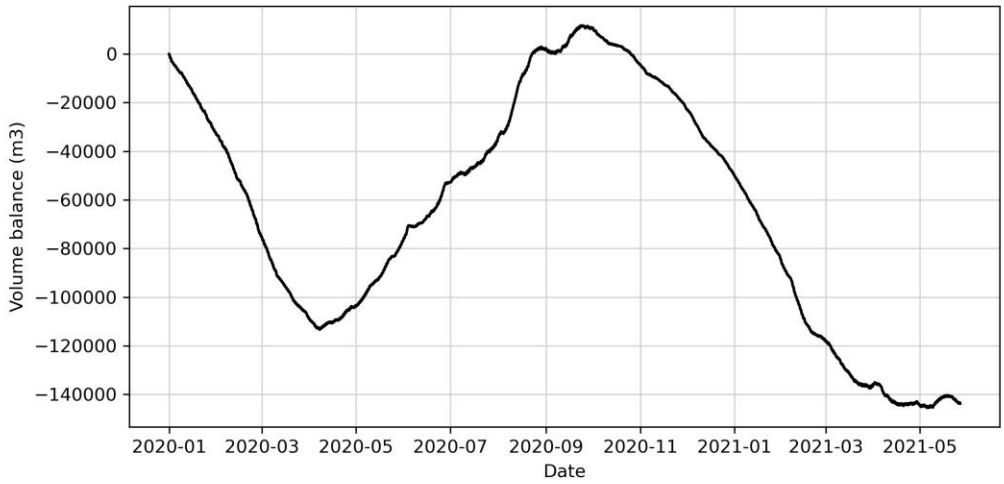


Figure 3.1.6: The volume balance of the ATES system. The warm well is loaded from April 2020 to September 2020 and unloaded in the months after that.

DTS data

The temperature in the subsurface is measured along 6 aquifer sections in the subsurface, 4 in the upper aquifer and 2 in the lower aquifer (Figure 3.1.8). From the measured temperature distribution along the vertical aquifer depth, the average aquifer temperature is calculated and shown Figure 3.1.8. The highest temperatures are observed at 2.5m from warm well 1. At larger distances from warm well 1, the temperature changes decrease, are timed differently and exhibit a more gradual course.

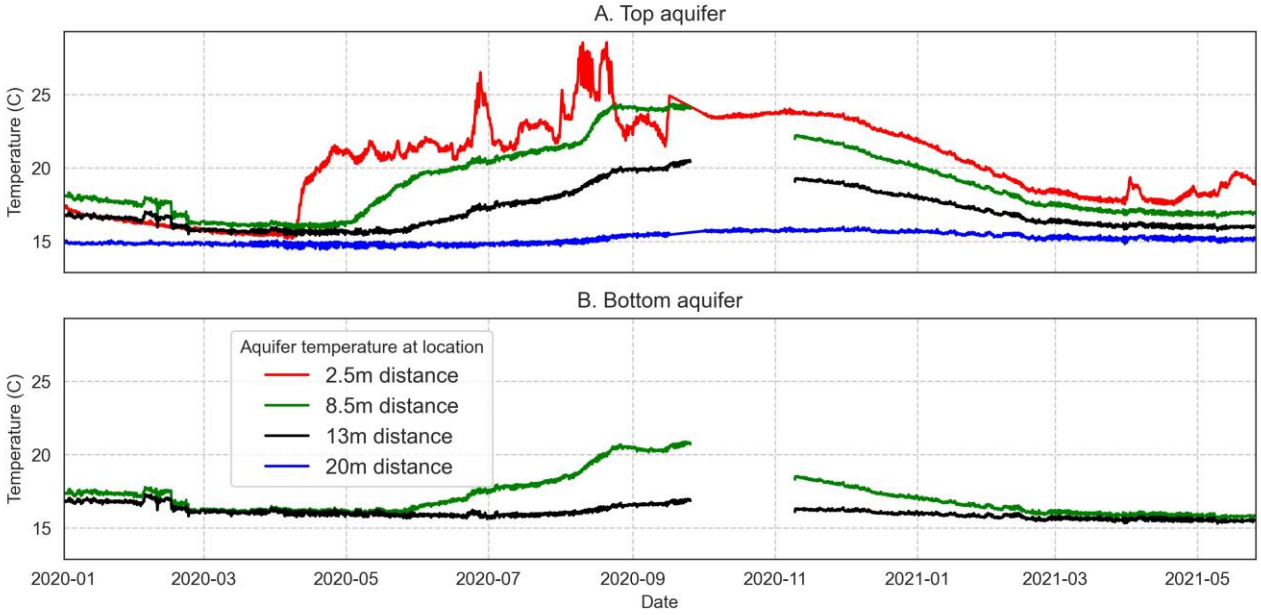


Figure 3.1.8: Average temperature of the aquifer section measured with the DTS cables in A. the top aquifer and B. the bottom aquifer. The missing data visible for the 8.5 and 13.5 m DTS locations is caused by a temporal break in the DTS cable in September to October 2020.

The DTS cables measure the subsurface temperature with a spatial resolution of 1 meter and a temporal resolution of 3 hours. This allows for analysis of the vertical temperature distribution along the vertical aquifer depth over time. The measurements indicate that the subsurface temperature is not equal for the total aquifer depth. In Figure 3.1.9 the temperature distribution in the shallow aquifer at 2.5m distance from warm well 1 is presented. In the middle, around 60m depth an anomaly in the aquifer composition is visible because lower temperatures are measured. This is visible during the periods when most hot water is stored (summer). Oppositely, when the aquifer cools down again after prolonged unloading, highest temperature are eventually

observed in the middle of the aquifer (e.g. December 2020 – January 2021 in Figure 3.1.9-A), again indicating a section of aquifer with a lower hydraulic conductivity.

After installation, the DTS equipment needs calibration to correct for the temperature off-set. At the time of installation in 2012, the first set of DTS cables was calibrated. However, over the years, multiple additions and changes were made to the DTS system (due to additions and breaks) making the calibration less reliable. The DTS is run double ended, the reverse signal is used to increase accuracy of the measured temperature, i.e. ~0.1°C.

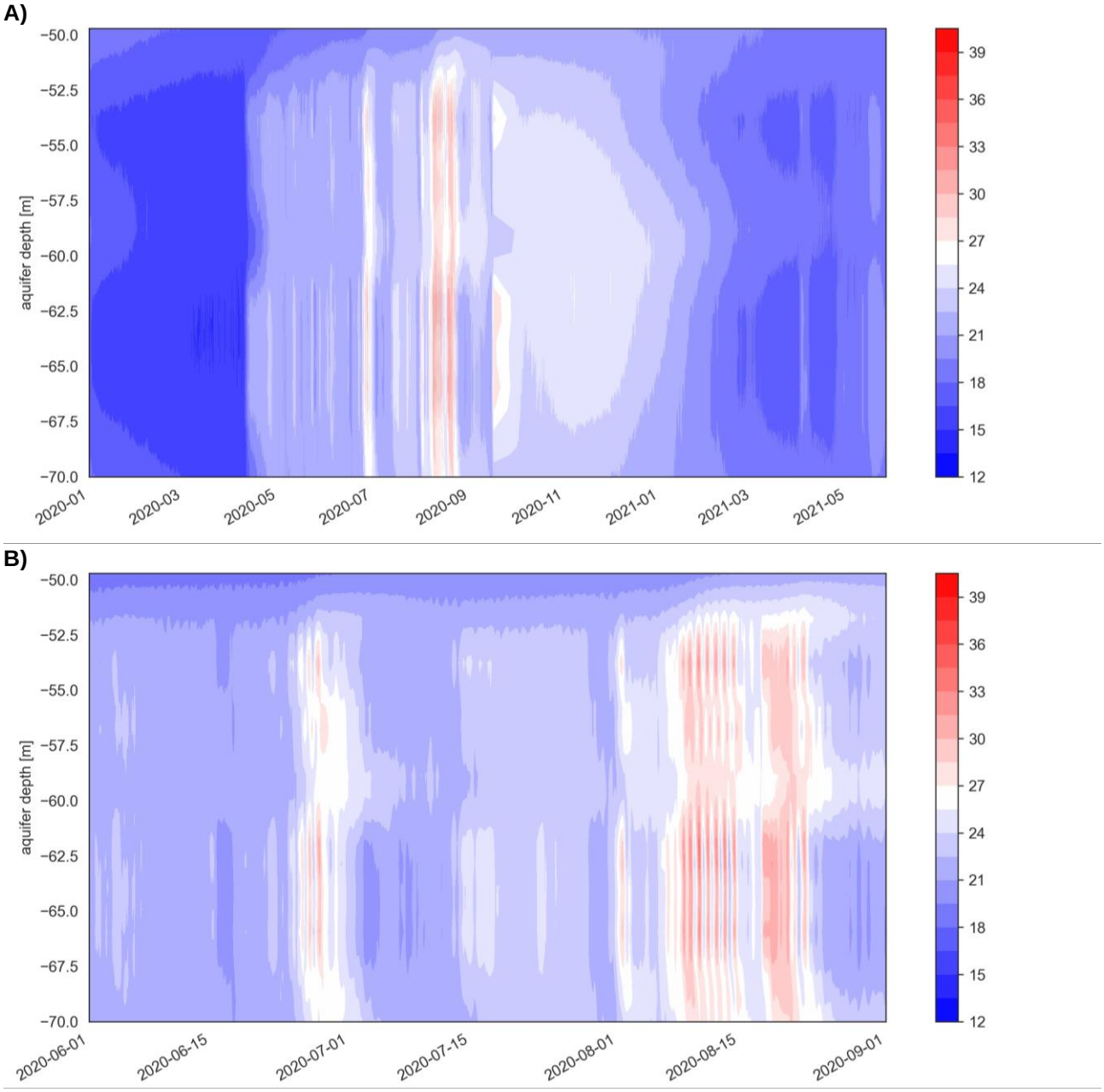


Figure 3.1.9: Vertical temperature distribution at the DTS cable in the upper storage aquifer at a depth of -50 to -70m (2.5m distance from warm well 1). A) Total analysed period. B) Zoomed in on the highest measured temperatures in summer, the period between 2020-06-01 and 2020-09-01.

3.1.3 Validation approach

3.1.3.1 Principles and assumptions

We model the performance and subsurface temperature at warm well 1. To do this, we make the following assumptions:

1. All wells inject/extract the same amount of volume. Hence, the volume for warm well 1 is calculated by dividing the total measured volume by 4.
2. The volume distribution between upper and lower aquifer is set to 40% and 60% respectively in the initial simulation. This was measured with a well flow test with a spinner in warm well 4 in 2017 (Bloemendal et al., 2020).
3. Initially assumed ambient temperature aquifer is 15 °C.
4. The subsurface is divided into aquifers and aquitards (layered vertically), all these formations are homogeneous in the initial simulation. (to improve model performance heterogeneity is added by varying the hydraulic conductivity of the model layers within these formations.)
5. Ambient groundwater flow is zero.

3.1.3.2 Assessment framework

Table 3.1.4: Parameters used for model validation.

Parameter	Description
The recovery efficiency (-)	$\eta = \frac{E_{out}}{E_{in}}$, calculated for the total loading/unloading cycle (07-04-2020 <> 07-04-2021) Recovery efficiency of ATES system = 0.65.
Extraction temperature warm well 1	Visual comparison measured and simulated extraction temperature & calculate the R^2
Temperature at DTS	Visual comparison of measured and simulated average subsurface temperature & calculate the R^2. DTS data is measured each 3 hours.

Linear regressions are used to verify the correspondence of the measured and modelled datasets. From these linear regressions, the slope, intercept and R^2 together provide insight to what extent the two datasets correspond. A slope and R^2 close to 1 and an intercept close to 0 are indicators for good correspondence and hence good model performance.

3.1.3.3 Validation parameters and steps

In this validation exercise, 3 main unknown parameters are identified: A) the ambient temperature and temperature off-set of the DTS measurements, B) the flow distribution of the different screens and C) heterogeneity/layering in the aquifer. A base case modelling scenario is defined following known and estimated model parameters. Based on the results of the base case scenario, the three parameters are varied in order to better match recovery efficiency, well extraction temperature and temperature distribution around the well. This results in the following scenarios:

Z. Base case scenario
A. Adjustment of ambient groundwater temperature and insight in the temperature off-set of the DTS
B. Flow distribution (upper and lower well screen)
C. Heterogeneity: effect of horizontal layering (higher/lower hydraulic conductivity)
D. Final model: optimal combination of above scenarios

3.1.4 Results

3.1.4.1 Case Z: Base case scenario

Well temperature and recovery efficiency

The modelled well temperature follows the main trends of the measured well temperature (Figure 3.1.10). The modelled well temperature has slightly lower peak values compared to the measured well temperature, meaning that the highest and lowest values are under-/overestimated respectively in the simulation. The modelled well temperature in winter (prolonged unloading) is decreasing slower compared to the measured well temperature. This indicates that the model overestimates the performance of the heat storage in warm well 1. From the correlation plot in Figure 3.1.10 we observe that the model, on average, overestimates the relatively low extraction temperatures (10-15 °C) and underestimates the relatively high extraction temperatures, hence the regression line intercepts at 4.7 °C and is at a slope of 0.76. Nevertheless, the general correlation between the measured and simulated extraction temperature is reasonable at R^2=0.71.

The modelled recovery efficiency is 0.91, while the measured recover efficiency calculated, based on measured pumping rates and well temperatures, is 0.65. This confirms that the average extraction temperature is modelled higher than was measured, in other words the performance is overestimated by the model, which could be due to an overestimation of the assumed ambient temperature.

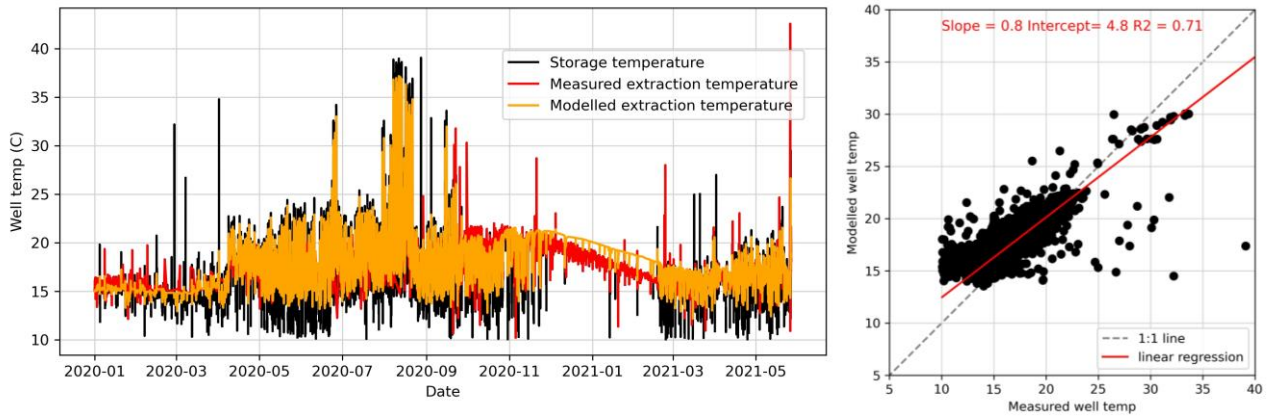


Figure 3.1.10: The measured and modelled hourly well temperature development in time (left) and correlation plot of the measured and modelled hourly extraction temperature (right) | Base case.

Subsurface temperature

The modelled subsurface temperature is compared to the DTS measured subsurface temperature at 2.5m, 8.5m, 13m and 20m from warm well 1 (Figure 3.1.11). Generally, the timing and characteristic subsurface temperature changes are represented well with the simulation model. At 2.5m, the temperature changes are most variable and highly influenced by the variations in injection temperatures, further away from warm well 1 the temperature variations smoothen and the also show lower temperature.

The temperature in the subsurface is underestimated in the shallow aquifer and overestimated in the deeper aquifer (Figure 3.1.11). Based on the volume-distribution (40% in shallow aquifer, 60% in deeper aquifer), the model simulates higher subsurface temperatures in the deeper aquifer as expected. The latter, and the timing off-set of the peaks indicates that the simulated injection volume in the deeper aquifer is too high.

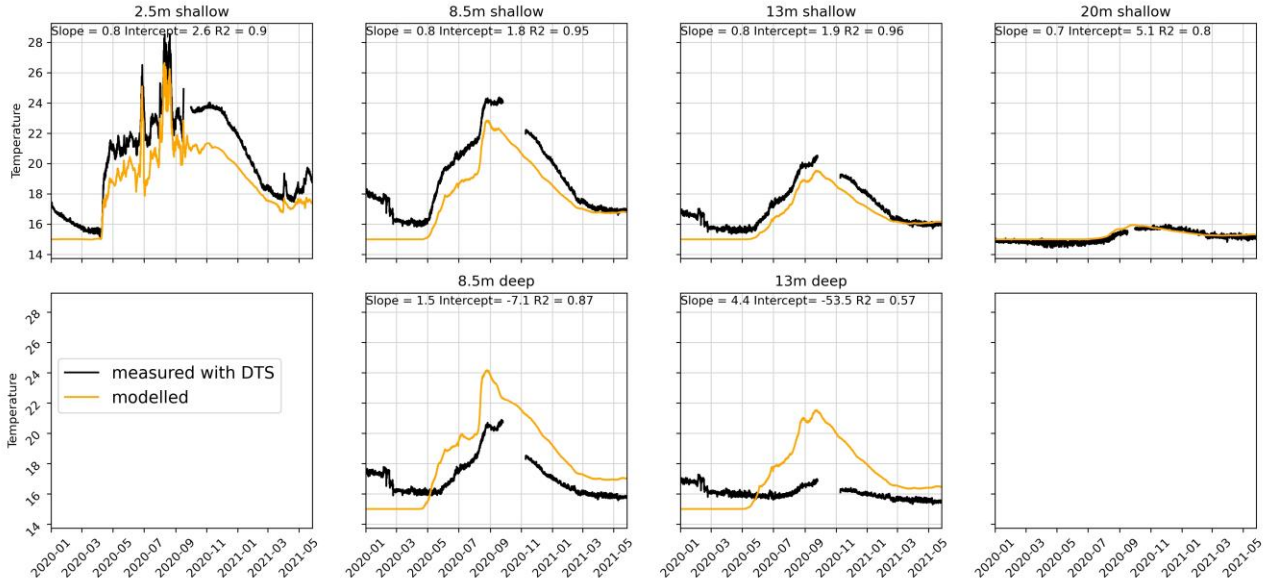


Figure 3.1.11: Comparison between modelled and measured (DTS) subsurface temperature in the shallow and deep aquifer, linear regression statistics are given at the top of each plot | Base case.

Main findings and next steps:

- The recovery efficiency is overestimated by the model. Moreover, the extraction temperature from the model when the well is depleted is overestimated. This indicates that the assumed ambient temperature of 15 °C based on initial DTS calibration is likely too high. This is validated in case A (section 3.1.4.2).
- The modelled thermal radius is smaller than measured in the shallow aquifer and larger than measured in the deeper aquifer. As the exact flow distribution between the two well screens is unknown, this is likely caused by an incorrect assumption for the flow-distribution. Initially, the shallow screen received 40% of the total flow and the deeper aquifer 60%. To test this in case B, we changed the flow distribution according to the observed distribution at 8.5 and 13m in Figure 3.1.11 (65% to shallow aquifer, 35% to deep aquifer).

3.1.4.2 Case A: Adjustment of ambient groundwater temperature

Well temperature and recovery efficiency

In this scenario, the ambient groundwater temperature in the model is adjusted to 12.5 °C (before 15 °C), indicating an offset from the DTS based temperatures of ~2.5 °C. As the temperature difference between the ambient groundwater and the stored groundwater increases, the losses increase (Figure 3.1.12). Thus, the modelled recovery efficiency decreased to 0.85, compared to 0.91 for storage with 15 °C ambient groundwater temperature.

Lowering the ambient groundwater temperature in the simulations also leads to a higher calculated recovery efficiency for the measured dataset. The measured extraction temperatures are relatively high when using 12.5 °C instead of 15 °C. As a results, the measured recovery efficiency, with T_{amb}=12.5 °C is 0.81, compared to 0.65 with T_{amb}=15 °C. With T_{amb}=12.5, the modelled recovery efficiency (0.85) is in better agreement with the measured recovery efficiency (0.81), illustrating the sensitivity for the ambient groundwater temperature used in both performance analysis and performance modelling. Most clearly, the good alignment of the modelled and measured temperature during prolonged unloading in the winter (December 2020 - February 2021) indicates the better fit with T_{amb}=12.5 °C (Figure 3.1.12, December 2020 – January 2021).

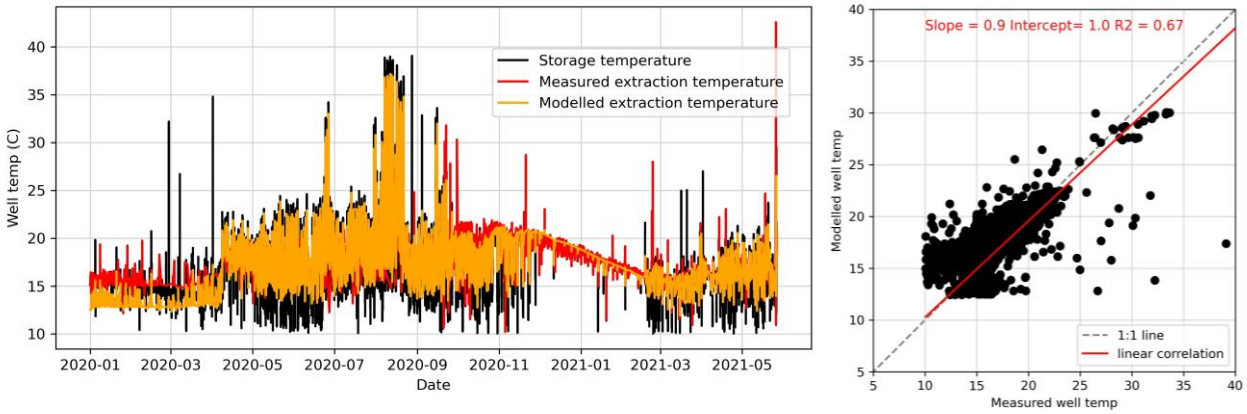


Figure 3.1.12: The measured and modelled hourly well temperature development in time (left) and correlation of the hourly average extraction temperature from the well (right) | Case 2.

Subsurface temperature

While the performance modelling was improved by the lowered ambient groundwater temperature, the modelled and measured subsurface temperature still do not correspond (Figure 3.1.13). High R^2 values for the 2.5m, 8.5m and 13m indicate that the variabilities are modelled well. However, the slope, and mainly the intercept (-3 to -83), indicate that there is an off-set in results. Based on the results of Figure 3.1.13, we observe that the DTS values are approximately 2.5 °C higher compared to the modelled results that are in agreement with the measured well temperature data. This may be caused by the DTS off-set, as was already pointed out in section 3.1.2.3. To validate this, the DTS temperature data is adjusted -2.5 °C in the final model and compared to the final model output.

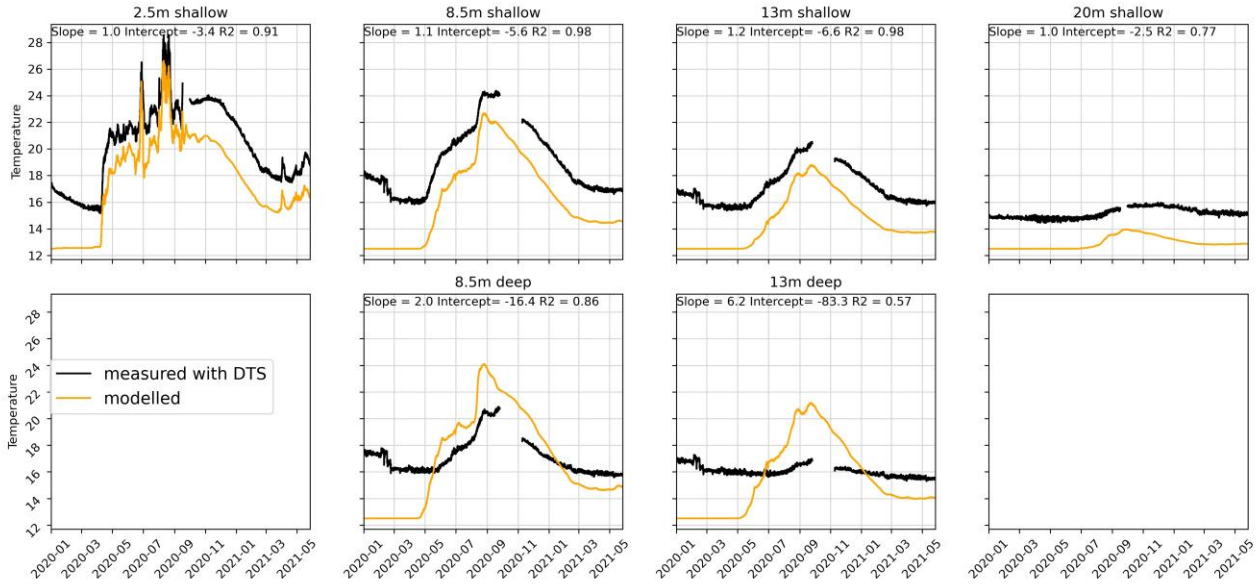


Figure 3.1.13: Comparison between modelled and measured (DTS) subsurface temperature in the shallow and deep aquifer, linear regression statistics are given at the top of each plot | Case 2.

Main findings

- Modelled recovery efficiency (decreased) is similar to the calculated measured recovery efficiency (increased) after assuming a lower ambient groundwater temperature of 12,5°C. This also is in line with the modelled well temperature being more in agreement with the measured well temperature (Figure 3.1.12).
- Subsurface temperature similarity did not improve. This may be caused by incorrect DTS-offset, which will be validated in the final model.

3.1.4.3 Case B: Optimization of volume distribution between aquifers

Well temperature and recovery efficiency

In this scenario, the volume distribution between the shallow and the deeper aquifer is altered to 65% in the shallow aquifer, and 35% in the deeper aquifer (compared to 40% / 60% respectively). This did not have any impact on the performance of the system, as the recovery efficiency of this scenario is equal to the base case scenario. Moreover, due to the identical properties of the shallow and deep aquifer, the modelled average well temperature in Figure 3.1.14 is equal to the base case model (Figure 3.1.14).

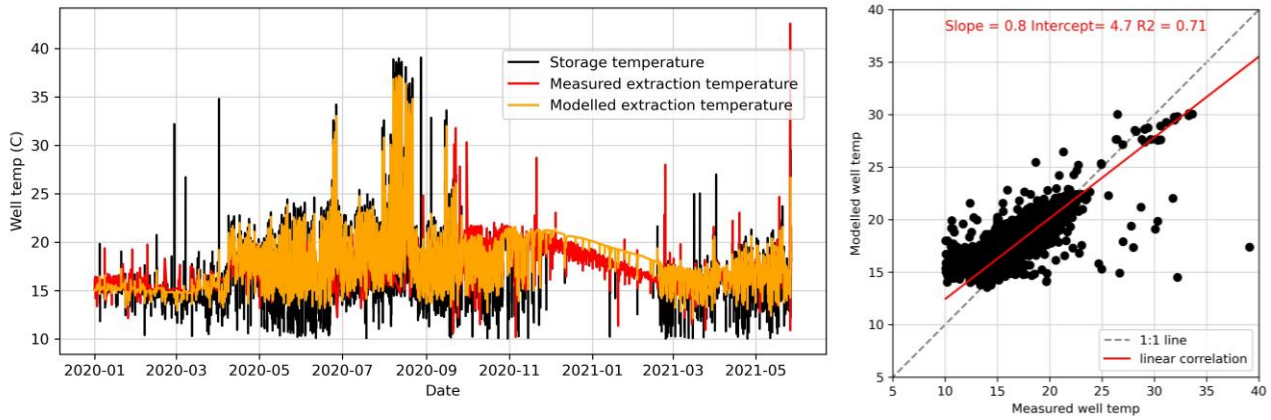


Figure 3.1.14: The measured and modelled hourly well temperature development in time (left) and correlation of the hourly average extraction temperature from the well (right) | Case 3.

Subsurface temperature

The agreement between the measured and modelled subsurface temperature increased, both in the upper and the lower aquifer. The largest temperature increase is now observed in the shallow aquifer, as is also the case in the measured dataset. As expected, the changed flow distribution leads to a better fit in absolute subsurface temperature (Figure 3.1.15). Especially at 8.5m and 13m this is very clear. At 2.5m, differences are still relatively large between the measured and modelled subsurface temperature. As the R^2 of all plots indicate, the proportion of variance that is explained between the two datasets does not improve compared to the scenario with lower ambient groundwater temperature), likely because this adjustment was not included in this model scenario.

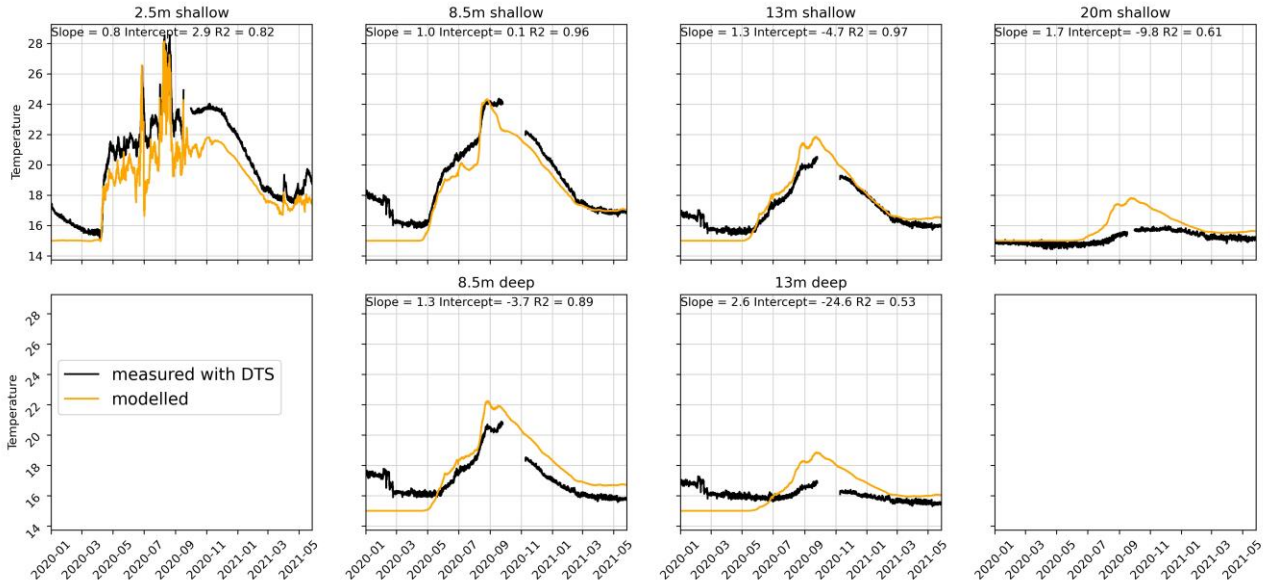


Figure 3.1.15: Comparison between modelled and measured (DTS) subsurface temperature in the shallow and deep aquifer, linear regression statistics are given at the top of each plot | Case 3.

Main findings

- The volume distribution adjustment does not have any effect on the energy performance
- The volume distribution adjustment showed improved agreement between the higher measured temperatures in the shallow aquifer and relatively lower temperatures in the deeper aquifer and absolute subsurface temperatures. However, the explained variance did not improve compared to base case scenario.

3.1.4.4 Case C: including aquifer heterogeneity in shallow aquifer

The sensitivity for the aquifer heterogeneity was tested using an adjusted hydrogeological schematization for the shallow aquifer (Table 3.1.5). Three layers of higher hydraulic conductivity were added, two small layers (2m) in the top part and one thicker layer (6m) in the bottom of the 20m shallow aquifer, based on Figure 3.1.9. As the total flow into the aquifer was not varied, the changed hydraulic conductivity only impacts the vertical flow distribution along the well screen.

Table 3.1.5: Schematic overview of the normal, homogeneous, hydrogeological model input and the adjusted heterogeneous horizontal layering.

	Depth of layer top	Hydraulic conductivity HOMOGENEOUS	Hydraulic conductivity HETEROGENEOUS
	0	0.05	0.05
	-20	35	35
	-50	0.05	0.05
Shallow storage aquifer	-55	35	20
	-57		40
	-59		20
	-61		40
	-63		20
	-66		40
	-72		20
	-75	0.05	0.05

The inclusion of aquifer heterogeneity does not have a considerable effect on the performance and average subsurface temperature, all energy performance results are equal to the base case scenario (Table 3.1.7). Higher temperatures are observed in the more permeable layers, Figure 3.1.16. Relatively more volume is pushed into the more permeable layers, meaning that a larger thermal radius develops compared to the less permeable layers (e.g. at 60m depth). The addition of heterogeneity improved the visual similarity between the measured and modelled results, Figure 3.1.16. The absolute temperatures are lower for the base case and heterogeneous scenario compared to the measured data, as previous adjustments (scenario A and scenario B) are not included in this heterogeneous model scenario.

Main findings

- The addition of heterogeneity in the model leads to spatial temperature variations, similar to the temperature variation measured with the DTS around the well.
- However, the addition of relatively limited variations in hydraulic conductivity does not affect the energy performance of the model (recovery efficiency, well temperature and average aquifer temperature). This is in correspondence with previous research on the effect of heterogeneity on ATES performance (Sommer et al., 2013).

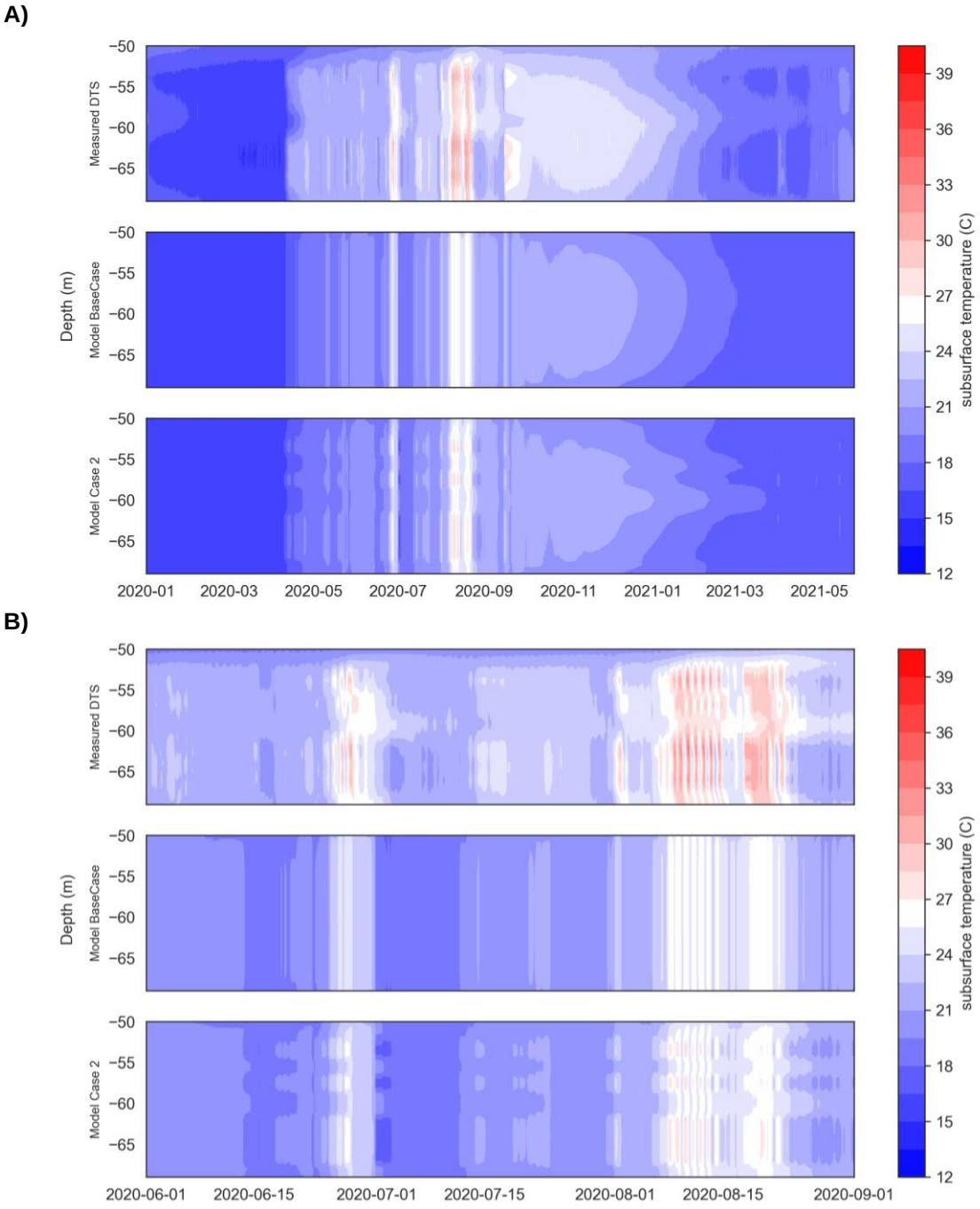


Figure 3.1.16: Vertical subsurface temperature distribution at 2.5m from warm well 1, comparison between measured and modelled results. A) Full analysis period, B) zoomed into the summer period 2020.

3.1.4.5 Case D: Final model

Well temperature and recovery efficiency

In this scenario all improvements obtained by the previous simulations are combined, Figure 3.1.17. The ambient groundwater temperature is set at 12.5, the measured DTS data is corrected by -2.5 °C and the changed volume balance was applied, equal to case B. Aquifer heterogeneity is not included as this does not affect the energy performance and the average subsurface temperature of the model.

The recovery efficiency of the final model was equal to case B, Table 3.1.7. This corresponds with the previous findings, the changed ambient groundwater has influence on the energy performance, the other adjustments do not. The modelled (0.85) and measured (0.81) recovery efficiency, i.e. the energy performance, are in reasonable agreement in the final model, Table 3.1.7.

Subsurface temperature

The combined effect of the adjusted model settings are presented in Figure 3.1.18.

- The corrected DTS temperature data (corrected with -2.5 °C) is reproduced using the final model.
- The measured subsurface temperature at 2.5m from warm well 1 corresponds well with the final model, as the slope (1,0) and intercept (-0.2) have improved towards close to optimal and the R^2 is also high (0.87).
- However, with increasing distance from warm well 1, the subsurface temperature simulated by the model is increasingly overestimated with respect to the measured subsurface temperature. Also, the final model, in combination with the corrected DTS temperature data, did not improve the similarity compared to the base case scenario (Table 3.1.6).

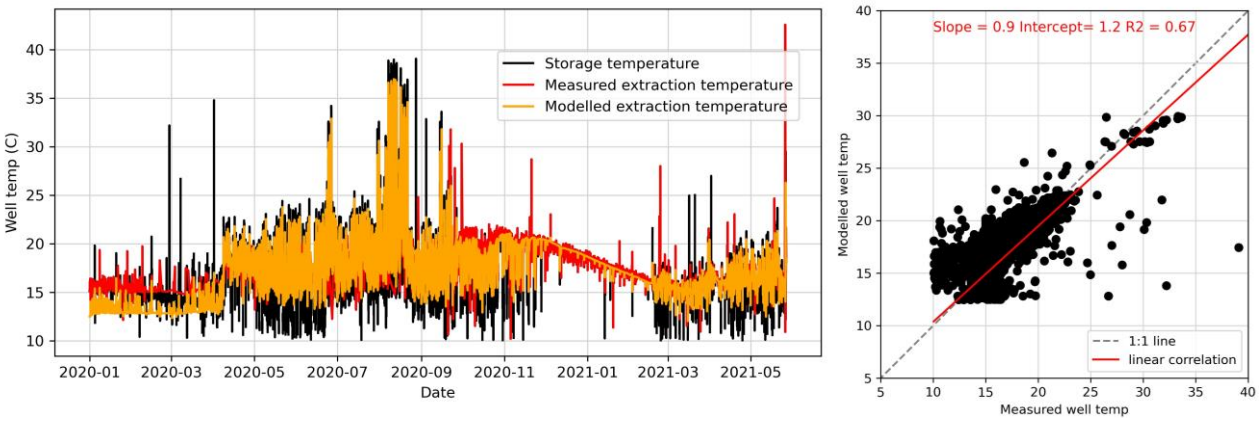


Figure 3.1.17: The measured and modelled hourly well temperature development in time (left) and correlation of the hourly average extraction temperature from the well (right) | Final model.

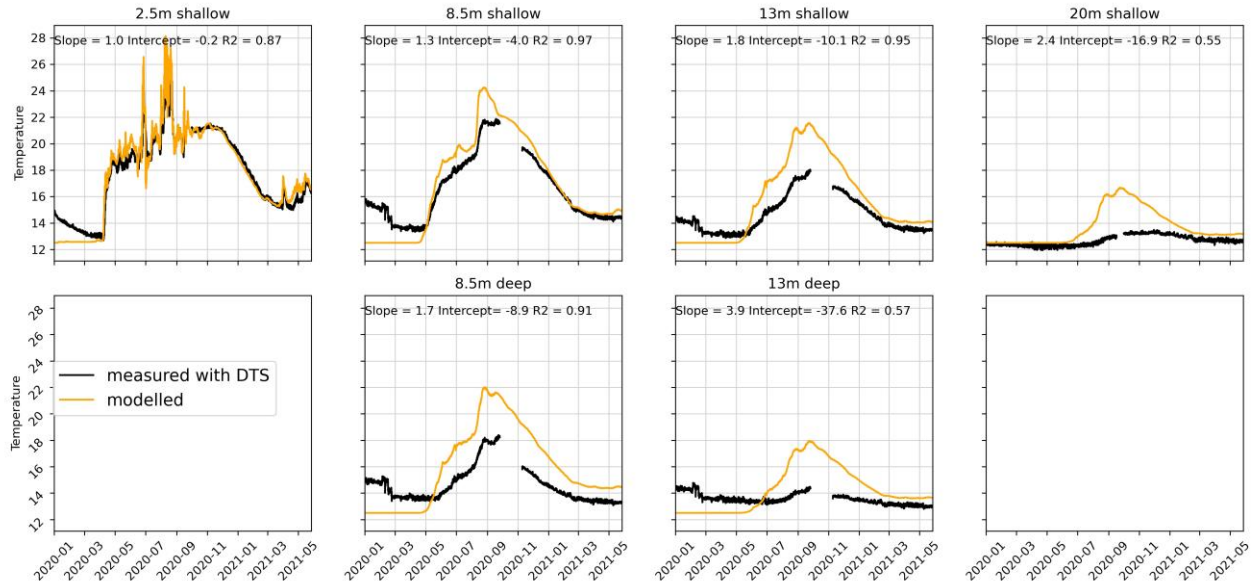


Figure 3.1.18: Comparison between modelled and adjusted measured (DTS) subsurface temperature in the shallow and deep aquifer, linear regression statistics are given at the top of each plot | Final model.

3.1.5 Discussion

With the validation steps carried out, the subsurface temperature that is modelled close to warm well 1 shows that detailed temperature data allows to identify several unknowns from the system. This resulted in improved model performance for the final model. However, at greater distance from the well, simulation results did not correspond well to the monitoring data. The following aspects may have affected this and/or point towards further improvement steps.

Aquifer properties

- The temperature level and propagation of the thermal front is overestimated by the model. This could mean that the volumetric heat capacity of the aquifer is larger than currently estimated. This could be caused by a higher porosity than expected or a higher heat capacity of the sand aquifer material (Anderson, 2005). This directly impacts the propagation of the temperature front around the storage well. However, values for aquifer material heat capacity do not vary so much, that may explain the large differences found.
- During storage, measured temperature levels of the ground(water) decreases faster than expected, indicating that the distribution of heat through the subsurface is faster than modelled. This could be caused by e.g:

  - A higher thermal conductivity than currently estimated.
  - Higher dispersivity, this would cause lower temperatures mainly at larger distance from the well, not near to the well, due to thermal retardation.

  However, the timing of the temperature changes is very similar for the modelled and measured scenario. This may also change when these hydrogeological/thermal properties are changed.

- It is currently estimated that ambient groundwater flow is negligible and hence has no influence on the energy performance. It could be that the ambient groundwater flow is faster than expected (in one of the /both aquifers), hence impacting the energy losses that occur in the aquifer domain around the well.
- Lateral heterogeneities could lead to preferential flow paths/zones, resulting in an unequal thermal radius around the well. Although lateral heterogeneities may be present in the storage aquifers, due to e.g. depositional changing environments due to ice-ages in the Pleistocene shallow sea setting (Noorbergen et al., 2015), it is likely that this is only of impact on a larger scale. We therefore don't expect that this has impact on the temperature distribution around warm well 1. However, ATES systems in more heterogeneous hydrogeological systems may be heavily affected by this (e.g. Winterleitner et al. (2018)).
  The effect of vertical heterogeneities did result in a better representation of the temperature field around the well. But the modelled production temperature and recovery efficiency did not improve as a result of that. As a result it is reasonable to expect that limited changes in the lateral extent of the temperature field also don't affect recovery temperature noticeably.

Modelling assumptions

- At greater distances from warm well 1 the hydraulic influence of neighbouring warm well 2 becomes more pronounced, which will cause thermal displacement by the simultaneous injection in well 2 (Figure 3.1.2). This will lead to the displacement of warm water that is injected at warm well 1 to the South-West side, and thus resulting in smaller thermal influence at the DTS locations, this would be in line with the model overestimation of the from DTS obtained temperatures. Given the respective locations of warm wells 1 and 2 and the DTS monitoring point, it makes sense that the mismatch becomes larger at larger distance. The hypothesis (section 3.1.2) that it is not expected that the influence of warm well 2 would affect the results of warm well 1 may be incorrect. To check this, a 3D model should be constructed to model all warm wells, to also take into account the hydrological effects of warm well 2.
- The measured data is the total volume of the 4 wells and an even distribution between these wells is assumed for extraction and injection (25% each). However, this distribution may be uneven due to differences in hydraulic conductivities at each well as well as difference in well skin due to well development and clogging variations. Future well tests could resolve this issue.

This validation effort shows that the monitored ATES system of Koppert Cress has challenging characteristics regarding analysis and modelling of the system;

- it has a highly short-cyclic and imbalanced character,
- it is gradually changing from a low-temperature ATES system to a high-temperatures ATES system,
- the hydrogeological setting is challenging, as 4 warm and 4 cold wells are used, which utilize 2 aquifers (total of 16 well screens),
- there is a large energy imbalance causing remarkable processes to occur, which make interpretation of the data difficult

When setting up the monitoring installation and research plan, it was expected that the system would store more heat at higher temperatures, and that the amount of stored heat would be in balance with the heat

demand. This more complex situation demands for monitoring of some more parameters to allow distinguishment between different processes that occur. In spite of the limitations encountered the following two main insights are obtained:

- Detailed monitoring parameters and detailed characterization of the hydrogeological and the ATES system are essential for robust and easy validation.
- On the other hand, this study also showed that it is possible to adequately simulate the performance of ATES systems using the constructed SEAWATv4 model, even under challenging circumstances, by assuming generic modelling principles. The correct ambient groundwater temperature is of high importance here.

3.1.6 Conclusions

In this modelling study, 5 models were built to optimize and validate model parameter and input settings. This was done by comparing the measured production temperature (hourly) and subsurface temperature changes (3-hourly) with their simulated counterparts. The results of these models are compared in Table 3.1.7 and Table 3.1.6. Each model provided insight that helped us to improve the modelling setup, which led to the final model (case D). The following conclusions are drawn from this study:

1. Performance of simulation is similar to measured dataset
The model validation allowed to optimize the model ambient groundwater temperature. This led to an increased fit between the modelled and measured recovery efficiency. The final model is adequately simulating the production temperature and recovery efficiency of the ATES system.

2. Modelled subsurface temperature improved compared to measured DTS temperature
The subsurface temperature distribution around warm well 1 was modelled. The following optimization were done, based on the validation simulations:

- Flow distribution between shallow and deep aquifer was altered
- Ambient groundwater temperature was identified
- The DTS dataset was calibrated to the measured well temperature and the model results

The final model is able to represent the characteristic temperature changes that is are observed with the DTS cable, the overall results increased compared to base case (A), Table 3.1.6. The model corresponds best to the measured DTS data close to the well, at 2.5m. At increasing distance from warm well 1, the mismatch increases between modelled and measured subsurface temperature. In this report, different aspects are discussed that could lead to the observed mismatch.

Table 3.1.6: Linear regressions results of the simulated subsurface temperature and DTS data for the base case (Z) and Final model (D).

	Location	2.5m			8.5m			13m			20m		
	Linear Regr	SLOPE	INTRCPT	R2	SLOPE	INTRCPT	R2	SLOPE	INTRCPT	R2	SLOPE	INTRCPT	R2
Base case model (Z)	Shallow	0.8	2.6	0.9	0.8	1.8	0.95	0.8	1.9	0.96	0.7	5.1	0.8
	Deep				1.5	-7.1	0.87	4.4	-53.5	0.57			
Final model (D)	Shallow	1	-0.2	0.87	1.3	-4	0.97	1.8	-10.1	0.95	2.4	-16.9	0.55
	Deep				1.7	-8.9	0.91	3.9	-37.6	0.57			

3. Heterogeneity has no noticeable influence on energy performance of model
In case C vertical heterogeneity (layering) was included. This had no effect on the energy performance of the model, Table 3.1.7. However, we showed that the constructed model adequately represented the heterogeneity that was observed with the DTS by including layers of higher permeability in the hydrogeological model setup.

Table 3.1.7: Recovery efficiency and linear regressions statistics of the simulated and measured well temperature for the five scenarios.

	Recovery efficiency modelled	Recovery efficiency measured	R^2	Slope	intercept
Base case (Z)	0.91	0.65	0.71	0.8	4.8
Case A: (T_{amb} = 12.5)	0.85	0.81	0.67	0.9	1.0
Case B: (Volume distribution shallow/deep)	0.91	0.65	0.71	0.8	4.7
Case C: (Heterogeneity)	0.91	0.65	0.7	0.8	4.9
Case D: Final model	0.85	0.81	0.67	0.9	1.2

3.2 ECW Energy, Middenmeer, the Netherlands

3.2.1 Introduction

The software code, model set-up and results of the thermal and reactive transport simulations performed for the Dutch HT-ATES site in Middenmeer (WP2) have been reported in deliverable D2.1. In the current deliverable, model results are compared to monitoring data to evaluate how the models perform. Adaptations to the models will be made, if necessary, based on the results of the comparison, and updated predictions will be reported. Note that the thermal simulations and predicted thermal evolution of the HT-ATES site reported in D2.1 were performed with the software code HST3D by IF Technology. These simulations focused on the long-term thermal impacts on the surroundings. Currently, only short-term monitoring results are available. The thermal model validation in this deliverable is applied to the DoubletCalc3D software by TNO, focusing on the short-term heat evolution within the aquifer. The thermal simulations will be performed using the actual injected volumes of hot water from the first few months of heat storage in the HT-ATES system, and the results of the thermal evolution will be validated using DTS data from the monitoring well.

3.2.2 Site description

The Agriport A7 horticultural area in the north-western part of the Netherlands is well on its way to become more sustainable and less dependent on fossil fuels. Their heat supply is provided for approximately 60% by three deep geothermal doublets, and in the spring of 2021 the first HT-ATES site has become operational to increase this percentage. The geothermal systems and HT-ATES site are operated by the energy company ECW Energy. A map of the area is shown in Figure 3.2.1. The HT-ATES site has been developed in the Maassluis Formation (early Pleistocene age) at a depth of 360 to 383 m. The analysis of the test drilling, which is completed for use as monitoring well, and the drillings of the hot and cold wells indicated very similar hydrogeological characteristics and it is assumed that the target formation is rather homogeneous. The locations of the three wells and the distances between them are shown in Figure 3.2.1. A detailed description of the site can be found in deliverable D2.1.

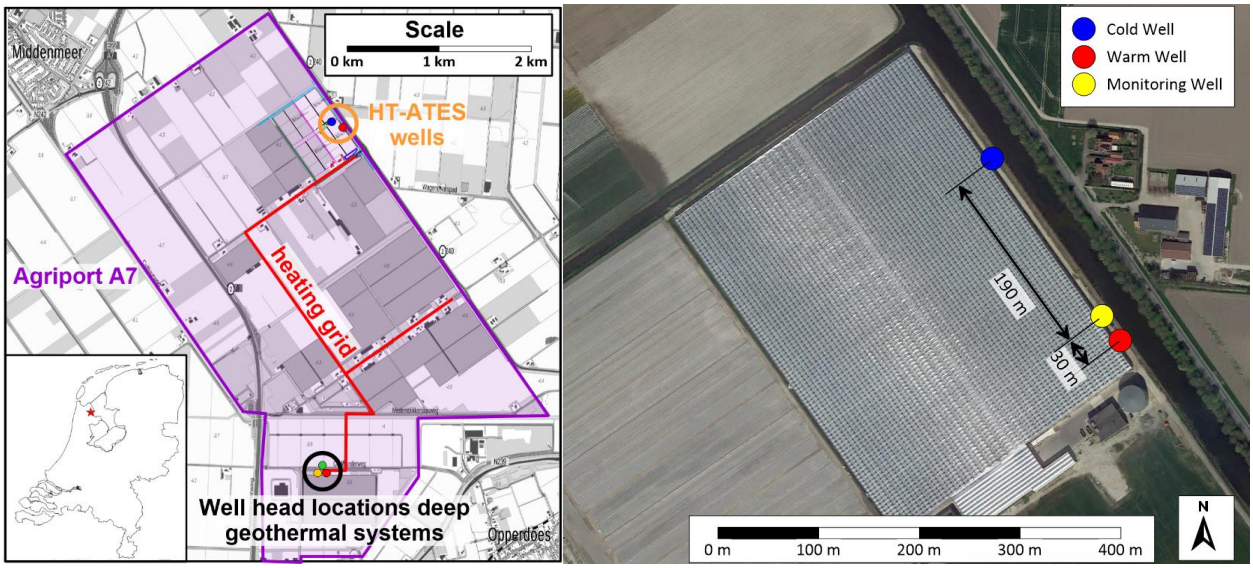


Figure 3.2.1: Left: Map of the Agriport area with heating grid and the locations of the HT-ATES and geothermal well heads. Right: Close up of the locations of and distance between the HT-ATES wells.

3.2.3 Model codes and data used

3.2.3.1 Thermal simulations

Thermal flow simulations were performed with ROSIM v0.1 (beta version, link www.thermogis.nl). ROSIM is an open source windows application developed by TNO, containing a workflow for 3D simulation of geothermal production and high-temperature aquifer thermal energy storage. With ROSIM, a static 3D subsurface simulation grid (including grid definition, layer properties and well data) and flow input deck (flow constraints) can easily be created. The dynamic flow simulation can be run with DoubletCalc3D (DC3D)[1] or OPM[2] (opm-project.org). For the model validation with monitoring data from the HT-ATES case study in Middenmeer, DC3D is used.

DC3D is an extension of DoubletCalc2D which is a software tool that is developed by TNO. DC3D is a dedicated single phase simulator based on SEAWAT algorithms. It enables the calculation of temperature and pressure development around two or more geothermal/HT-ATES wells in three dimensions over time. The input of DC3D consists of fixed reservoir parameter values for, and/or 3D maps representing temperature, aquifer depth, aquifer thickness, porosity, net-to-gross, permeability and salinity. The output includes:

- graphs showing pressure, flow rate and temperature at the wells against time,
- 3D grids of pressure and temperature per time step

Contrary to OPM, DC3D includes temperature-dependent density, which is crucial for the modelling of HT-ATES systems due to the importance of density-driven flow for the HT-ATES performance (Table 3.2.1).

[1] https://www.nlog.nl/tools
[2] OPM is a fully implicit, thermal black-oil simulator (Eclipse E100 clone) capable of running industry standard models. The software is open source and freely available. It is developed and maintained by an international community and grows rapidly, the online manual can be found online: https://opm-project.org/?page_id=955

Table 3.2.1: Differences between OPM and DoubletCalc3D simulation software.

	OPM	DoubletCalc3D
Temperature-dependent water viscosity	✗	✓
Temperature-dependent water density (HT-ATES)	✗	✓
Simultaneous Rate & Pressure flow constraint	✓	✗
Faulted grids	✓	✗

The hot, cold and monitoring well data provide sufficient information to develop a suitable subsurface model. Uncertainties will be addressed in section 3.2.4. Table 3.2.2 and Table 3.2.3 give an overview of the required parameters and the data used for the subsurface model and the flow simulation respectively.

Table 3.2.2: Required input subsurface model and used data.

Required input subsurface model	**Used data**
Depth sand and clay layers	Heatstore D2.1
Thickness sand and clay layers	Heatstore D2.1, later adapted to fit DTS data
Initial temperature subsurface	DTS data from monitoring well
Porosity	Athy's law
Horizontal permeability	Heatstore D2.1
Vertical permeability	Heatstore D2.1
Matrix rock conductivity	Calculated with porosity and bulk matrix conductivity (Hantschel, Kauerauf 2009)
Water conductivity	0.6 W/mK
Salinity	Data
Heat capacity sediment	850 J/kgK (Hantschel, Kauerauf 2009)
Density sediment	2000 kg/m3 (Fjaer, Holt et al. 2008)
Net to gross	1
Amount of wells	2, coordinates see Figure 3.2.3
Filter depth	362.5 – 279 m-mv
Well diameter	31 inch
Skin	0

Table 3.2.3: Required input for flow simulation and used data.

Required input flow simulation	
Flow rate hot well	Flow meter, Figure 3.2.5
Injection temperature hot well	DTS data, Figure 3.2.5

Spatial discretization

The model grid consists of a 600 x 600 x 96 m grid containing 190.800 grid cells of 10 x 10 m horizontally. The vertical grid size varies per vertical layer, the cell size is smaller (1 m) at the target aquifer. Horizontal local grid refinement of 5 m and later of 2 m is applied around the wells, to model more accurately the temperature and pressure development around the well.

Layering

A layer cake model with homogeneous layers has been created based on the well data and the estimated rock properties have been aligned with the thermal model in HST3D reported in D2.1 (Table 3.2.4, Figure 3.2.2).

Table 3.2.4: Hydrogeological and thermal properties used in the model.

Layer nr	Interpretation	Base depth	Thickness	Perm xy	Perm z	Temperature	Salinity	Porosity	Rock matrix conductivity
		mbgl[3]	m	mD	mD	°C	ppm	-	W/mK
Overburden	top layer	4	4	142	14	12.0	0	0.60	9.29
	aquifer 1	17	13	21254	5314	12.1	0	0.41	6.26
	aquitard 1	33	16	141	14	12.3	0	0.59	8.62
	aquifer 2	47	14	14047	3512	12.5	0	0.40	6.16
	local aquitard	53	6	140	14	12.6	574	0.58	7.99
1	aquifer 2	67	14	14020	1402	12.7	1148	0.40	6.11
2	aquitard 2	74	7	14	1	12.8	1751	0.57	7.54
3	aquifer 3	229	155	55236	13809	13.7	6400	0.39	5.85
4a*	aquitard 3	329	100	13	1	15.2	13719	0.48	4.88
4b*	aquitard 3	360	31	13	1	16.0	17478	0.45	4.43
5**	aquifer 4	383	23	16862	5616	16.3	19028	0.37	5.33
6	aquitard 4	425	42	13	1	16.6	20894	0.43	4.11
7	aquifer 5	460	35	8523	2131	17.1	23104	0.36	5.19

* this layer is split in two to increase the amount of cells right above the target aquifer and simulate this more accurately.

** adapted to fit DTS data:

Layer nr	Interpretation	Base depth	Thickness	Perm xy	Perm z	Temperature	Salinity	Porosity	Rock matrix conductivity
		mbgl	m	mD	mD	°C	ppm	-	W/mK
5a	aquifer 4	367	5	26592	6648	16.2	18569	0.37	5.35
5b	aquifer 4	368	1	133	13	16.2	18856	0.44	4.30
5c	aquiter 4	380	12	16847	5611	16.3	19315	0.36	5.32

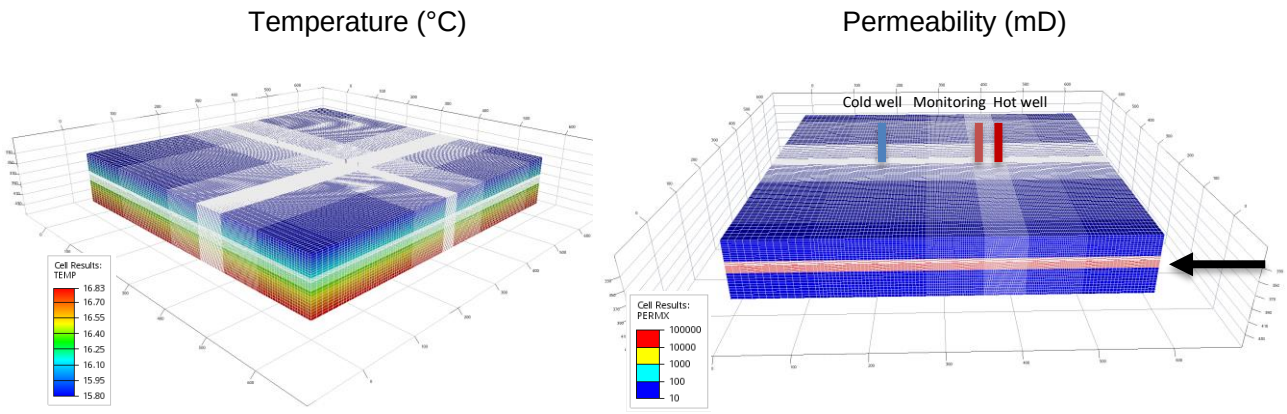


Figure 3.2.2: Subsurface model including parameters, an example of temperature and permeability.

[3] Meter below ground level

Well configuration

The hot and cold wells are 220 m apart. The monitoring well is located in between, at 30 m distance from the hot well (Figure 3.2.3).

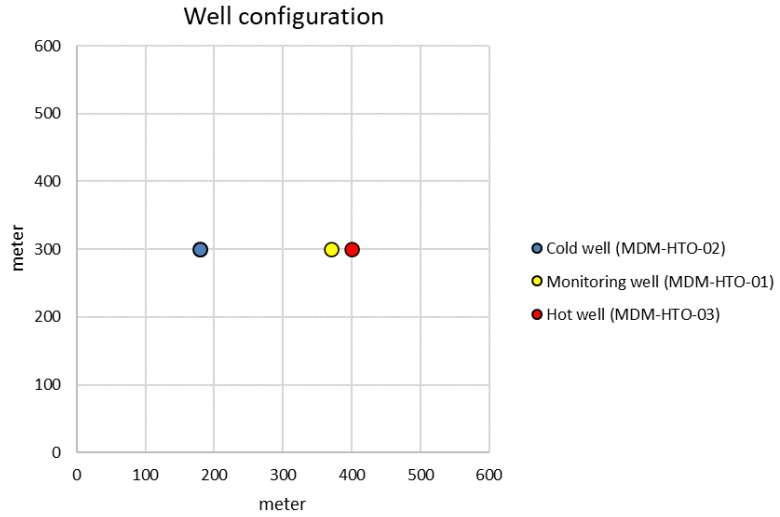


Figure 3.2.3: Location of wells in the model.

Pumping scheme

Figure 3.2.4 shows the pumping scheme as derived from the monitoring data, it shows the temperature and flow rate in the hot well. Measurement frequency is 10 minutes, both for flow rate and temperature (see also D5.2). In Figure 3.2.5 the monitoring is adapted to use as model input by calculating daily averages.

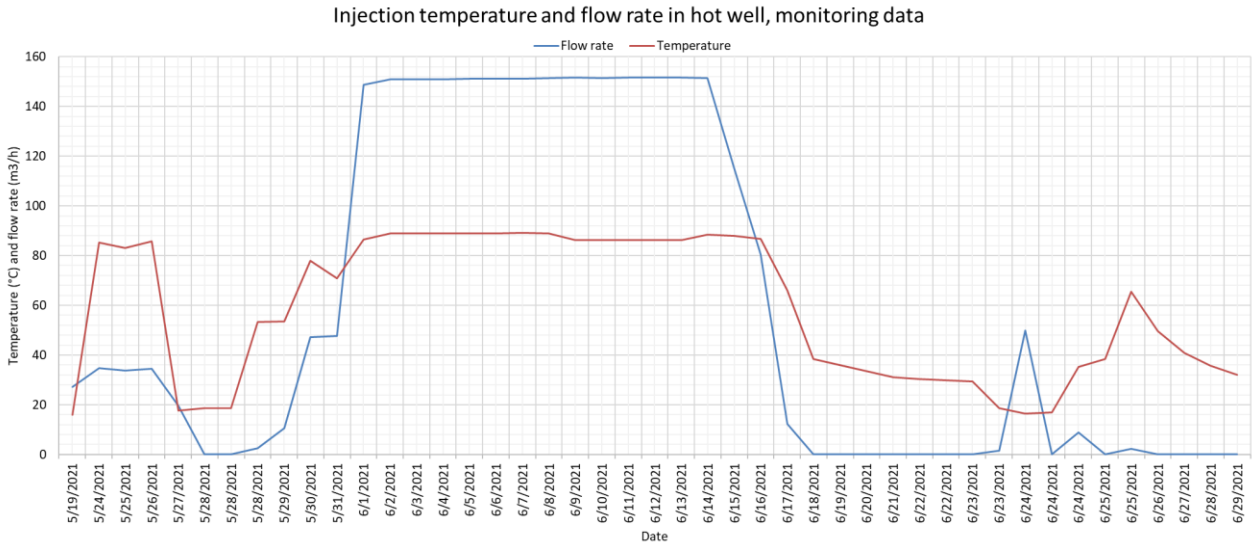


Figure 3.2.4: Injection temperature and flow rate in the hot well from the monitoring data. Note that some dates include several flow rates and temperatures data points within one day, this has been changed to daily averages for modelling (Figure 3.2.5).

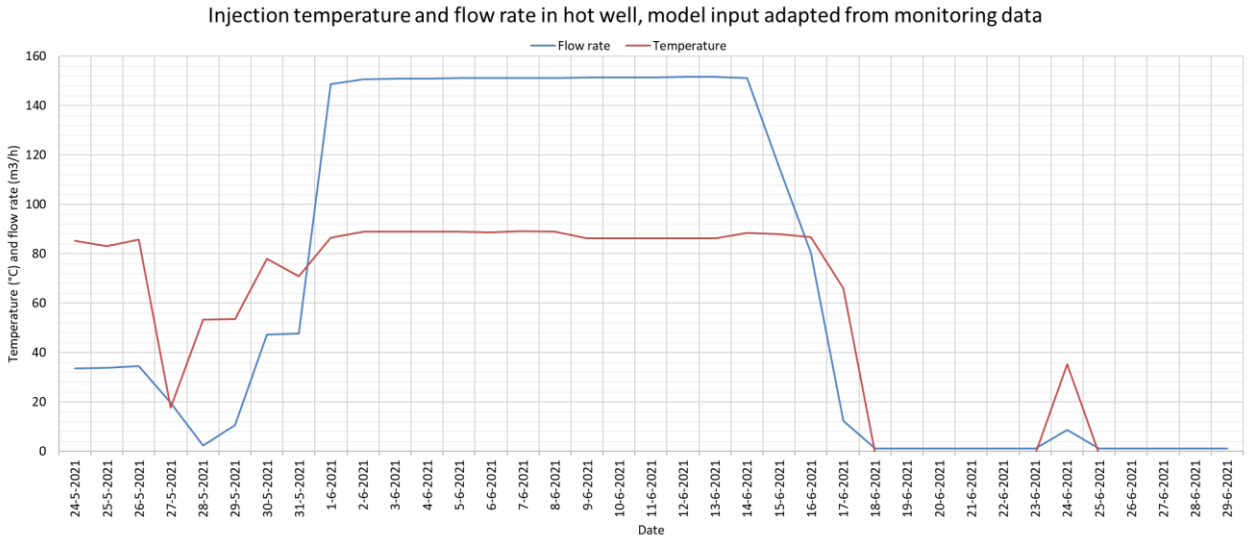


Figure 3.2.5: Injection temperature and flow rate in the hot well, daily averages from monitoring data for model input. This scheme is used in the thermal simulations.

3.2.3.2 Reactive transport simulations

The software code TOUGHREACT (reactive transport software), and the models developed with this code are described in detail in deliverable D2.1. The reactive transport model was updated to match the final design of the HT-ATES site and the thermal model:

- The aquifer thickness in the model was increased from 18 to 23 m by increasing the thickness of the 8 layers from 2.25 m to 2.875 m
- The distance between the hot and cold well was increased from 195 to 220 m
- Initial temperature is changed to 15.5 °C
- Permeability of aquifer and over- and underburden is in line with the thermal model input
- Rock density, conductivity and heat capacity are in line with the thermal model input
- Initial geochemistry of the groundwater is based on average values from 5 baseline measurements

The temperature evolution predicted by the reactive transport model has been compared to the DTS data for the initial stage of heat storage at the HT-ATES site. Geochemical monitoring data, e.g. fluid analyses from the monitoring well will be compared to the simulation results described in D2.1, once they are available. Unfortunately, the sampling from the monitoring well, and sample analyses has not yet taken place at the time of writing.

3.2.4 Model validation

3.2.4.1 Monitoring data

The temperature is monitored in the hot, cold and monitoring well using distributed temperature sensors (DTS) and recorded every 10 minutes. The temperature at the hot well is used as input for the simulations, and model validation for the first loading phase is performed using the temperature data from the monitoring well. The arrival of the heat front and the rate of heating at a specific level of the monitoring well and differences of arrival times along the length of the monitoring well (which is an indication of the buoyancy effect, e.g. the angle of the heat front) were used to check the performance of the thermal and reactive transport simulations. Figure 3.2.6 shows a visualization of the temperature within the monitoring well during the first ~20 days of heat storage. The heat front arrives at the top of the reservoir after ~11-12 days of the start of heat storage. Figure 3.2.7 shows temperature profiles along the monitoring well for several points in time.
The reactive transport model will also be validated by geochemical analyses of groundwater samples from the monitoring well, once available. The water samples still need to be taken and analyzed and it is currently unclear when this will be done.
The monitoring data, both DTS and geochemical analyses will be discussed in detail in deliverable D5.2.

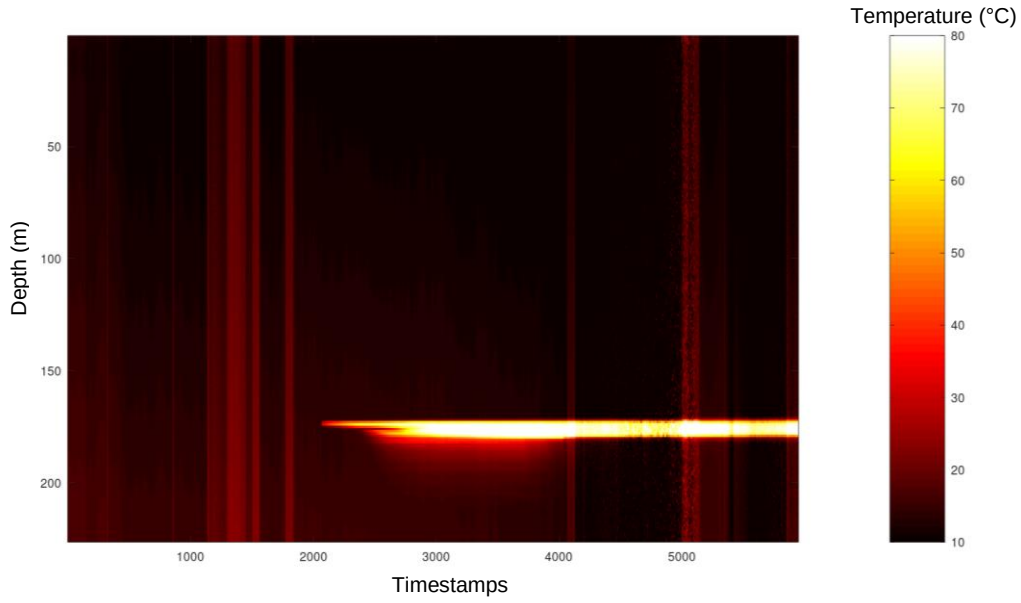


Figure 3.2.6: Visualisation of DTS temperature data in the monitoring well. Y-axis shows depth in meters (not from the surface, the reservoir is located at 360-383 mbgl), the x-axis represents timestamps (recorded measurements), between 19 May 2021 and 30 June 2021. Note that the timestamps do not represent a continuous timeline due to gaps in the recorded data.

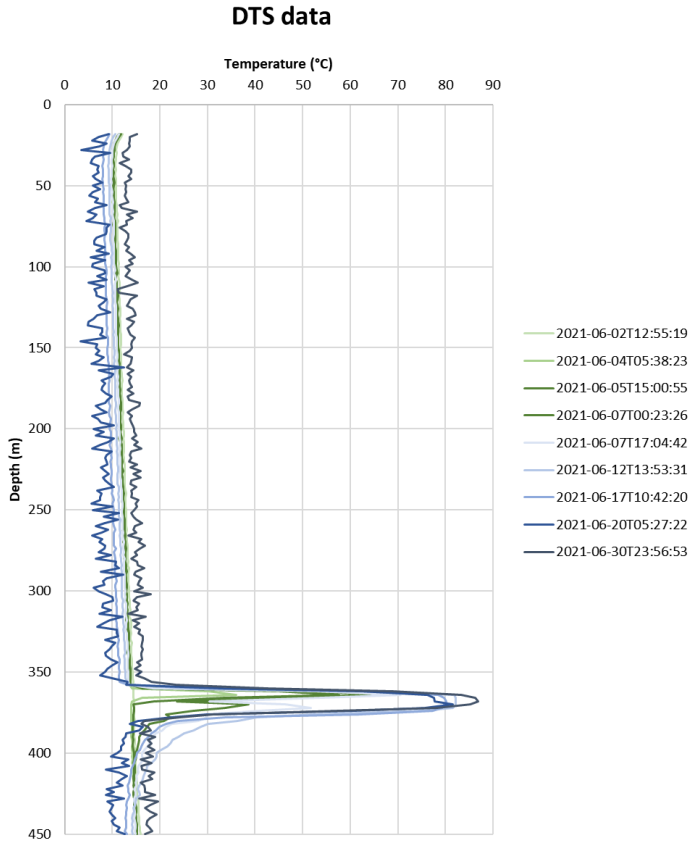


Figure 3.2.7: DTS data in the monitoring well at nine selected timestamps.

3.2.4.2 Thermal model

The model input data (Table 3.2.4) is used to create the hydrogeological subsurface model. Several simulations with changing parameters have been performed in order to fit the DTS data, results are shown in Figure 3.2.8. It should be taken into account that the plotted temperature in the monitoring data is taken from the upper part of the aquifer, while the modelled temperature is the average of the filter interval. Furthermore, an offset of several degrees in temperature in the monitoring data is possible due to uncertainties in the measured temperature. Therefore, this exercise is a first attempt to compare the monitoring data to the model results and show the effects of changing aquifer characteristics instead of aiming for an exact fit between the two.

The first simulation was done with a lateral cell size of 5 m and a vertical cell size of 1 m (see Figure 3.2.8, “cell size xy = 5m”). When compared to the DTS data (blue curve), the modelled temperature shows a more gradual curve compared to the steeper DTS curve. This is caused by numerical diffusion due to the large lateral grid cell size. Therefore, the lateral grid cell size has been reduced to 2 m (vertical grid cell size in the target aquifer is still 1 m). When a lateral cell size of 2 m has been applied, the shape of the temperature curve fits the DTS data better (see Figure 3.2.8, “cell size xy = 2m”).

The DTS data (Figure 3.2.6) clearly shows that the thermal front moves faster in the upper part of the aquifer than in the lower part. This coincides with the well description, showing coarser sand in the upper part of the aquifer and finer sand in the lower part of the aquifer. In the middle of the aquifer seems to be a thin layer that lags behind on the thermal front. A possibility for this could be a high(er) permeability streak (HPS) in the upper part of the aquifer, and a thin, less permeable (clay-rich) layer in the middle part. However, at first sight, the well description and log data do not show this possible less permeable layer or any clear local heterogeneities. These features are implemented in the model to try and fit the shape of the DTS data (Figure 3.2.8, “cellsize xy = 2m, HPS” & Figure 3.2.9).

In Figure 3.2.8 (“cell size xy = 2m, HPS”) is seen that the hot water in the model arrives in the monitoring well later than in the DTS data. This has likely to do with the thickness of the aquifer layer. From the well data, the total aquifer thickness was estimated at 23 m (360-383 mbgl), with a well screen installed from 362-379 mbgl. To fit the DTS data, the aquifer thickness in the model has been reduced to 20 m (from 361-381 mbgl). The curve shows an improved fit with the DTS data (Figure 3.2.8, “cell size xy = 2m, HPS, thickness 20m”).

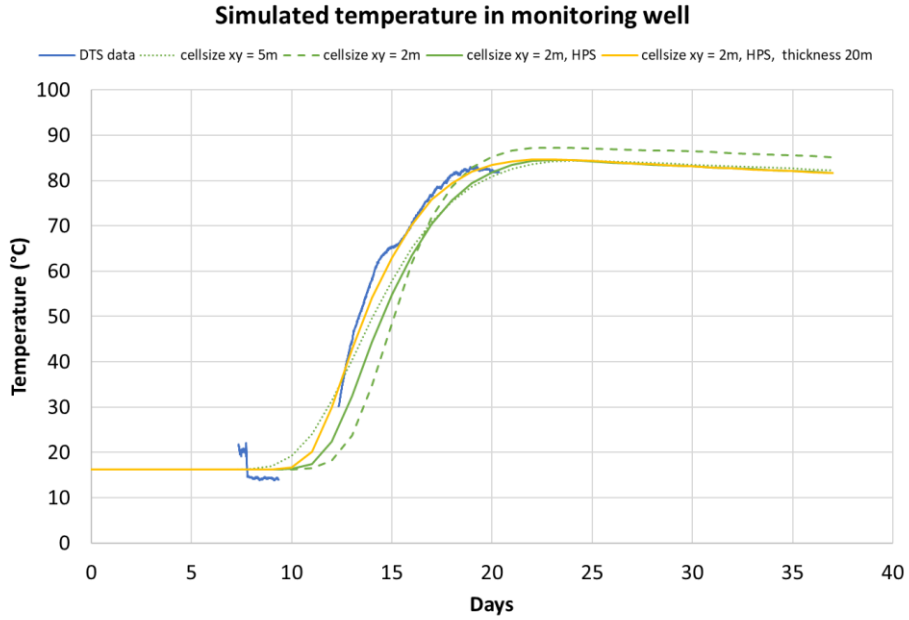


Figure 3.2.8: Simulated temperature in monitoring well for different model settings over time (green and yellow curves). The blue line shows the DTS data, note that the exact depth of the measured data is not known, but somewhere in the upper part of the reservoir. The yellow line shows the model (average temperature of the aquifer) with a good fit to the monitoring data. HPS is ‘High Perm Streak’, referring to the adaptation in the model regarding the more permeable upper part of the aquifer, a low permeable thin layer and the less permeable lower part of the aquifer.

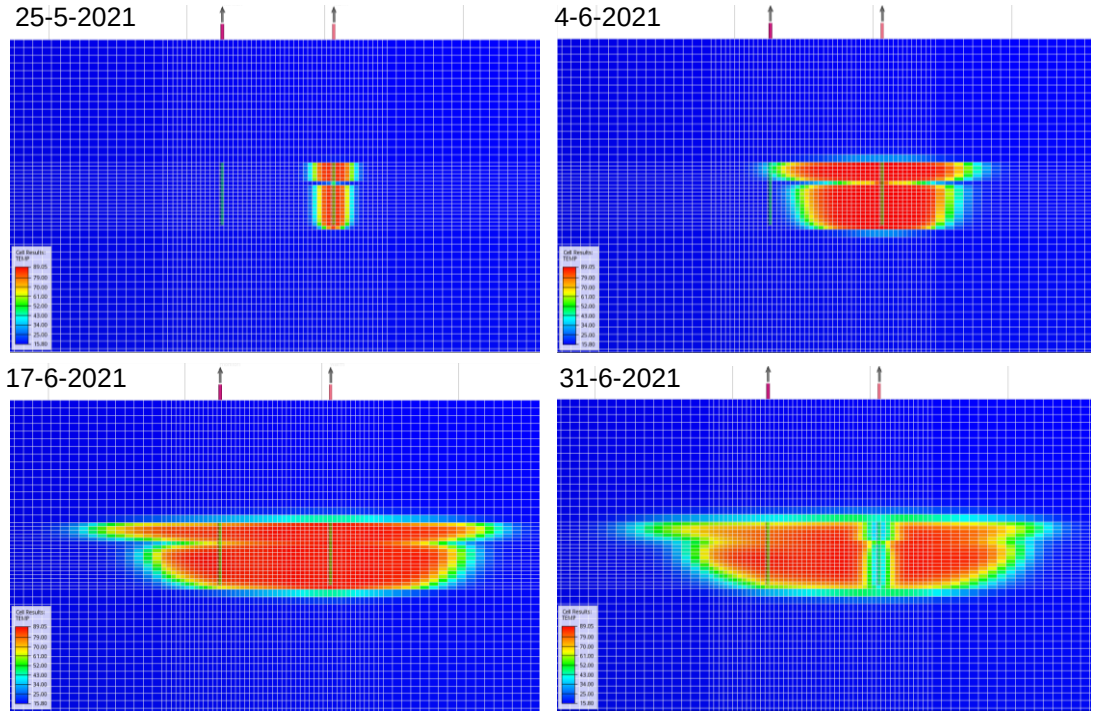


Figure 3.2.9: Cross section 3D thermal model. Four timesteps are shown to visualise the development of the shape of the thermal volume for the scenario with a high permeable streak and an underlying low permeable streak.

Figure 3.2.9 shows the development of the shape of the thermal volume in the subsurface for the scenario with a low permeable streak in the storage aquifer. This is shown in more detail in Figure 3.2.10, in which the temperature along the monitoring well in the aquifer is plotted against the depth of the filter (362-379 mbgl) for nine selected timestamps. The first five timestamps show great similarities between the simulation and the DTS data. The heat front arrives in the upper part of the aquifer first, which can be related to the combined effect of buoyancy and the higher permeable streak implemented in the model. At a depth of 367-368 mbgl the heating is slightly lagging behind, and the implementation of the low permeable streak in the model leads to similar results. In the last four timestamps it is seen that the DTS data shows a more gradual temperature course in the middle part and lower temperatures in the lower part of the aquifer, while the simulation data still shows a clear difference in temperature in the middle part of the aquifer and a higher temperature in the lower part. It is therefore possible that the low permeable streak has a higher permeability than estimated in the model, which will also allow for more buoyancy flow towards the upper part of the aquifer. This will probably better fit the DTS data in the last four timestamps of Figure 3.2.10. Uncertainty remains in reservoir parameters such as the kh/kv ratio, heterogeneity, thickness of sand and clay layers, inflow characteristics, salinity and bulk heat capacity. Also technical issues related to the well might impact the water flow and hence the temperature evolution within the aquifer. These can be investigated in further research with model validation by data assimilation and optimization of the system. In addition, the DTS data needs a detailed evaluation to assess the technical robustness and exclude technical issues.

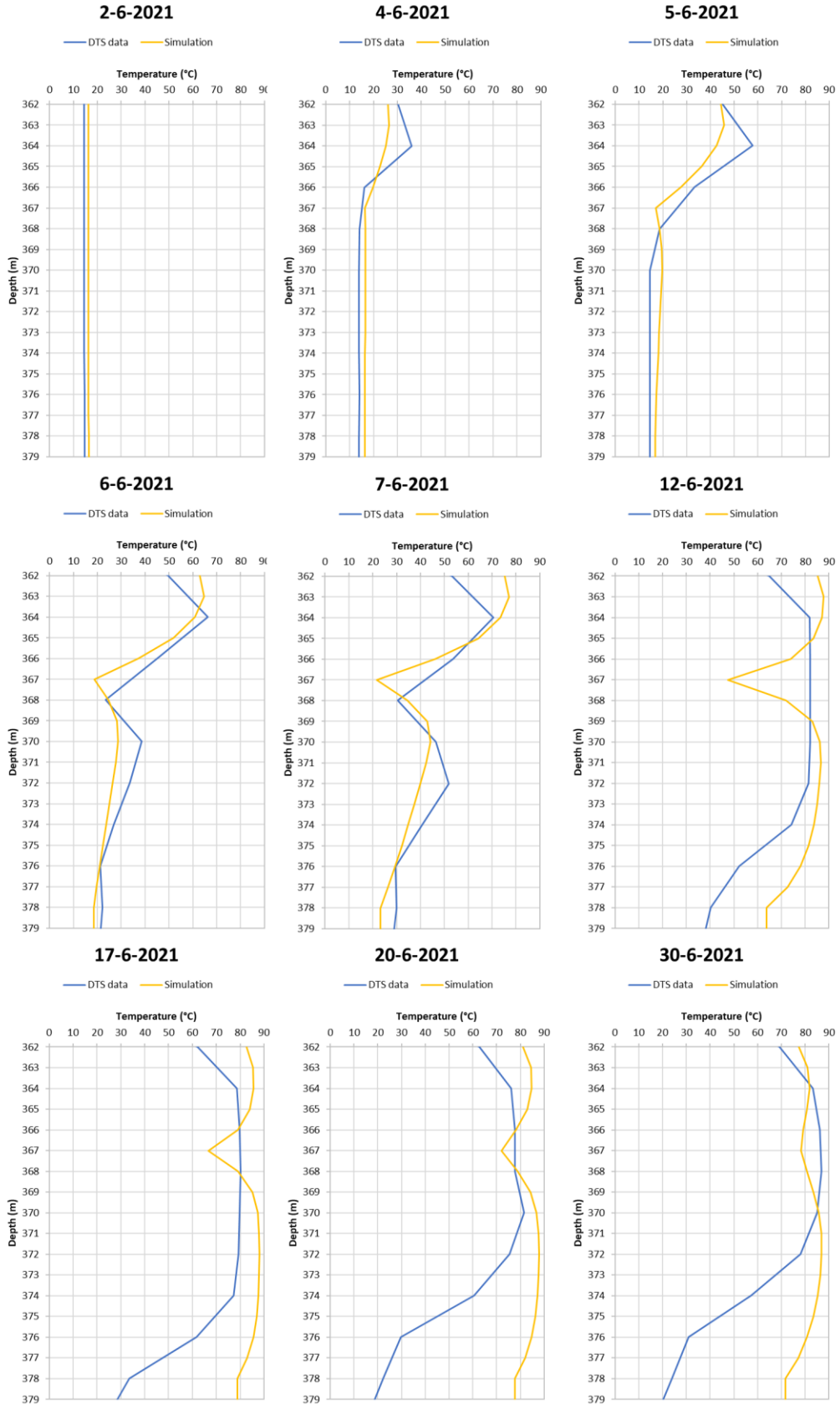


Figure 3.2.10: Temperature profiles along the monitoring well in the target aquifer for nine different moments in time. The blue line shows the DTS data, the yellow lines the simulated temperature for the scenario with adapted cell size, the high permeability streak and change in thickness.

3.2.4.3 Geochemical and reactive transport

First, the thermal evolution prediction of the reactive transport model was compared to the DTS measurements in the monitoring well for the short term, the first weeks of heat storage. For this purpose, only the short term is relevant since the DTS data showed that the heat front had already passed the monitoring well after a few days of loading. The DTS measurements in the monitoring well were compared to the predicted temperature at the location of the monitoring well in the model. For this purpose, the first weeks of the hot water injection (May/June 2021) was simulated based on the true injected volumes. Since TOUGHREACT cannot easily implement variable injection volumes, the injected volumes were approximated by two phases:

- ~7.5 days of constant injection at 44 m^3/h;
- ~16 days of constant injection at 150 m^3/h

The final injected volumed after 23.5 days in the model is equal to the true injected volume; 66,420 m^3. The average injection temperature is 82.5°C. The results of the simulated temperature and the DTS data is shown in Figure 3.2.3. The arrival of the heat front is slightly later in the field than predicted, whereas the heating is faster than predicted, e.g. the curve of the DTS data is steeper. This result is very similar to the results of the initial thermal simulation results. Decreasing the cell size in the thermal model improved the comparison with the DTS data. In TOUGHREACT smaller grid cells will further increase computation times, which is already very long. Overall, the model prediction is quite good and good enough for its purpose, considering that the model set-up is focused on the geochemical reactions rather than an accurate thermal prediction.

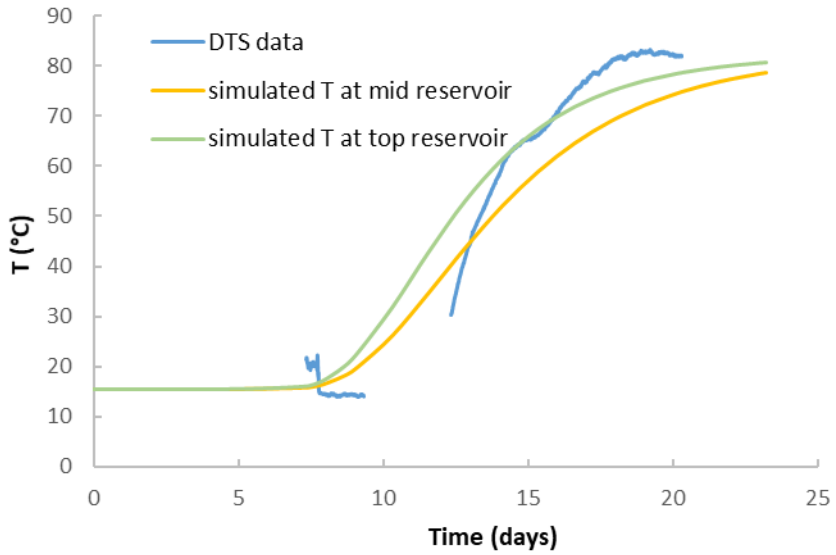


Figure 3.2.11: Simulated and measured temperature at the monitoring well for the first 23.5 days of heat storage. Note that the exact depth of the measured data is not known, but somewhere in the upper part of the reservoir.

Eight scenarios were run for the reactive transport simulation (Table 3.2.5); see D2.1 for a detailed description. Four of these scenarios simulate the HT-ATES system with water treatment, e.g., the addition of CO_2 to the hot water to prevent carbonate scaling. Water treatment is applied in the HT-ATES and therefore these four scenarios are relevant for model validation. The predicted evolution of the pH and calcium and magnesium concentration of the water at the monitoring well is different for the four scenarios (Figure 3.2.4). Geochemical monitoring data from the monitoring well will, as soon as the data is available, be compared with the predicted values to distinguish between the scenarios. The measured variations in the various baseline groundwater analyses are small, and especially the distinction between scenario 3 and the other three scenarios based on monitoring data should be possible.

Table 3.2.5: Reactive transport simulations scenarios performed with TOUGHREACT.

Scenario	With/without CO2	Mineral reactions	SI calcite 0.3 to 0*	SI dolomite 0.9 to 0*
1	no water treatment	Calcite and dolomite	19 to 80°C	19 to 150°C
2	no water treatment	Calcite and dolomite	80 to 150°C	80 to 150°C
3	with CO_2 added	Calcite and dolomite	19 to 80°C	19 to 150°C
4	with CO_2 added	Calcite and dolomite	80 to 150°C	80 to 150°C
5	no water treatment	Calcite precipitation only	19 to 80°C	-
6	no water treatment	Calcite precipitation only	80 to 150°C	-
7	with CO_2 added	Calcite precipitation only	19 to 80°C	-
8	with CO_2 added	Calcite precipitation only	80 to 150°C	-

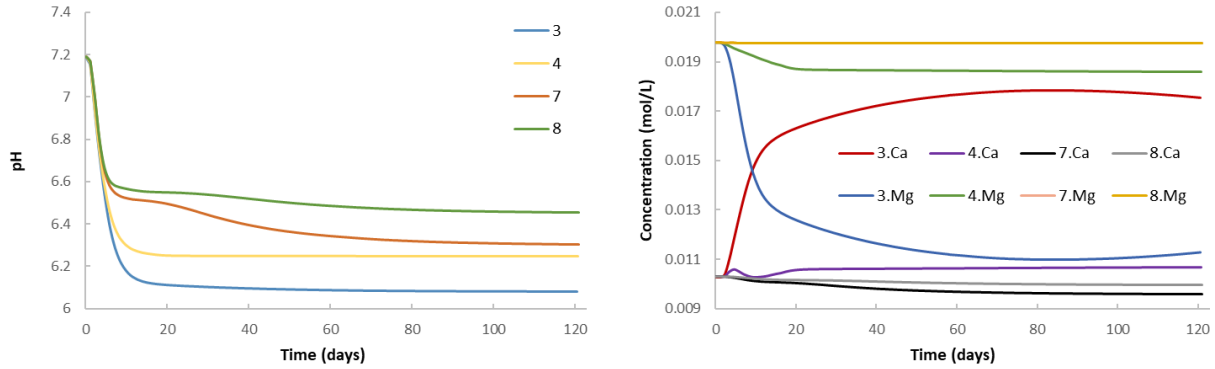


Figure 3.2.12: Simulation results at the monitoring well for the four scenarios with water treatment.

3.2.5 Future performance prediction

3.2.5.1 Thermal model

The model that fitted best to the DTS data (Figure 3.2.8, “cell size xy = 2m, HPS, thickness 20m”), is used to perform a thermal simulation for five years. Flow input data is used from the pumping scheme described in D2.1 Figure 3-4 (200 m^3/h case). In year 1-3 the flow rates are lower than in year 4 and 5 due to the start-up of the system. An injection temperature of 90°C and a cut-off temperature of 55°C was applied to the hot well. The injection temperature in the cold well is set to 30°C.

The results are shown in Figure 3.2.13 and Figure 3.2.14, and are similar to the results obtained in HEATSTORE D2.1. In the last two years, more water is produced from the hot well (see pumping scheme D2.1), therefore the efficiency is higher in these years, as more heat is recovered. The efficiency in year 4 is higher than in year 5 because in year 3 the production rate is lower and more heat is left behind, causing the aquifer to be warmer at the start of year 4. In year 4 the production rate is higher and the cut-off temperature is reached, causing the aquifer to be cooled down more which explains the lower efficiency in year 5.
In year 1, 4 and 5 the cut-off temperature is reached, therefore the production is stopped. Large timesteps of 1 month (chosen to reduce computation time), cause that for these years a whole month of production is missing which is seen in the temperature development of the hot well. In future simulations, a smaller timestep can improve simulation results.

In WarmingUP (Netherlands national research project) or possible future research projects, further performance prediction will be done with data assimilation in where all available monitoring data will be used to update the model state and reflect the observations from the DTS data. Model variations concerning several subsurface parameters will be applied and tested against the available data.

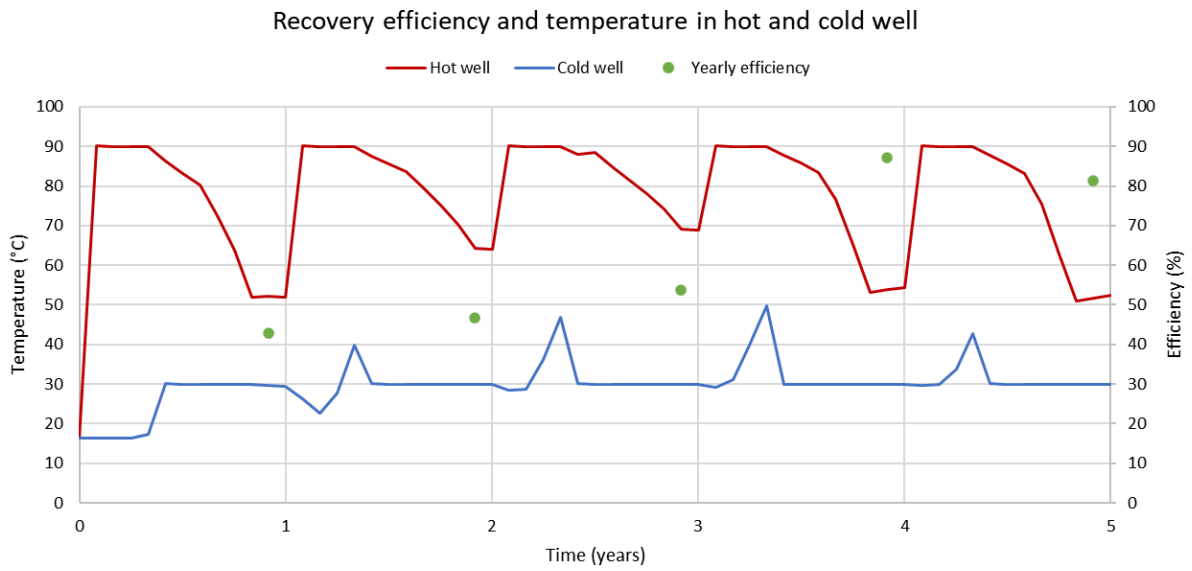


Figure 3.2.13: Calculated temperatures in the wells for the first 5 years.

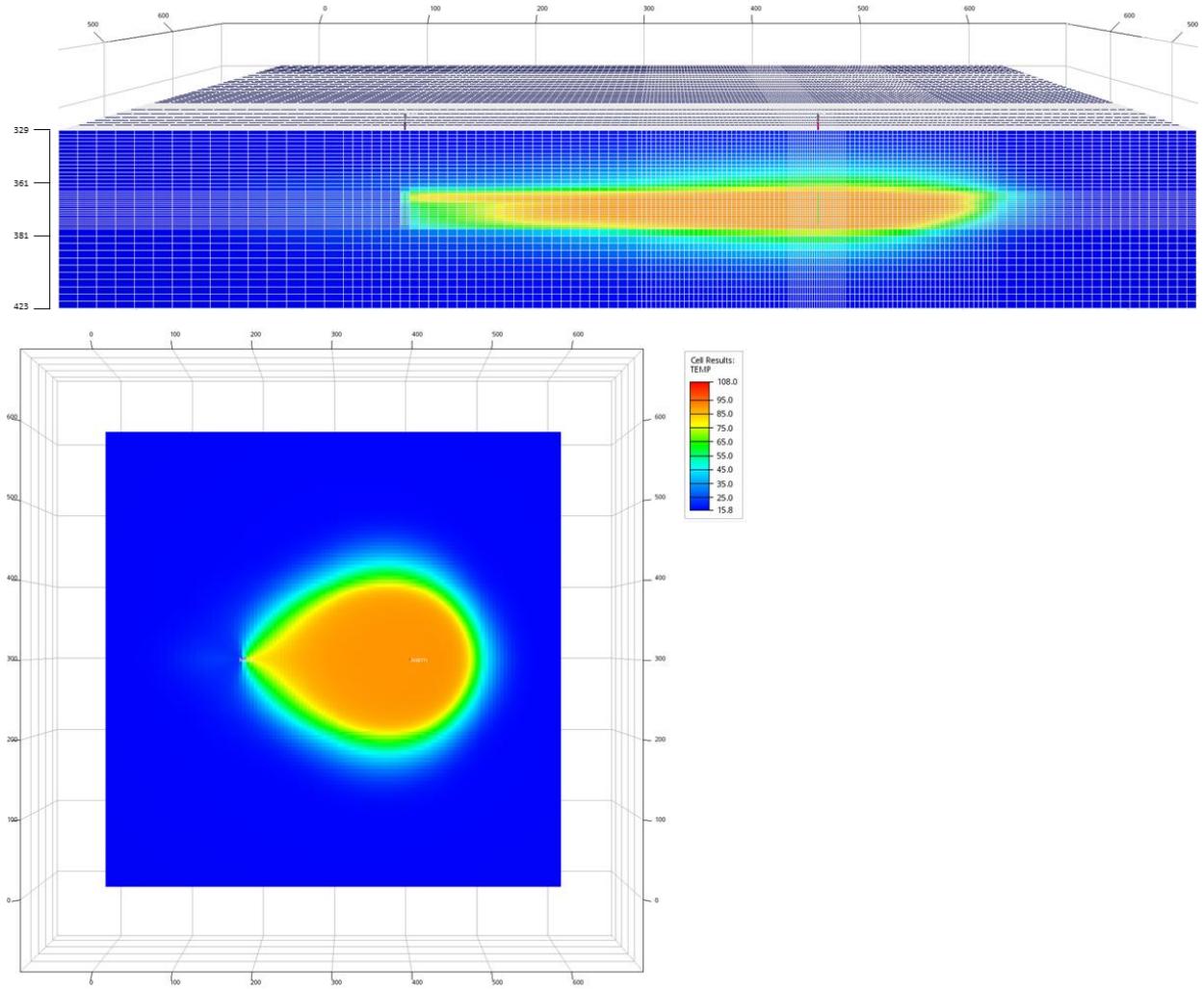


Figure 3.2.14: Vertical and horizontal cross section of the calculated temperatures at the end of the last heat injection period of the 5th year.

3.2.5.2 Reactive transport model

Future performance prediction will be done after model validation by geochemical monitoring data. Unfortunately, this can only be done after the HEATSTORE project has ended since the geochemical monitoring data is not yet available.

3.3 Geneve, Switzerland

This section aims at presenting the results of the subsurface modelling activities carried out for the Geneva case study and implemented in the framework presented in Figure 3.3.1. In particular, the focus is directed to the evolution of TH-THM and HM models produced before and after site-specific data from Geo-01 and GEo-2 data become available.
The overall approach has always been to combine subsurface and energy system data to produce in a first step static 3D models and then 3D dynamic models to predict the performance of the HT-ATES system object of the study.
In a first phase, subsurface literature data were used to provide constraints with respect to petrophysical and hydraulic properties of the potential reservoir. A set of simplified geological models were created to run dynamic TH and THM models. In a second phase, once new data were available from the GEo-01 and GEo-2 wells, site-specific modelling, were performed resulting in improved and calibrated version. The results also were used for T6.6 about assessment of the environmental effects of HT-ATES in Geneva and upscaled for T1.3 about favourability assessment of the potential of HT-ATES implementations in Switzerland.

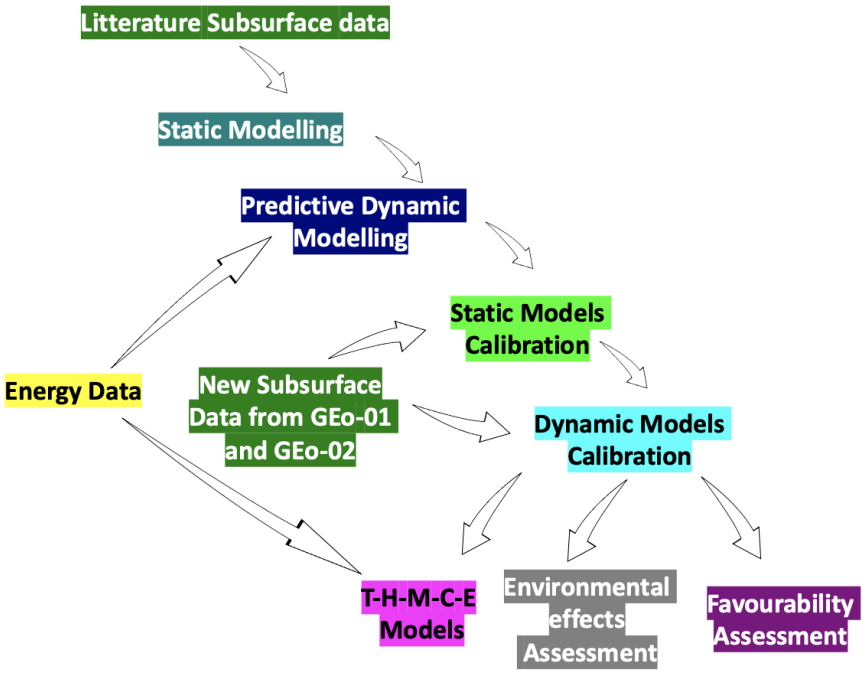


Figure 3.3.1: Modelling framework implemented for the Geneva case study

The Geneva pilot aims at assessing the feasibility of seasonal storage of up to 50 GWh/a waste heat from the Cheneviers incinerator (Canton of Geneva), using a high temperature aquifer thermal energy storage concept. Several target aquifers exist at different depths and are currently being explored and characterized phase to better constrain options for HT-ATES. The modelling inputs of HEATSTORE WP2 therefore also aim at providing improved guidance for optimal site selection in a complex geologic situation.
The Western Geneva basin is geologically complex and potentially "challenging" for HT-ATES, due to significant topography at the boundaries potentially imposing artesian hydraulic conditions, strike-slip and thrust faults leading to compartmentalization of aquifers and providing potential leak-off structures, inclination of aquifers raising the question of the role of buoyancy in storage, and variable sedimentary facies within the aquifers leading to heterogeneity. Modelling therefore starts with addressing the question of how these complexities can that potentially affect possible storage by simulating generic scenarios of geologic situations to be encountered in the Geneva underground. Results are expected to provide a foundation for selecting the "best" underground volumes within the Cenozoic and Mesozoic geologic units. Such characterization will allow the industrial partner to plan future activities such as drilling additional wells, design the development of the

district heating network, define business models, and evaluate the overall sustainability of the ATES system in Geneva.
Since the availability of suitable drilling sites is relatively low on the Geneva territory, the choice of geological sites is strongly influenced by surface boundary conditions such as proximity to the waste water network, suitability of nearby crops, and SIG's future plans to increase the district heating network coverage. Drilling targets have nevertheless been defined according the available geological and geophysical data, which consists mainly of 2D reflection seismic data. This allowed the identification of the Mesozoic units as the most favourable targets for geothermal development. With these targets in mind, it is important to have a broader view of the surrounding geology to be able to construct geometrically relevant simulation models.

3.3.1 GEO-series Wells

The GEo-01 and GEo-02 wells are in the western part of the Geneva Canton and were drilled as a exploration wells to characterize the geothermal conditions in the Upper Mesozoic carbonates and assess their geothermal potential for applications as hydrothermal resources or storage (Figure 3.3.2).
GEo-01 well is 744 m deep and drilled into the Cenozoic Units down to 407 MD (Measured Depth) to then enter the Mesozoic carbonates of the Lower Cretaceous unit down to 648 m MD and eventually penetrated the Upper Jurassic carbonates until bottom hole. Artesian fluid flow condition characterizes this well, with a flow rate of 55l/s, 32.4℃ wellhead temperature and 8bars wellhead pressure, representing a very encouraging geothermal resource suitable for heat production and direct uses. Geothermal fluids rise towards the surface along a strike-slip fault structure cutting through the Lower Cretaceous and Upper Jurassic carbonates, being the former responsible of more than 70% of the total mass discharged (Guglielmetti et al., 2020). The natural recharge of the system here is from the Jura Mountain chains and circulation at depth is related to the hydraulic gradient. The faults encountered in the Lower Cretaceous are most likely open faults, laterally confining and vertically promoting localized fluid circulation.
GEo-02 well is 1456 m deep and drilled the same units as GEo-01 reaching the top of the Lower Cretaceous at 769.9m MD. Artesian fluid flow condition characterizes this well, with low flow rate of 0.3-0.6l/s, 18˚C even if the measured bottomhole temperature observed at the end of the drilling operations was about 55˚C and about 12 bars stabilized wellhead pressure and 8bars. As per GEo-01 the natural recharge of the system is expected to be dominated by meteoric waters infiltrating in the Jura Mountain chains and circulation at depth is related to the hydraulic gradient. The fractures encountered in the Lower Cretaceous and Upper Jurassic are most likely mineralized and tight, preventing large fluid flow in this region.

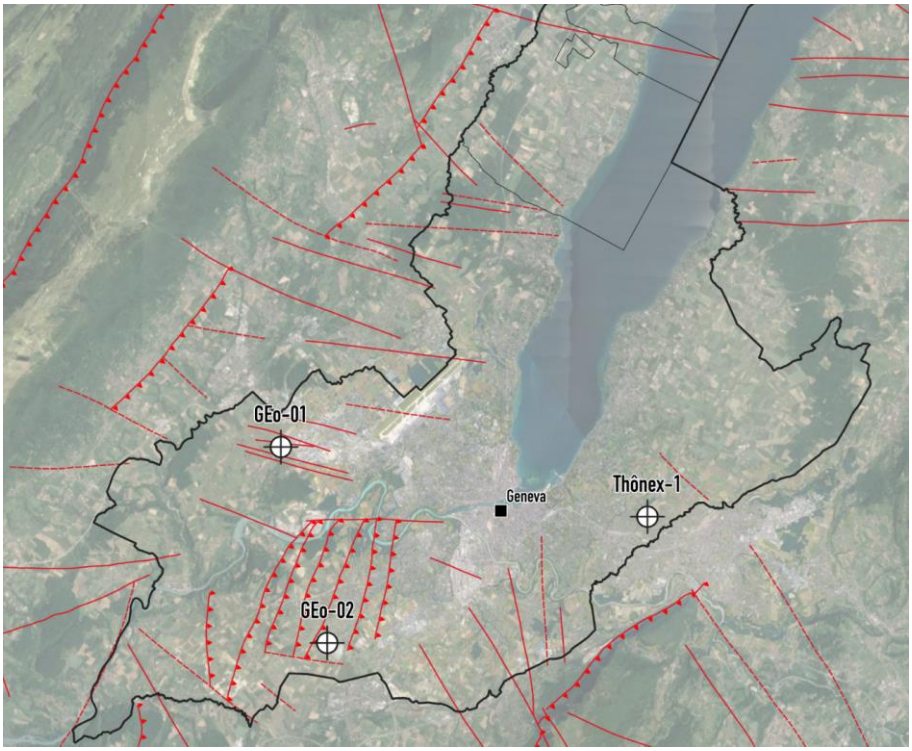


Figure 3.3.2: Location of the GEo-01 and GEo-2 wells (modified from Clerc & Moscariello, 2020).

3.3.2 Simulation models

Our aim is to simulate underground fluid flow, heat transport, geomechanics, and chemical reactions to estimate the efficiency, feasibility, and safety of using the Geneva subsurface as an HT-ATES site. For conceptual, mathematical, and computational convenience, we have divided our simulation efforts into three main groups: TH = Thermal-Hydrological, THM = Thermal-Hydrological-Mechanical, and THC = Thermal-Hydrological-Chemical. TH simulations will focus on (a) assessing thermo-hydrological challenges to heat storage in the complex subsurface of the Geneva Basin and (b) on quantifying overall thermal efficiency plausible-yet-simplified realizations of the underground heterogeneity (i.e. formation layers, faults and fractures) as well as pre-existing hydrological conditions (e.g. ground water flow). TH simulations will also essentially act as a screening process to determine scenarios to be further simulated by THM models, which in turn will focus primarily on locating and quantifying mechanical-related safety issues (e.g. ground surface deformation and subsurface stresses and strains). THC-based simulation scenario development will also be aided by insight from TH and THM models, and the results will help to anticipate issues related to mineral dissolution and precipitation reactions occurring in the entire ATES system (e.g. porosity and permeability decrease due to carbonate precipitation, carbonate scaling at the heat exchanger, etc.).

3.3.3 Model verification and benchmarking

HT-ATES model verification and benchmarking has been extensively covered in HEATSTORE D2.3 (Peter; Alt-Epping and Mindel, 2020). The Heatstore Task 2.3 benchmarking initiative essentially showed that all participating teams are able to simulate basic TH and THC processes related to thermal underground storage using their code of choice. In the majority of cases, the results were reasonably similar. The lessons learned lead us to formulate the following recommendations for the use of numerical TH modelling in the design, operation and optimization of HT-ATES systems and other geothermal applications:

- A necessary requirement is that the simulator is able to account for first order effects of temperature- and pressure dependent fluid properties in the simulated physical processes. A particular learning here is that the use of sufficiently accurate models is needed to correctly capture the important effects of temperature and pressure dependence on fluid properties. For HT-ATES and other geothermal applications, accurate treatment of this is critical to assess the economic viability and potential of projects.
- Depending on the question at hand, the availability and usage of realistic options for initial and boundary conditions is essential to arrive at a sufficiently accurate solution. Depending on the question at hand, the availability and usage of realistic options for initial and boundary conditions is essential to arrive at a sufficiently accurate solution. Proxy setups such as constant temperature instead of enthalpy-rate injection in wells, for example, may lead to inaccurate simulations. More generally, for commercial simulators, a ‘licensing factor’ comes into play: the lack of an appropriate license for a feature typically disallows its usage and the applicability to the problem of interest may be affected.
- The person responsible for carrying out the simulation should be experienced in the use of the simulator for the particular problem at hand. Some simulators are well documented for certain types of problems, but it is often the case that only developers or very experienced users may know how to apply some features to particularly uncommon problems. Even experienced users of widely used and well tested simulators may obtain inaccurate results on relatively simple problems when the simulator is applied outside their normal area of simulation expertise.

3.3.3.1 TH models

3.3.3.1.1 Simulation scenarios (carried out by Mindel et al. - ETHZ)

Within the HEATSTORE project objectives, a TH reservoir simulations were initially performed for Geneva to produce predictive general sub-scenarios of the performances of HT-ATES systems using the available data from literature and in a second phase using the site-specific geologic, hydraulic and petrophysical data from GEo-01 and GEo-02. Each sub-scenario input is summarized in Table 3.3.1. The meaning and values associated to each code is explained throughout this section (Mindel et al, 2020).

Table 3.3.1: Parameters and values used in TH simulations. A factorial design for all parameters is used resulting in 1152 3D reservoir simulations. All simulations use a temperature of 90 ℃ for HT-ATES charging and a well spacing of 141m.

Aquifer Permeability	Aquifer Thickness	Well Strategy	Groundwater	Fracture Configuration	Aquifer Dip
K13	L200	single	YGW	F0	FLAT
5K13	L300	doublet	NGW	FU	INCL
K12	L400	5spot		FD	

3.3.3.1.2 Geometrical-geological model

Based on the detailed geological study carried out by the University of Geneva (UniGe) and Services Industriels de Genève (SIG) (see Figure 3.3.1), a cube-shaped 1 $km3$ geological model was constructed by UniGe using the original analysis of a collection of subsurface datasets. The latter included 2-D seismic reflection data, petrophysical data, and well reports from the recently drilled GEO-01 well. The analysis was coupled to the interpretation of data available from the GEOMOL 3D Project (GeoMol Team 2015), where a 3D geological model of the Geneva area can be extracted. All horizons were initially considered as horizontal surfaces, simplifying structures in preparation for simulation work, however a dip angle of 15º was introduced in the simulation scenarios to include its possible effects on the thermal efficiency.
While keeping the supplied material properties in mind, we have further simplified the model geometry to the basic necessary elements for an ATES, consisting of a single permeable aquifer rock layer confined between two layers with lower permeability and porosity (see Figure 3.3.2). This characteristic and idealized configuration tends to prevent heat and fluid loss, as well as heat contamination to the surroundings (Dincer & Rosen, 2011), and its simplification was assumed sufficient for a first design iteration given the relatively large amount of simulations needed. As part of one of the possible variants, we have also introduced a model containing a 15° dip angle (see Figure 3.3.3c).

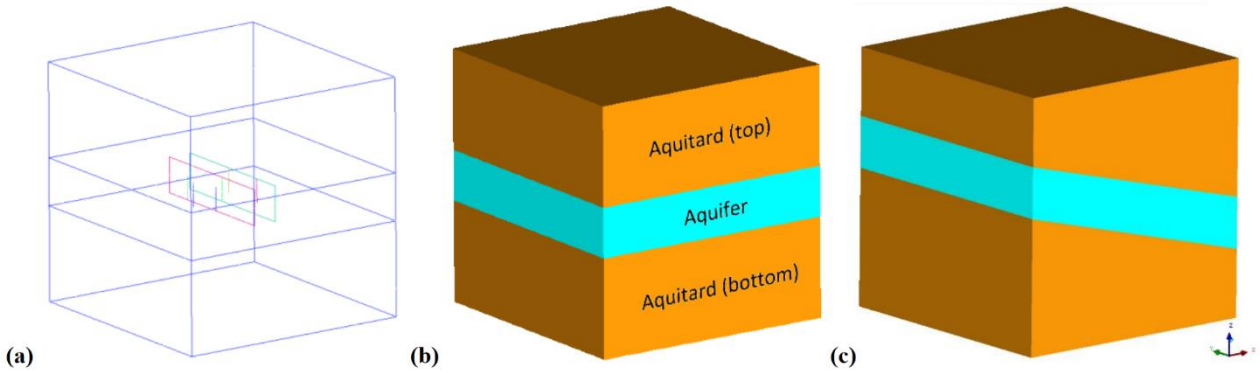


Figure 3.3.3: Geometrical/Geological model representing the basic elements of an ATES, depicting (a) possible well and fracture locations, (b) a flat version of the model, and (c) a version of the model possessing an aquifer with a 15º angle of dip.

Analysis of 2D reflection seismic data has revealed a considerable presence of fault corridors in the studied area, and since their role is still uncertain (Rusillon, 2017) they could be acting as groundwater channels as proved by GEo-01, or have become impermeable obstacles and compartmentalized target aquifers, as shown by the results of GEo-02. A summary of the values used, including permeability and thickness values related for variants K13, 5K13, K12, L200, L300, and L400 is presented in Table 3.3.2 (Mindel et al, 2020).

Table 3.3.2: Summary rock material parameters

Parameter	Units	Aquitard (top)	Aquifer	Aquitard (bottom)
Density (ρ_r)	[kg/m^3]	2450	2450	2680
Permeability (k) (original matrix)	[m^2]	10^{-17}	10^{-15}	10^{-17}
Permeability K13 (k) (fractured, effective)	[m^2]	10^{-17}	10^{-13}	10^{-17}
Permeability 5K13 (k) (fractured, effective)	[m^2]	10^{-17}	$5{\cdot}10^{-13}$	10^{-17}
Permeability K12 (k) (fractured, effective)	[m^2]	10^{-17}	10^{-12}	10^{-17}
Porosity (ϕ) (matrix, effective)	[-]	0.01	0.2	0.01
Permeability (k) (fracture, effective)	[m^2]	N/A	10^{-11}	N/A
Porosity (ϕ) (fracture, effective)	[-]	N/A	0.5	N/A
Fracture thickness	[m]	N/A	0.1	N/A
Specific Heat Capacity ($c_{p,r}$)	[J/(Kg·K)]	860.2	832.9	849.9
Thermal Conductivity λ_r (λ_r)	[W/(m·K)]	2.275	2.806	2.692
Thickness L200 (L)	[m]	400	200	400
Thickness L300 (L)	[m]	350	300	350
Thickness L400 (L)	[m]	200	400	400
Groundwater velocity (v_{gw}) (assumed)	[m/yr]	N/A	2	N/A

3.3.3.1.3 Fractures

Due to the geologic setting, the hydraulic properties of the aquifers targe ted in the study area are controlled by fracture conditions. As shown in Table 3.3.2, smaller scale fractures and cracks are taken into account via an “effective” matrix porosity and permeability in the aquifer region. In terms of scenario variants, the modelled option F0 denotes a model without any other explicit fractures. Using the ground water flow direction as a reference, a single fracture is located 50 [m] in front or Upwind of GW_1 (i.e. the main charging well) for variant FU, while an identical fracture set 50 [m] behind or Downwind of GW_1 for variant FD. These last two cases implement a single fracture as a zone of specific assumed width and properties (also shown in Table 3.3.2) and therefore while porosity is high, it is not equal to 1, which would be the case for a perfectly void fracture.

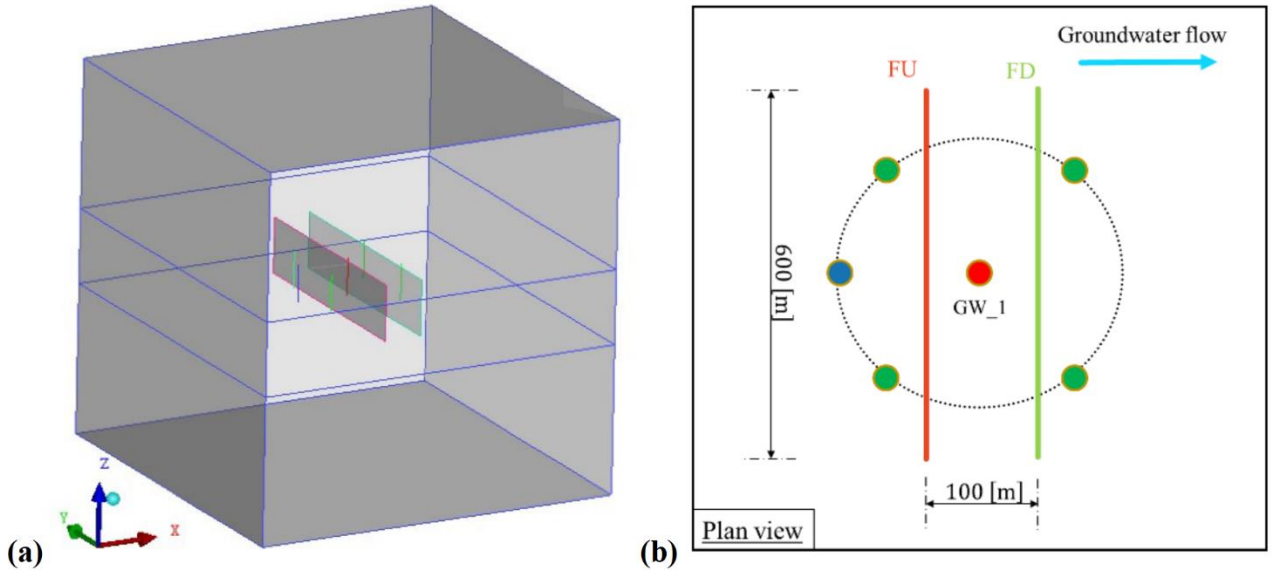


Figure 3.3.4: (a) Three-dimensional geological model depicting fracture locations used in the simulations and (b) their specific x-y plane view location with respect to the wells.

3.3.3.1.4 Operational design

The HT-ATES system will be integrated into Geneva's district heating network through the Les Cheneviers incineration plant (Driesner et al., 2019), and thus some reasonably accurate input can be obtained to run predictive models over its projected lifetime. Thermal power output from the waste incineration plant, demand, and temporal load/unload cycles over a one year time frame were provided by UniGe and SIG based on SIG's district heating systems development plan. With this information we obtained the necessary parameters for all the simulation configurations needed (Table 3.3.3).

Table 3.3.3: Summary of basic input parameters

Total Energy to Inject (E_{GWh}) (per year cycle)	50 [GWh]
Temperature of injection (T_i)	90 [℃]
Period of injection/charge (t_c)	120 [days]
Period of production/discharge (t_d)	120 [days]
Period of storage (2) ($t_{s,1}$, $t_{s,2}$)	60 [days], 65.25 [days]
Volume injected (V_i)	622080 [m³]
HT-ATES life time	15 [yr]

The basic operational yearly cycle strategy consists of a period of continuous charge for 120 [days], followed by storage for 60 [days], discharge for 120 [days], and further storage for 65.25 [days]. To prevent the thermal front from reaching the auxiliary wells (Ganguly and Kumar, 2015) and assuming the possibility of a single, a doublet, and a 5-spot well strategy, a basic table of essential simulation input parameters was developed were the single, both doublet wells and the main well in the 5-spot have rates of 0.06m³/s and the aux wells in the 5-spot have rates of 0.0015 m³/s. Recommendations for well spacing in the literature vary between a single thermal radius r_{th} and three thermal radii, and consequently an intermediate value was chosen.
A basic well design and the corresponding well names to be used in each scenario variant corresponding to well strategy, can be observed in Figure 3.3.5Figure 3.3.5: Well pattern (a) 3D design, and (b) plan view of their locations and names.. The 'single' case assumes that an auxiliary well exists to satisfy environmental re-injection requirements, but it is sufficiently far from the main well so that its effects can be neglected. For any simulation, all wells involved are fully vertical and have the same screen length, which is vertically centred on the aquifer layer. The screen length was designed as one half of the aquifer thickness.

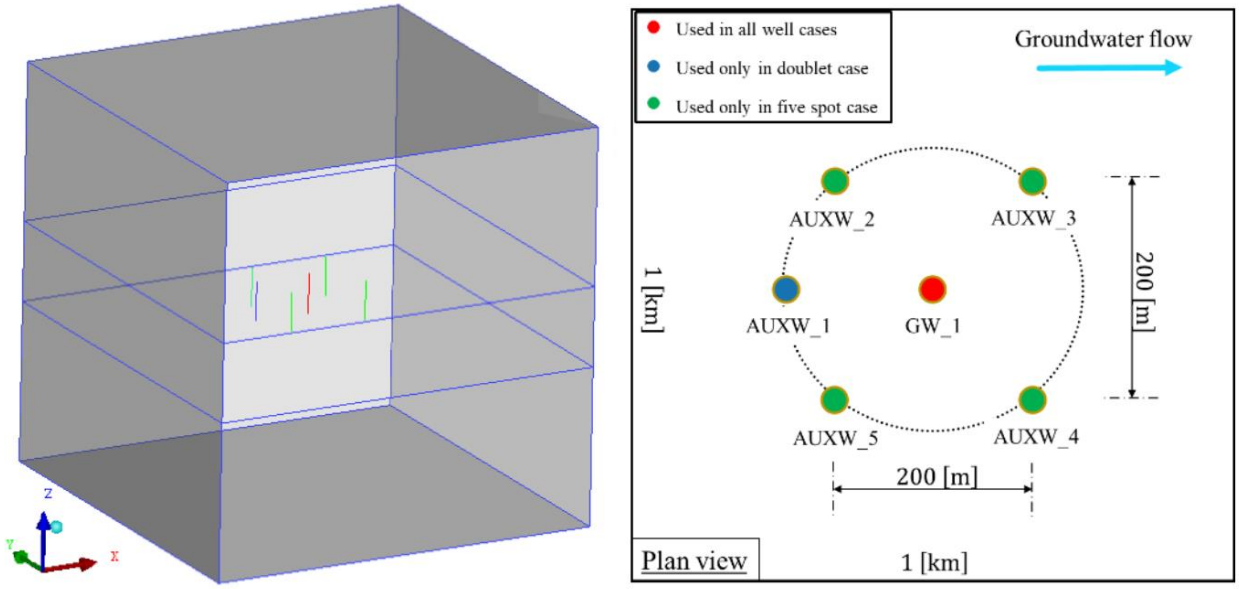


Figure 3.3.5: Well pattern (a) 3D design, and (b) plan view of their locations and names.

3.3.3.1.5 Results and discussion

We obtained results for 324 simulations produced by the combinations of simulation variant codes provided in Table 3.3.1. To assess the effects of the varying parameters on the cyclic efficiency using this relatively large volume of data, we applied an exergetic analysis for the HT-ATES expected lifetime. This type of analysis is favorable with respect to an energetic basis, since it accounts for the temperature at which water is produced from the aquifer (Dincer and Rosen, 2011).

Exergy efficiency tends to be lower than energy efficiency in ATES systems, given that as T approaches Tref the exergy contribution approaches zero. If T < Tref at any point in time, the amount of exergy for that period is negative. As a result, the amount of energy input, stored, and discharged from the ATES will only be useful, or of good enough quality, as long as the temperature of the flow can be maintained above Tref. Over subsequent cycles and theoretically depending on the discharge time and overall length of each cycle, the exergy lost to the aquifer surroundings results in an increase of the temperature near the well at the end of each cycle, thus increasing η_{ex} and η_{ex} over time as shown in Figure 3.3.6.

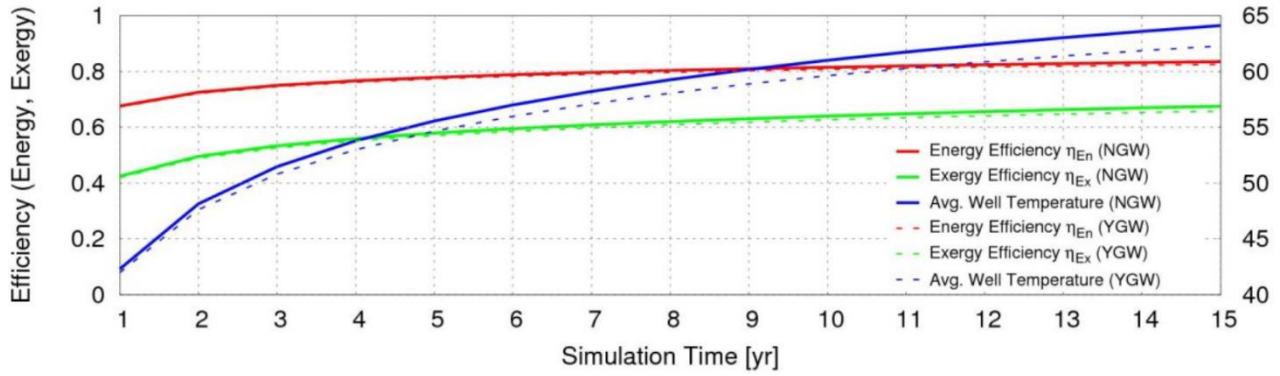


Figure 3.3.6: Energy and Exergy efficiency variation with time for the expected ATES lifetime comparing a case with and without groundwater flow: 'L400_K13_5spot_NGW_F0_FLAT', and 'L400_K13_5spot_YGW_F0_FLAT'. The right-hand axis depicts temperature measurement at well GW_1 for the end of each cycle.

3.3.3.1.6 Conclusions

We have carried out a numerical study of various HT-ATES system realizations based on parameters of aquifer permeability, aquifer thickness, well pattern, groundwater conditions, dip angle, and fracture configurations. By simplifying an originally complex geological situation, we obtained a series of scenarios aimed at a fundamental understanding of how ATES systems respond to their settings and surroundings, and how to best design them. Considering the geology, material properties, fluid properties, and industry-based operational conditions we have also investigated the plausibility storing hot water and recovering it at the maximum temperature possible via an exergetic analysis of a large number of simulations.

Our study further confirms some observations that have already been made in the literature, particularly with respect to groundwater drift and buoyancy effects present in high permeability aquifers. We have also observed that when active, auxiliary wells help mitigate pressure-peak related effects, improve the thermal front sweep, and also provide some measure of shielding against the drift due to the flow of groundwater.

In particular, we observed that although a permeability design sweet-spot could be numerically found for a particular geologic/geometric configuration, the design process is rather driven by the geo-availability in the prospective site, thus highlighting the importance of a thorough and continued (hydro-)geological study. In a similar light, lower aquifer thicknesses seem to be a favorable configuration, although to inject an equal amount of energy through a water volume rate at the same temperature, higher injection pressures are required which may impose mechanical limitations. Furthermore, particularly in terms of simulations when flow-rate values are equivalent, permeabilities do impose a numerical limitation since resulting injection and production pressure values can be unrealistically high and low, respectively. Further work should be carried out in expanding this study, particularly to better contextualize the geological configurations with particular emphasis on fractures and faults, analyze the effects of aquifer depth, and assess the influence of surface temperature conditions.

3.3.4 TH models site-specific (carried out by Daniilidis et al. - UniGe)

3.3.4.1 Thermal effects on the reservoir natural state temperature

Thermo-Hydraulic (TH) models have been performed to optimise the design of a preliminary HT-ATES system based on a under different configurations of subsurface conditions (Mindel & Drienser, 2020). The goal for the Thermal-Hydraulic (TH) modeling of Geo-01 and Geo-02 wells is to understand the performance of the considered aquifers for heat storage and to assess the extent of the thermal radius after 15 years of operation. This assessment makes us of simple layer-cake models following the insights and overall design as presented by (Mindel and Driesner, 2020).

3.3.4.2 Methodology

The Delft Advanced Research Terra Simulator (DARTS) (DARTS, 2021)(DARTS, 2021) is used to perform the simulations, using the Operator Based Linearization approach (Khait and Voskov, 2018) that has been shown to be accurate and fast against other simulators (Wang et al., 2020). Water properties are based on IAPWS97 (Huber et al., 2009) as implemented in the python package IAPWS (Romera, 2020).

3.3.4.3 Model Setup

The simulation domain is comprised of the respective reservoir, confined by two 50 m thick bounding layers at the top and the bottom. Spatial discretization is kept constant throughout, with a horizontal resolution of 10 m and a vertical resolution of 2 m. Boundary conditions are implemented with the use of large volume cells at the top and bottom layers, as well as the north and south vertical layers of the modelling domain. The well spacing is kept constant at 150 m for all simulations. The wells are rate-controlled, and the systems are operated for 15 years. The hot well uses an injection temperature of 90 ℃ while the cold well has an injection temperature of 50 ℃. Common input parameters are summarized in Table 3.3.4. The scenarios considered in the models are listed in Table 3.3.5.

Table 3.3.4: Input parameters shared across all models.

Parameter	Value
Well spacing	150 m
Charge – Store – Discharge - Rest	120 – 60 – 120 – 65.25 days
Simulation time	15 years
Porosity	10 %
Pressure gradient	10 MPa / km
Temperature gradient	30 ℃ / km
Reservoir permeability kv / kh	0.1
Hot / Cold well injection temp	90 / 50 ℃
Confining layers permeability / porosity	0.5 mD / 0.1 %

Table 3.3.5: Modelling scenarios.

Well and Scenario	Targeted Reservoir	Reservoir Permeability (m^2)	Reservoir Thickness (m)	Injection Rate (kg/s)
GEO-01	LC + Siderolitic	$3{\cdot}10^{-13}$	350	60
GEO-02 SC1	LC-UJ	$7{\cdot}10^{-16}$	700	3.9
GEO-02 SC2	Siderolitic	$7{\cdot}10^{-16}$	150	0.7
GEO-02 SC3	Siderolitic	$3{\cdot}10^{-13}$	150	60

3.3.4.4 Geo-01

Figure 3.3.7 and Figure 3.3.8 show the temperature distribution in the production interval after 15 years of production, using the inputs from Table 3.3.6. The large reservoir thickness combined with a relatively high permeability results in the hot plume being more prominent in the upper part of the aquifer. Additionally, due to the interaction between the hot and the cold well, the hot plume is asymmetrical, having a reduced extent between the wells and a larger extent away from the wells. The thermal radius is therefore slightly larger at

shallower depths and exceeds the 3˚C at 100m radius distance from the well as defined by the Swiss water protection law (OEaux, annexe 2 chapitre 21 alinéa 3) at the end of the 15 years operation period simulated.

Table 3.3.6: Input parameters Geo-01 model.

Parameter	Value
Domain x, y, z	500 m, 500 m, 454 m
Discretization dx, dy, dz	10 m, 10 m, 2 m
Cell count	567,500
Well rates	60 l/s - 5184 m^3/day
Permeability	$3 \cdot 10^{-13}$ m^2 - 304 mD

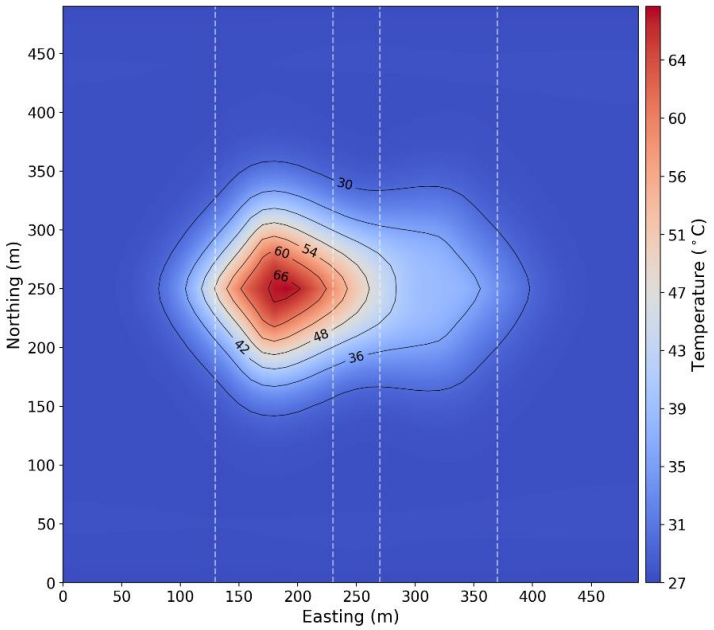


Figure 3.3.7: Geo-01 map view at a depth of 575 m after 15 years of operation. The vertical dashed white lines represent a distance of 100m on each side of the hot and cold well respectively.

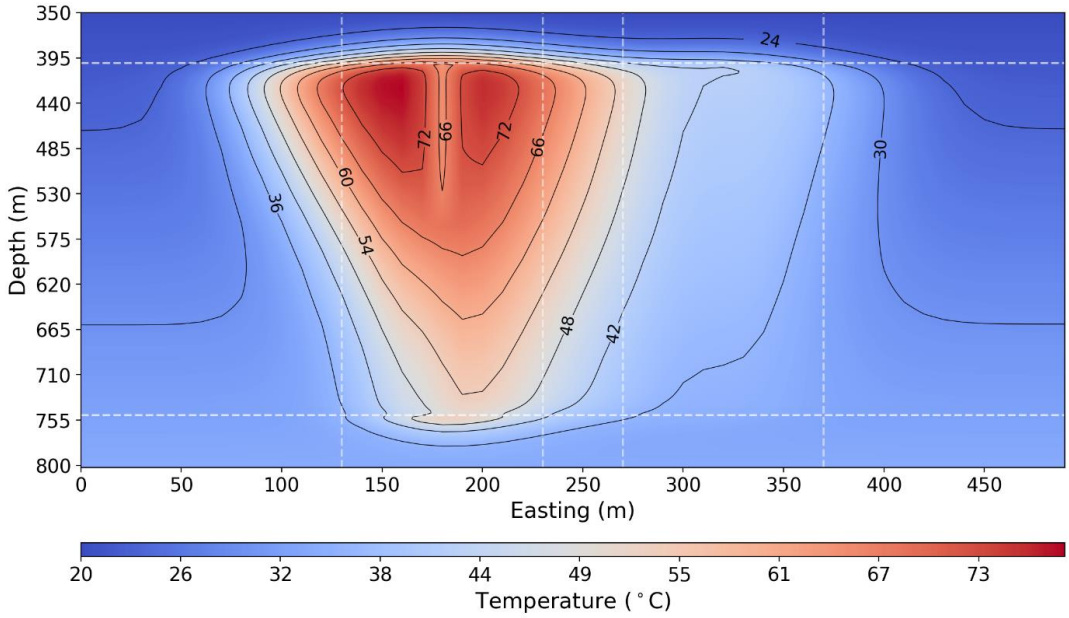


Figure 3.3.8: Geo-01 vertical section across the wells (northing = 500 m) after 15 years of operation. The horizontal dashed white lines mark the production interval, while the vertical dashed white lines represent a distance of 100m on each side of the hot and cold well respectively.

3.3.4.5 Geo-02 SC1

Figure 3.3.9 and Figure 3.3.10 show the temperature distribution in the production interval after 15 years of production, using the inputs from Table 3.3.7. The low permeability and low rates result in a very small thermal radius. Differences between the bottom and the top of the hot plume are attributed mostly to conduction, which is laterally favorable at shallower depths due to the increased temperature gradient with the undisturbed temperature field. Reservoir temperature does not exceed the 3˚C at 100m radius distance from the well as defined by the Swiss water protection law (OEaux, annexe 2 chapitre 21 alinéa 3) at the end of the 15 years operation period simulated.

Table 3.3.7: Input parameters Geo-02 SC1 model.

Parameter	Value
Domain x, y, z	500 m, 500 m, 704 m
Discretization dx, dy, dz	10 m, 10 m, 2 m
Cell count	1,005,000
Well rates	0.7 l/s – 60.5 m^3/day
Permeability	$7 \cdot 10^{-16}$ m^2 – 0.7 mD

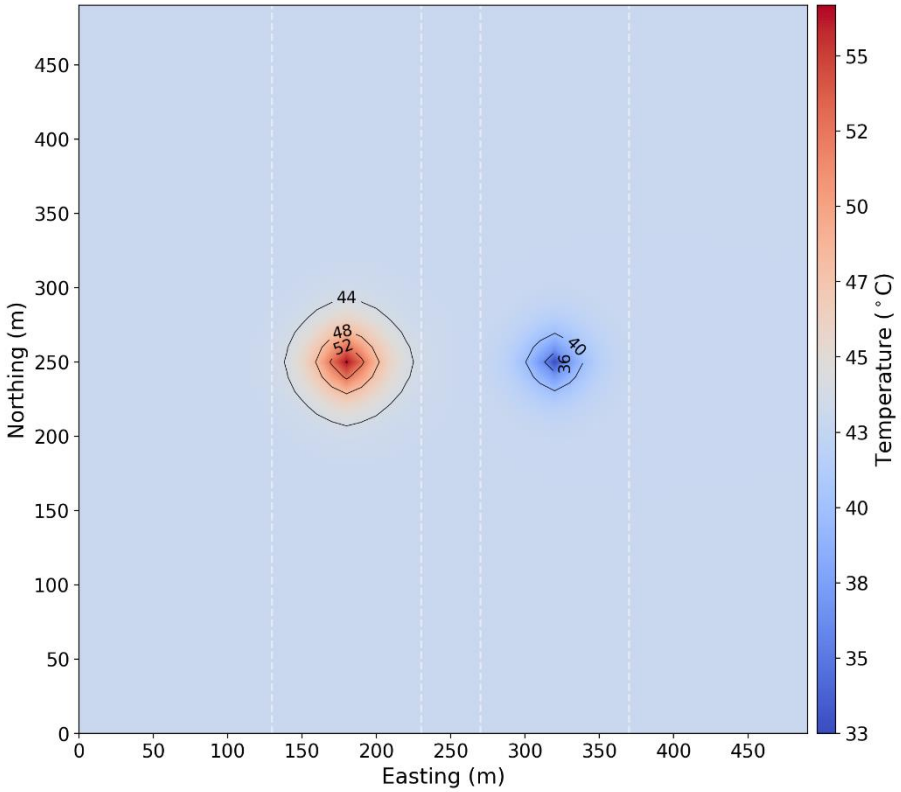


Figure 3.3.9: Geo-02 SC1 map view at a depth of 1100 m after 15 years of operation. The vertical dashed white lines represent a distance of 100m on each side of the hot and cold well respectively.

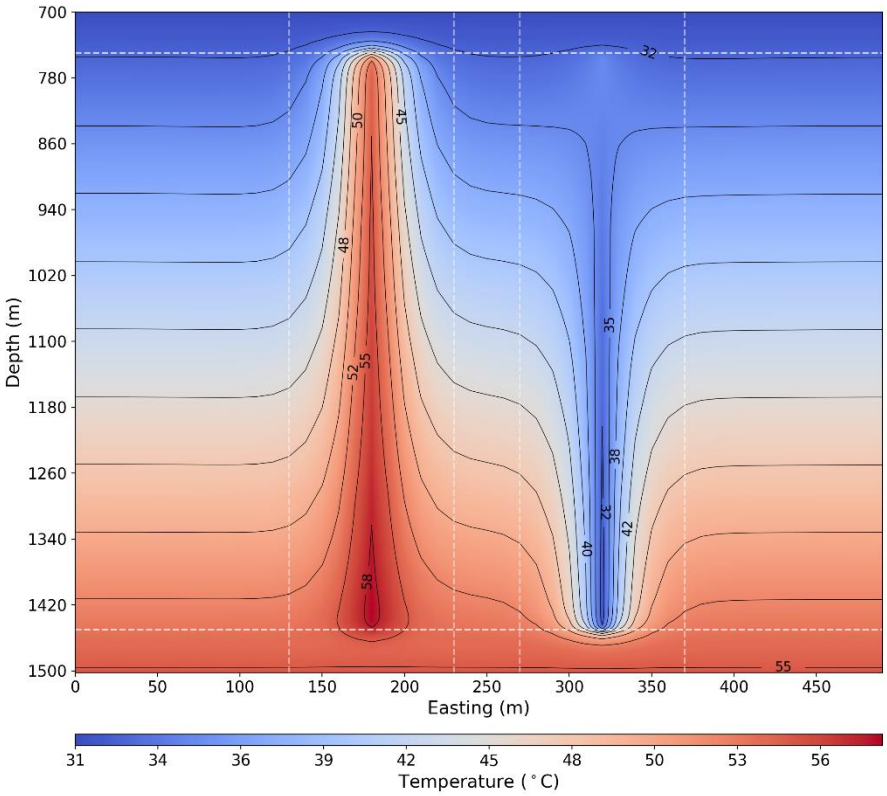


Figure 3.3.10: Geo-02 SC1 vertical section across the wells (northing = 500 m) after 15 years of operation. The horizontal dashed white lines mark the production interval, while the vertical dashed white lines represent a distance of 100m on each side of the hot and cold well respectively.

3.3.4.6 Geo-02 SC2

Figure 3.3.11 and Figure 3.3.12 show the temperature distribution in the production interval after 15 years of production, using the inputs from Table 3.3.8. Compared to SC1 using the same rates with a significantly reduced reservoir thickness results in a slightly larger extent of the hot plume. Additionally, the smaller temperature difference at the top and bottom of the domain (due to the lower thickness) results in a more homogeneous lateral extent of the cold plume. The shallower parts of the domain remain slightly less extensive compared to the deeper ones but differences are minor. Reservoir temperature does not exceed the 3˚C at 100m radius distance from the well as defined by the Swiss water protection law (OEaux, annexe 2 chapitre 21 alinéa 3) at the end of the 15 years operation period simulated.

Table 3.3.8: Input parameters Geo-02 SC2 model.

Parameter	Value
Domain x, y, z	500 m, 500 m, 154 m
Discretization dx, dy, dz	10 m, 10 m, 2 m
Cell count	317,500
Well rates	0.7 l/s – 60.5 m^3/day
Permeability	$7 \cdot 10^{-16}$ m^2 – 0.7 mD

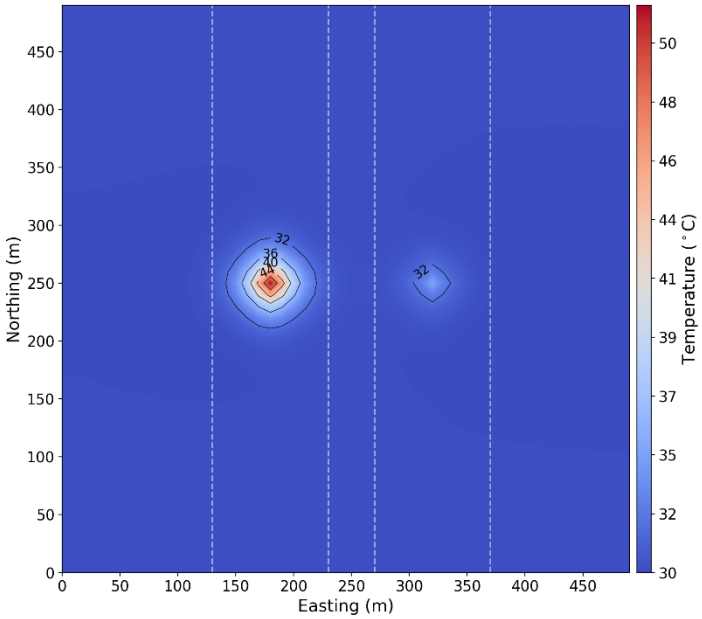


Figure 3.3.11: Geo-02 SC2 map view at a depth of 675 m after 15 years of operation. The vertical dashed white lines represent a distance of 100m on each side of the hot and cold well respectively.

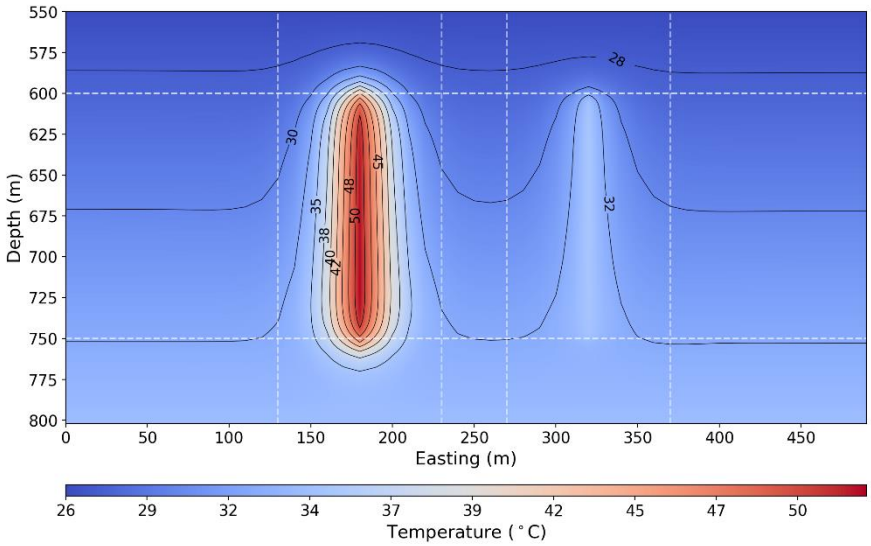


Figure 3.3.12: Geo-02 SC2 vertical section across the wells (northing = 500 m) after 15 years of operation. The horizontal dashed white lines mark the production interval, while the vertical dashed white lines represent a distance of 100m on each side of the hot and cold well respectively.

3.3.4.7 Geo-02 SC3

Figure 3.3.13 and Figure 3.3.14 show the temperature distribution in the production interval after 15 years of production, using the inputs from Table 3.3.9. Increasing both the rate and permeability with the same reservoir thickness as SC2 results in a noticeably larger hot plume laterally and a larger thermal radius. Moreover, the vertical shape of the hot plum is now more pronounced in the shallower part. This is attributed to the large contribution of convection in the temperature field and resembles qualitatively the Geo-01 model. Similarly to the Geo01 model, the hot plume is asymmetric and extends farther away from the wells compared to the space between the two wells. Reservoir temperature exceeds the 3˚C at 100m radius distance from the well as defined by the Swiss water protection law (OEaux, annexe 2 chapitre 21 alinéa 3) at the end of the 15 years operation period simulated.

Table 3.3.9: Input parameters Geo-02 SC3 model.

Parameter	Value
Domain x, y, z	1000 m, 1000 m, 154 m
Discretization dx, dy, dz	10 m, 10 m, 2 m
Cell count	317,500
Well rates	60 l/s – 5184 m^3/day
Permeability	$3 \cdot 10^{-16}$ m^2 – 304 mD

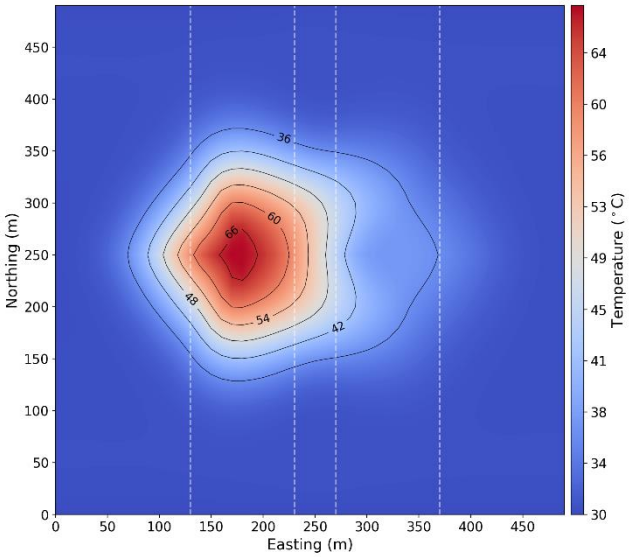


Figure 3.3.13: Geo-02 SC3 map view at a depth of 675 m after 15 years of operation. The vertical dashed white lines represent a distance of 100m on each side of the hot and cold well respectively.

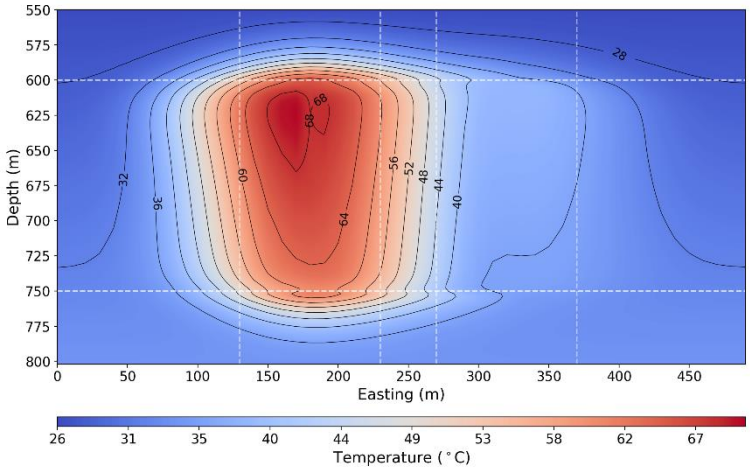


Figure 3.3.14: Geo-02 SC3 vertical section across the wells (northing = 500 m) after 15 years of operation. The horizontal dashed white lines mark the production interval, while the vertical dashed white lines represent a distance of 100m on each side of the hot and cold well respectively.

3.3.5 THM models (carried out by Birdsell et al. - ETHZ)

THM and HM modeling were carried out by ETHZ with the goal to predict the potential ground deformation effects associated with repeated seasonal cycles on injection and extraction. Two are mechanisms that cause this expansion: (1) thermal expansion, and (2) gradients in pore pressure, which act like a body force. These deformations can affect fluid flow by altering the porosity and permeability of the porous media as it deforms. Furthermore, the motion of the solid matrix means that fluid velocity must be considered as a velocity relative to the solid rock. The deformations can also alter the heat transfer by “advecting” the heat with solid grain motion. Finally, the thermal and hydrological systems are also coupled through the equation of state and the

fluid motion, which advects heat. The most notable aspect of the mechanical expansion in the context of HT-ATES may be the potential for the ground surface to deform. While ground surface deformation has been studied in other contexts, it has received very little attention in the context of HT-ATES, and it could lead to regulatory and/or geotechnical challenges.
As in the TH model, we assume viscous dissipation has a negligible contribution to the heat equation since large temperature differences are injected and extracted, an assumption which has precedent (Bear and Corapcioglu, 1981). Single-phase, single-component fluid flow is assumed, which neglects the possibility of dissolved salts to alter the fluid density. This is an acceptable assumption since the THM results are focused primarily on mechanical effects, and fluid flow is of secondary interest. Plastic deformations are not accounted for because we assume that plastic (i.e. large) deformations would occur only under unacceptable UTES operating conditions. If plastic deformations are shown to be important, they can be incorporated into simulations later in the project.

3.3.6 THM models generic

The THM modelling group focused primarily on the Geneva project provided early results on two aspects of the project: (a) the uplift due to the first loading/injection stage of heat storage, and (b) the subsidence due to the planned pumping test of GEo-01. The subsidence simulations were motivated by a desire to inform the monitoring activities during the pumping test. Two-dimensional THM simulations have been performed for relatively simple geological conceptual models. Three-dimensional hydromechanical (HM) simulations have also been conducted, using more information about material properties and stratigraphy provided by UniGe. The base case scenario shows that substantial surface deformation occurs. For the injection scenario, a simple sensitivity analysis shows that uplift is diminished marginally for stiffer rock and/or if a deeper target formation is selected. Uplift is substantially diminished if auxiliary well(s) are included to balance the reservoir pressure. Figure 3.3.15 shows example results for a scenario with and without an auxiliary well. Based on these preliminary results, the placement and operation of auxiliary well(s) is very important and needs to be carefully considered as the project progresses. When selecting the reservoir(s) for heat storage, rock properties (e.g., transmissivity and elastic parameters), should also be considered to ensure that the reservoir(s) can store the desired amount of heat and fluid without resulting in excessive pore pressure or surface deformation.

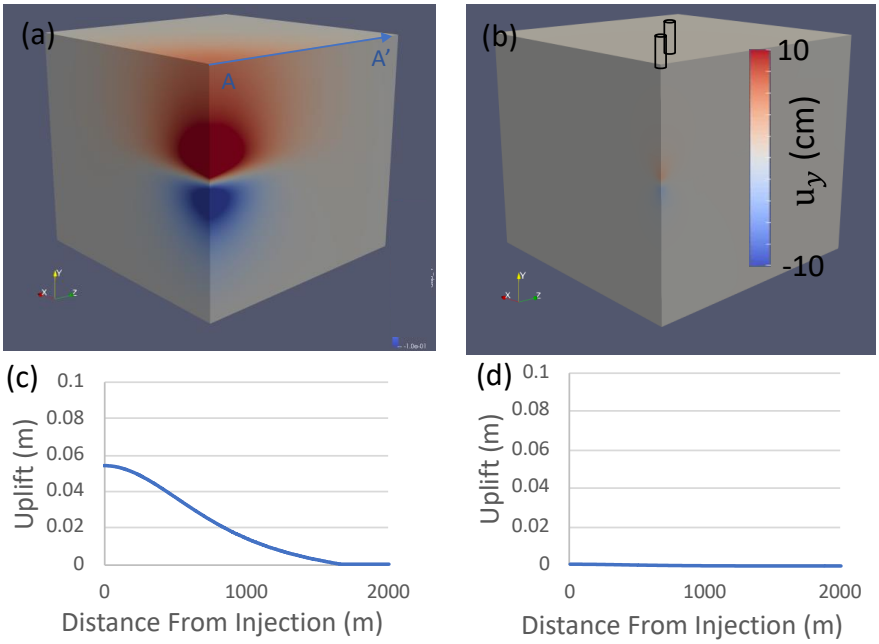


Figure 3.3.15: THM model results. Figures (a) and (b) show the vertical displacement for the base case scenario and a scenario where auxiliary wells balance the pressure. The approximate surface location of the main well (red) and auxiliary well (green) are shown in (b). The simulations take advantage of symmetry. Figures (c) and (d) represent the surface uplift along line AA' shown in (a).

3.3.7 THM models site-specific (ETH)

3.3.7.1 Soil mechanics effects

We explore two aspects of ground surface deformation at the Geneva wells using a hydro-mechanical (HM) model. Firstly, we model ground deformation during the pumping test at GEO-01 and compare to deformation measurements from Work Package 5 (see D5.2). Secondly, we perform predictive simulations of the potential ground deformation resulting from HT-ATES to explore the question: what HM ground deformation could we expect if GEO-01 or GEO-02 were used as one well in an HT-ATES doublet? This report is a summary of work performed to understand ground deformation, but more detailed methodology and results are available in the Month 35 version of Deliverable 2.1 (Driesner et al., 2019).

3.3.7.2 Ground deformation results

The following two sub-sections focus on results for the GEO-01 pumping test and the predictive HT-ATES modelling.

3.3.7.3 GEO-01 pumping test

Figure 3.3.16 shows ground surface deformation during the GEO-01 pumping test. Deformation data comes from two sources: (a) GPS monitoring near GEO-01 and (b) the HM numerical model. There is not a clear trend of subsidence (or uplift) in the GPS data. The GPS data was provided by Nicolas Houlié Geologie GmbH and Services Industriels de Genève (SIG) and was collected as part of Work Package 5. The ground deformation was both positive (upwards) and negative (downwards), depending on the time and the GPS station. For the most part, the magnitude of deformation was less than the size of the error bars, so we cannot interpret any significant deformation from the GPS data. In contrast, the numerical model shows a clear trend of subsidence that increases with time. We perform a sensitivity analysis on the Young's modulus, and find that smaller Young's modulus corresponds to a larger magnitude of subsidence. This makes intuitive sense because Young's modulus is a measure of the strength of the rock, and a weaker rock will deform more under the same pressure change. Subsidence would be clearly seen in the GPS data if the field-scale Young's modulus were below 0.35 GPa, and therefore we infer 0.35 GPa as the lower bound of Young's modulus for the HT-ATES predictive modeling.

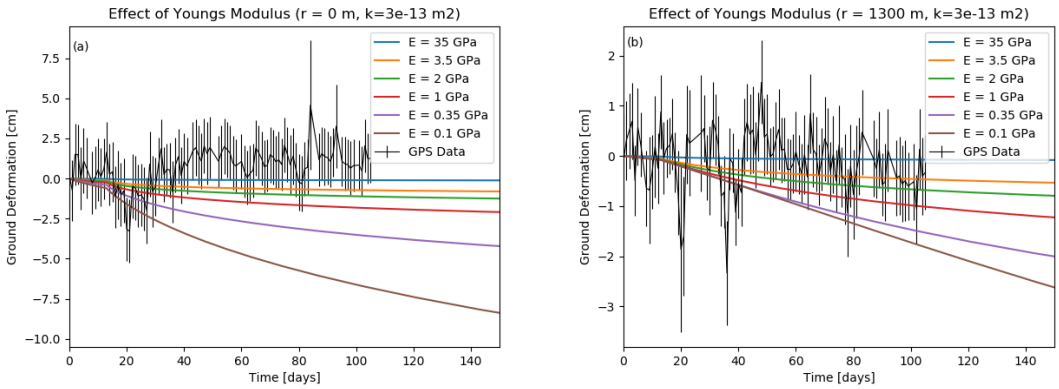


Figure 3.3.16: Ground deformation versus time (a) at the well, representing the first GPS station, and (b) 1300 m from the well, representing the second GPS station. Black lines represent GPS data with error bars, and colorful lines represent HM model results for a sensitivity analysis on Young's modulus.

3.3.7.4 HT-ATES predictive modelling

Table 3.3.10 summarizes the predictive modeling results for a number of scenarios and parameter values. The GEO-01 scenario targets the LC-UJ and uses the permeability inferred form the GEO-01 pumping test (i.e., $3*10^{-13}$ m^2). GEO-02 Scenario 1 targets the LC-UJ, which is deeper at GEO-02, and uses a smaller permeability (i.e., $7*10^{-16}$ m^2), which is in-line with values observed from a pumping test at GEO-02 and observed at the Thônex well. GEO-02 Scenarios 2 and 3 target the shallower Siderolitic rock, and use the permeabilities matching GEO-02 Scenario 1 and the GEO-01 scenario, respectively. When the lower value of permeability is used, the flow rate is curtailed below 60 kg/s, due to the HF constraint.

Table 3.3.10: Predictive Simulation Scenarios and Results

Well and Scenario	Targeted Reservoir and Depth [m]	Reservoir Permeability [m^2]	Reservoir Thickness [m]	Flow Rate [kg/s]	Young modulus [GPa]	Years simulated	Maximum ground deformation [cm]
GEO-01	LC-UJ (400 – 750)	$3*10^{-13}$	350	60	35	15	<0.01
					2	1	0.10
					0.35	15	0.49
GEO-02 Scen. 1	LC-UJ (750-1450)	$7*10^{-16}$	700	3.9	35	15	0.015
					2	1	0.053
					0.35	1	0.055
GEO-02 Scen. 2	Siderolitic (600-750)	$7*10^{-16}$	150	0.7	35	1	<0.01
					2	1	0.015
					0.35	1	0.016
GEO-02 Scen. 3	Siderolitic (600-750)	$3*10^{-13}$	150	60	35	1	<0.01
					2	1	0.097
					0.35	1	0.40

Figure 3.3.17 shows the aquifer pore pressure and ground surface deformation at the end of the injection stage. We find that pore pressure does not change dramatically from year to year, and the magnitude of the ground surface deformation tends to be largest in the first year. Therefore, we only present the first year of each scenario in the figure. Pore pressure is elevated near the injection well and depleted near the production well. The largest change in aquifer pressure is reached in GEO-02 Scenario 1, followed by GEO-02 Scenario 2. This makes sense because these are the scenarios where the flow rate is limited by the hydraulic fracturing constraint. Surface deformation is positive (upward) for the right portion of the plot (i.e., x>2000 m), whereas it is negative (downward) for the left portion of the plot (i.e., x<2000 m). The ground surface deformation is sensitive to the Young's modulus. The largest modelled ground surface deformations are 0.49 and 0.40 cm, which occur when Young's modulus is 0.35 GPa for GEO-01 and GEO-02 Scenario 2, respectively. For these scenarios, there may be some boundary effects, and it is possible that the predicted uplift would be smaller if a larger mesh were used. All other scenarios have deformation ≤0.1 cm.

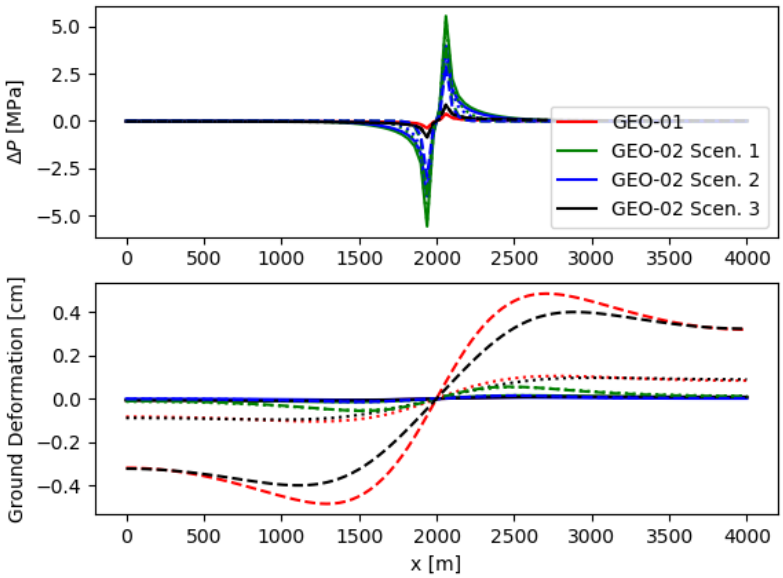


Figure 3.3.17: (Top) Difference in aquifer pressure from the initial condition versus spatial coordinate *x* along a line that intersects the two wells. (Bottom) Ground surface deformation versus *x* at the end of Year 1 injection for each scenario. Solid lines use Young's modulus equals 35 GPa, dotted lines use 2 GPa, and dashed lines use 0.35 GPa. The cold/production well and hot/injection well are located at x = 1925 m, and x = 2075 m, respectively.

3.3.8 Conclusions

In this section, we provided results about the modelling results evolution with respect to Th and THm/HM prediction before and after the acquisition on site-specific data from GEo-01 and GEo-02 wells.
TH models carried out in a first phase by ETHZ (Mindel et al. 2020) provided a general overview of the main mechanisms controlling fluid flow during HT-ATES operations and affecting the overall performances of the modelled scenarios. Once new data from GEo-01 and GEo-2 data become available the static and dynamic models were refined by UniGe allowing a more accurate prediction of the performances of a set of potential HT-ATES configurations at the two study sites.
With respect to soil mechanics and effects on ground deformation, the initial study carried out by ETHZ revealed that in certain general conditions significant ground deformation of tens of centimetres can be predicted. However, the site-specific hydraulic results produced after production tests provided accurate hydraulic parameters for the two wells resulting in poro-elastic deformation in the order of a few millimetres.

3.4 Forsthaus, Bern, Switzerland

3.4.1 Site description

3.4.1.1 Conceptualization

3.4.1.1.1 Location & UTES concept and specifications, scope and aims of the studyTES concept and specifications, scope and aims of the study

The Forsthaus Heat Storage project is run by Geo-Energie Suisse AG (GES) on behalf of Energie Wasser Bern (EWB). It is supported by the Swiss Federal Office of Energy and is part of the Swiss contribution to the European Geothermica-HEATSTORE project.

The Forsthaus project is located in the northern part of the city of Bern (Switzerland) next to EWB's power production site "Energiezentrale Forsthaus" (Figure 3.4.1).

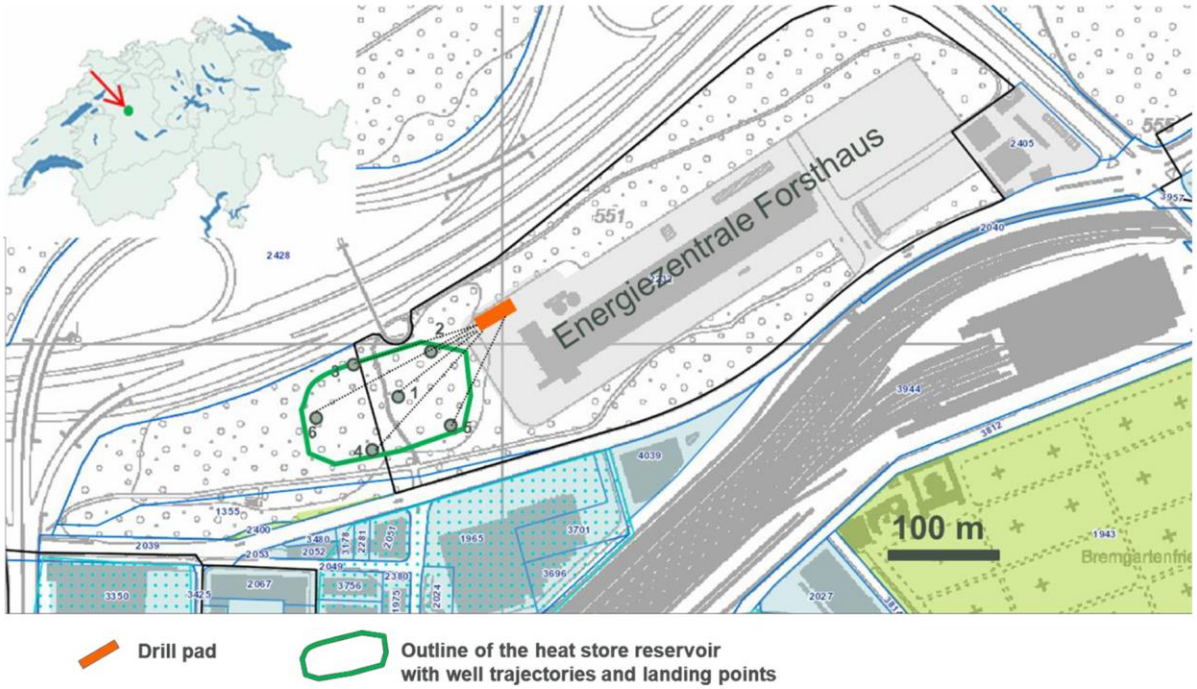


Figure 3.4.1: Location of the Aquifer Thermal Energy Storage project "Bern Forsthaus".

The purpose of the Forsthaus project is to create an Aquifer Thermal Energy Storage (ATES) where waste heat from the "Energiezentrale Forsthaus" will be stored during the summer instead of being dissipated into the atmosphere. That heat will be back-produced during the wintertime to feed into a district heating network.

The project design anticipates a main well at the centre of the system and peripheral auxiliary wells. The main well is used to inject and produce the energy in the form of hot water. The auxiliary wells are used to regulate the flow at the boundary, maintain the desired aquifer reservoir pressure and connect to the surface system, so that the underground geological formation, the wells and the surface facilities are acting as a closed loop system (Figure 3.4.2 and Figure 3.4.6).

3.4.1.1.2 Preliminary operation mode

The HT-ATES Bern-Forsthaus is seasonally operated with loading cycles during summer time and unloading cycles during winter time (Figure 3.4.2) according to a preliminary schedule summarized in Table 3.4.1.

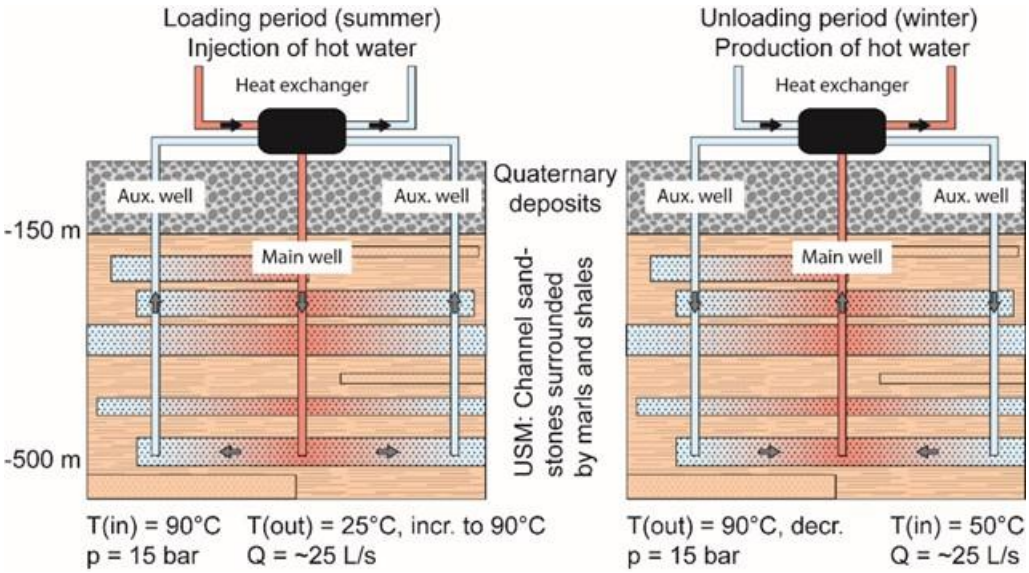


Figure 3.4.2: Preliminary operation parameters during loading cycles (left) and unloading cycles (right).

Table 3.4.1: Preliminary operation parameters for loading- & unloading-cycles.

	Loading cycle	Unloading cycle
Temperature:	90°C	Starting from 90°C down to 50°C
Duration:	216 days	149 days
Circulation rate:	25 L/s	
Heat losses:	ca. 40% (based on coupled thermos-hydraulic modelling	
Running time UTS:	20 years	

Under these preliminary assumptions the energy balance for the UTES Bern-Forsthaus was calculated:

- Total amount of lost heat stored in the reservoir: ca. 21.3 GWh/a
- Total amount of heat gained from the reservoir: ca. 12.8 GWh/a
- Reduction of CO_2–output: 2,531 tons/a

3.4.1.2 Geology and reservoir model

3.4.1.2.1 Geology

The reservoir of the UTES project Bern-Forsthaus is located within the Lower Freshwater Molasse (USM) and belongs to the Swiss Molasse Basin (Figure 3.4.3). The Swiss Molasse is a thick Tertiary sedimentary body created by the detrital filling of a subsidence basin that was caused by the uplift of the Alps. At the project site the USM is covered by quaternary unconsolidated deposits (gravels, sands, clays) of about 150 m thickness. These unconsolidated deposits comprise a shallow freshwater aquifer from 8 to 10 m below surface.

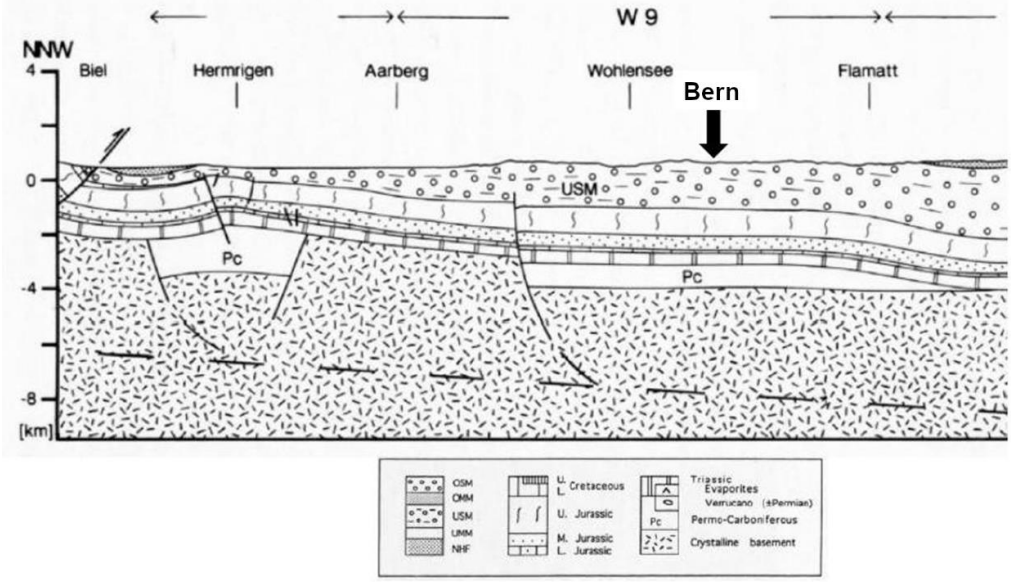


Figure 3.4.3: Regional NNW–SSE geological cross section across the project site in Bern. The acronym of the Lower Freshwater Molasse is "USM" (from Pfiffner et al., 1997).

3.4.1.2.2 Conceptual reservoir model

In the past the USM was the subject of detailed sedimentological and hydrogeological studies: NAGRA NTB 90-41, 1990; Platt et al., 1992; Keller, 1992; NAGRA NTB 92-03, 1993; Küpfer, 2005; Hölker, 2006. These defined specific architectural/facies elements illustrated in Figure 3.4.4. Their geometrical and hydrogeological properties are summarized in Figure 3.4.5.

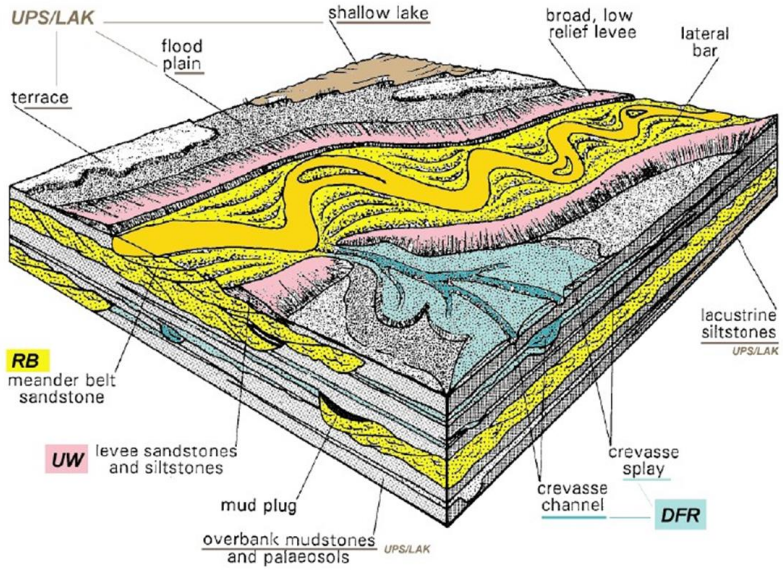
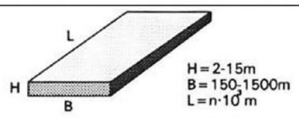
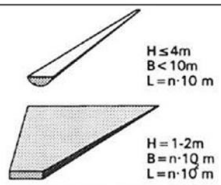
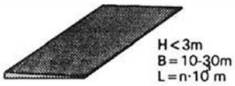


Figure 3.4.4: Summary facies model for the deposition of the USM (from Hölker, 2006).

Architectural element	Simplified geometry	Hydraulic conductivity Porosity (median)
RB (meander belt **sandstones**)		**1E-04 to 1E-07 m/s 19.6%**
DFR (crevasse splay and channel **sandstones and siltstones**)		**1E-06 to 1E-08 m/s 14.2%**
UW (levee **sandstones and siltstones**)		**1E-07 to 1E-10 m/s 8.2%**
UPS (overband **mudstones and palaeosols**)		**1E-08 to 1E-11 8.1%**
LAK (lacustrine **siltstones**)		**No data, but very low values**

(Modified from Keller, 1992)

Figure 3.4.5: Architectural elements of the USM.

Based on the current state of knowledge, a conceptual reservoir model for the UTES Bern-Forsthaus was established consisting of mainly two elements: Porous and permeable sandstone-layers (RB-elements) embedded within a low-porosity and low-permeability matrix (UW-, UPS & LAK-elements). The permeable sandstone-layers are used for fluid and heat transport whereas heat will be stored within the sandstone layers and the surrounding matrix composed of marl- and mudstone (Figure 3.4.6). Physical properties and the mineralogical composition of the rock units in the stratigraphic succession at Forsthaus are summarized in Table 3.4.2.

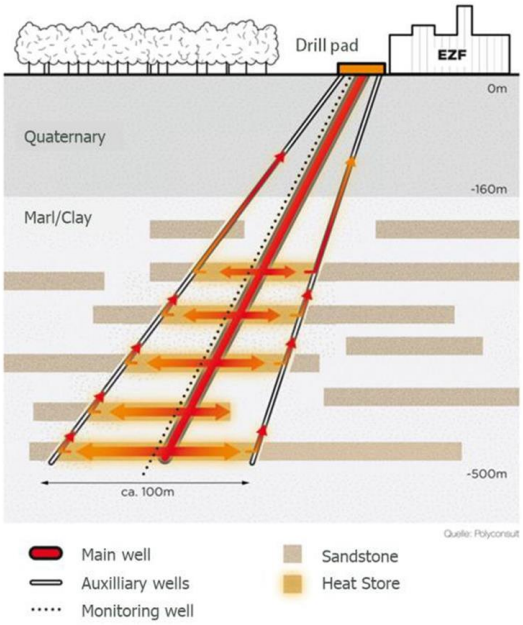


Figure 3.4.6: Expected conceptual reservoir model for UTES Bern-Forsthaus showing the RB-sandstones embedded within the matrix composed of marl- and claystone (UW-, UPS & LAK-elements).

Table 3.4.2: Specific material properties for "sandstones" and "matrix" of the USM and the overlaying unconsolidated sediments from the literature

	Heat capacity	Thermal conductivity	Density	E-Moduli	UCS	Mineralogy
	J/kg/K	W/m/K	kg/m³	GPa	MPa	wt %
Sandstones	1037	2.67	2300	1 ± 0.3	6.5 ± 4	Quartz 35% Albite 18% Felspar 11% Calcite 12% Dolomite 2% Clay minerals, Mica 22% C_{org} 0.1%
Matrix	1037	2.67	2500	0.5 ± 0.5	23 ± 6 (mudstone) 8 ± 5 (marlstone)	Quartz 18% Albite 10% Felspar 7% Calcite 15% Dolomite 5% Clay minerals, Mica 45% C_{org} 0.1%
Quaternary unconsolidated deposits	840 - 1000	0.4 - 2.3	1800 - 2500	0.02 - 0.05	-	-

3.4.1.2.3 Local geothermal and groundwater conditions

The heat store reservoir will be placed between 200 to 500 m below surface (Figure 3.4.6). Applying the mean geothermal gradient of 3°C/100 m, the natural reservoir temperature will be in the range of 17°C to 26°C. Horizontal groundwater flow is restricted to the discontinuous permeable sandstone layers. Vertical groundwater flow is highly inhibited due to almost impermeable matrix. The USM is therefore regarded as an aquitard. From a near-by offset well (about 30 km to the east) it is known that in-situ pore pressures are different in different sandstone layers. All of them were found to be confined, but not artesian. Formation water found in the USM can be classified as Na-HCO_3 to Na-Cl-type. NAGRA (NTB 88-25) reported the hydrochemical composition for formation water from the USM as "Referenzgrundwässer USM". So far, no results from hydrochemical water-sample analysis are available for the project site. Water samples will be taken during the course of drilling and testing.

3.4.1.2.4 Planned Well Design and Testing

3.4.1.2.4.1 Well design

The main well and auxiliary wells are all foreseen with the same standard design, and are therefore interchangeable. The monitoring well architecture has not yet been defined, but will be lighter than that of the main/auxiliary well design. The construction of the wellbore will follow 3 phases of drilling and casing as described in Table 3.4.3 and Figure 3.4.7

Table 3.4.2: Well design.

	Phase 1	Phase 2	Phase 3
Hole size	26 in	20 in	14 ¾ in
Measured Depth	22 m	200 m	560 m
Vertical depth	20 m	180 m	500 m
Casing size	22 in	16 in	10 ¾ in
Casing weight	114.8 lbs/ft.	75 lbs/ft.	40.5 lbs/ft.
Connections	STC	LTC	Premium/LTC
Casing material	J55	J55	N80
Special equipment	Cemented with a stab-in shoe Cement basket for a potential top job	Cemented with a stab-in shoe Cement basket for a potential top job	Top 200 m equipped with thermally insulated casing and premium coupling. Bottom part with normal LTC using A fibre optic runs along the casing. The fibre is prepared for oriented perforating away from the fibre.
Comments	Section is slightly slanted to reach target	Main deviation section to reach the top of the reservoir at the set slant angle	Section will be entirely cored and successive testing will take place after each cored section.

3.4.1.2.4.2 Coring and selective testing of the reservoir section (Phase 3)

Phase 3 focuses on the reservoir section. This section will be entirely cored and slanted towards the target zone and total depth within the USM at 500 m vertically. The coring and testing operations will be performed sequentially. Where the cored section exhibits good reservoir properties, the specially designed testing equipment (wireline packer system) will be run across that section through the coring bit, and a selective hydraulic test will be performed in order to characterize the test interval with respect to in-situ formation pressure, transmissivity, hydraulic boundaries and as soon as more than one well is available wellbore interconnectivity (hydraulic tomography). The procedure will allow establishing a geological and transmissivity profile for the entire reservoir section. This profile will be later used to select the zones to be perforated and used for the heat storage volume.

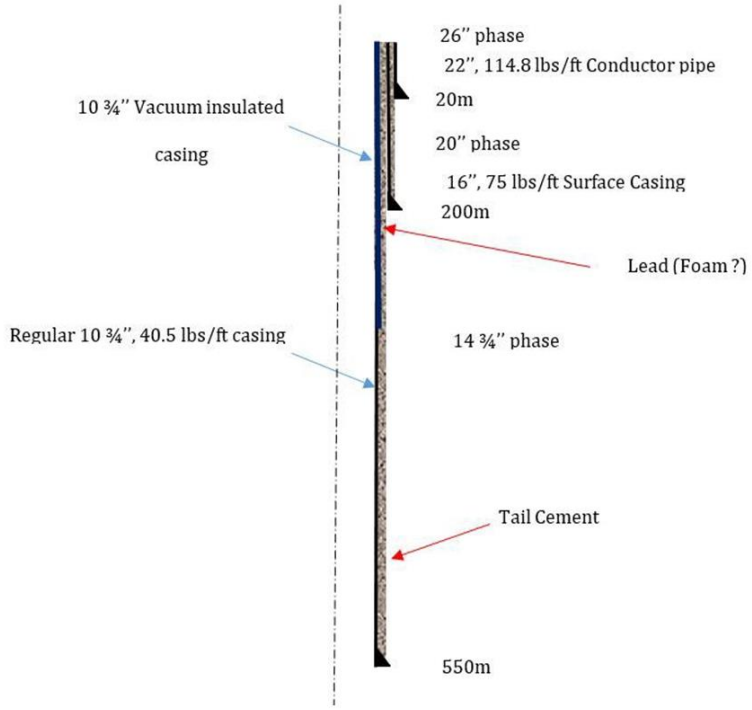


Figure 3.4.7: Planned well design

3.4.2 Models, model codes and data used

3.4.2.1 University of Bern, batch chemical and reactive transport modelling

Modelling carried out at the University of Bern focuses on geochemical water-rock interaction that is expected to occur during the operation of the Forsthaus HT-ATES. Geochemical reactions are computed either in stand-alone batch simulations using PHREEQC (*https://www.usgs.gov/software/phreeqc-version-3*) or in a coupled thermal-hydraulic-chemical (THC) mode using PFLOTRAN (*www.pflotran.org*). Both codes are free and open source software packages.

There are currently no data available from the Forsthaus site because drilling is scheduled to start at the end of the year 2021. Limited published information is available on the properties of the target sandstones in the area of Bern City. As a means of filling this knowledge gap for the Geospeicher Forsthaus project, drill core samples were obtained from a new (2017) 35 m deep borehole that was drilled approximately 2 km to the southeast of the planned HT-ATES. Analyses of groundwaters from six nearby wells were also obtained. These samples constitute the closest first-hand indications on the mineralogy of the reservoir rock and the pore water chemistry of the USM to the planned HT-ATES site.

The sample analyses yielded mineral and porewater compositions (Tables 3.4.3 and 3.4.4, respectively) and estimates for porosity and permeability (Table 3.4.5). In addition, several experiments were carried out on the cores to better understand chemical processes taking place in the reservoir and in the heat exchanger during operation under expected thermal and hydrogeological conditions. These investigations comprised three types of experiments: 1) infiltration of artificial pore water into rock samples at 60 °C and 90 °C, batch water-rock reactions at 20 °C, 60 °C and 90 °C and precipitation/nucleation experiments involving calcite.
Results from these experiments provide constraints for the type of mineral reactions (e.g. silicate and carbonate dissolution/precipitation, surface reactions in clay minerals) that are expected to occur in the reservoir during HT-ATES, and for their rates.
These experimental constraints are used in system-scale reactive transport simulations to assess the potentially detrimental impact that chemical processes may have on the efficiency of the HT-ATES. Potential risks induced by chemical reactions include clogging of flowpaths in the reservoir due to mineral precipitation, the release of unwanted chemical compounds from dissolving minerals and their transport into aquifers used for other purposes, mineral scaling in the wells and heat exchanger and corrosion of the installation. (e.g. mineral scaling in the heat exchanger, clogging of fluid pathways in the reservoir due to mineral precipitation).

Table 3.4.3: Quantitative XRD-analyses (reported as wt.%) performed at the University of Bern on drill corefrom the USM

	Wt%
Quartz	41
K-feldspar	8
Plagioclase	20
Calcite	8
Dolomite	3
Illite/Musc./Biotite	6
Smectites	12
Kaolinite	0
Chlorites	2

Table 3.4.4: Chemical composition of USM groundwaters collected near the main railway station of Bern.

		GW1	GW2	GW3	GW4	GW5	GW6
Water table	m a.s.l.	515	522	523	518	513	510
Temperature	°C	15.6	15.3	13.7	14.1	13.4	13
EC (sampl.)	µS/cm	633	1,110	580	714	660	641
pH (sampl.)	-	7.79	7.76	8.05	7.51	7.90	8.33
Ca2+	mg/L	71.5	118	26.6	72.6	62.9	17.5
Mg2+	mg/L	29.5	37.5	19.7	41.7	36.2	10.9
Na+	mg/L	13.3	49.3	75.7	23	21.3	108
K+	mg/L	1.8	3.7	2.8	3.1	2.6	2.6
HCO3-	mg/L	231	190	298	244	237	243
Cl-	mg/L	43.7	151	30.8	104	65.5	56
NO3	mg/L	59.1	123	0.4	20.7	23.3	12.2
SO42-	mg/L	28.2	80	22.5	39.3	48.2	41.6
DOC	mg/L	< 1.0	1.3	< 1.0	< 1.0	2.3	3.5
TDS	mg/L	478	753	476	548	497	492
Simplified water type[1]		Ca-HCO3	Ca-Cl	Na-HCO3	Ca- HCO3	Ca-HCO3	Na-HCO3
S.I. calcite (sampling)		0.2	0.3	0.3	0.2	0.3	0.1

[1]Only major cation and anion are given.

Table 3.4.5: Porosities, grain densities and Klinkenberg-corrected permeability results from a USM core sample.

Porosity (vol.%)	18.3
Grain density (g/cm^3)	2.67
Permeability (mD)	370

Numerical simulations are carried out in two steps. First, PHREEQC is used to design a chemical model that reproduces the experimental results. This involves 1) constructing a representative geochemical reaction network, comprising reacting primary and secondary minerals, aqueous species and possibly other reactive compounds such as organic substrates. 2) identifying and incorporating the relevant reaction processes (mineral dissolution/precipitation reactions, ion exchange, aqueous complexation and redox reactions) and 3) calibrating the mineral reaction rates against the time series of water compositions extracted from the experiments. Once satisfactory agreement between the PHREEQC model and experimental results has been achieved, the calibrated chemistry model is translated into PFLOTRAN format.

In the second step, a 3D coupled thermal-hydraulic-chemical (THC) model of the Forsthaus HT-ATES is constructed in PFLOTRAN. The chemical model in PFLOTRAN corresponds to the PHREQQC model that fits the experimental data. Owing to the lack of site-specific information, in the current preliminary model the USM is represented as a generic succession of alternating low-permeability clay and high-permeability sandstone layers of variable thickness. The total thickness of the reservoir is 350 m. The USM is overlain by a 150 m thick layer of Quaternary sediments (Figure 3.4.8 and Figure 3.4.9). The well arrangement is patterned after that in Figure 3.4.1, the distance between supporting and main well is 50 m (Figure 3.4.9). The system operates according to the parameters and the loading/unloading schedule summarized in Figure 3.4.2 and Table 3.4.1, respectively.

The system-scale reactive transport model provides a theoretical framework in which the results from the experiments are coupled to the thermal-hydraulic (TH) processes taking place during repeated loading and unloading cycles. Feedbacks between chemically induced porosity and permeability changes and the thermal-hydraulic model can be implemented as an option. The model results provide spatial distributions of thermal, hydraulic and chemical properties and their evolution in time. Aside from identifying potentially adverse chemical processes mentioned above, these simulations also track the shape and extent of the thermal plume over time and can thus aide with the design of the well arrangement and the loading/unloading schedule for optimal heat exploitation.

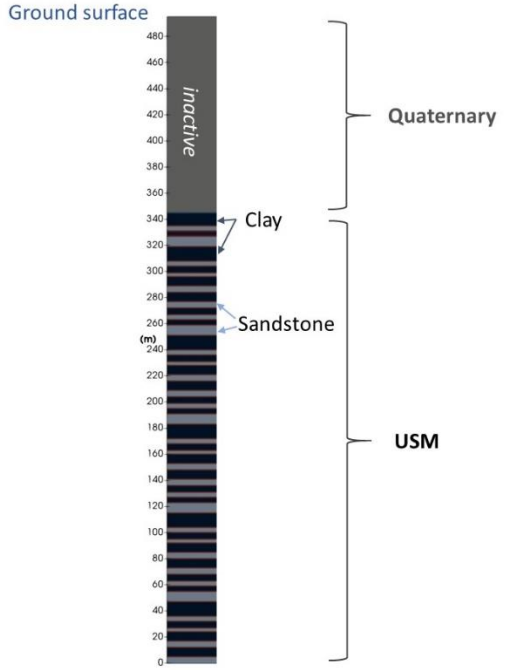


Figure 3.4.8: Generic representation of the stratigraphy at the Forsthaus site.

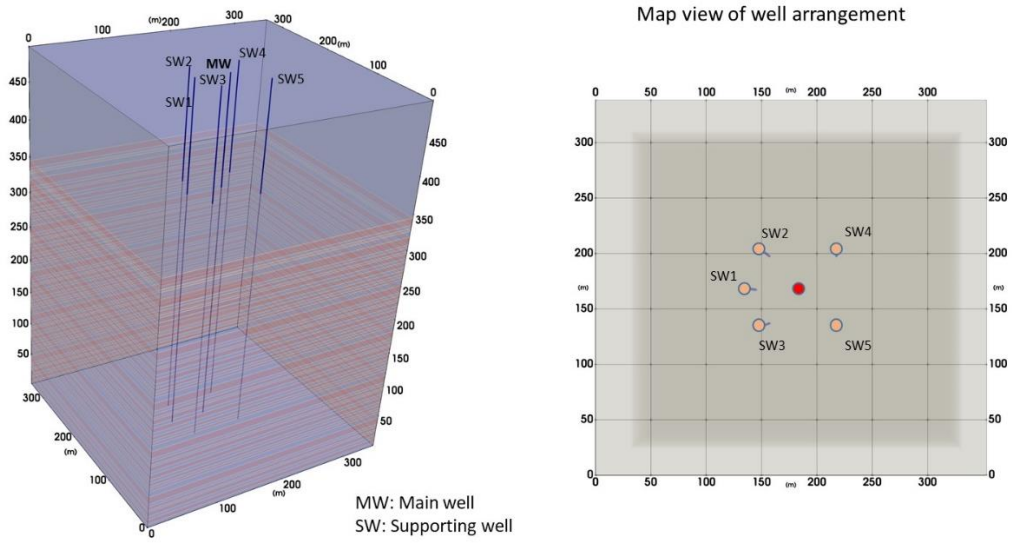


Figure 3.4.9: Model domain and well arrangement (MW: Main well, SW: Supporting well).

3.4.3 Validation

A chemical batch model using PHREEQC is designed to reproduce the experimental data carried out on drill cores from a borehole some 2 km away from the Forsthaus site and using water compositions consistent with pore waters sampled from various wells drilled into the USM. The PHREEQC model is thus constrained by field data and “validated” against experimental results.

However, one should be aware that here are several sources of uncertainty even if the model is validated against experiments. For instance, the chemical reaction network of the model is restricted to the availability of phases and their thermodynamic properties in the database and may not include the phase that occur in reality. Moreover, given the large number of independent parameters, the choice of phases and parameters may not be unique and the same results can be obtained by using a different combination of phases and parameters. In the end, the choice of phases and parameters, if not constrained by experimental results, depend on the user’s judgement and expertise.

There are currently no data from the Forsthaus site that can be used to validate the PFLOTRAN site-scale model. Nevertheless, PFLOTRAN has been verified against other reactive transport codes in a benchmarking project within the HEATSTORE framework (Task D2.3) (Alt-Epping and Mindel, 2020). The benchmark problems presented in Alt-Epping and Mindel (2020) were loosely based on the Forsthaus system using preliminary chemical and mineralogical data from the analyses of the USM drill cores and pore water samples. This benchmarking exercise showed excellent agreement between participating codes for simple simulation cases (Figure 3.4.10). The agreement was still very good for more complex reactive-transport problems (Figure 3.4.11 and Figure 3.4.2). Three modelling groups from different institutions participated in this benchmark exercise using three different codes: University of Bern (PFLOTRAN), BRGM (MARTHE-PHREEQC) in France and UPC in Spain (RETRASO).

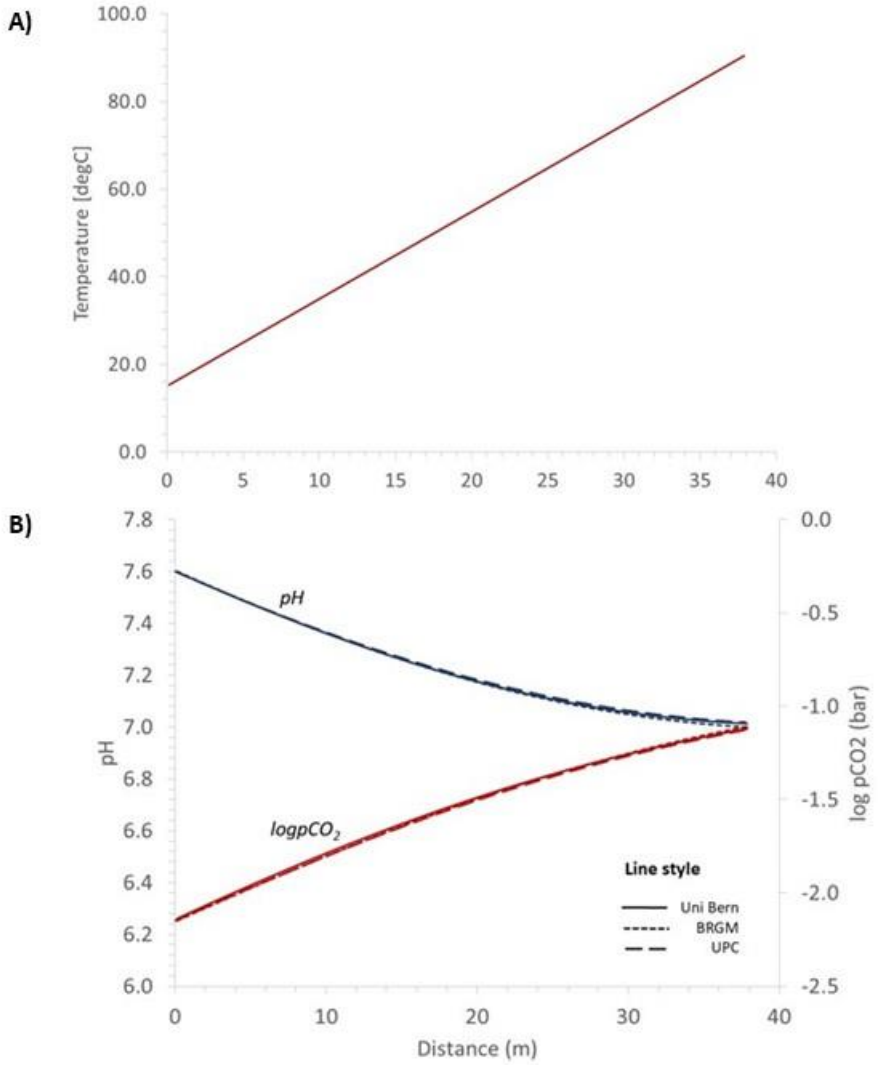


Figure 3.4.10: A simple benchmark problem in Alt-Epping and Mindel (2020) involved up-temperature flow along a 1D flowpath allowing carbonate minerals to precipitate. Panel A: temperature profile along the flowpath, flow is from left to right, implying a temperature increase from 15 °C to 90 °C. Panel B: example output and comparison of results (pH and log(pCO_2) from the benchmark participants. The agreement between the three participating codes is excellent.

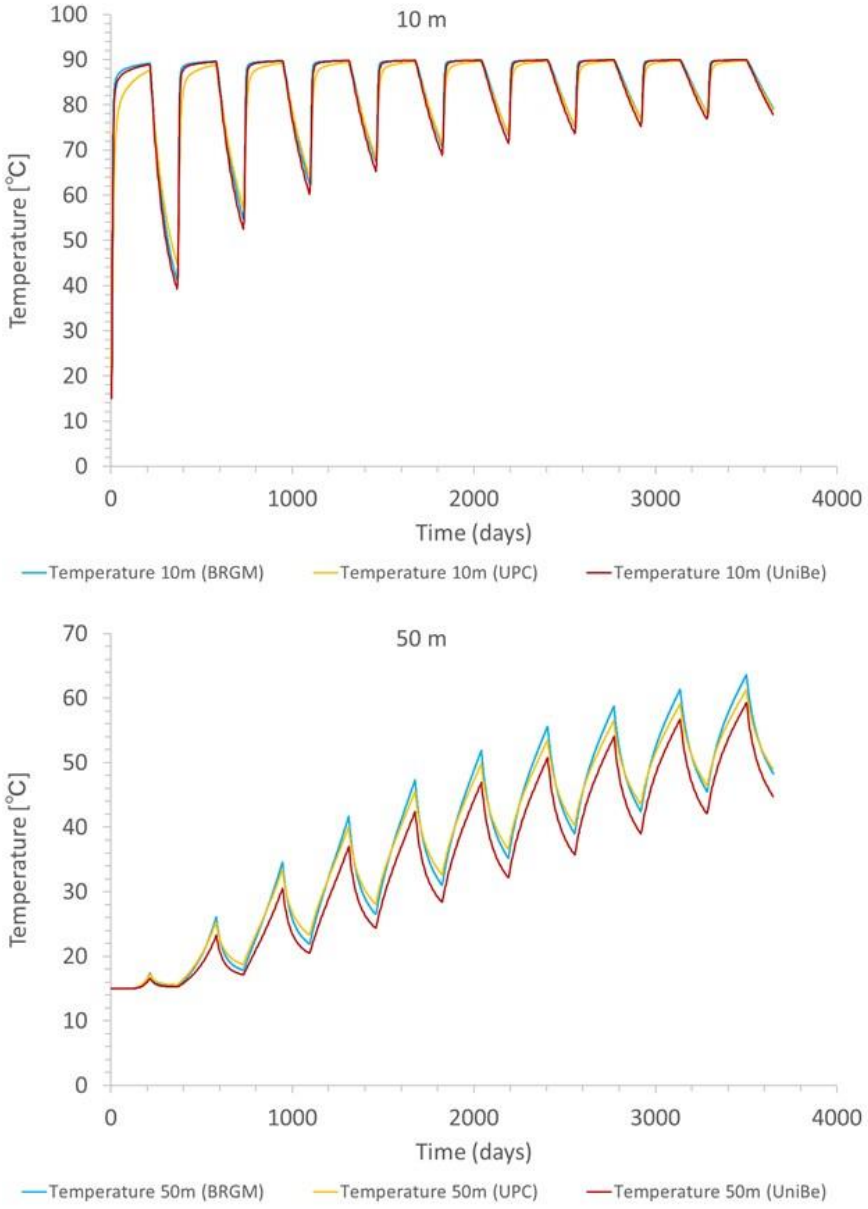


Figure 3.4.11: Temperature evolution in an axisymmetric model over 10 years of repeated loading and unloading, 10 m and 50 m away from the injection/extraction well. The code intercomparison shows very good agreement.

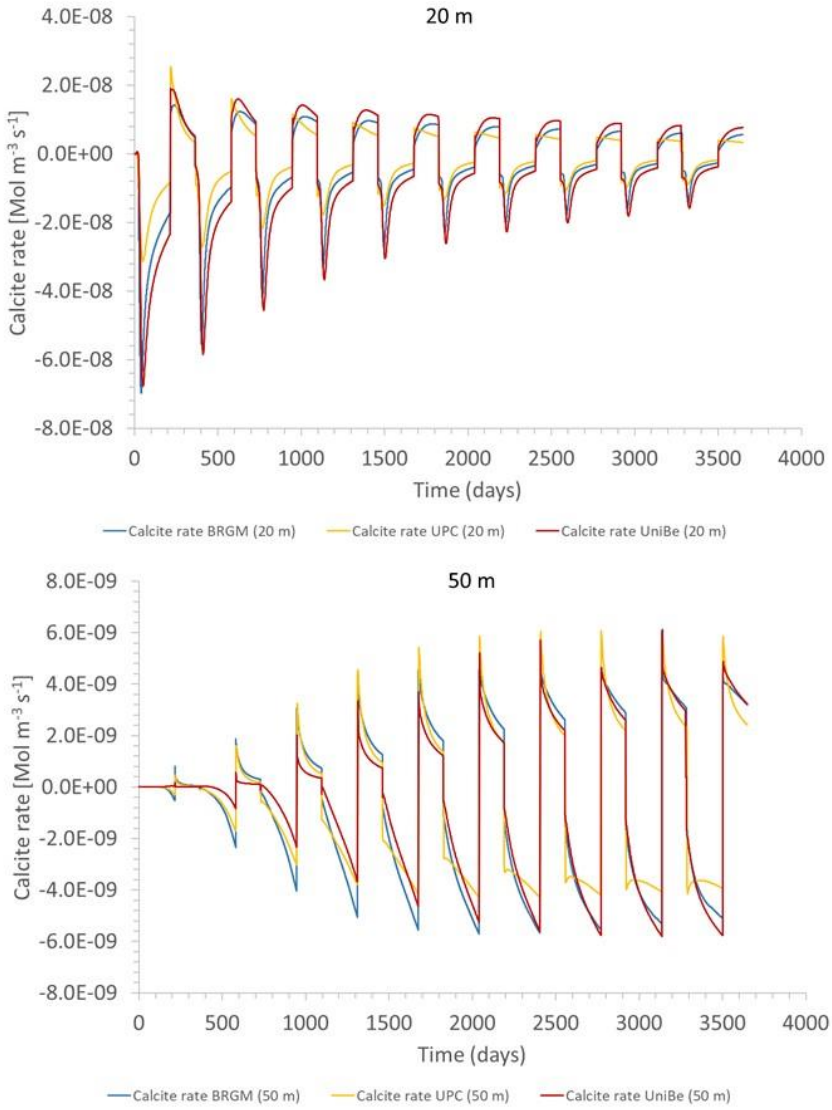


Figure 3.4.12: Calcite precipitation (> 0) and dissolution (< 0) rates corresponding to the temperature evolution in Figure 3.4.11, 20 and 50 m away from the injection/extraction well. The code intercomparison shows good agreement.

As soon as drilling starts in late 2021, we expect to learn more about the site-specific stratigraphic succession of sandstone and clay units, their hydraulic properties (permeability, porosity), mineralogy and pore water composition. These are all crucial information for the parameterization of the system-scale reactive transport model. Furthermore, the final arrangement of the wells, the characteristics of the heat exchanger (volume, surface area) and the pumping schedule will be essential input to the model. The existing preliminary model can be easily adapted as soon as new information and data become available. A rigorous validation of the model can only be performed when the project has been launched.

3.5 Reykir/Hengill, Iceland

The two Icelandic case studies are not specifically in this task and therefore do not have a specific validation exercise. Standard calibration procedures were however used to calibrate numerical models for both the Hengill high temperature site and the Reykir/Reykjahlíð low temperature site within tasks 2.1 and 2.2 in WP2. The model setup and calibration results for both case studies are described in detail in project deliverable 2.1 (Driesner, in prep.) but due to their connection to the topic of calibration/validation discussed here they are shortly listed here as well.

3.5.1 Site description Hengill

The Hengill area is located in SW-Iceland, about 30 km east of Reykjavík (Figure 3.5.1). It is a volcanically and tectonically active area built up of basaltic rocks, both hyaloclastite formed during glacial periods and interglacial lavas. The area has been studied intensively in connection with geothermal utilization (see e.g. Árnason et al., 2010; Franzson, 1998; Franzson et al., 2010, 2005; Helgadóttir et al., 2010; Saemundsson, 1995; Sinton et al., 2005 and references therein). Two co-generative geothermal power plants are operated in the Hengill area: The Nesjavellir Power Plant, commissioned in 1990, in the northern part of the volcanic complex and The Hellisheiði Power Plant, commissioned in 2006, located in the southern part of the volcano. The combined installed capacity of both power plants is 423 MW_e and 540 MW_{th}. The geothermal reservoirs supplying fluid for the power plants are 230-330°C. In total 88 high temperature production wells, and 26 injection wells, the deepest more than 3000 m, have been drilled into the geothermal systems in the Hengill volcano to supply steam and water for the power plants and receive the spent fluid for reinjection.

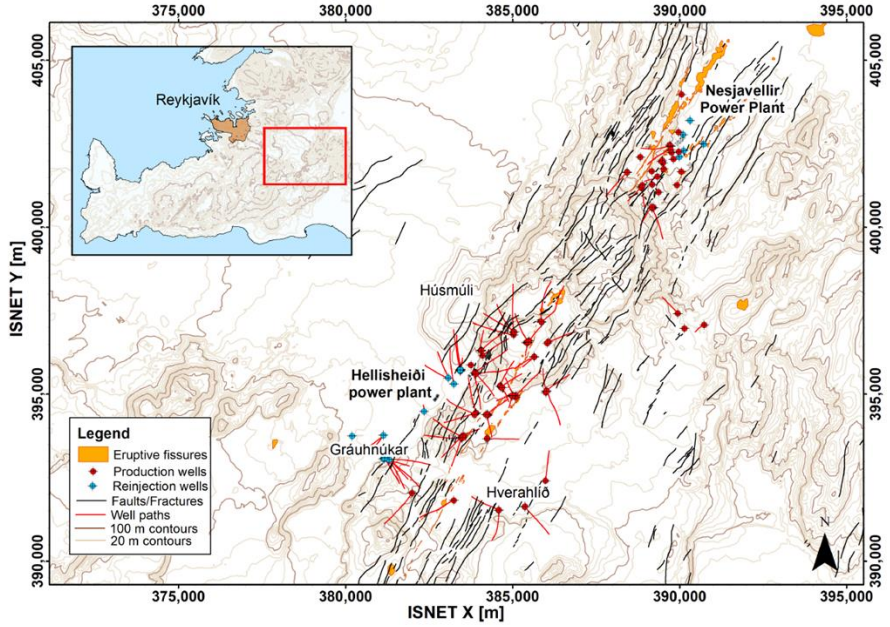


Figure 3.5.1: A map of the Hengill area showing elevation contours, surface fractures, eruptive fissures, production and reinjection wells as well as well paths projected to the surface. 1:90000. The inset shows the location of the area in SW-Iceland (Data Source: Reykjavík Energy and Nation Land Survey of Iceland) (Map from Driesner (in prep.)).

3.5.2 Models, model codes and data used for Hengill

A field-scale hydrothermal numerical model exists covering the whole Hengill area. It is run using the numerical simulator TOUGH2 (Pruess et al., 2012) as implemented in forward and inverse mode within the iTOUGH2 code (Finsterle, 2007). Within the HEATSTORE project, the model was updated, recalibrated and deepened to include deeper layers underneath the conventional geothermal system to prepare for deeper drilling within the field. A detailed description of the model setup is presented in Driesner (in prep.).
Extensive monitoring data is collected from the wells to monitor the behavior and response of the field to utilization and this data is used to calibrate the numerical model. Well head pressure is continuously logged. Every six months tracer fluid tests are performed on the wells. Results from these tests give enthalpy data and are used to construct productivity curves that are then used to continuously estimate flow from each well from

pressure measurements. Temperature and pressure profiles with depth are measured annually in monitoring wells and the data used to estimate pressure drawdown in the system in response to production. Numerous pressure and temperature profiles are also taken following the drilling of each well to be able to estimate formation temperature and pressure. Flow of water into reinjection wells is continuously logged.

3.5.3 Calibration for Hengill

The model was calibrated by fitting observed data and production history to simulation results. The initial state of the model was calibrated against estimated formation temperature and pressure profiles and the production history was calibrated against drawdown in monitoring wells and enthalpy measurements from production wells for the whole production history. An example of calibration results for formation pressure and temperature for well HE-42 in Hellisheiði is shown in **Figure 3.5.2**a and an example of calibration results for pressure drawdown in three monitoring wells in Hellisheiði is shown in **Figure 3.5.2**b. A detailed description of the calibration results, both for the shallower and deeper version of the model, is presented in Driesner (in prep.). New data is added to the model annually.

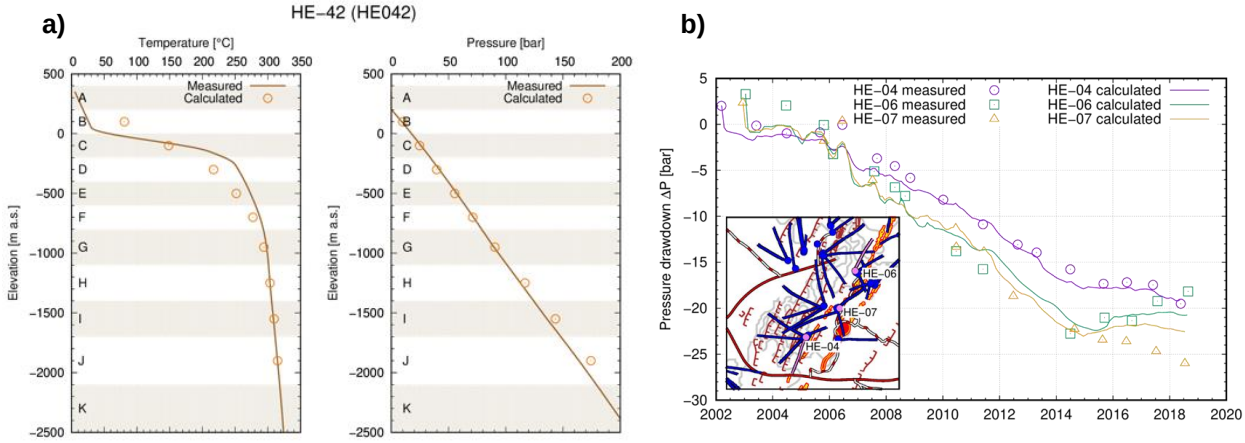


Figure 3.5.2: Examples of a) initial state comparisons between measured and calculated formation temperature and pressure for well HE-42 in Hellisheiði; and b) comparison between measured and calculated drawdown in monitoring wells in Hellisheiði (figures from Driesner (in prep.)).

3.5.4 Site description Reykir/Reykjahlíð

The Reykir/Reykjahlíð geothermal system covers about 10 km^2 and is located within Iceland´s capital area. The system is separated into two subareas, Reykir and Reykjahlíð. They are both at an elevation of about 20-80 m above sea level (m a.s.l.). The stratigraphy in the area is characterized by alternating sequences of subaerial basaltic lava flows and hyaloclastite formed during glacial periods (Tómasson, 1997). Production from deep wells in the fields started in 1971 and today the average combined production is about 1000 L/s of 86 °C warm water which is supplied to the district heating system. Deep production caused a decline in system pressure, free flowing from older shallower wells stopped and water level dropped down to a depth of 50-100 m below sea level (Björnsson and Steingrímsson, 1995). Today, active production wells are 34, 22 in Reykir and 12 in Reykjahlíð (Figure 3.5.3).

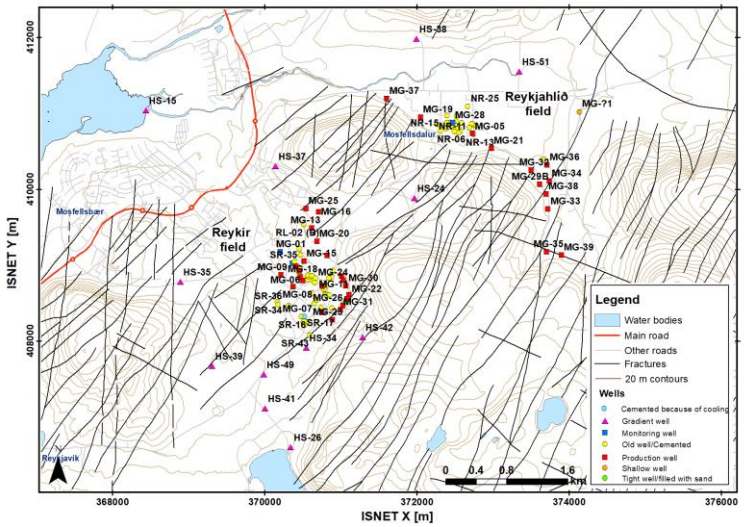


Figure 3.5.3: Map of the Reykir/Reykjahlíð fields showing wells, roads, elevation contours, and fractures that have been mapped on the surface for this specific area. Active production wells are shown with red squares (Data source: National Land Survey of Iceland, Reykjavík Energy and Iceland Geosurvery, ÍSOR) (Map from Driesner (in prep.)).

3.5.5 Models, model codes and data used for Reykir/Reykjahlíð

A field-scale hydrothermal numerical model covering both subareas was constructed and calibrated within the HEATSTORE project. Similarly as for the Hengill model, it is run using the numerical simulator TOUGH2 (Pruess et al., 2012) as implemented in forward and inverse mode within the iTOUGH2 code (Finsterle, 2007). A detailed description of the model setup is presented in Driesner (in prep.). Monitoring data is collected from the wells to monitor the behavior and response of the field to utilization and this data was used to calibrate the numerical model. Flow and temperature of produced water from the wells is monitored every two weeks. Water level in monitoring wells was monitored monthly but in 2019 loggers were placed in the wells that record water level every 10 min. Sensors that log water level every hour were additionally placed in two wells in Reykir. Temperature profiles with depth were taken following the drilling of each well, both production, monitoring and research wells, to be able to estimate formation temperature in and around the fields.

3.5.6 Calibration for Reykir/Reykjahlíð

The model was calibrated by fitting observed data and production history to simulation results. The initial state of the model was calibrated against estimated formation temperature profiles. The production history was calibrated against drawdown in monitoring wells and temperature measurements of produced fluid from production wells for the whole production history. An example of calibration results for formation temperature for wells MG-05 and MG-36 in Reykjahlíð is shown in **Figure 3.5.4a and an example of calibration results for pressure drawdown in three monitoring wells in the fields is shown in Figure 3.5.4b. A detailed description of the calibration results is presented in Driesner (in prep.).**

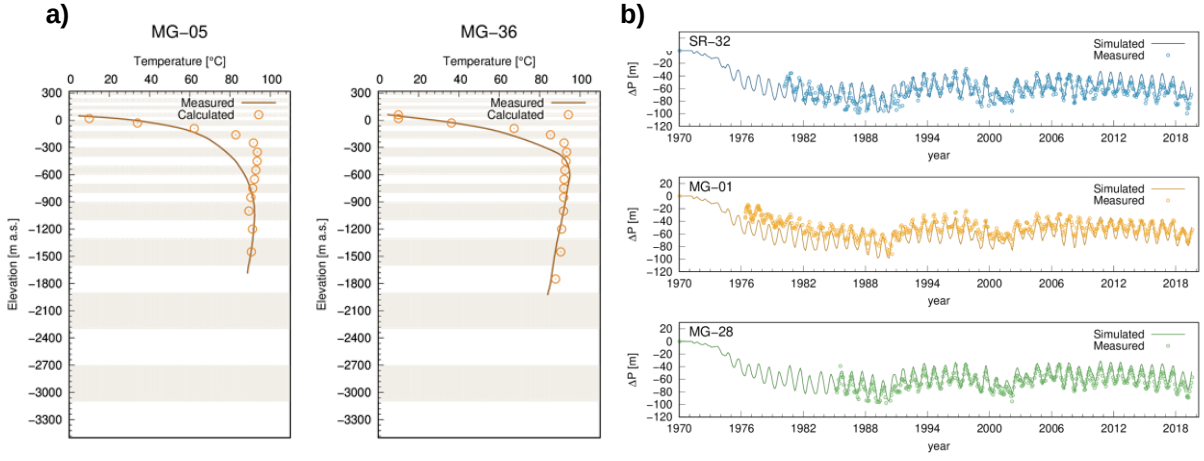


Figure 3.5.4: Example of a) initial state comparisons between measured and calculated formation temperature for wells MG-05 and MG-36 in Reykjahlíð and b) comparison between simulated and measured draw-down for monitoring wells SR-32, MG-01 and MG-28 during the period 1970-2019.

3.6 BTESmart Vallin fier, Annecy, France

3.6.1 Site description

The new version of the BTESmart project is located in Annecy, France. A geothermal borehole heat exchangers field (18 BHEs/100 m each) was built in 2012 under the playgrounds of the school “Vallin-Fier”, to heat and cool the buildings. After nearly 10 years, the temperature measurements showed that the soil temperature was decreasing for many reasons, including the non-use of geocooling, as planned. The BTESmart Vallin-Fier projects consists in converting this geothermal facility into a real heat storage, by connecting it to solar thermal panels to recharge the underground, mainly during summer. In addition, two wells (130 m) were drilled inside the field. Temperature sensors were installed inside them to give more insight in the conversion process, in August 2021. On this new pilot site, it was however not possible to install lateral recovery boreholes.

Figure 3.6.1: Geothermal BHEs filed of “Vallin-Fier” school and location of the two monitoring wells of the BTESmart project.

3.6.2 Models, codes and data used

In the final version of deliverable 2.1, the initial version of the BTESmart project was modelled using two different approaches: numerical (using FEFLOW) and analytical.
As agreed with the National Funding Agency, due to the change of pilot site during the Heatstore project, the tasks and deliverable 5.3 will be done in the frame of a national extension.
The work to be done will consist in:

- Validate the analytical model written in Matlab in the case there is circulation in the lateral boreholes (against numerical model),
- Integrate this analytical model in a new TRNSYS Type.
- Make the models fit the data by reasonably changing key parameters.

3.6.3 Validation

As the national extension will last until January 2023, there will be data to use for comparison for a full winter (2021/2022), a full summer (2022) and the beginning of the following winter (October/November 2022). The data to be used for model validation will be: temperatures at the inlet and outlet of the geothermal field, flow inside the BHEs and temperatures measured along the 2 monitoring wells (32 sensors per well).

3.7 Fraunhofer IEG colliery, Bochum, Germany

3.7.1 Site description

The Bochum pilot plant aims at utilizing an abandoned coal mine, which is directly located under the premises of the Fraunhofer IEG Campus in Bochum, as a seasonal mine thermal energy storage. Seasonal surplus heat from solar thermal collectors will be stored during the summer within the mine and utilized for heating the Fraunhofer buildings on the site.in winter
The MTES demo-site is located at the Fraunhofer IEG where, among other infrastructures and laboratories, a 10.000 m² drill site with existing research, observation and production wells, in conjunction with the approved authorized 50 km² mining area "Future Energies", allows further drill tests by using the Bo.REX (Bochum Research and Exploration Drilling Rig). This led the way of a very cost-effective exploration of the flooded coal mine in a depth of approx. 63 m below ground.
The Markgraf II mine produced 37.043 tons of coal during 1953 to 1958. Based on a calculation with a coal density of 1,35 g/cm^3, we can assume a void volume of approx. 27.439 m^3. This volume does not include any drifts and shafts, which need to be analysed based on the mine layout. Considering the effect of mine subsidence, the remaining void volume will most likely be in the range of approx. 10 %. Utilizing a ΔT of 50 K within the mine water, a heat capacity of approx. 165 MWh, which resembles the yearly heat demand of the Fraunhofer IEG compound, could be stored within dedicated drifts of the small colliery for the heating season. Based on this first evaluation the yearly Fraunhofer IEG heat demand could be substituted by emission free solar thermal energy.

The site is located at the southern edge of the northerly dipping “Münster Cretaceous Basin”. It is one of the biggest continuous sedimentary basins in Germany with sediments consisting primarily of Upper Cretaceous layers (Figure 3.7.1). Among these, the argillaceous marls of the Emscher Formation are of particular importance because they seal the upper aquifer. The Emscher Formation comprises Campanian, against the lower aquifer which lies within Upper Carboniferous and Cenomanian/Turonian strata. The Emscher Formation shields because of its sealing characteristics, the underlying hard-coal deposit of Carboniferous age. Close to the pilot plant location Carboniferous rocks are cropping out at the surface.

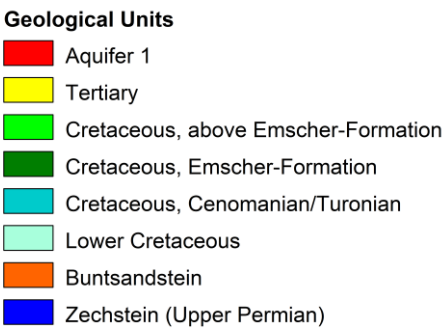


Figure 3.7.1: Geological units of the “Münster Cretaceous Basin”

The groundwater flow in the “Münster Cretaceous Basin” can be schematized considering two different types of aquifers.
The first aquifer type is constituted by shallow aquifers: they are spatially discontinuous if the whole basin is considered; they are generally outcropping but can be locally overlapping and vertically bounded by impervious strata e.g. Emschermergel (Coniac/Santon).
The second is a deep aquifer, hundreds of meters thick in the central region, which corresponds to the intensely fissured Cenomanian-Turonian carbonate basement, which extends over the whole basin; it is hereinafter called deep aquifer. This aquifer is affected by a southward directed regional flow coming from the “Teutoburger Wald” mountains.

3.7.2 Model codes and data used

The software code chosen for the numerical finite-element modelling work was the 3D groundwater flow model SPRING, developed by the delta h Ingenieurgesellschaft mbH, Germany (König et al., 2020). The program was first published in 1970, and since then has undergone a number of revisions.
SPRING is widely accepted by environmental scientists and associated professionals. The software allows the simulation of steady and non-steady flow in aquifers of irregular dimensions as well as confined, unconfined and unsaturated flow, or a combination thereof. It is also possible to integrate model layers of varying thickness or to let certain layers pinch out. The software code also supports coupled density-dependent flow of groundwater and mass transport processes in fractured systems with discrete fractures as well as stochastically generated fracture networks. SPRING uses the finite-element approximation to solve flow and transport equations. This means that the model area or domain is represented by a number of nodes and elements. Hydraulic properties are assigned to these nodes and elements and an equation is developed for

each node, based on the surrounding nodes. A series of iterations are then run to solve the resulting matrix problem utilising a pre-conditioning conjugate gradient (PCG) matrix solver for the current model.
The model is said to have “converged” when errors reduce to within an acceptable range. SPRING is able to simulate steady and non-steady flow, in aquifers of irregular dimensions, as well as confined, unconfined and unsaturated flow, or a combination thereof.
Different model layers with varying thicknesses as well as out pinching model layers are possible. The edges of the model domain, or boundaries, typically need to be carefully defined, and fall into several standard categories. SPRING is a fully integrated solution; it comes with an integrated pre- and post-processor, aimed at making data input and 2-D and 3-D visualisation faster and simpler.

3.7.3 Calibration for Bochum

A steady-state calibration of the site groundwater flow model was performed using 8 groundwater level data points within the model domain. Only water levels observed in groundwater monitoring boreholes were considered representative of the shallow and deep aquifers and used for the calibration. Figure 3.7.2 shows the modelled groundwater surface of the upper aquifer (steady state).

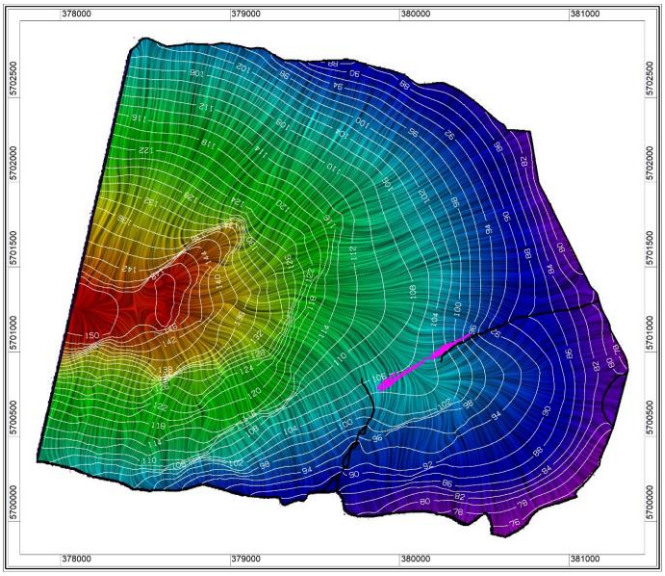


Figure 3.7.2: Calculated groundwater surface of the upper aquifer (white contour lines) and flow field coloured by potential head from red=high to purple=low, mine site in magenta.

3.7.4 Validation

In December 2020 in-situ tests were carried out to predict the plant operation accurately and update the model. The provided data sets were used to calibrate the SPRING model. Over a period of 7 days 46 °C warm water was pumped into to the system with a flow rate of 5.8 m^3/h. It was circulated between injection well MO1 and production well MP1 (Ort 4 - scenario C). After 7 days, pumping was stopped, and the system was continuously monitored. The injection test, or transient state data set, reflects the desired response to injection.
The SPRING model should ideally be able to follow the steady-state data and also accurately model the change in temperature during the transient state.
For the transient state the proposed injection scenario was simulated using the model by keeping infiltration and production rates constant. The model was set to inject 46 °C warm water at the injection well and to extract the same amount of water at production well. The initial temperature of the system was set at 11 °C.
Figure 3.7.3 shows the modelled temperature distribution for the injection test between MO1 (infiltration) and MP1 (production) wells. Figure 3.7.4 shows the temperature response of the calibrated model at the production well.

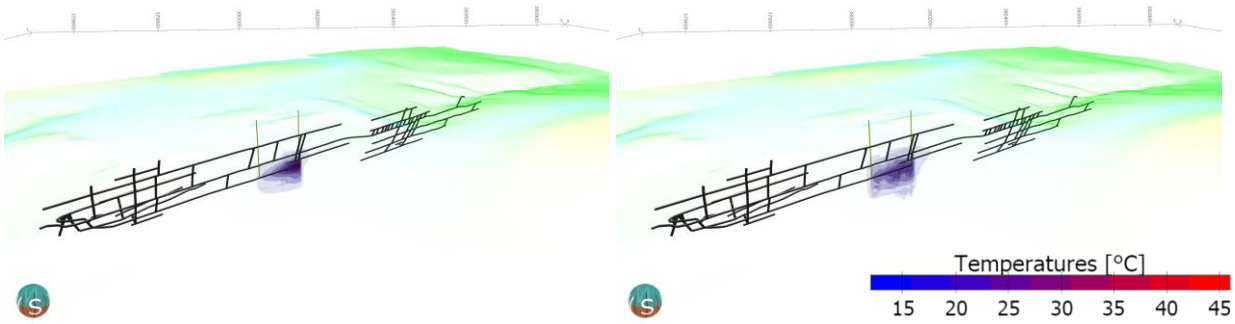


Figure 3.7.3: Temperature distribution model after 3 days (left) and 7 days (right) of the in-situ test.

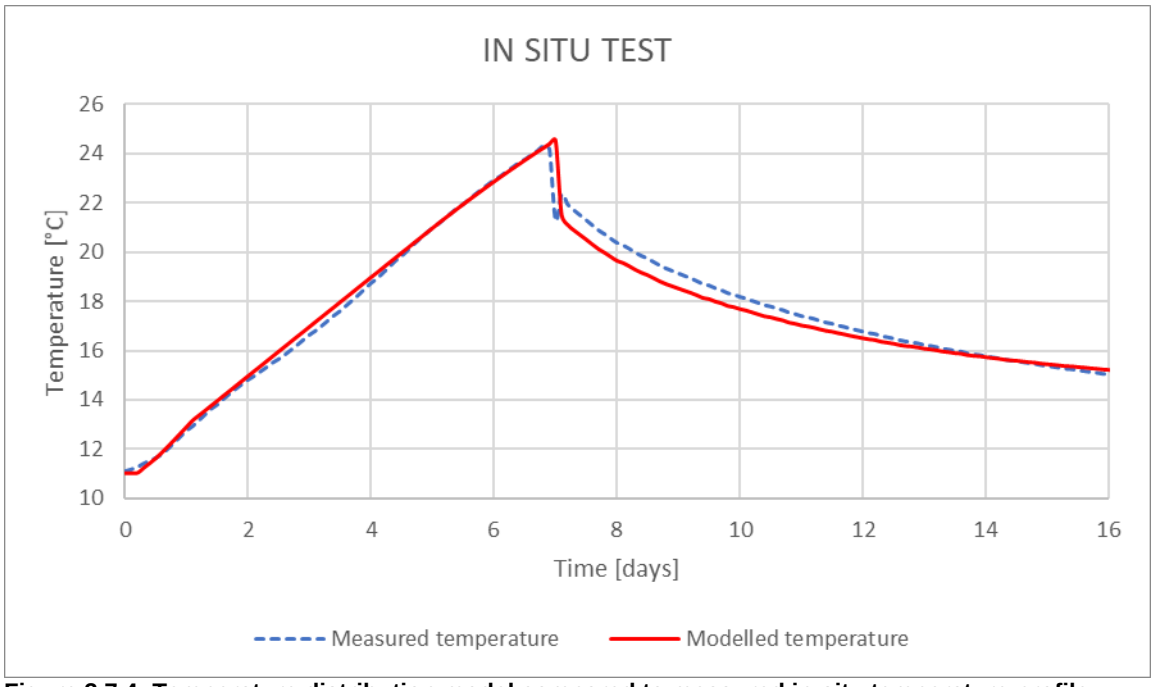


Figure 3.7.4: Temperature distribution model compared to measured in-situ temperature profile.

As the injection rate is much higher in comparison to the sensitivity analysis, temperature at the production well increases more rapidly. After 7 days, the infiltration was stopped and temperatures decreased as expected. In comparison to the infiltration phase the system reacts with a slower temperature decrease. A detailed description of the MTES calibration results is presented in Driesner (in prep.).

3.8 Dronninglund, Denmark

The site of Dronninglund has been extensively presented as a part of WP2 in D2.1 and in TR2.3. The main point of the present study is to make use of the latest measurements gathered from the Pit Thermal Energy Storage (PTES) in Dronninglund to further validate the models that have been previously developed and gather information about the evolution of the PTES over more recent years.

3.8.1 Site description

This Danish case study is that of a PTES, set up as a part of the heating system that combines 37'573 m^2 of solar collectors, a 60'000 m^3 PTES and a 2.1 MW (cooling capacity) absorption heat pump (until mid-2020). The system supplies heat to a district heating network located in Dronninglund, Denmark.

This site has been extensively presented in TR2.3 and remains unchanged for the current work. As a reminder of the main geometry and thermal parameters from the PTES used as a reference for the modelling part, and more specifically for the present work, it is considered that:

- The soil around the storage has a density of 2'000 kg/m^3, a thermal conductivity of 0.4 W/(m·K), and a specific heat of 900 J/(kg·K)
- The storage is in the shape of an inverted truncated pyramid, has a height of 16 m, a bottom square area of 26 x 26 m^2, a top square area of 90.38 x 90.38 m^2, and a slope 1:2 (26°) to the horizontal (see Figure 3.8.1)

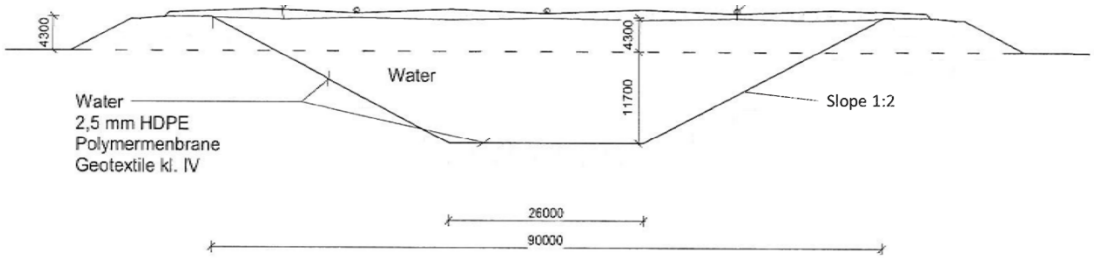


Figure 3.8.1: Dronninglund PTES main geometry parameters.

The same measurements as the ones used for TR2.3 will be reused for the comparison of the present report, but in a slightly different way (see following section about data). Additional available measurements have been used, and some measurements are no longer used. The list of measurements used is the following:

- 32 water temperature sensors, one for every 0.5 m
- 3 temperature sensors in the pipes leading to the top, middle and bottom diffusers of the PTES
- Flow measurement for the middle pipe, and calculated flows for the top and bottom pipes of the PTES
- Flow measurements for the 2 solar heat exchangers inlet, on the secondary side (side of the PTES and the DHN)
- Temperature measurements in the pipes at the inlet and outlet of each solar heat exchangers, on the secondary side
- Flow measurement at the outlet of the transmission line to the DHN
- Temperature measurements in the pipes (forward and return) of the transmission line to the DHN
- Ambient air temperature

Several extra measurements have been used compared with TR2.3: the ambient air temperature, the solar heat exchanger's flows and temperatures, and the transmission line flow and temperatures. The ambient air will be used to set the temperature above the storage cover and the temperature above the soil in the TRNSYS model. The solar heat exchangers and transmission line measurements will be used as a reference for the energy balance of the system.

3.8.2 Models, model codes and data used

3.8.2.1 Model used

The models chosen for the present study are the same as those used in TR2.3. Only the specifics for this study will be described here.

One of the conclusions that can be taken from TR2.3 is that Type 1300-1301 is most appropriate for the comparison between available measurements and calculations. The reason for that is:

- Type 1300-1301 is a good compromise compared with Type 342 and Type 1322: it fast in terms of computational time, and its geometry is close to the actual geometry of the PTES, thus providing good accordance with measurements once calibrated
- The available measurements being incomplete and imprecise, having such a precise model as Type 1322 doesn't improve the agreement between calculations and measurements, even after calibration

It was therefore chosen to use only Type 1300-1301 to make calculations for the present study.

3.8.2.2 Data used

The data used for the present study comes from the same dataset as the one used in TR2.3, using not only data from 2017, but also data from 2014, 2015, 2016, 2018 and 2019. The main updates on the data used for this study are presented below.

3.8.2.2.1 Averaged values

The conclusion of TR2.3 regarding the choice between 10' data and 60' data was that 10' inputs did not improve accordance between model calculation and measurements, while increasing the total calculation time for all models. The use of 60' averaged data can therefore be recommended to obtain faster results.

Another benefit of using 60' averaged values as inputs is that the measurement inaccuracies are absorbed. For the present study, it was chosen to make an arithmetic average of the measured flows and for the temperatures measured inside the PTES, but use a different kind of average for the temperatures measured inside pipes. For those, a flow-weighted average was chosen, because the measured temperature of the water flowing through a pipe is most representative when the flow is high.

Here is an example to illustrate why: let's consider a given theoretical pipe. For a given hour, during the first ten minutes a water flow of 100 kg/s flows through the pipe, and during this time the measured temperature is 10°C. During the remaining fifty minutes, the flow is 10 kg/s and the measured temperature is 15°C. Calculating the energy input (with a reference temperature of 0°C) gives: $\dot{m} \cdot \Delta t \cdot Cp \cdot \left(T_{in} - T_{ref}\right) = \left((100 \cdot 60 \cdot 10 \cdot 4180 \cdot 10) + (10 \cdot 60 \cdot 50 \cdot 4180 \cdot 15)\right) = 4.389\ GJ$. If we use arithmetic averaged values for all parameters, the average temperature in the pipe is 14.17°C, and the average flowrate is 25 kg/s. The energy input is then: $\dot{m} \cdot \Delta t \cdot Cp \cdot \left(T_{in} - T_{ref}\right) = (25 \cdot 3600 \cdot 4180 \cdot 14.17) = 5.331\ GJ$. If we use a flow-weighted average for the pipe temperature, then the pipe temperature becomes 11.67°C, and the energy input is: $\dot{m} \cdot \Delta t \cdot Cp \cdot \left(T_{in} - T_{ref}\right) = (25 \cdot 3600 \cdot 4180 \cdot 11.67) = 4.390\ GJ$, which is the same as the energy input measured with the 10' data. This clearly shows the interest of using flow-weighted average temperatures for the pipes.

The measured data has been pre-processed from the available 10' measurements into 60' measurements using different averages. Table 3.8.1 presents which averages have been used for which measurements. Figure 3.8.2 to Figure 3.8.7 also show the different energy balances obtained using only arithmetic averages and the ones obtained using the flow-weighted averages. For system energy balance, using flow-weighted averages increases the calculated total heat loss for every year.

Table 3.8.1: Measured data and pre-processing treatment.

Type of data	Arithmetic averaged data	Flow-weighted averaged data
Temperatures	32 PTES temperature measurements Ambient air temperature Temperature above/under the PTES lid cover	Temperatures of the top-mid-bottom PTES inlet/outlet pipes Temperatures of forward/return temperatures of the transmission line to the DHN Inlet/outlet temperatures of the secondary side of the solar heat exchangers
Flows	Flow through top-mid-bottom PTES inlet/outlet pipes Flow through secondary side of solar heat exchangers Flow through the transmission line to the DHN	-

3.8.2.2.2 Flow values

The flow measurements for the PTES inlet/outlet pipes are based on three principles:

- Middle pipe volume flowrate is measured
- Bottom and top pipes' volume flowrates are calculated (by the SCADA system) based on volume flow balance and knowledge of the other volume flowrates measured by the system (secondary side of the solar heat exchangers and transmission line to the DHN)
- Directions of the flows are determined by the SCADA system

In TR2.3, it was already mentioned that the SCADA-calculated bottom and top flows were flawed, as they were mostly based on volume flow balance and not mass balance. For some timesteps, even the volume flow balance isn't respected. The annual energy charged/discharged resulting in using the measured or SCADA-calculated flowrates for the inlet/outlet PTES pipes with their corresponding flow directions are presented in Figure 3.8.2 and Figure 3.8.3.

Using the PTES water temperature distribution sensors, the internal energy content of the PTES can also be calculated, and with knowledge of this, total heat losses can be calculated using the energy balance applied to the PTES:

$$Q_{in} - Q_{out} - Q_{losses,total} = \Delta E_{content}$$

Where:

- Q_{in} is the heat charged into the PTES
- Q_{out} is the absolute value of the energy discharged from the PTES
- $Q_{losses,total}$ is the absolute value of the total PTES heat losses
- $\Delta E_{content}$ is the PTES internal energy variation

Total heat losses are presented in Figure 3.8.5. The energy balance obtained with the provided flow measurements indicates a total PTES heat loss of about 1'050 MWh in 2017, 850 MWh in 2018, 1'300 MWh in 2019 and 1'600 MWh in 2020. The energy charged is in the range of 11'150 MWh for 2017, 14'000 MWh for 2018, 12'450 MWh for 2019 and 12'900 MWh for 2020.

When using the flowrates reported by the SCADA system, the calculated total losses do not follow the right trend between 2017 and 2019. In 2017 the PTES has been much less charged than the other years. In 2018, it has been much more charged. This means that the soil around the PTES should be heated up in 2018 and the bottom and side losses should be higher in 2018 than in 2017. The same increase should happen to the PTES lid losses, due to higher top PTES temperatures (related to the increased charge of the PTES). PTES losses should therefore increase between 2017 and 2018, and using the available measured PTES flows data, the opposite is observed.

Then from 2018 to 2019, the PTES is being less charged, which means that the soil losses should decrease, and the lid losses as well, since the PTES top temperature should be lower than the previous year. The total losses should therefore decrease, and the opposite is observed.

On this last point, it should however be noted that with time, the lid insulation properties gradually decrease, which means that lid losses should gradually increase with time. But the jump in total heat losses from 2018 to 2019 cannot be explained by this, since it is more significant than the increase in total heat losses observed

the following year (increase of 450 MWh total heat losses from 2018 to 2019, whereas the increase is of 300 MWh from 2019 to 2020). The total heat losses in 2020 should increase even more compared with 2019, since charge of the PTES in 2020 was higher than that of 2019.

These factors combined show that the available inlet/outlet PTES flow measurements from Dronninglund don't describe the physical behaviour of the PTES. Therefore, another reference data should be used for the inputs to the TRNSYS model. The error induced by the measurement comes most likely from the volume balance assumed by the SCADA system, together with flaws in the calculated flows for the top and bottom pipes of the PTES.

In the present study, mass flow balance was therefore assumed for the flows going into and out of the PTES and flow measurements for the top and bottom PTES inlet/outlet pipes were adapted based on measured middle mass flowrate. The procedure for calculating new flows based on available data is described below.

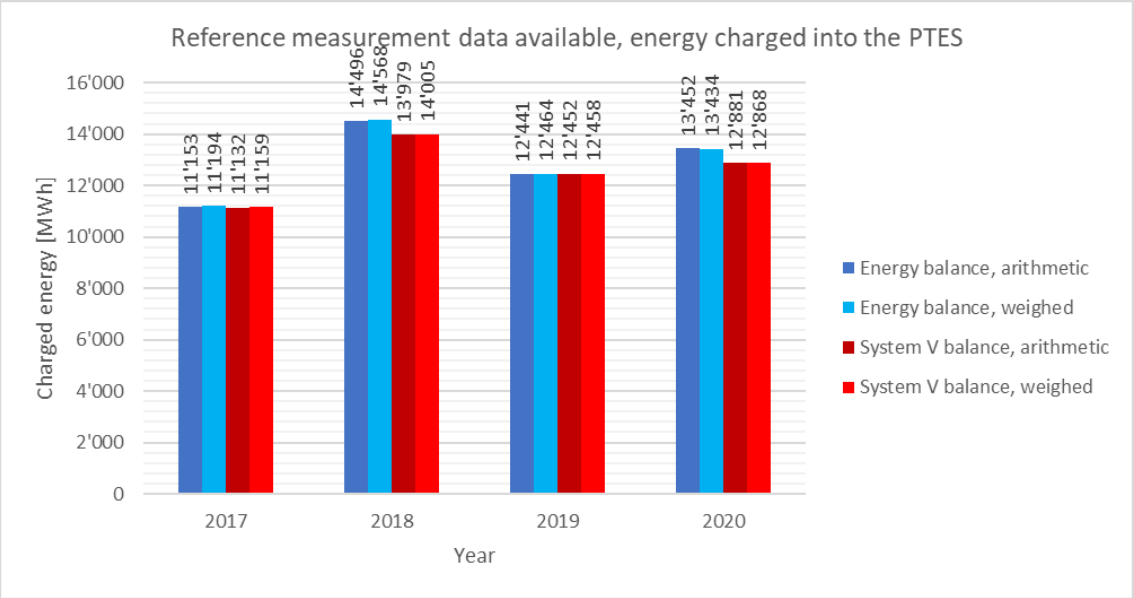


Figure 3.8.2: Yearly energy charged into the PTES, calculated based on measurements made available for the PTES ('System V balance') and for the rest of the solar heating system ('Energy balance'). Both results from using 60' arithmetic averages (darker colours) and 60' flow-weighted averages (lighter colours) are shown.

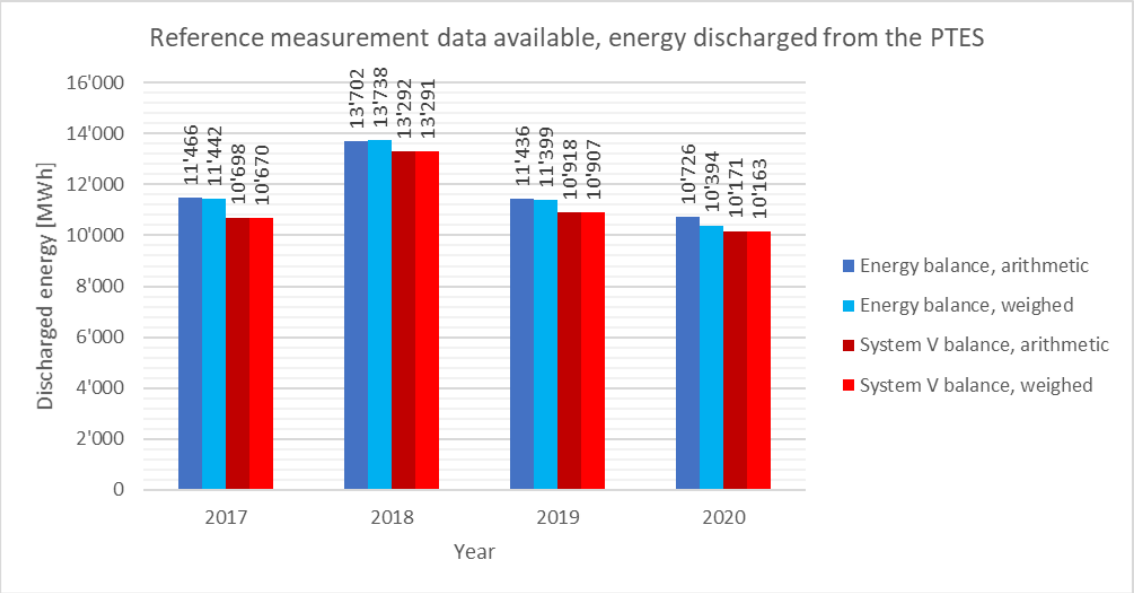


Figure 3.8.3: Yearly energy discharged from the PTES, calculated based on measurements made available for the PTES ('System V balance') and for the rest of the solar heating system ('Energy balance'). Both results from using 60' arithmetic averages (darker colours) and 60' flow-weighted averages (lighter colours) are shown.

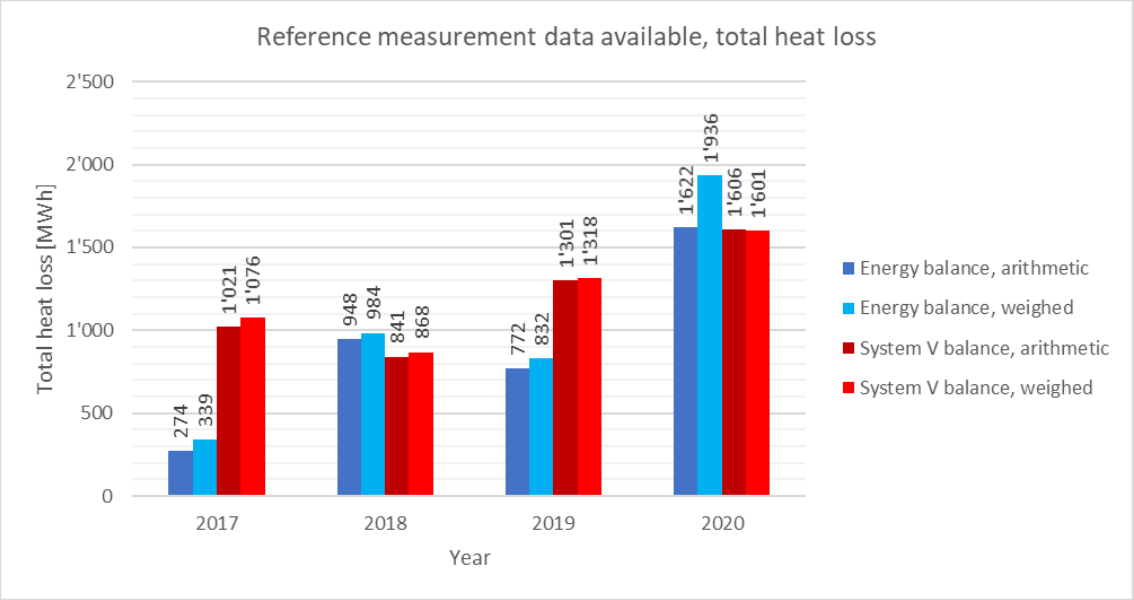


Figure 3.8.4: Calculated yearly total PTES heat losses, based on measurements made available for the PTES ('System V balance') and for the rest of the solar heating system ('Energy balance'). Both results from using 60' arithmetic averages (darker colours) and 60' flow-weighted averages (lighter colours) are shown.

First of all, 10' inputs are averaged into 60' data using the convention presented in Table 3.8.1. The flow directions determined by the SCADA system are used to set the direction of the flow for each pipe for each 10' timestep, and then the arithmetic average volume flowrate is calculated for each pipe. All volume flowrates are then converted into mass flowrates using temperature measurements closest to where the flows are measured/calculated.

Using knowledge of the energy balance outside the PTES (energy input or output from the solar collector field and the district heating network), it is possible for each timestep to estimate the energy input/output to/from the PTES. This provides the results presented in Figure 3.8.2, Figure 3.8.3 and Figure 3.8.4.

This calculation method based on system heat balance measurements, although flawed as well (it indicates a total heat loss of 339 MWh for 2017, which is impossible because it's less than half the heat losses calculated

for all previous and following years) can be used to determine the direction of the flows of each inlet/outlet pipes of the PTES, for each timestep. The method is the following:

- For each timestep, we use the absolute value of the mass flowrate obtained doing the 60' averages for the top and bottom pipes of the PTES, and the arithmetic value of the mass flowrate obtained doing the 60' average of the middle pipe of the PTES
- Then we assume the flow through each pipe is the result of a combination of inlet/outlets and mass flow balance for each timestep (see Table 3.8.2)

Since the middle flowrate is measured, we assume it's always correct in absolute value. The direction of the middle flow, however, is assumed to be either positive or negative. This makes a total of 8 possible flow combinations, based on measured and SCADA-determined flows (for lack of only measured flows, SCADA-determined flowrates for the top and bottom PTES inlet/outlet pipes are, in turn, used as a reference for the mass balance).

The system energy balance measurements provide the right trend in terms of total heat losses: they increase from 2017 to 2018, then decrease the next year, and increase again the year after. Although the absolute values vary significantly (from 339 MWh to 1'936 MWh for the flow-weighted measurements), the general trend seems reliable, and using these measurements as a reference to re-calculate the PTES pipe flows is therefore justified.

Once the 8 flow combinations have been calculated, the energy input/output (resulting from the calculated flowrates with the measured pipe temperatures) is calculated for each timestep and compared to the system energy balance measurements. The energy difference is used as a "score" for each combination. The combination that provides the smallest deviation compared with the reference is chosen as the new measured flowrates for this timestep. An "advantage" is given to the cases where the actual measured middle flowrate is being used (and not its opposite), as this measurement corresponds to the only PTES flow that is being measured and not calculated. The "advantage" given is that the corresponding score is divided by 10 (combinations numbers which are uneven in Table 3.8.2). This procedure provides the results presented in Figure 3.8.6 and Figure 3.8.7.

Table 3.8.2: Flow combination possibilities for the determination of the flow directions for the PTES inlet/outlet pipes.

Combination number	Top mass flowrate	Middle mass flowrate	Bottom mass flowrate
1	$-\dot{m}_{mid} - \lvert\dot{m}_{bot}\rvert$	$\dot{m}_{mid}$	$\lvert\dot{m}_{bot}\rvert$
2	$\dot{m}_{mid} + \lvert\dot{m}_{bot}\rvert$	$-\dot{m}_{mid}$	$-\lvert\dot{m}_{bot}\rvert$
3	$-\dot{m}_{mid} + \lvert\dot{m}_{bot}\rvert$	$\dot{m}_{mid}$	$-\lvert\dot{m}_{bot}\rvert$
4	$\dot{m}_{mid} - \lvert\dot{m}_{bot}\rvert$	$-\dot{m}_{mid}$	$\lvert\dot{m}_{bot}\rvert$
5	$\lvert\dot{m}_{top}\rvert$	$\dot{m}_{mid}$	$-\dot{m}_{mid} - \lvert\dot{m}_{top}\rvert$
6	$-\lvert\dot{m}_{top}\rvert$	$-\dot{m}_{mid}$	$\dot{m}_{mid} + \lvert\dot{m}_{top}\rvert$
7	$-\lvert\dot{m}_{top}\rvert$	$\dot{m}_{mid}$	$-\dot{m}_{mid} + \lvert\dot{m}_{top}\rvert$
8	$\lvert\dot{m}_{top}\rvert$	$-\dot{m}_{mid}$	$\dot{m}_{mid} - \lvert\dot{m}_{top}\rvert$

In Table 3.8.2, $\lvert\dot{m}_{top}\rvert$ and $\lvert\dot{m}_{bot}\rvert$ are respectively the absolute value of the top and bottom PTES pipes' measured mass flowrates, and $\dot{m}_{mid}$ is the arithmetic value (positive or negative) of the measured middle PTES pipe's mass flowrate.

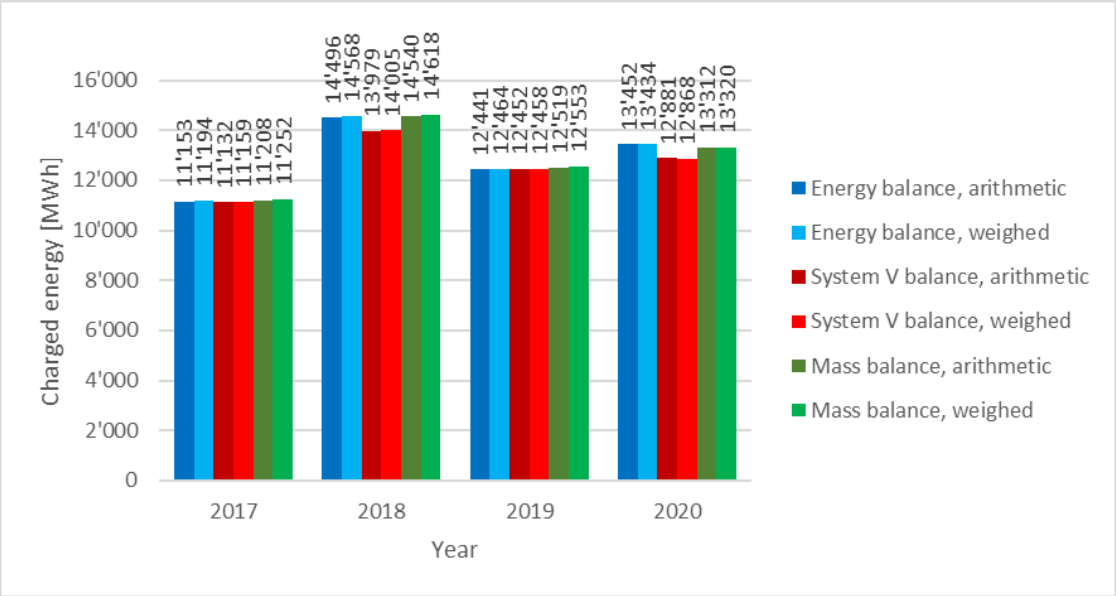


Figure 3.8.5: Yearly energy charged into the PTES, calculated based on measurements made available for the PTES ('System V balance'), measurements for the rest of the solar heating system ('Energy balance') and calculated using mass balance according to described procedure ('Mass balance'). Both results from using 60' arithmetic averages (darker colours) and 60' flow-weighted averages (lighter colours) are shown.

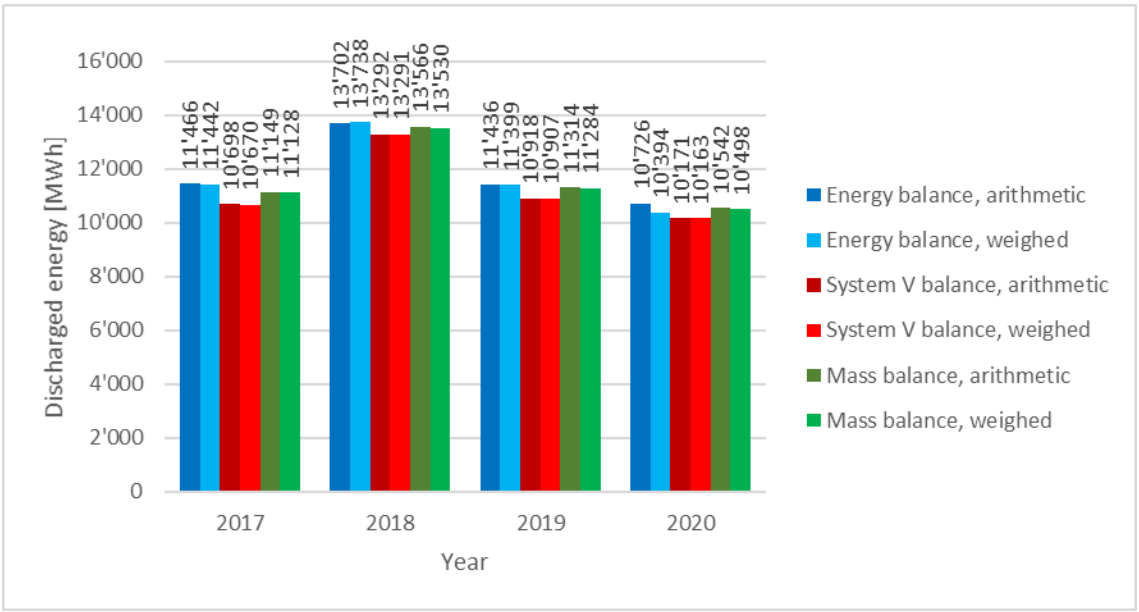


Figure 3.8.6: Yearly energy discharged from the PTES, calculated based on measurements made available for the PTES ('System V balance'), measurements for the rest of the solar heating system ('Energy balance') and calculated using mass balance according to described procedure ('Mass balance'). Both results from using 60' arithmetic averages (darker colours) and 60' flow-weighted averages (lighter colours) are shown.

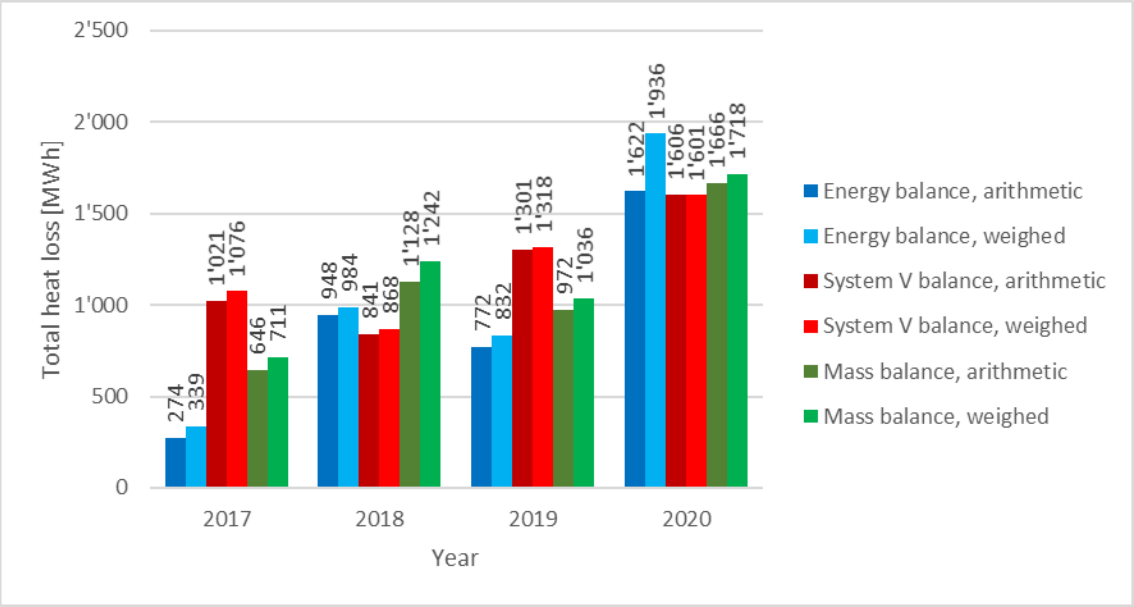


Figure 3.8.7: Calculated yearly total PTES heat losses, based on measurements made available for the PTES ('System V balance'), measurements for the rest of the solar heating system ('Energy balance') and calculated using mass balance according to described procedure ('Mass balance'). Both results from using 60' arithmetic averages (darker colours) and 60' flow-weighted averages (lighter colours) are shown.

Results of this data processing procedure provides satisfying results in terms of calculated charged and discharged energy (results are in the same order of magnitude as the reference system energy balance measurements). As for the calculated total heat losses, the expected trend is met between 2017 and 2020, and the results for 2017 are no longer unrealistic (711 MWh total yearly heat losses, compared with 339 MWh for the system energy balance measurements). This method also exhibits higher heat losses for 2020.

3.8.2.2.3 Methodology – root mean square error analysis

In TR2.3 the coefficient of determinations R^2 was used to evaluate model accuracy. In the present report, the Root Mean Square Error (RMSE) was preferred to R^2 to evaluate accuracy of the model. RMSE is defined as the following:

$$RMSE(y) = \sqrt{\frac{\sum_{i=1}^{n}(\hat{y}_i - y_i)^2}{n}}$$

Where:

- $\hat{y}_i$ is the measured value at timestep i of parameter y
- y_i is the calculated value at timestep i of parameter y
- n is the total number of measurement points used as a reference for the calculation

These RMSE coefficients will be calculated for several parameters, and will be divided by the standard deviation of the given parameter, to provide coefficient noted C:

$$C(y) = \frac{RMSE(y)}{\sigma(y)}$$

Where $\sigma(y)$ is the standard deviation of measured parameter y. These C coefficients provide an evaluation of the relative RMSE for the given parameter and can therefore be used as a comparison metric for different parameters.

Additionally, annual sum of different parameters will be compared, and the relative difference between the measured and the calculated annual sums will be evaluated as another way to evaluate model accuracy. The relative annual sum difference will be noted D:

$$D(y) = \frac{\left|\sum_{i=1}^{n} \hat{y}_i - \sum_{i=1}^{n} y_i\right|}{\left|\sum_{i=1}^{n} \hat{y}_i\right|}$$

The nomenclature is unchanged compared with the equation of RMSE. These coefficients will also be used to assess model prediction accuracy. It is important to mention already here that these coefficients are the results of sums, which will be calculated only for the years with valid reference data available (see explanation below).

3.8.2.2.4 Reference parameters to evaluate model prediction accuracy and cost function

The methodology presented in the previous section will be applied the following way:

- Top, middle, bottom and average PTES temperatures will be evaluated with C coefficients (and give an overview of the error, on average, of calculations compared with measurements at each timestep)
- Energy inlet or outlet for each timestep will also be evaluated with a C coefficient
- Total annual energy charged into the PTES, energy discharged from the PTES and total heat losses will be evaluated with D coefficients

Annual energy balance can be done by evaluating annual energy charged into the PTES, discharged from the PTES, and PTES internal energy variation. With these parameters, it is possible to evaluate the total annual heat losses using the following heat balance:

$$\Delta E_{internal\ energy} = Q_{charged} + Q_{discharged} + Q_{total\ heat\ losses}$$

Where:

- $Q_{charged}$ is the annual heat charged into the PTES, and is positive
- $Q_{discharged}$ is the annual heat discharged from the PTES, and is negative
- $Q_{total\ heat\ losses}$ is the total annual heat losses
- $\Delta E_{internal\ energy}$ is the PTES internal energy variation between the beginning and the end of each year

When the model runs, C and D coefficients are calculated and used to make a global cost function, which evaluates global prediction accuracy of the model. The cost function is the average of a global C coefficient and a global D coefficient, as follows:

$$C_{global} = \frac{C(T_{top,PTES}) + C(T_{mid,PTES}) + C(T_{bot,PTES}) + C(T_{avg,PTES}) + C(Q_{charged/discharged})}{5}$$

Where:

- $T_{top,PTES}$, $T_{mid,PTES}$ and $T_{bot,PTES}$ are respectively the PTES top, middle and bottom temperatures, which are at the same height as the top, middle and bottom diffusors in the TRNSYS model and in the actual PTES in Dronninglund
- $T_{avg,PTES}$ is the PTES average temperature, and is calculated from the PTES internal energy content: it is the temperature the PTES would have if it had a homogeneous temperature and the same energy content as the one measured for the PTES at every timestep
- $Q_{charged/discharged}$ is the heat either charged or discharged from the PTES at every timestep (can be positive or negative)

$$D_{global} = \frac{D(Q_{charged}) + D(Q_{discharged}) + D(Q_{total\ heat\ losses})}{3}$$

$$Cost\ Function = \frac{C_{global} + D_{global}}{2}$$

This cost function is minimized in the model calibration, where model parameters are varied and calibrated to obtain the best possible accuracy. The chosen cost function is a good compromise between:

- Obtaining a realistic behaviour of the thermal storage, through the PTES temperature distribution evaluation (with C coefficients)
- Obtaining good annual thermal energy balance, through the evaluation of D coefficients

The calibrated model will provide good annual heat balance (which is important for energy system calculations) as well as good temperature distribution inside the PTES (which is important to make sure the behaviour of the PTES is realistic and able to predict the actual physical behaviour of a PTES, when using different design assumptions).

3.8.2.2.5 Reference sensors for top, middle and bottom PTES temperatures

For the temperature distribution of the water inside the PTES, different temperature sensors can be used as a reference for the comparison between measurements and modelling. The temperature sensors are equally distributed inside the PTES, whereas the TRNSYS PTES model is using equal volume segments, which means that a given sensors doesn't exactly match a model segment (see Figure 3.8.23).

The sensor placed most at the top of the PTES is sensor number 1, and the one closest to the bottom is number 32. For the bottom PTES temperature, the last PTES sensor is being used as a reference, for the middle PTES temperature, both sensors number 11 and 12 are used. For the top PTES temperature, sensors 1 and 2 will be used. For top and middle PTES temperatures, 2 sensors are being used as a reference, as they are the sensors closest to the height of the top and middle diffusors. Having a combination of two sensors gives the opportunity to have a reference temperature that adapts to the model's specific segment distribution. The proportion in which each temperature sensor is used is determined by a coefficient (see following Section 3.8.3.1)

3.8.2.2.6 Reference years for the measurements

The aim of this study is to test further the TRNSYS component previously tested in TR2.3 against new measurements but also to use modelling to assess the accordance between calculations and measurements, and therefore also evaluate the PTES measurements.

Before 2017, the available measurements contain a lot of data errors: sensors that stopped working during long periods, impossible measurements, etc. (see Figure 3.8.8). This is especially true for the year 2014 where the PTES was charged for the first time, but also, in a less significative way for 2015 and 2016. Therefore it is preferable not to use those years of data as a reference for the comparison with modelling, and to use monitored years from 2017 as a reference instead. Figure 3.8.9 presents the monitored temperatures inside the PTES for 2017 to 2019. Monitored years before 2017, however imprecise, can be used to simulate the preheating of the soil around the storage, by charging and discharging the PTES during those years, using measured flows as inputs to the TRNSYS model.

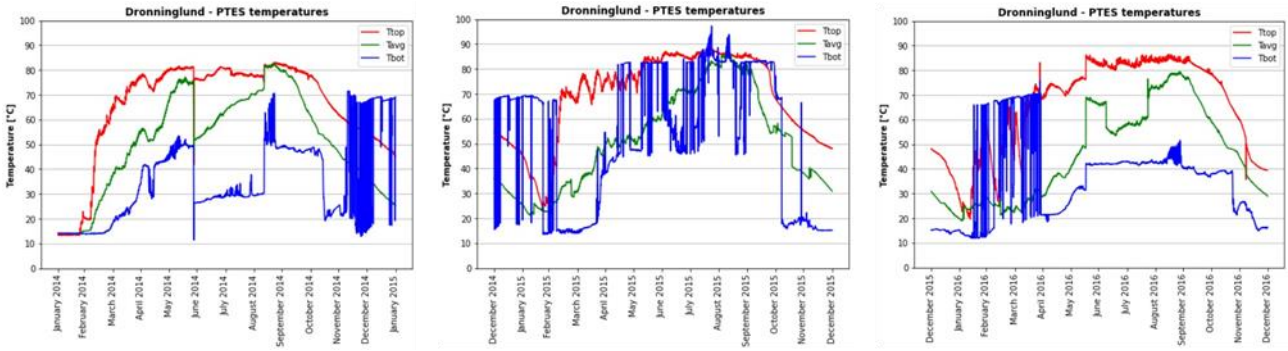


Figure 3.8.8: Top, average and bottom temperatures measured inside the PTES in Dronninglund for the years 2014 to 2016.

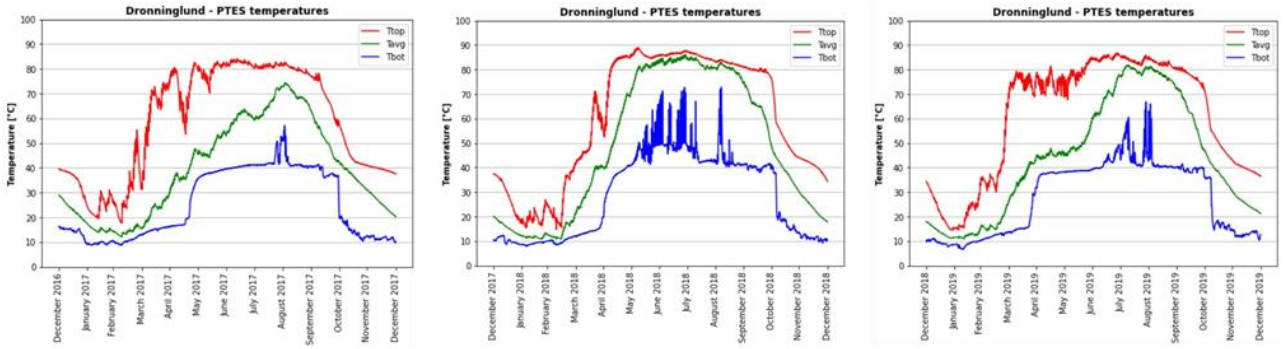


Figure 3.8.9: Top, average and bottom temperatures measured inside the PTES in Dronninglund for the years 2017 to 2019.

As mentioned in section 3.8.2.2.2, recalculated flow values provide satisfying results in terms of annual heat balance and calculated total heat losses. They follow the expected trend from 2017 to 2020. In absolute value, Figure 3.8.7 shows that no matter which measurements are used, the total heat losses for 2020 are significantly higher than any other year, while the heat charged into the PTES, isn't (see Figure 3.8.6). This indicates that the lid insulation likely was less effective during that year and was an early sign of lid degradation that led to a change of the lid in 2021[4]. Since these high total heat losses are unusual and related to a strong degradation of the lid, it has been chosen to exclude measurements from 2020 from the current model calibration. These extraordinary losses would have been complicated to account for in the model. Data from 2017, 2018 and 2019 will therefore be used as reference years for the measurements, and 2020 will not be used for the comparison.

Then, it was chosen to run model calibration using three modelling configurations:

- The first modelling configuration uses input data from 2014 to 2016 for the preheating of the soil, and input plus reference data from 2017. Total simulation time is four years
- The second modelling configuration uses input data from 2014 to 2016 for the preheating of the soil, and input plus reference data from 2017 to 2018. Total simulation time is five years
- The third modelling configuration uses input data from 2014 to 2016 for the preheating of the soil, and input plus reference data from 2017 to 2019. Total simulation time is 6 years

This way, it will be possible to compare the accuracy of model prediction, using reference measurements over one, two or three years, and extract some information about the evolution of the PTES during those years.

3.8.3 Validation

3.8.3.1 Calibration results

Table 3.8.3 presents the results of model calibration when using 2017, 2017 to 2018, and 2017 to 2019 as reference years for the simulation. For each simulation, the cost function value is given, together with the lid insulation conductivity corresponding to the top loss coefficient (assuming a homogeneous 0.24 m insulation thickness).

[4] https://www.aalborgcsp.com/news-events/newstitle/news/aalborg-csp-secures-new-contract-with-the-ptes-low-cost-energy-storage-technology/

Table 3.8.3: Calibration results using the TRNSYS PTES model with Dronninglund data.

Number of years simulated	Model name	T_{gi} [°C]	Top loss coefficient [kJ/hr/m²/K]	Side loss coefficient [kJ/hr/m²/K]	Bottom loss coefficient [kJ/hr/m²/K]	S1 coefficient	S11 coefficient	Cost function value	Calculated lid insulation conductivity [W/(m·K)]
4	1300_1_opt	1.5	0.7043	0.0138	54.5	0.5930	0.5437	0.01531	0.047
4	1300_1_ini	1.5	0.6	2.4	2.4	0.5930	0.5437	0.04091	0.04
5	1300_2_opt	2.4	0.8655	0.0128	90.0	0.6067	0.2543	0.03008	0.0577
5	1300_2_ini	2.4	0.6	2.4	2.4	0.6067	0.2543	0.09079	0.04
6	1300_3_opt	1.5	0.9042	0.0707	75.0	0.5689	0.2636	0.04128	0.0603
6	1300_3_ini	1.5	0.6	2.4	2.4	0.5689	0.2636	0.1043	0.04

Where:

- T_{gi} is the initial soil average temperature
- Top, side and bottom loss coefficients are the constant and uniform loss coefficients for the top, sides and bottom of the PTES used as a parameter for each simulation
- S1 and S11 coefficients are respectively the proportion of the temperature measured by PTES sensors number 1 and 11 used as a reference for top and middle temperature PTES model temperatures. The reference temperatures are defined by: $T_{ref,top} = S1 \cdot T_1 + (1 - S1) \cdot T_2$ and $T_{ref,mid} = S11 \cdot T_{11} + (1 - S11) \cdot T_{12}$

Where T_1 is the temperature measured by sensor number 1, T_2 by sensor number 2, etc.

For each simulation duration, the calibration was first carried out using the program GenOpt, which launches TRNSYS simulations automatically and tests different parameters until it reaches the lowest value for the cost function (refer to Section 3.8.2.2.4).

Then, after each calibration, another run was carried out using the same optimized initial soil temperature, S1 and S11 coefficients, but using different top, sides and bottom loss coefficients. For the top coefficient, the theoretical value of insulation conductivity is used. The side and bottom loss coefficients are assumed to be equal to four times the top loss coefficient. These extra runs are used to see the effect of calibration.

Figure 3.8.10 shows the detail of the C and D coefficients for each simulation.

Further detailed results from the calibration using four, five and six-year simulations are presented and analysed below.

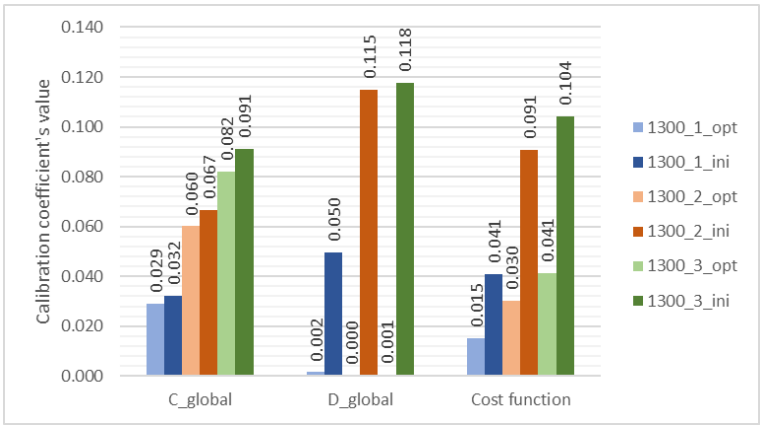


Figure 3.8.10: Cost function, C and D coefficients for each simulation using calibrated (light colours) and non-calibrated (dark colours) heat loss coefficients.

3.8.3.1.1 Data from 2017: simulation results description

For the calculations using reference data from the year 2017 (and input data from 2014 to 2017), the simulation time is 4 years, but the calibration (calculation of C and D coefficients) is made based on the sum over the last year only.

The calibrations results (see Table 3.8.3) show that the model top loss coefficient that provides the best agreement with measurements is higher than the theoretical top loss coefficient: 0.7043 kJ/(hr·m^2·K) instead of 0.6 kJ/(hr·m^2·K), which respectively correspond to thermal conductivities of 0.047 W/(m·K) and 0.04 W/(m·K). This is as expected since the insulating properties of the lid decrease with time.

Figure 3.8.11 shows measured and calculated top, middle and bottom water temperatures inside the PTES, as well as measured and calculated energy content (which is proportional to the PTES average temperature). Good accordance between calculations and measurements is found (the simulation captures the right trend for the energy content and the top, middle and bottom temperatures).

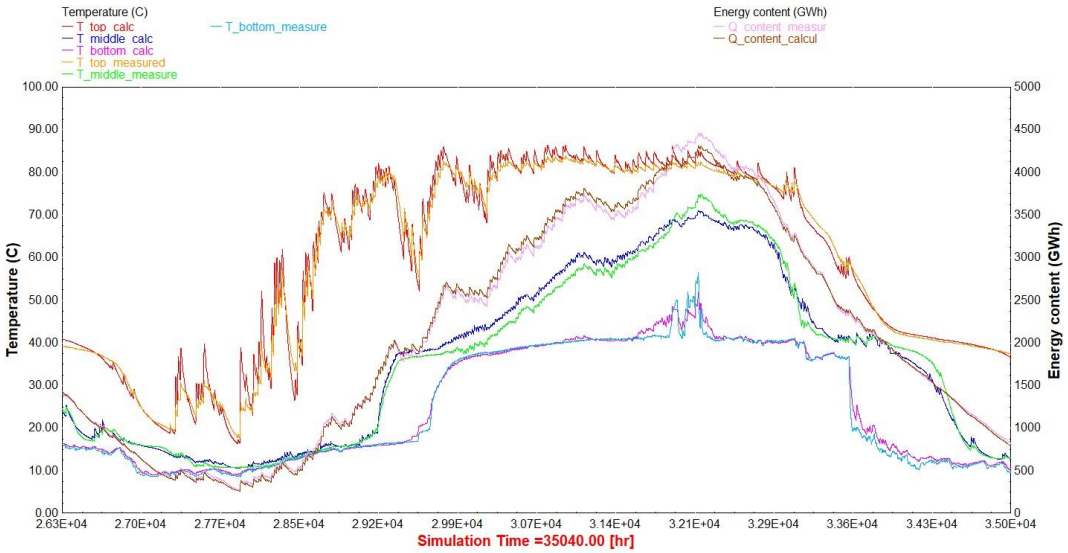


Figure 3.8.11: Measured and calculated hourly energy content, top, middle and bottom temperatures inside the PTES in 2017, using TRNSYS model parameters calibrated from reference year 2017.

Figure 3.8.10 provides further insight into the difference between the calibrated and non-calibrated model accuracies. The main difference is observed for D_{global} coefficients: for the non-calibrated model simulation, the D_{global} coefficient is 0.050, while it is 0.002 for the calibrated model simulation. The C_{global} for the calibrated model simulation is also slightly lower than the C_{global} for the non-calibrated model simulation. This means that the calibration improves mostly the yearly energy balance accuracy of the model. The same effect is observed for the calibration done with the two and three-year calibrations.

Figure 3.8.12 shows the comparison between calculated and measured charged and discharged heat for the calibrated and non-calibrated models. Figure 3.8.13 shows the comparison between calculated and measured total PTES yearly heat losses and internal energy variation. The main difference observed is that the calculated total yearly heat losses for the calibrated model are equal to the reference data total heat losses. The calibrated model is also closer to measurements for charged energy, discharged energy and internal energy variation.

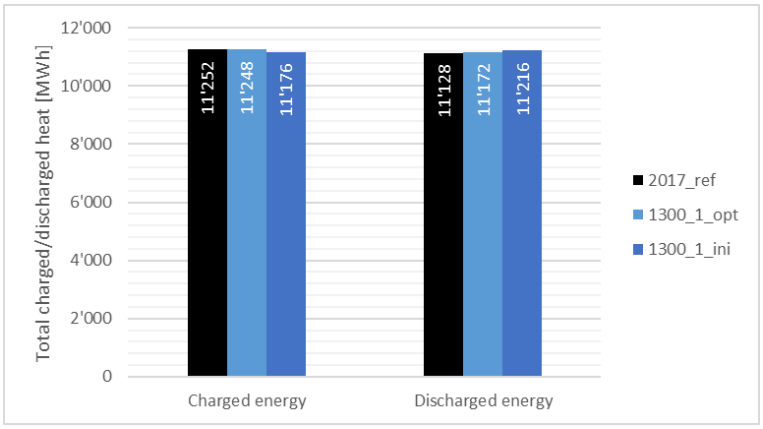


Figure 3.8.12: Measured and calculated annual charged and discharged energy into the PTES for the year 2017. Light and dark coloured bars are used respectively for the calibrated and non-calibrated model results. The black bars are representing the reference data.

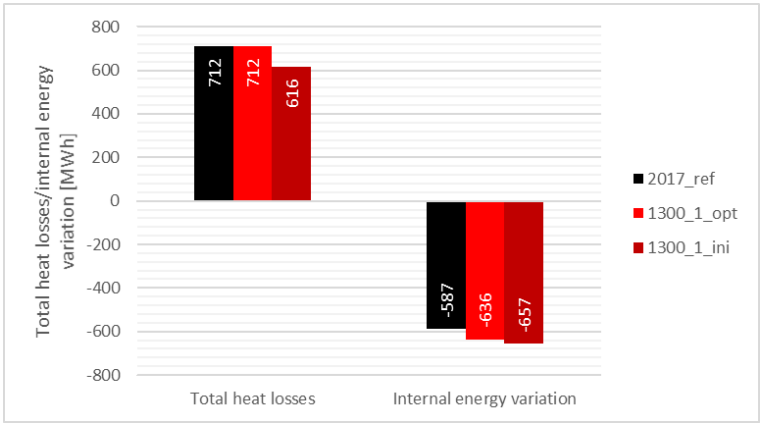


Figure 3.8.13: Measured and calculated PTES internal energy variation for the year 2017, and corresponding total annual heat losses. Light and dark coloured bars are used respectively for the calibrated and non-calibrated model results. The black bars are representing the reference data.

On all aspects, the calibrated version of the model is performing better than the non-calibrated version of the model, but the non-calibrated version isn't too far off compared with the reference either (less than 15% deviation for the yearly energy balance parameters, and an hourly PTES water temperature accuracy close to that of the calibrated model).

3.8.3.1.2 Data from 2017 to 2018: simulation results description

For the calculations using reference data from 2017 to 2018 (and input data from 2014 to 2018), the simulation time is 5 years, but the calibration is made based on the sum over the last two years only.

The calibrations results (see Table 3.8.3) show that the model top loss coefficient that provides the best agreement with measurements is higher than the previous calibrated top loss coefficient: 0.8655 kJ/(hr·m^2·K) instead of 0.7043 kJ/(hr·m^2·K), which respectively corresponds to thermal conductivities of 0.0577 W/(m·K) and 0.047 W/(m·K). This is again as expected since the insulating properties of the lid decrease with time.

Figure 3.8.14 shows measured and calculated top, middle and bottom water temperatures inside the PTES, as well as measured and calculated energy content. Good accordance between calculations and measurements is found (the simulation captures the right trend for the energy content and the top, middle and bottom temperatures).

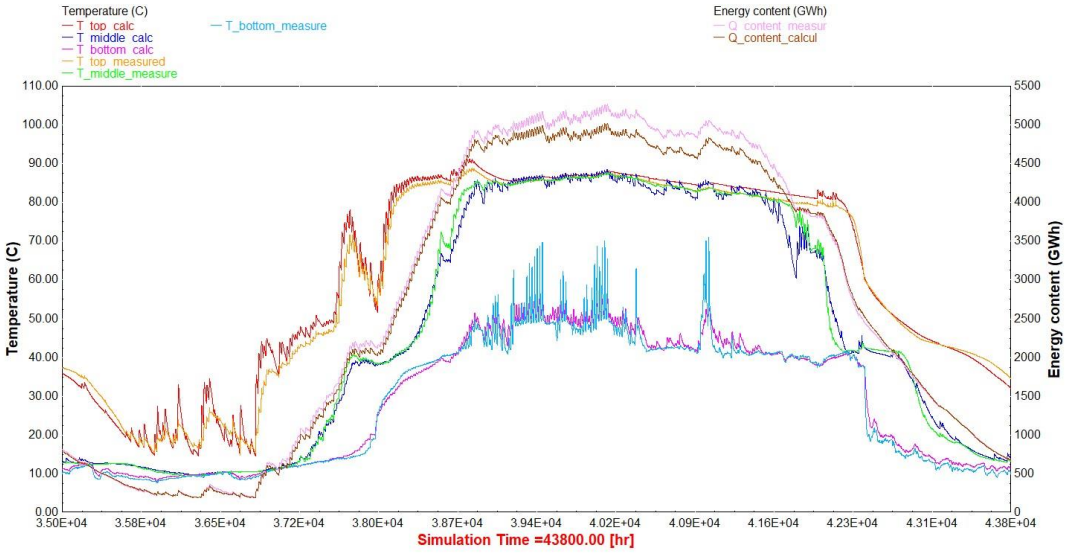


Figure 3.8.14: Measured and calculated hourly energy content, top, middle and bottom temperatures inside the PTES in 2018, using TRNSYS model parameters calibrated from reference years 2017 to 2018.

Figure 3.8.10 shows that the coefficient D_{global} is still close to 0 for the calibrated model with reference data from 2017 to 2018, while C_{global} has significantly increased both for the calibrated and the non-calibrated models compared with the models using reference data from only 2017. This is most likely because the top loss coefficient is fixed for the whole simulation. For 2017, the top loss coefficient providing the best accuracy was 0.7043 kJ/(hr·m^2·K), and is now 0.8655 kJ/(hr·m^2·K). This has a direct impact on the modelled temperatures inside the PTES for 2017 and how close they are to the measurements (especially the top temperature), hence the increase in C_{global}. The calibrated model's C_{global} coefficient remains lower than for the non-calibrated model, showing the improvement in hourly accuracy when making the calibration.

Figure 3.8.15 shows the comparison between calculated and measured charged and discharged heat for the calibrated and non-calibrated models. Figure 3.8.16 shows the comparison between calculated and measured total PTES heat losses and internal energy variation over the period. The main difference observed is that the calculated total heat losses for the calibrated model are equal to the measured total heat losses. The calibrated model is again also closer to measurements for charged energy, discharged energy and internal energy variation.

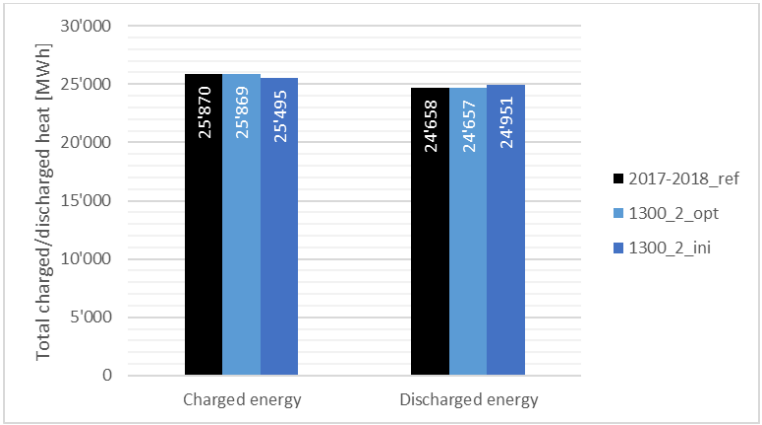


Figure 3.8.15: Measured and calculated total charged and discharged energy into the PTES for the years 2017 to 2018. Light and dark coloured bars are used respectively for the calibrated and non-calibrated model results. The black bars are representing the reference data.

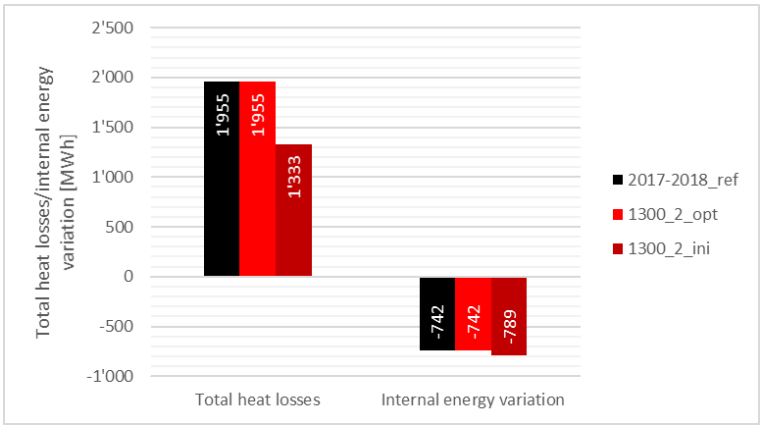


Figure 3.8.16: Measured and calculated PTES internal energy variation for the years 2017 to 2018, and corresponding total heat losses. Light and dark coloured bars are used respectively for the calibrated and non-calibrated model results. The black bars are representing the reference data.

On all aspects, the calibrated version of the model is performing better than the non-calibrated version of the model. The non-calibrated version is further off from the reference data for two-year comparison in terms of total calculated heat losses (over 30% difference) compared with the one-year comparison. This is as expected, as the non-calibrated model doesn't account for the degradation of the lid's insulating properties over time.

3.8.3.1.3 Data from 2017, 2018 and 2019: simulation results description

For the calculations using reference data from 2017, 2018 and 2019 (and input data from 2014 to 2019), the simulation time is 6 years, but the calibration is made based on the sum over the last three years only.

The calibrations results (see Table 3.8.3) show that the model top loss coefficient that provides the best agreement with measurements is higher than the previous calibrated top loss coefficient: 0.9042 kJ/(hr·m^2·K) instead of 0.8655 kJ/(hr·m^2·K), which respectively corresponds to thermal conductivities of 0.0603 W/(m·K) and 0.0577 W/(m·K). This is once again as expected, since the insulating properties of the lid decrease with time, but the increase is less important between 2018 and 2019 than between 2017 and 2018 calibrations.

Figure 3.8.17 shows measured and calculated top, middle and bottom water temperatures inside the PTES, as well as measured and calculated energy content. Good accordance between calculations and measurements is found (the simulations capture the right trend for the energy content and the top, middle and bottom temperatures).

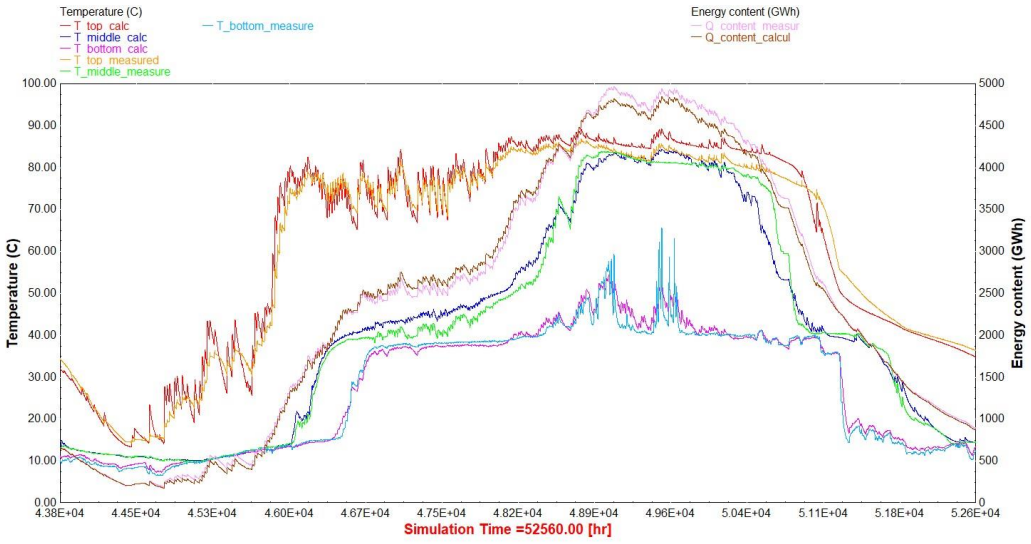


Figure 3.8.17: Measured and calculated hourly energy content, top, middle and bottom temperatures inside the PTES in 2019, using TRNSYS model parameters calibrated from reference years 2017 to 2019.

Figure 3.8.10 shows that the coefficient D_{global} is still close to 0 for the calibrated model with reference data from 2017 to 2019, while C_{global} has increased both for the calibrated and the non-calibrated models compared with the models using reference data from 2017 to 2018. This is most likely because the top loss coefficient is fixed for the whole simulation which makes hourly prediction for the previous years less accurate. For 2017 to 2018 reference data, the top loss coefficient providing the best accuracy was 0. 8655 kJ/(hr·m^2·K), and is now 0.9042 kJ/(hr·m^2·K). The calibrated model's C_{global} coefficient remains lower than the non-calibrated model's, showing the improvement in hourly accuracy when making the calibration.

Figure 3.8.18 shows the comparison between calculated and measured charged and discharged heat for the calibrated and non-calibrated models. Figure 3.8.19 shows the comparison between calculated and measured total PTES yearly heat losses and internal energy variation over the period. The main difference observed is that the calculated total heat losses for the calibrated model are equal to the measured total heat losses. The calibrated model is again also closer to measurements for charged energy, discharged energy and internal energy variation.

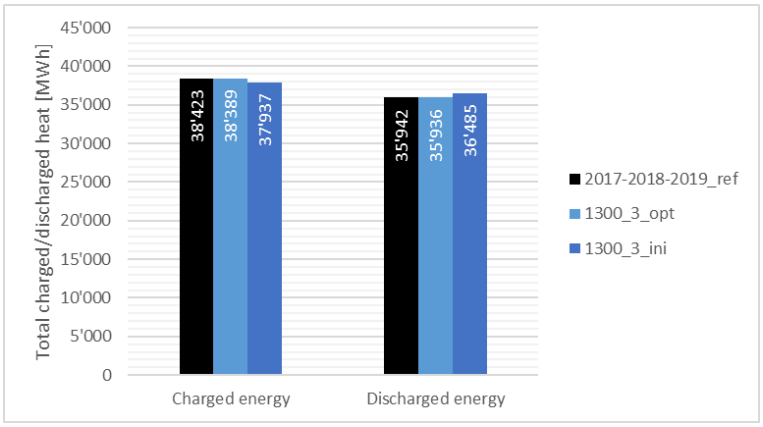


Figure 3.8.18: Measured and calculated total charged and discharged energy into the PTES for the years 2017 to 2019. Light and dark coloured bars are used respectively for the calibrated and non-calibrated model results. The black bars are representing the reference data.

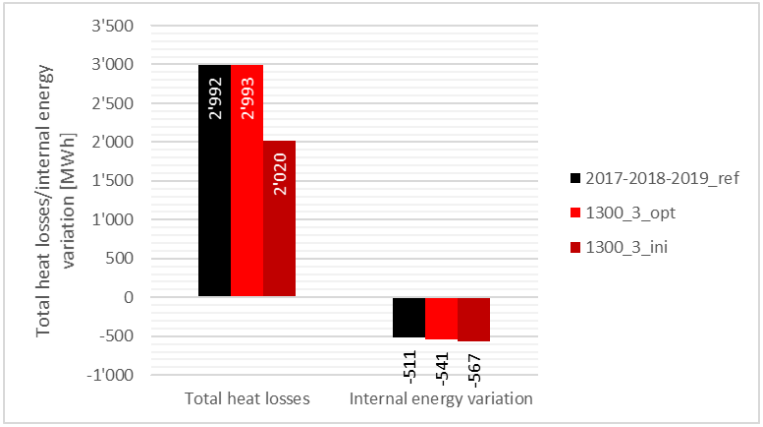


Figure 3.8.19: Measured and calculated PTES internal energy variation for the years 2017 to 2019, and corresponding total heat losses. Light and dark coloured bars are used respectively for the calibrated and non-calibrated model results. The black bars are representing the reference data.

On all aspects, the calibrated version of the model is performing better than the non-calibrated version of the model. The non-calibrated version is similarly off from the reference data for three-year comparison in terms of total calculated heat losses (over 30% difference) compared with the two-year comparison.

Some more information can be obtained from the calibration model obtained using reference data from 2017 to 2019 (see parameters of model "1300_3_opt" from Table 3.8.3). The corresponding calibration is the model setup that provides the best overall energy balance for the PTES, with the best possible hourly accuracy over the period. It behaves on average over the three years as close to the actual PTES as possible.

Figure 3.8.20 shows the comparison between calibrated model results and measurements for the yearly energy charged and discharged from the PTES. For every year, a good accordance is found.

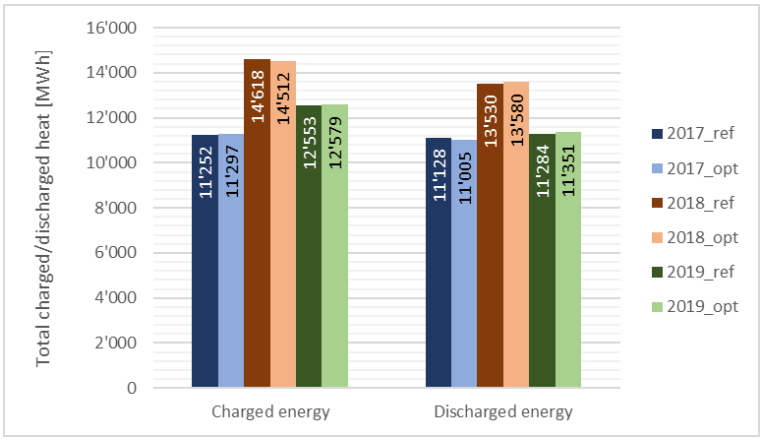


Figure 3.8.20: Measured and calculated yearly charged and discharged energy into the PTES for the years 2017 to 2019. Dark coloured bars are used for the reference data and light colours are used for the calibrated model results. Blue bars correspond to 2017, brown to 2018 and green to 2019.

Figure 3.8.21 shows the comparison between calibrated model results and measurements for the yearly total heat losses and PTES internal energy variation. The trend from 2017 to 2019 (increase, then decrease of total heat losses) is still observed with the modelled total heat losses, although the absolute values for the total heat losses are diverging. The maximum deviation for yearly total heat losses is +205 MWh, for 2017, and can be explained by the fixed top loss coefficient of 0.9042 kJ/(hr·m^2·K) used in model "1300_3_opt", which is higher than the calibrated top loss coefficient of 0.7043 kJ/(hr·m^2·K) used in model "1300_1_opt" (see Table 3.8.3).

For 2018, the total heat losses deviation is of -184 MWh. The model calculations give us access to the heat losses contribution (see Figure 3.8.22): top losses gradually increase from 802 MWh to 841 MWh between 2017 and 2019, while side and bottom losses increase from 2017 to 2018 then decrease from 2018 to 2019. This is as expected since in 2018 the PTES has been more charged than during the other years. Although the trend is following the expected behaviour, the absolute value of the bottom and sides losses is most likely wrong if we consider the following factors:

- The model calibration with reference data from 2017 to 2019 ended with very low side losses coefficient of 0.0707 kJ/(hr·m^2·K) and very high bottom losses coefficient of 75.0 kJ/(hr·m^2·K) (the possible reasons for such values are discussed in the conclusion)
- The PTES Geometry is such that the bottom surface area is much smaller than the sides surface area
- The stratification inside the PTES also is such that the water in contact with the sides of the PTES is warmer that the water in contact with the bottom

This means that the heat losses to the sides of the PTES (which should be the biggest contribution to the bottom and sides heat losses) are way underestimated, and the bottom losses on the other side (which should be low because of the low water temperature and smaller surface area) are overestimated. The bottom losses can most likely not compensate for the underestimated sides heat losses, which means that the overall bottom and sides heat losses (and heat losses variation) between 2017 and 2019 are most likely underestimated. This can explain why total heat losses are underestimated for 2018.

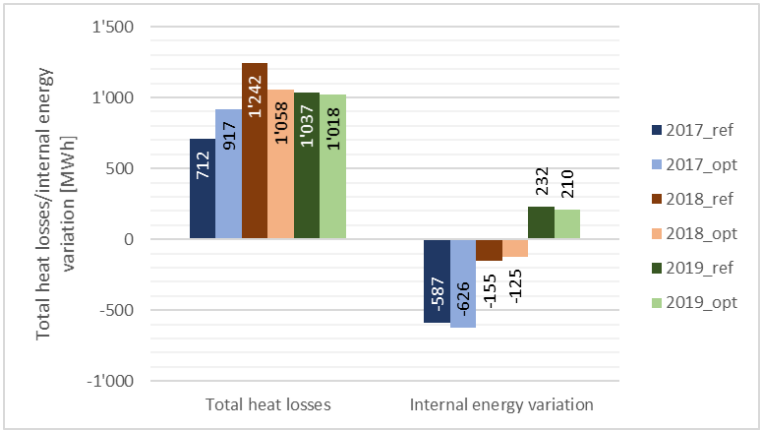


Figure 3.8.21: Measured and calculated yearly total heat losses and PTES internal energy change for the years 2017 to 2019. Dark coloured bars are used for the reference data and light colours are used for the calibrated model results. Blue bars correspond to 2017, brown to 2018 and green to 2019. The TRNSYS model used for the calculations is the calibrated model 1300_3_opt.

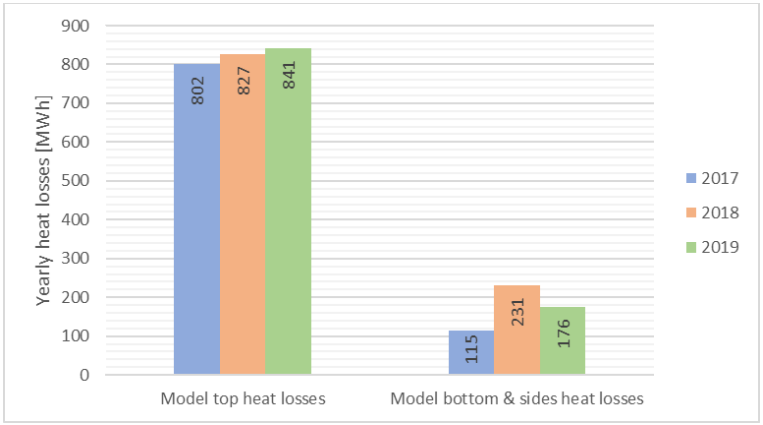


Figure 3.8.22: Calculated top and bottom+sides heat losses for the years 2017 to 2019, using calibrated model 1300_3_opt. Blue bars correspond to 2017, brown to 2018 and green to 2019.

3.8.3.2 Results analysis and conclusions

The calibration parameters providing the lowest value of the cost function offer some interesting insights on the modelling but also on the actual PTES in Dronninglund:

- The calibrated top loss coefficient increases between 2017 and 2019 from 0.7043 kJ/(hr·m^2·K) to 0.9042 kJ/(hr·m^2·K). This translates the increase of lid insulation conductivity at Dronninglund over time (deterioration of the lid's insulating properties)
- Once calibrated, the model captures accurately the PTES temperature distribution as well as the yearly heat balance
- The calibrated initial soil temperature (1.5°C for 2017, 2.4°C for 2017 to 2018, 1.5°C for 2017 to 2019) seems to be coherent with the fact that the PTES was put in operations during early 2014 (when the ambient air temperature was low, and therefore the temperature of the surrounding soil as well)
- The calibrated models give us access to the evolution of the PTES lid losses, while the available measurements did not provide a reliable value for those (789 MWh for 2017, 737 MWh for 2018 and 704 MWh in 2019 according to the heat flux sensor placed on the lid)
- The calibrated models give us access to an estimate of the distribution of heat losses between lid and other (sides and bottom) heat losses (see Figure 3.8.22)

The calibration parameter results provide however some unrealistic results regarding the side and bottom losses coefficients. Side loss coefficients are relatively low, and bottom loss coefficients are relatively high. This could be due to several factors:

- The presence of groundwater beneath the bottom of the PTES could explain the high value of bottom loss coefficient

- The fact that the side and bottom losses are significantly smaller than the lid heat losses (in the calculations with model 1300_3_opt, they represent 21.9% of the total heat losses for year 2018). A variation in sides and bottom heat losses therefore has a limited impact on the total heat losses and are harder to calibrate

Moreover, the lack of reliable data regarding lid, sides and bottom losses, as well as soil temperature measurements, makes it impossible to determine which of those factors has the most influence on the results. This lack of data also makes it more complex to properly calibrate the side and bottom loss coefficients.

In TR2.3, model "1300_3" was calibrated with 60' averaged data from 2017, using TRNSYS components Type 1300 & Type 13001 and provided the best global accuracy. The calibrated parameters are shown in Table 3.8.4. These results seem uncoherent compared to the results of the present study but can be explained. They are partly calibrated on lid losses which were underestimated because they had been calculated from a theoretical lid conductivity, lid thickness and on measured temperatures above and under the lid, which are unreliable. The present study provides more satisfying results regarding the analysis of the lid losses, without using estimated lid losses as a reference for calibration, which is an interesting result.

Table 3.8.4: Calibration results from TR2.3 using the TRNSYS PTES Types 1300 and 1301 with Dronninglund data for 2017 exclusively.

Number of years simulated	Model name	T_{gi} [°C]	Top loss coefficient [kJ/hr/m²/K]	Side loss coefficient [kJ/hr/m²/K]	Bottom loss coefficient [kJ/hr/m²/K]	Total heat losses [MWh]	Internal energy change [MWh]
1	1300_3	23.5	0.47	3.7	16.8	908	-692

Therefore, we can conclude that the developed method:

- For pre-treatment of input data
- For calibration of the model based on a heat balance for inlet energy, outlet energy, and internal energy variation to calculate total heat losses
- For calibration of the PTES over several years of reference and input data

Yields interesting results about Dronninglund's PTES lid degradation over time and can be used as a tool to follow up on the evolution of its insulating properties. This information is precious as most of the heat lost by the PTES happens through the lid and is crucial to properly model system integration using this component.

Further work should be done to properly calibrate the model regarding sides and bottom heat losses, but to do so, further study should be done. Such work would be facilitated by the use of extra data regarding soil temperatures and/or heat losses to the sides and bottom of the PTES, which are unfortunately unavailable.

Dronninglund being a full-scale pilot plant, proper validation (as defined in Figure 2.1.1) could not be carried out because the PTES is used for operations and not for lab testing. A subset of the data could be used to validate the model during a given period (charging and/or discharging of the PTES), but this was not carried out for the present work, as this would pose some other issues (related to initialisation of ground temperatures, amongst other things) and due to a lack of time.

One last point is that since the insulation conductivity has been increasing with time in Dronninglund, while the TRNSYS PTES component Type 1300 and Type 1301 uses the lid loss coefficient as a fixed parameter for the duration of the simulation, it is impossible to properly calibrate the model over a long period of time. A partial solution for that is to make the calibration of the model over shorter periods of time. The fixed coefficient is the same for the entire simulation, also for the period where there is no comparison (soil and storage preheating), but then for the comparison period, the losses will be better estimated.

This solution has been implemented for 2018, 2019 and 2020: instead of using several years of reference data for the different years, only the last year was used as a reference for the calibration. This approach is similar to what was done for 2017 with model 1300_1 and provides the calibration results shown in Table 3.8.5.

Table 3.8.5: Calibration results using the TRNSYS PTES model with Dronninglund data, last-year comparison.

Number of years simulated	Model name	T_{gi} [°C]	Top loss coefficient [kJ/hr/m²/K]	Side loss coefficient [kJ/hr/m²/K]	Bottom loss coefficient [kJ/hr/m²/K]	S1 coefficient	S11 coefficient	Cost function value	Calculated lid insulation conductivity [W/(m·K)]
4	1300_1_opt	1.5	0.7043	0.0138	54.5	0.5930	0.5437	0.01531	0.0470
5	1300_4_opt[5]	1.5	0.960	4.48	22.9	0.713	0.060	0.02712	0.0640
6	1300_5_opt[5]	1.5	0.912	1.76	46.0	0.437	0.247	0.01630	0.0608
7	1300_6_opt[5]	1.5	1.288	5.42	34.6	0.658	0.000	0.02585	0.0859

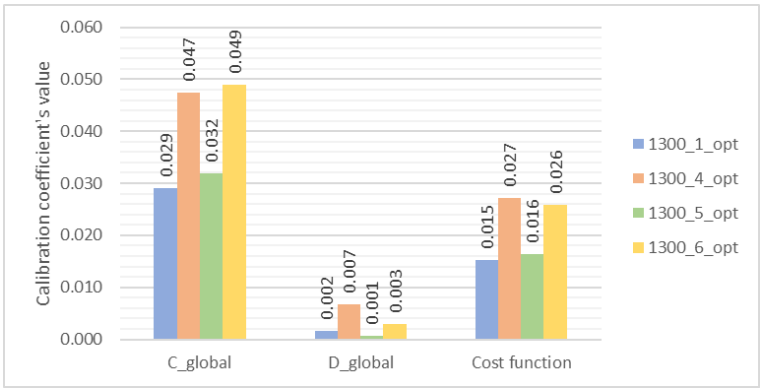


Figure 3.8.23: Cost function, C and D coefficients for the different last-year calibrated model simulations.

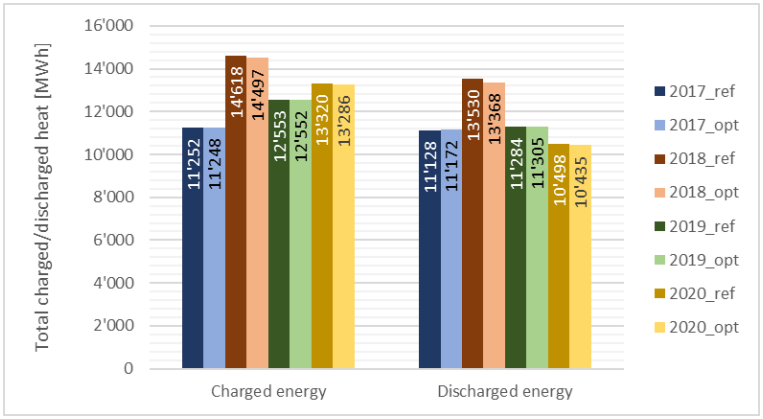


Figure 3.8.24: Measured and calculated yearly charged and discharged energy into the PTES for the years 2017 to 2020. Dark coloured bars are used for the reference data and light colours are used for the calibrated model results. Blue bars correspond to 2017, brown to 2018, green to 2019 and yellow to 2020.

[5] Models 1300_4 to 1300_6 are the ones using reference data for respectively 2018, 2019 and 2020 (one-year comparison only), as opposed to models 1300_2 and 1300_3 which were using 2017 to 2018 and 2017 to 2019 respectively as a reference (two-year and three-year comparison).

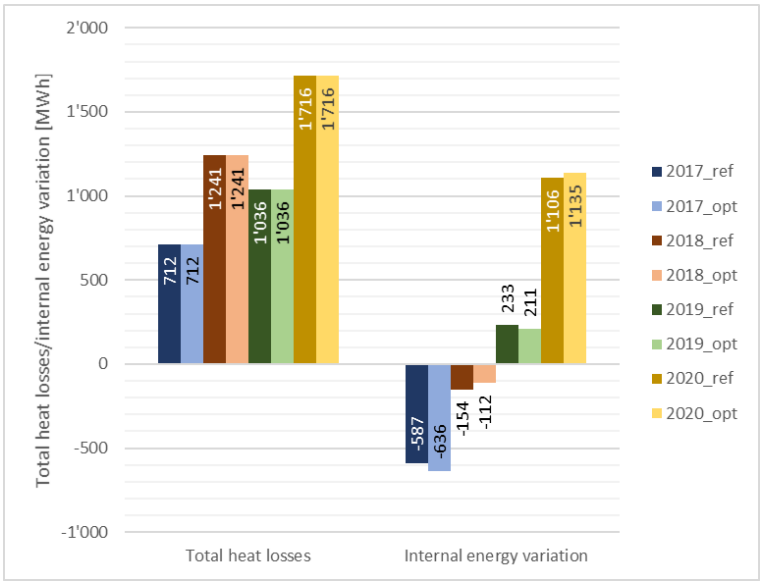


Figure 3.8.25: Measured and calculated yearly total heat losses and PTES internal energy change for the years 2017 to 2020. Dark coloured bars are used for the reference data and light colours are used for the calibrated model results. Blue bars correspond to 2017, brown to 2018, green to 2019 and yellow to 2020. The TRNSYS models used for the calculations are the calibrated models 1300_1_opt, and 1300_4_opt to 1300_6_opt.

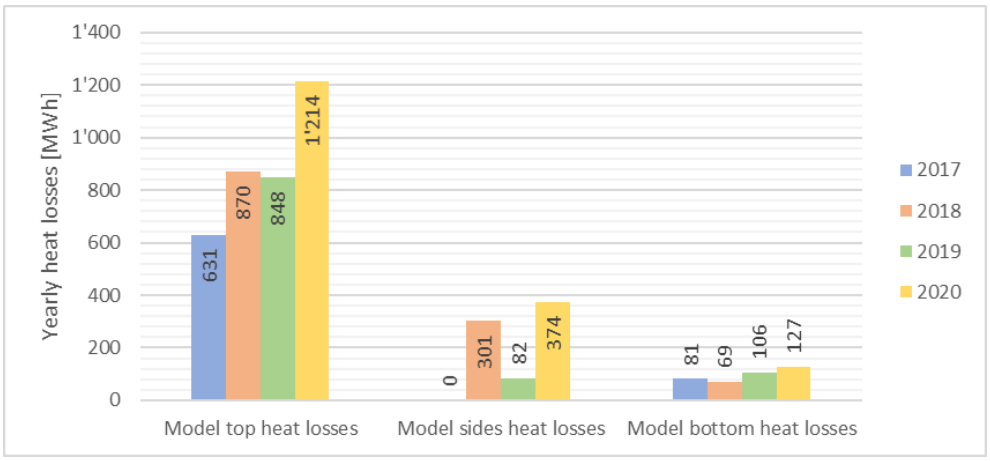


Figure 3.8.26: Calculated top, sides and bottom heat losses for the years 2017 to 2020, using calibrated model 1300_1_opt, and 1300_4_opt to 1300_6_opt. The colour code is the same as for the previous figures.

The results obtained with the last-year comparison calibration method no longer present unrealistic values for the calibrated sides and bottom loss coefficients of 2018 and 2019. Table 3.8.5 presents the calibrated coefficients, and Figure 3.8.26 shows the distribution of the calculated heat losses for the last year of simulation using the last-year calibrated models. There is a clear correlation between the value of the sides heat loss coefficient and the calculated yearly heat losses through the sides of the PTES. This of course was to be expected, but it's also the parameters which provide the best global accuracy of the model (both hourly and yearly accuracy). For all years, the bottom loss coefficient is rather high (above 22 kJ/hr/m^2/K), and the yearly calculated yearly heat losses through the bottom of the PTES are rather stable around 100 MWh ± 30 MWh. The calculated total heat losses correspond perfectly to the measured total heat losses (see Figure 3.8.25) for all years.

The calibrated top loss coefficient (see Table 3.8.5) gradually increases with time, which is as expected with the progressive degradation of the lid. Figure 3.8.23 shows the evolution of the C and D coefficients for the different simulations. As in the first calibration (using several years of reference data), the yearly coefficient D_{global} is close to 0 for all calibrated models, but this time the coefficient C_{global} doesn't gradually increase (compared with the results presented in Figure 3.8.10) for the different comparison years. It increases from 2017 to 2018, then decreases between 2018 and 2019, and increases again.

This is due to the change of lid insulation conductivity from one year to the other. Between 2017 and 2018, the lid insulation conductivity increased (as pointed out by both calibration methods), which means that model 1300_4, which uses a high top loss coefficient for the entire simulation, also uses a high top loss coefficient when simulating the year 2017 (which doesn't correspond to the actual situation). This means that the year preceding 2018 in the simulation for model 1300_4 isn't representative of the behaviour of the storage during that year, and therefore the storage soil and water preheating is off. This results in a less accurate behaviour of the model in 2018, and thus in a higher C_{global} coefficient. This effect is also observed between 2019 and 2020, where the lid loss coefficient increases further and the C_{global} coefficient too. What has just been described for coefficients C_{global} is also observed (but to a lesser extent) with coefficients D_{global}.

This interpretation of the results is further confirmed by the fact that when the lid insulation conductivity is similar from one year to the other (e.g. between 2018 and 2019, but also most likely between 2016 and 2017), the calibrated model performs better in terms of hourly accuracy: lower coefficients C_{global} are obtained for those years (see Figure 3.8.23).

The current extra set of model calibration (using only the last year as a refence for the comparison) has validated the approach of making a calibration of the model over shorter periods of time, but has also confirmed the inconvenient of the fixed loss coefficient values used in TRNSYS Types 1300 and 1301. The monitoring data and the calibrations made with the data (after post-processing) provides a good overview of the evolution of the lid loss coefficient with time, but is limited by the fact that the TRNSYS model uses a constant lid loss coefficient (which isn't the case in real life).

If one calibrated model should be selected for validation, it would be the one calibrated using data from 2019 (and input data for preheating from 2014 to 2018), as it was obtained for a period where the lid losses were rather stable, and the side-bottom losses were not negligible. Calibrated model 1300_1 is influenced by the extraordinarily low sides-bottom losses observed in 2017, which can explain the obtained very low sides losses coefficient, and high bottom losses coefficient. But as mentioned previously, validation is more complex to make for this specific case, as the PTES in Dronninglund is a full-scale operating pilot plant, and not a test plant.

4 Discussion and conclusions

Section 0 presented nine examples of model validation of subsurface heat storage dynamics in the content of the HEATSTORE case studies. Each case study provided (1) a short description of the site and its subsurface and/or geology; (2) the type(s) of models (e.g., hydrothermal, geomechanical, hydrogeochemical), model codes and data (e.g., temperature, chemical data, groundwater head) used for model validation; as well as (3) a description of the model validation approach and a self-assessment of the extent to which the model can be considered validated. Table 4.1 summarizes the validation tests and experiments performed for each one of the subsurface heat storage dynamics models used in the HEATSTORE case studies presented in section 0 and in relation to the definitions and validation methods presented in section 0.

Table 4.1: Validation of subsurface dynamics models used for the HEATSTORE case studies.

Sect.	Case study	Type of UTES (1)	Model(s) and model code(s) used	Validation approach(es) and parameter(s)	Model / Validation type (2)	Predictive capacity (3)
[3.1]	Koppert-Cress, Monster, the Netherlands	HT-ATES	TH (SEAWAT)	Comparison to operating and monitoring data (temperature distribution, flow distribution, hydraulic conductivity)	Strong model / Strong validation	unspec.
[3.2]	ECW Energy, Middenmeer, the Netherlands	HT-ATES	TH (HST3D and DC3D) THC (TOUGHREACT)	Comparison to monitoring data (temperature) Comparison to monitoring data (geochemical data) – *planned*	Strong model / Strong validation	High Unknown
[3.3]	Geneve, Switzerland	HT-ATES	TH and THM models (unspec.)	Comparison to monitoring data (temperature, energy/exergy efficiency)	Strong model / Strong validation	unspec.
[3.4]	Forsthaus, Bern, Switzerland	HT-ATES	THC (PFLOTRAN)	Comparison to other models (MARTHE-PHREEQC-BRGM and RETRASO-UPC) Model-fitting to monitoring data ('ad hoc' modifications) – *planned*	Strong model / Weak validation	n/a
[3.5]	Reykir/Hengill, Iceland	HT-ATES	TH (TOUGH2)	Calibration of TH model, but no model validation step	Strong model / Weak validation	n/a
[3.6]	BTESmart Vallin fier, Annecy, France	BTES	TH (Matlab analytical model) TH (FEFLOW numerical model)	Comparison of analytical model to numerical model Comparison to monitoring data (flow, temperature)	Weak model / Weak validation Weak model / Weak validation	unspec. unspec.
[3.7]	IEG colliery, Bochum, Germany	HT-MTES	TH (SPRING)	Comparison of short-term temperature distribution of numerical simulations from calibrated model to field data during a tailored injection scenario	Strong model / Strong validation	High

[3.8]	Dronninglund, Denmark	HT-PTES	TH (TRNSYS)	Calibration of TH model, but no model validation step	Strong model / Weak validation	Unknown

Notes: (1) as defined in Kallesøe and Vangkilde-Pedersen (2019); (2) as per Figure 2.1.1; (3) as evaluated by the modeler(s) under specific operating/monitoring conditions.
Abbreviations: ATES, aquifer thermal energy storage; BTES, borehole thermal energy storage; HT, high-temperature; MTES, mine thermal energy storage; PTES, pit thermal energy storage; TH, thermal-hydrological; THC, thermal-hydrological-chemical; THM, thermal-hydrological-mechanical; UTES, underground thermal energy storage.

Table 4.1 shows a broad range of approaches to model validation of subsurface dynamics across the HEATSTORE case studies. Most case studies performed a model calibration step using field or literature data followed by a model validation step. In most cases the validation step consisted of a comparison of numerical simulations to field data from a tailored experiment or from monitoring and operating data. In a few cases, however, no specific experiment was performed for the purpose of model validation so the predictive capacity could not be assessed. In other cases, despite performing a validation step, an assessment of the predictive capacity of the model(s) was not provided by the modelers. Finally, one case performed a numerical model validation through comparison with other model codes and is planning to use monitoring data to update the model ('ad hoc' modifications). This approach however is not a recommended validation strategy because it requires to perform further validation tests of the "calibrated-modified" model using an additional independent set of field data (see section 0). Only four case studies followed a 'strong' model validation of a field-scale model (cases 3.1 to 3.3 and 3.7).

The report showed different validation approaches of numerical models used to simulate UTES systems. The level of details and depth of the validation process of subsurface dynamics models depend on the type of UTES, the type of data available, and the type of physical and chemical processes under study. In general, robust calibration and validation of models can be obtained by comparison of numerical simulations with field data obtained from tailored experiments. A rigorous 'strong' validation process requires new field data collected during the monitoring of the UTES system operation, thus *after* the project has started its commercial operation. However, in practice, given the actual performance of UTES systems can be known at a much shorter time scale (approximately one year) than with other geoengineering systems (e.g., deep geological repositories for radioactive waste), it does not seem necessary to perform a full (strong) model validation process in UTES applications. In that case, information about the partial plant operation can be obtained so they can be used to decide whether to proceed with the full deployment of the UTES project.

5 References

Alt-Epping, P., Mindel, J.E. (Eds.), 2020. HEATSTORE – Benchmarking and improving models of subsurface heat storage dynamics. GEOTHERMICA – ERA NET Cofund Geothermal, 104 pp.

Anderson, M. P., 2005. Heat as a Ground Water Tracer. *Ground water, 43*(6), 951-968. doi:10.1111/j.1745-6584.2005.00052.x

Anderson, M.P., Woessner, W.W., 1992. The role of the postaudit in model validation. *Advances in Water Resources, Validation of Geo-hydrological Models 15*, 167–173. doi:10.1016/0309-1708(92)90021-S

Árnason, K., Eysteinsson, H., Hersir, G.P., 2010. Joint 1D inversion of TEM and MT data and 3D inversion of MT data in the Hengill area, SW Iceland. *Geothermics, The European I-GET Project: Integrated Geophysical Exploration Technologies for Deep Geothermal Reservoirs 39*, 13–34. doi:10.1016/j.geothermics.2010.01.002

Babuška, I., Nobile, F., Tempone, R., 2008. A systematic approach to model validation based on Bayesian updates and prediction related rejection criteria. *Computer Methods in Applied Mechanics and Engineering, Validation Challenge Workshop 197*, 2517–2539. doi:10.1016/j.cma.2007.08.031

Bear, J., Corapcioglu, M.Y., 1981. A mathematical model for consolidation in a thermoelastic aquifer due to hot water injection or pumping. *Water Resources Research 17*, 723–736. doi:10.1029/WR017I003P00723

Björnsson, G., Steingrímsson, G., 1995. Hitalíkan af Reykjasvæðunum í Mosfellsbæ [Temperature model of the Reykir areas in Mosfellsbær] (No. OS-95016/JHD-02). Orkustofnun, Reykjavík.

Bloemendal, M., Beernink, S., Bel, N., Hockin, A. E., & Schout, G., 2020. *Transitie open bodemenergiesysteem Koppert-Cress naar verhoogde opslagtemperatuur. Evaluatie van energiebesparingen en grondwatereffecten*. Retrieved from Nieuwegein: https://library.kwrwater.nl/publication/61755396/

Bloemendal, M., Beernink, S., Hartog, N., & Van Meurs, B., 2019. *Transforming ATES to HT-ATES. Insights from dutch pilot project*. Paper presented at the European Geothermal Congress EGC, Den Haag.

Bloemendal, M., & Hartog, N., 2018. Analysis of the impact of storage conditions on the thermal recovery efficiency of low-temperature ATES systems. *Geothermics, 17*, 306-319. doi:10.1016/j.geothermics.2017.10.009

Clerc N. and Moscariello A., 2020. A revised structural framework for the Geneva Basin and the neighboring France region as revealed from 2D seismic data: implications for geothermal exploration Swiss Bull. angew. Geol. Vol. 25/1+2, 2020 S. 109-131.

DARTS, 2021. Delft Advanced Research Terra Simulator (DARTS) [WWW Document]. URL https://darts.citg.tudelft.nl/

Davis, P.A., Goodrich, M.T., 1990. A proposed strategy for the validation of groundwater flow and solute transport models, in: Proceedings of GEOVAL90 Symposium on Validation of Geosphere Performance Assessment Models. Swedish Nuclear Power Inspectorate, Stockholm, Sweden, pp. 580–588.

Dincer, I. and Rosen, M.A., 2011. *Thermal energy storage systems and applications*. John Wiley & Sons.

Driesner, T. (Ed.), in prep. HEATSTORE – Final report on tools and workflows for simulating subsurface dynamics of different types of High Temperature Underground Thermal Energy Storage. GEOTHERMICA – ERA NET Cofund Geothermal.

Driesner, T., Drijver, B., Mathijssen, H., Struijk, E.L.M., Van Wees, J.D.A.M., Griffioen, J., … Zanon, V., 2019. HEATSTORE – Initial report on tools and workflows for simulating subsurface dynamics of different types of High Temperature Underground Thermal Energy Storage. GEOTHERMICA – ERA NET Cofund Geothermal.

Eidsvik, J., Mukerji, T., Bhattacharjya, D., 2016. *Value of information in the earth sciences: Integrating spatial modeling and decision analysis*. Cambridge University Press.

Fetter, C. W., 2001. *Applied Hydrogeology* (4th ed.). Upper Saddle River, NJ, USA.

Finsterle, S., 2007. iTOUGH2 User's Guide. (LBNL-40040). Lawrence Berkeley National Laboratory (LBNL), Berkeley, CA (United States).

Flavelle, P., 1992. A quantitative measure of model validation and its potential use for regulatory purposes. *Advances in Water Resources, Validation of Geo-hydrological Models Part 1 15*, 5–13. doi:10.1016/0309-1708(92)90028-Z

Franzson, H., 1998. Reservoir geology of the Nesjavellir high-temperature field in SW-Iceland, in: Proceedings 19th Annual PNOC-EDC Geothermal Conference. pp. 13–20.

Franzson, H., Gunnlaugsson, E., Árnason, K., Sæmundsson, K., Steingrimsson, B., Hardarson, B., 2010. The Hengill geothermal system, conceptual model and thermal evolution, in: Proceedings World Geothermal Congress. p. 1.

Franzson, H., Kristjánsson, B.R., Gunnarsson, G., Björnsson, G., Hjartarson, A., Steingrímsson, B., … Gíslason, G., 2005. The Hengill-Hellisheiði geothermal field. Development of a conceptual geothermal model, in: Proceedings World Geothermal Congress. Antalya, pp. 1–7.

Ganguly, S., Kumar, M.S.M., 2015. A numerical model for transient temperature distribution in an aquifer thermal energy storage system with multiple wells. Lowland Technology International 17, 179–188. doi:10.14247/LTI.17.3_179

Gens, A., Sánchez, M., Guimarães, L.D.N., Alonso, E.E., Lloret, A., Olivella, S., … Huertas, F., 2009. A full-scale in situ heating test for high-level nuclear waste disposal: observations, analysis and interpretation. *Géotechnique 59*, 377–399. doi:10.1680/geot.2009.59.4.377

GeoMol Team, 2015. GeoMol – Assessing subsurface potentials of the Alpine Foreland Basins for sustainable planning and use of natural resources – Project Report, 188 pp. (Augsburg, LfU)

Guglielmetti, L., Eichinger, F., Moscariello, A., 2020. Geochemical Characterization of Geothermal Waters Circulation in Carbonate Geothermal Reservoirs of the Geneva Basin (GB), in: Proceedings World Geothermal Congress 2020. Reykjavik, Iceland.

Harbaugh, A. W., Banta, E. R., Hill, M. C., & McDonald, M. G., 2000. Modflow-2000, the u.S. Geological survey modular ground-water model—user guide to modularization concepts and the ground-water flow process Virginia: US Geological Survey.

Hassan, A.E., 2004. Validation of Numerical Ground Water Models Used to Guide Decision Making. *Groundwater 42*, 277–290. doi:10.1111/j.1745-6584.2004.tb02674.x

Hecht-Mendez, J., Molina-Giraldo, N., Blum, P., & Bayer, P., 2010. Evaluating MT3DMS for Heat Transport Simulation of Closed Geothermal Systems. *Ground water, 48*(5), 741-756. doi:10.1111/j.1745-6584.2010.00678.x

Helgadóttir, H.M., Snaebjörnsdottir, S., Níelsson, S., Gunnarsdóttir, S.H., Matthíasdóttir, T., Hardarson, B., … Franzson, H., 2010. Geology and hydrothermal alteration in the reservoir of the Hellisheidi high temperature system, SW-Iceland, in: Proceedings, World Geothermal Congress, Bali, Indonesia. pp. 25–29.

Hölker, A. (2006). Lower Freshwater Molasse. Stochastic Modelling of hydraulic conductivity. Report for Hauptabteilung für die Sicherheit von Kernanlagen CH-5232 Villigen-HSK by Andreas Hölker Proseis AG.

Huber, M.L., Perkins, R.A., Laesecke, A., Friend, D.G., Sengers, J. V., Assael, M.J., … Miyagawa, K., 2009. New international formulation for the viscosity of H2 O. *Journal of Physical and Chemical Reference Data 38*, 101–125. doi:10.1063/1.3088050

Kallesøe, A.J., Vangkilde-Pedersen, T. (Eds.), 2019. Underground Thermal Energy Storage (UTES) – state-of-the-art, example cases and lessons learned. GEOTHERMICA – ERA NET Cofund Geothermal.

Keller, B. (1992). Hydrogeologie des schweizerischen Molasse-Beckens: aktueller Wissensstand und weiterführende Betrachtungen. *Eclogae Geologicae Helvetiae, 85*.

Küpfer, T. (2005). Arbeitsbericht NAB 04-09. Hydrogeologie der Unteren Süsswassermolasse.

Khait, M., Voskov, D., 2018. Operator-based linearization for efficient modeling of geothermal processes. *Geothermics 74*, 7–18. doi:10.1016/j.geothermics.2018.01.012

Konikow, L.F., 1992. DISCUSSION OF “The modeling process and model validation” by Chin-Fu Tsang, November-December 1991 issue, v. 29, no. 6, pp. 825-831. *Groundwater 30*, 622–623. doi:10.1111/j.1745-6584.1992.tb01543.x

Konikow, L.F., 1978. Calibration of ground-water models, in: Verification of Mathematical and Physical Models in Hydraulic Engineering. ASCE, pp. 87–93.

Konikow, L.F., Bredehoeft, J.D., 1992. Ground-water models cannot be validated. *Advances in Water Resources, Validation of Geo-hydrological Models Part 1 15*, 75–83. doi:10.1016/0309-1708(92)90033-X

Langevin, C. D., 2008. Modeling Axisymmetric Flow and Transport. *Ground water, 46*(4), 579-590. doi:10.1111/j.1745-6584.2008.00445.x

Langevin, C. D., Dausman, A. M., & Sukop, M. C., 2010. Solute and heat transport model of the Henry and hilleke laboratory experiment. *Ground water, 48*(5), 757-770. doi:10.1111/j.1745-6584.2009.00596.x

Langevin, C. D., Thorne, D. T., Dausman, A. M., Sukop, M. C., & Guo, W., 2008. *SEAWAT Version 4: A computer program for simulation of multi-Species Solute and heat transport*. Retrieved from Reston, Virginia:

Louwyck, A., Vandenbohede, A., Bakker, M., & Lebbe, L., 2014. MODFLOW procedure to simulate axisymmetric flow in radially heterogeneous and layered aquifer systems. *Hydrogeology Journal, 22*(5), 1217-1226. doi:10.1007/s10040-014-1150-0

Luis, S.J., McLaughlin, D., 1992. A stochastic approach to model validation. *Advances in Water Resources, Validation of Geo-hydrological Models Part 1 15*, 15–32. doi:10.1016/0309-1708(92)90029-2

Mindel, J. E., Alt-Epping, P., Landes, A. A. L., Beernink, S., Birdsell, D. T., Bloemendal, M., ... Driesner, T., 2021. Benchmark study of simulators for thermo-hydraulic modelling of low enthalpy geothermal processes. *Geothermics, 96*, 102130. doi:10.1016/j.geothermics.2021.102130

Mindel, J.E., Driesner, T., 2020. HEATSTORE: Preliminary Design of a High Temperature Aquifer Thermal Energy Storage (HT-ATES) System in Geneva Based on TH Simulations, in: Proceedings World Geothermal Congress 2020. pp. 1–12.

Mummert, M.C., 1996. Model validation and uncertainty analysis: An example using a nitrate percolation model, in: Subsurface Fluid-Flow (Ground-Water and Vadose Zone) Modeling, ASTM STP 1288. American Society for Testing and Materials, Philadelphia, Pennsylvania, pp. 187–200.

NAGRA NTB 88-07 (1990). Hydrochemische Synthese Nordschweiz: Tertiär- und Malm-Aquifere.

NAGRA NTB 88-25 (1988). Sedimentstudie – Zwischenbericht 1988.

NAGRA NTB 90-41 (1990). Sedimentäre Architektur der distalen Unteren Süsswassermolasse und ihre Beziehung zur Diagenese und den petrophysik. Eigenschaften am Beispiel der Bohrungen Langenthal.

NAGRA NTB 92-03 (1993). Untere Süsswassermolasse im Erdsondenfeld Burgdorf. Charakterisierung mittels Geologie, Petrophysik und Fluid Logging.

Nishimura, S., Gens, A., Olivella, S., Jardine, R.J., 2009. THM-coupled finite element analysis of frozen soil: formulation and application. *Géotechnique 59*, 159–171. doi:10.1680/geot.2009.59.3.159

Noorbergen, L. J., Lourens, L. J., Munsterman, D. K., & Verreussel, R. M. C. H., 2015. Stable isotope stratigraphy of the early Quaternary of borehole Noordwijk, southern North Sea. *Quaternary International, 386*, 148-157. doi:10.1016/j.quaint.2015.02.045

Olivella, S., Gens, A., 2005. Double structure THM analyses of a heating test in a fractured tuff incorporating intrinsic permeability variations. *International Journal of Rock Mechanics and Mining Sciences, Research results from the Decovalex III & Benchpar projects 42*, 667–679. doi:10.1016/j.ijrmms.2005.03.007

Oreskes, N., Shrader-Frechette, K., Belitz, K., 1994. Verification, Validation, and Confirmation of Numerical Models in the Earth Sciences. *Science 263*, 641–646. doi:10.1126/science.263.5147.641

Pfiffner, O.A., Lehner, P., Heitzmann, P., Müller, St. & Steck, A. (1997). Results of NRP 20. Deep Structure of the Swiss Alps. Birkhäuser.

Platt, N.H., Keller, B. (1992). Distal alluvial deposits in a foreland basin setting – the Lower Freswater Molasse (Lower Miocene), Switzerland: sedimentology, architecture and palaeosols. Sedimentology 39, 545-565.

Popper, K., 1959. *The Logic of Scientific Discovery*, 2005 rev. ed. Routledge.

Pruess, K., Oldenburg, C.M., Moridis, G., 2012. TOUGH2 User's guide, Version 2. (LBNL-43134). Lawrence Berkeley National Laboratory (LBNL), Berkeley, CA (United States).

Richardson, R.A., Wright, D.W., Edeling, W., Jancauskas, V., Lakhlili, J., Coveney, P.V., 2020. EasyVVUQ: A Library for Verification, Validation and Uncertainty Quantification in High Performance Computing. *Journal of Open Research Software 8*, 11. doi:10.5334/jors.303

Robinet, J.-C., Ducoulombier, A., Plas, F., 1999. Etude qualitative de codes de calcul THM européens par intercomparaisons. *Revue Française de Génie Civil 3*, 673–691. doi:10.1080/12795119.1999.9692673

Rohmer, J., Perreaux, M., Perozzi, L., Maragna, C., 2020. Theoretical framework for the representation of uncertainties. GEOTHERMICA – ERA NET Cofund Geothermal.

Romera, J.J.G., 2020. Python library for IAPWS standard calculation of water and steam properties. doi:10.5281/ZENODO.3734292

Roy, C.J., Oberkampf, W.L., 2011. A comprehensive framework for verification, validation, and uncertainty quantification in scientific computing. *Computer Methods in Applied Mechanics and Engineering 200*, 2131–2144. doi:10.1016/j.cma.2011.03.016

Rusillon, E., 2017. Characterisation and rock typing of deep geothermal reservoirs in the Greater Geneva Basin (Switzerland & France). University of Geneva, Geneva. doi:10.13097/archive-ouverte/unige:105286

Rykiel, E.J., 1994. The Meaning of Models. *Science 264*, 330–331. doi:10.1126/science.264.5157.330

Saemundsson, K., 1995. Geological map of the Hengill area 1: 50,000. Orkustofnun, Reykjavík.

Scheidt, C., Li, L., Caers, J. (Eds.), 2018. *Quantifying Uncertainty in Subsurface Systems*. Wiley & the American Geophysical Union, New York, N.Y.

Sinton, J., Grönvold, K., Sæmundsson, K., 2005. Postglacial eruptive history of the Western Volcanic Zone, Iceland. *Geochemistry, Geophysics, Geosystems 6*. doi:10.1029/2005GC001021

Sommer, W., Valstar, J., van Gaans, P., Grotenhuis, T., & Rijnaarts, H., 2013. The impact of aquifer heterogeneity on the performance of aquifer thermal energy storage. *Water Resources Research, 49*(12), 8128-8138. doi:10.1002/2013wr013677

Sterman, J.D., 1994. The Meaning of Models. *Science 264*, 329–330. doi:10.1126/science.264.5157.329-b

Thorne, D. T., Langevin, C. D., & Sukop, M. C., 2006. Addition of simultaneous heat and solute transport and variable fluid viscosity to SEAWAT. *Computers & Geosciences, 32*, 1758-1768. doi:10.1016/j.cageo.2006.04.005

TNO, 2019. *Totstandskomingsrapport hydrogeologisch model (REGIS II)* (TNO 2019 R11654). Retrieved from Utrecht: https://www.dinoloket.nl/en/regis-ii-hydrogeological-model

Tomasdottir, S., & Gunnarsson, G., 2019. HEATSTORE – Interim report on UTES-type/site-specific simulators based on academic/research codes. GEOTHERMICA – ERA NET Cofund Geothermal.

Tómasson, J., 1997. Megin jarðlagasyrpur í borholum á Reykjasvæðunum í Mosfellsbæ [Main lithological sequences in wells in the Reykir areas in Mosfellsbær] (No. JT-97-02). Orkustofnun, Reykjavík.

Tsang, C.-F., 1992. REPLY TO the preceding Discussion of “The Modeling Process and Model Validation.” *Groundwater 30*, 623–624. doi:10.1111/j.1745-6584.1992.tb01544.x

Tsang, C.-F., 1991. The Modeling Process and Model Validation. *Groundwater 29*, 825–831. doi:10.1111/j.1745-6584.1991.tb00568.x

van Lopik, J. H., Hartog, N., & Zaadnoordijk, W. J., 2016. The use of salinity contrast for density difference compensation to improve the thermal recovery efficiency in high-temperature aquifer thermal energy storage systems. *Hydrogeology Journal, 24*(5), 1255-1271. doi:10.1007/s10040-016-1366-2

van Lopik, J. H., Hartog, N., Zaadnoordijk, W. J., Cirkel, D. G., & Raoof, A., 2015. Salinization in a stratified aquifer induced by heat transfer from well casings. *Advances in Water Resources, 86*, 32-45. doi:10.1016/j.advwatres.2015.09.025

Voss, C.I., 1984. A finite-element simulation model for saturated-unsaturated, fluid-density-dependent ground-water flow with energy transport or chemically-reactive single-species solute transport. *Water Resources Investigation Report*, *84*, p.4369.

Wang, Y., Voskov, D., Khait, M., Bruhn, D., 2020. An efficient numerical simulator for geothermal simulation: A benchmark study. *Applied Energy 264*, 114693. doi:10.1016/j.apenergy.2020.114693

Winterleitner, G., Schütz, F., Wenzlaff, C., & Huenges, E., 2018. The Impact of Reservoir Heterogeneities on High-Temperature Aquifer Thermal Energy Storage Systems. A Case Study from Northern Oman. *Geothermics, 74*, 150-162. doi:10.1016/j.geothermics.2018.02.005

Younker, J.L., Boak, J.M., 1994. Geological Models. *Science 264*, 1065–1065. doi:10.1126/science.8178159

Zheng, C., & Wang, P. P., 1999. *MT3DMS: A Modular Three-Dimensional Multispecies Transport Model for Simulation of Advection, Dispersion, and Chemical Reactions of Contaminants in Groundwater Systems; Documentation and User's Guide*.