Ref. Ares(2024)3394508 - 10/05/2024

Reporting and monitoring geothermal heat pumps in Europe

Deliverable D2.2

Version: Final
Date: 03/05/2024
Main Authors:

Marlon Brancher (GeoSphere Austria)

Cornelia Steiner (GeoSphere Austria)

GeoBOOST

Co-funded by the European Union

Review

- R. Pasquali (GeoServ Solutions)
- A. Cunningham (GeoServ Solutions)
- P. Schnabl (GeoENERGIE Konzept)
- J. Chocobar Villegas (Technical University Munich)

Contributions

- B. Thelin (Rototec)
- H.J.L. Witte (Groenholland)
- J. Koczorowski (PORTPC)
- J. Ozimek (PORTPC)
- K. Zosseder (Technical University Munich)
- P. Schnabl (GeoENERGIE Konzept)
- R. Pasquali (GeoServ Solutions)

Table of Contents

1. Introduction

1.1. Strategic impact and deliverable rationale

The transition towards renewable energy sources within the European Union (EU) is a key aspect of its strategy to achieve carbon neutrality and meet ambitious climate goals. Geothermal heat pumps (GHPs) emerge as a vital technology in this transition, offering efficient and sustainable heating and cooling solutions. The present deliverable has been produced in the context of the EU LIFE21 GeoBOOST project, which aims to catalyse GHP adoption by overcoming major obstacles to the technology.

This document deals with the challenge of augmenting data collection, monitoring, and standardisation practices for GHP installations across Europe. Currently, most EU Member States do not maintain detailed records or databases for installations equipped with GHPs. This shortcoming negatively impacts important aspects, including the visibility of the technology, the ability to discern current and emerging market trends between different GHP systems, and the thorough understanding of the implementation of regulatory requirements. In other words, more detailed and standardised data are essential for multiple reasons:

- Such information can offer a clearer picture of the current state and potential growth of the GHP market, inform the development of more efficient and cost-effective designs, and facilitate the spatial identification of optimal deployment based on specific geological and hydrogeological contexts.
- Understand the deployment patterns of the technology which typically has spatial heterogeneity across different regions within countries.
- Support the evaluation of environmental impacts and potential studies, including the possible negative thermal interference between adjacent systems and the broader implications for subsurface thermal regimes.
- Better data can give stakeholders the ability to make more informed decisions, advance technological innovations, and fully realise the potential of GHPs in contributing to the EU's renewable energy goals.

1.2. Background and project presentation

The transition towards renewable energy sources for energy consumption in the EU is crucial for achieving carbon neutrality. This strategy is related to the ambitious targets set by the "Fit for 55 Package", aiming to reduce greenhouse gas emissions by 55% by 2030, and the EU Green Deal, which sets the vision for a climate-neutral EU by 2050.

The updated Renewable Energy Directive (EU/2023/2413) increases the EU's renewable energy target for 2030 from 32% to 42.5%, with the potential to further raise this target to 45%. In 2022, renewable energy sources constituted 23% of the gross final energy consumption in the EU. To meet the revised 2030 target, EU member states must collectively make enhanced efforts.

Heating and cooling accounts for about half of the EU's total gross final energy consumption. The proportion of renewables used in heating and cooling in the EU showed a continued increase, reaching an average of approximately 25% in 2022. Within the EU, Sweden is currently the leading country, with a 69% reliance on renewables for heating and cooling, followed by Estonia at 65%. On the other hand, countries with the lowest percentages of renewables in heating and cooling are Ireland (6%), the Netherlands (9%), and Belgium (10%) (**Fig. 1**). Whilst considerable progress has been made, much more remains to be done.

Among the renewable energy solutions, geothermal heat pumps (GHPs) are a highly efficient and cost-effective technology for both residential and non-residential applications (Bayer et al., 2019; Lund et al., 2022). Despite their proven efficiency, the deployment of GHPs within the EU's energy portfolio is still relatively marginal. This underscores the urgency to address longstanding obstacles hindering their widespread adoption. The challenges are diverse, primarily stemming from a lack of thorough analysis of the specific regulatory, financial, and policy hurdles facing GHPs (García-Gil et al., 2020; Somogyi et al., 2017; Tsagarakis et al., 2020).

For example, the cost ratio between electricity and gas significantly influences consumer choices regarding heating solutions. Historically, gas has been cheaper than electricity, which has made gas boilers a popular choice for heating (**Fig. 2**). The push towards greener alternatives, such as GHPs, depends on altering this cost dynamic to favour electricity. Generally, a ratio of less than around 2.0-2.5 supports the argument that heat pumps could become financially more attractive than gas boilers over their lifespan (Sarsentis and Orso, 2023). Yet, it should be grasped that the electricity-to-gas price ratio is less meaningful for the Nordic countries. This is because the gas grid penetration there is relatively low. Biomass may be a better indicator than gas in this case. The same reasoning also applies to countries where other types of heating fuels like coal and oil are prevalent (Sarsentis and Orso, 2023).

It should also be noted that GHPs have commonly been seen as a technology for reducing energy consumption in buildings. However, recent evidence indicates that the widespread deployment of GHPs, coupled with retrofitting solutions in single-family homes, can primarily serve as a means to reduce grid costs on a national scale (Liu et al., 2023). This is highly pertinent nowadays with the electrification.

Fig. 1. *Evolution of the share of renewable energy sources in the heating and cooling sector from 2004 to 2022 across countries in Europe besides the EU27 (2020 onwards). Data source: Eurostat (nrg_ind_ren).*

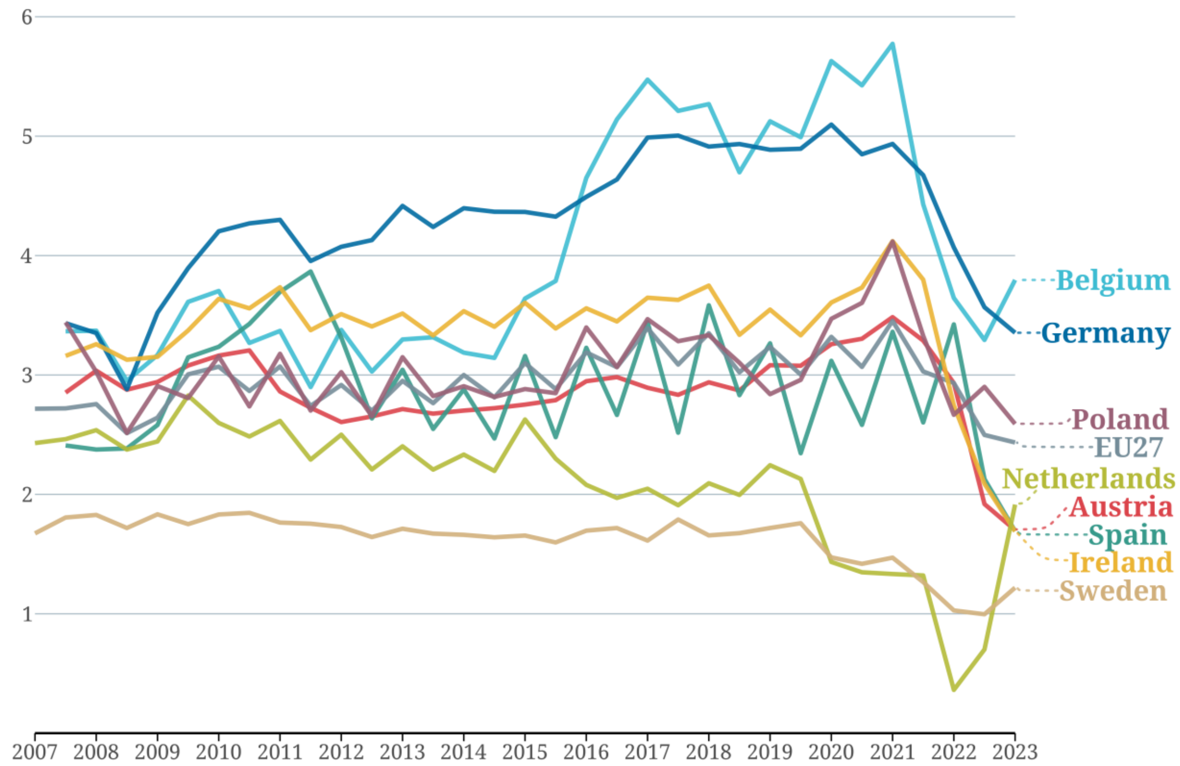

Fig. 2. *Evolution of electricity-to-gas price ratios (y-axis) for household costumers in the GeoBOOST partner countries besides the EU27 (2020 onwards) over the 2007-2023 period (x-axis). This non-dimensional ratio has been calculated based on bi-annual data prices in Euro for an annual electricity consumption from 2,500 kWh to 4,999 kWh (band DC) and an annual gas consumption from 20 GJ to 199 GJ (band D2). For 2023, data was available for the first semester as of the data download. Data source: Eurostat (nrg_pc_202 and nrg_pc_204).*

In response to these challenges, the EU LIFE21 GeoBOOST project, titled "*Boosting geothermal heat pumps to mainstream cost-effective and efficient renewable heating and cooling in buildings*", aims to catalyse the adoption of GHPs across the EU. The project is supported by a consortium of research institutes, companies, and organisations from eight countries: Austria, Belgium, Germany, Ireland, the Netherlands, Poland, Spain, and Sweden.

By leveraging the collective expertise of this multidisciplinary consortium, the main goal of GeoBOOST is to overcome market obstacles for GHPs. Despite being the most energy-efficient and cost-effective renewable energy solution for heating and cooling, GHPs have frequently been overlooked by policymakers and research proposals focused on market adoption.

GeoBOOST is centred on the following aspects:

- The absence of comprehensive data and monitoring standards for GHPs.
- Revalorisation of high upfront capital expenditure costs.

- Deficient business models and adequate financing mechanisms.
- The need for regulatory alignment and the streamlining of current processes for authorisation, certification, and licensing.
- The enlargement of a skilled workforce.
- The lack of awareness of the significant benefits, cost-effectiveness, and energy efficiency of GHP equipment and systems.

1.3. Geothermal systems classification

There are many approaches in use for the classification of geothermal energy systems. Among the most used classification principles are depth, temperature levels (enthalpy), and the state of the circulation path of the heat transfer fluid (loop type).

Depth classification usually defines three categories: shallow geothermal energy, exploiting the subsurface up to several hundred meters; medium depth, spanning approximately 500 to 2000 m; and deep geothermal, extending beyond 2000 m. This categorisation reflects the principle that the Earth's temperature increases by roughly 30 °C per km of depth (so-called geothermal gradient), influencing the potential applications of the geothermal source.

Temperature level or enthalpy can also be used to classify geothermal technologies. Moderate- and high-enthalpy systems operate above approximately 60 °C, typically with sources above 90 °C possessing sufficient exergy for power generation. Conversely, low-enthalpy systems, often shallow, operate below about 35 °C. A heat pump is then required to upgrade the more moderate temperatures found closer to the surface to levels suitable for space heating and cooling and domestic hot water production. For example, the temperatures at the source rock of GHPs should range between 2 °C and 18 °C for central Europe and it might reach lower temperatures in the northern countries and higher temperatures in the South of Europe.

In this work, the GHP term is used in connection with "shallow" and "low temperature" systems. GHPs are well-suited for a vast array of heating and cooling applications across different environments. These include urban centres, villages, and rural areas, ranging from single-family houses (with capacities between 3–15 kW) to industrial and agricultural applications, as well as heating and cooling networks (with capacities exceeding 2500 kW) (Witte, 2023). Cooling can also be provided directly to buildings. This process, commonly referred to as free cooling, bypasses the need for refrigeration cycles and the operation of a heat pump.

The classification can be extended to the state of the circulation path of the heat transfer fluid (i.e., loop type). This principle is frequently used for shallow geothermal energy systems for which two primary configurations are predominant:

- ***Closed systems*** are characterised by the use of a heat transfer fluid circulating within a network of pipes to exchange heat with the ground. These can be further categorised based on their installation orientation and depth. Vertical or inclined configurations involve one or more borehole heat exchangers (BHEs) that are drilled into the ground to depths that can reach a few hundred meters below the surface. In contrast, horizontal configurations are positioned at very shallow depths, typically of a few meters below the ground.

- ***Open systems*** can directly exploit groundwater, if the hydrogeological and hydrochemical conditions are appropriate, as the heat exchange medium in an application called groundwater heat pump (GWHP) systems. Most of these installations operate based on a pair of wells within shallow aquifers. In this well-doublet configuration, groundwater is extracted from the aquifer, circulated through the heat pump system for energy transfer, and subsequently reinjected in the same aquifer. Despite being less common, it is important to note that at times the extracted groundwater is also discharged into nearby rivers, lakes, or other surface water bodies, for instance.

Borehole Thermal Energy Storage (BTES) and Aquifer Thermal Energy Storage (ATES) systems serve as specialised subsets of Underground Thermal Energy Storage (UTES) systems, mirroring the general functional principles of BHE and GWHP systems, respectively. ATES and BTES systems are primarily designed for the subsurface storage of thermal energy for later use, leveraging the inherent thermal stability of the underground environment. More on ATES and BTES is given in **Appendix 1**.

In addition to these more traditional systems, thermoactive geostructure (TAG) such as energy piles, diaphragm walls, and thermally activated slab floors are becoming important. These structures can be framed in the closed-system category. TAG integrates the functional requirements of building foundations and underground walls with the thermal exchange capabilities of geothermal systems, offering an interesting use of subsurface space and resources. By employing the thermal mass of structural elements, TAG systems provide an alternative method for harnessing shallow geothermal energy while supporting structural loads.

These diverse GHP system configurations collectively offer a versatile range of solutions for sustainable building climate control. Their application and effectiveness are influenced by local geological and hydrogeological conditions, system design, and building energy requirements. The diversity of GHP systems shows the importance of tailored approaches in the rollout of geothermal technologies.

1.4. Scope of the deliverable and important definitions

The overarching challenge addressed in the present report is the lack of comprehensive, harmonised, and accessible data on GHP system construction, performance, and operational practices across different European countries. Thus, the scope of this deliverable is focused on data generation aspects, monitoring standards, and reporting practices.

Herein, ***"data"*** refers to the qualitative and quantitative information linked to the design, implementation, and operation of a GHP installation.

The term "*design, implementation, maintenance, and operation*" is intended to comprise the components and aspects of a GHP system in the broadest sense. This covers everything from the physical parts of the system to the operational approaches employed to ensure efficiency and to comply with regulations. The qualitative and quantitative information will therefore provide information on how different parts function within the system, its performance indicators, and any other relevant operational data.

An "**installation**" in the scope of this work is defined as the entire system set up, including all its constituents and infrastructure, designed to provide sustainable energy for heating and cooling. As a result of these definitions, "**GHP installation data**" means the collection of information and figures related to the setup and implementation of shallow geothermal systems.

1.5. Main objectives

Within the focal aspects of GeoBOOST described earlier, this deliverable is connected to Task 2.2 ("*Improving market monitoring and reporting*") within Work Package 2 ("*Market Topologies*"). Task 2.2 is centred on one of the barriers to the widespread adoption of GHPs: the deficiency of standardised data generation and reporting standards.

Hence, the ambition of this deliverable is to contribute to the streamlining of data collection and monitoring approaches for GHP systems. Achieving a systematic approach to data collection and monitoring will be essential for overcoming the fragmented and often inconsistent reporting practices of GHP installation data prevalent in the field.

Main objectives of the report are:

i. To outline the current mechanisms for reporting and monitoring of market indicators at the European level referring to available European market reports.

ii. To provide a critical analysis of the EU's methodologies for accounting renewable energy contributions from heat pumps, with a focus on GHPs for heating and cooling.

iii. To conduct an evaluation of the situation for reporting and monitoring within the GeoBOOST partner countries to capture specific insights within national markets.

iv. To identify and catalogue essential parameters for improved granularity and reliability of GHP installation data.

v. To propose a standardised framework for capturing GHP installation data for key system types in order to facilitate consistent and clear data sharing and analysis with third parties.

2. Reporting and monitoring mechanisms for the GHP market in Europe

The present section examines the mechanisms currently in place for the reporting and monitoring of market indicators related to heat pumps, with a specific emphasis on GHPs. It outlines the existing situation of market data collection at the European scale.

This analysis was done by prioritising sources of information that are recognised for their relevance to the sector and policymaking. Namely, the scrutiny is based on industry-standard market reports, specifically those published by the European Geothermal Energy Council (EGEC) and the European Heat Pump Association (EHPA), as an attempt to ensure consistency. This is because these reports are not only benchmarks within the industry but also serve as critical resources for Member States and the European Commission, highlighting their significance.

The following presents the insights gathered from the examination, distinguishing between those obtained from EGEC reports and those sourced from the EHPA reports. Common and uncommon points are then discussed.

2.1. EHPA Market Report

The EHPA Market Report is a comprehensive document that provides data on various types of heat pumps in 21 countries (EHPA, 2023). The primary market indicator tracked by the EHPA is related to sales data, that is, the number of units sold. From this, the market development, segmentation and share is explored.

EHPA sales data acquisition and processing methodology is described in Annex 5 of the latest available version of the report (EHPA, 2023). A standardised questionnaire is used to obtain data on units sold. This questionnaire is distributed to national heat pump associations, statistics bureaus, and research facilities. Encompassed are heat pumps providing heating, heating and cooling, sanitary hot water, and process heat.

The EHPA report primarily focuses on heat pumps that provide a heating function and/or sanitary hot water, aligning with the requirements of the Renewable Energy Directive. In order to account for this, the EHPA has developed a methodology based on correction factors.

By accounting for air/air, air/water, brine/water, hybrid, and variable refrigerant flow (VRF) units, the correction methodology encompasses a spectrum of heat pump technologies. Specifically, the methodology considers how different types of heat pumps (air/air, air/water, brine/water, hybrid, and VRF units) are accounted for in cold and warm climate countries. For air/air units, adjustments are made to exclude AC-only units, with special mention of Sweden's data collection practices. In warm climates, only a portion of sales is reported, based on the assumption of their use as the primary heat source, exemplified by an Italian market study (Gazzetta Ufficiale, 2012; Pieve and Trinchieri, 2019). Reversible heat pumps connected to hydronic systems are fully counted due to their primary heating function. A broad definition is adopted for hybrid heat pumps to incorporate them into the database: "*A hybrid heat pump is the combination of a heat pump and a fossil fuel based boiler that has a controller between both heaters and is designed to be sold together under one commercial reference*". VRF systems are included with a 90% consideration rate to account for potential use discrepancies.

Further, EHPA estimates the environmental benefits of the heat pump stock, deriving additional market indicators. In particular, the EHPA report shows how the (*i*) renewable energy contribution, (*ii*) the energy savings, and (*iii*) the greenhouse gas emissions savings of heat pumps are calculated according to the EU methodology (**Sec. 3** of the present document shows this methodology) and the EHPA assumptions.

In our opinion, the strengths of the EHPA's approach are the following:

- **Specificity and adaptation to climate zones:** The methodology acknowledges the difference in heat pump usage between cold and warm climates, tailoring the data collection to reflect these variations. This specificity can help to derive a more accurate representation of heat pump usage across different geographical areas.

- **Correction factors and estimations:** The use of correction factors to exclude AC-only units and the estimation methods for missing data (e.g., Sweden's air/air units) demonstrate a proactive approach to achieving comprehensive and accurate data. These adjustments help mitigate potential biases or inaccuracies in the raw sales data.

- **Inclusion of a broad range of heat pump types:** By accounting for air/air, air/water, brine/water, hybrid, and VRF units, the methodology encompasses a wide spectrum of heat pump technologies. This inclusivity allows for a more complete picture of the market and technology adoption.

- **Flexibility in definitions:** The adoption of a broad definition for hybrid heat pumps and the pragmatic approach to counting VRF systems indicate flexibility. This adaptability is in practice needed for incorporating a variety of technologies and their evolving applications in the market.

On the other hand, perceived limitations include:

- **Data segregation:** While GHPs can be more broadly counted, it is nearly unmanageable to identify and segregate all types GHP systems when the data is reported by energy source and distribution medium, as is the case for EHPA. A clear distinction of the GHP systems according to their primary source of energy (e.g. groundwater, boreholes, topmost soil) would be beneficial to get a better understanding of GHP market dynamics. As seen in **Sec. 1.3**, GHP systems come in various configurations, each with unique characteristics and installation requirements.

- **Reliance on assumptions:** While necessary, the reliance on assumptions (e.g., the share of total sales included for warm climate zones based on the Italian market study) introduces a degree of uncertainty. The applicability of these assumptions to other countries or over time may vary, potentially affecting the accuracy of the data.
- **Potential for data gaps:** The methodology's effectiveness can be compromised by data collection challenges, such as the gap in Sweden's air/air sales data collection from 2011 to 2021. While estimations are made, the absence of direct data collection may lead to inaccuracies, which are hard to quantify.
- **Complexity in implementation:** The approach, while thorough, can lead to complexity in data analysis pipelines. Ensuring consistency and accuracy across different national heat pump associations may present additional challenges.
- **Granular data:** The data underpinning all of the analyses is more or less only available and/or offered at the national level. More granular data is necessary to understand sub-national patterns. Unquestionably, the methodology's European-wide scope must be recognised. The EHPA's approach, driven by national-level data, meets its objectives, so the potential for incorporating more detailed sub-national insights should be seen as a future opportunity.
- **Memberships:** There could be an under-reporting issue due to the lack of representation of some manufacturers in national heat pump associations. Relying solely on data from associations that link to the European Heat Pump Association (EHPA) data gathering methodology may lead to some data limitation. This can only be resolved when sub-national patterns are better assessed, as not all manufacturers are members of these associations.

In conclusion, overall, the methodology for counting heat pump sales and usage demonstrates a considerate approach to addressing the complexities of the market and technology use in different climates. The strengths lie in its specificity, adaptability, and inclusivity. However, the reliance on assumptions and potential data gaps are points where further refinement or additional data collection methods could bring more accuracy. The careful balance between detailed data collection and practical challenges of implementation indicates the effort to provide meaningful insights into the heat pump market.

2.2. EGEC Market Report

The EGEC Market Report aims to collect and analyse data specifically related to geothermal energy applications across the European market. This geothermal focus is critical for better delineating the current situation and trends within the sector. The EGEC report covers geothermal power plants, geothermal district heating and cooling, and GHPs. Furthermore, there is strong emphasis on GHPs used in district heating or cooling, industry supply, and individual buildings.

EGEC methodology is described in the "Data Collection and Methodology" section of the latest available version of the report (EGEC, 2023). The methodology initiates with a straightforward approach to data gathering, soliciting information from national coordinators across the European Union. These coordinators are responsible for submitting the most recent data for the year preceding the report's publication (EGEC, 2023).

Regarding GHPs, the main market indicator EGEC seeks to report is the number of units sold, alongside the installed stock up to the year preceding the report release. Average system capacities are also asked for when existing. In addition, EGEC looks for market data on very large systems, both for closed and open systems. For closed systems, "very large' is defined as systems with total borehole heat exchanger length larger than 10 km. For open systems, "very large" means systems with installed capacity above 100 kW. For these very large systems, there are requests for detailed information on the number of boreholes, borehole depth, wells depth, installed capacity for heating and/or cooling, and year of commissioning. This highlights a commitment to capturing details of the GHP market. The interest in reporting large systems seems to come from the perception that such installations are becoming more relevant, by increasing in number and size.

Thus, the methodology attempts to provide a more holistic understanding of the GHP market, from individual residential installations to large-scale industrial and district heating systems. This separation is in fact important for distinguishing between trends across various market segments and understanding technological advancements and growth.

In our view, the strengths of the EGEC's approach are:

- **Inclusiveness:** The methodology's broad scope and focus on detailed aspects of GHP systems can give a more comprehensive representation of the GHP market, acknowledging its complexity and the importance of newer installations and technological upgrades.

- **Data sources:** By leveraging regional expertise through data collection from national coordinators, the methodology attempts to enhance the accuracy and relevance of the collected data, providing a well-founded view of the market's current state across the EU. Indeed, with data sourced directly from national coordinators, the methodology can offer the most recent information available up to the year preceding the report's publication.

Conversely, perceived limitations comprise:

- **Potential for inconsistency:** Variability in data quality and completeness across different countries, due to reliance on national coordinators, may impact the consistency of market analysis and it is not addressed. In fact, the effectiveness of the methodology is entirely dependent on the reliability and availability of submitted data, which may vary.

- **Limited scope for differentiation:** Further refinement would still be needed to distinguish between specific GHP systems to better understand their market dynamics.

- **Granular data:** The data underpinning the report is more or less only available and/or offered at the national level, which means that without more refined data sub-national diffusion patterns cannot be examined. Again, this point is meant in context of the potential for incorporating more detailed sub-national insights as a future opportunity.

In conclusion, the EGEC's methodology for GHP data collection and analysis stands as a simple but robust framework that can offer a more detailed overview of the European GHP market. It strikes a balance between pragmatism and depth, capturing important points for the understanding of the GHP market. Nonetheless, refining the methodology to address data consistency, enhance the differentiation between system types, and achieve more granular data to identify sub-national diffusion patterns, could expand its utility. Enhancements in these areas could advance market analysis, facilitating informed decision-making, and supporting the strategic development of the GHPs in Europe. We believe this effort would mark a significant step toward mapping the GHP market in a more methodical way, providing a valuable resource for all involved stakeholders.

2.3. Final remarks

The number of units sold is currently the primary market indicator used at the European level to demonstrate technology growth. But this is not enough to reflect the required heating and cooling services in different types of buildings and consumers, nor to fully distinguish between the different types of GHP systems.

While the diversity and adaptability of GHP systems are strengths from an environmental and energy efficiency standpoint, it should be acknowledged that these same characteristics pose significant challenges for data collection and segregation efforts.

The variety of GHP configurations, coupled with the lack of standardised reporting criteria, makes it difficult to uniformly classify and report these systems based on their energy source and distribution medium alone. This translates into practical difficulties typically came across for an accurate and comprehensive reporting of GHP systems. Once more, it highlights the urgent need for more standardised data collection and reporting frameworks in this field.

3. Accountability of energy from heat pumps

3.1. Policy background

The EU's intensified commitment to renewable energy has necessitated the development of robust methodologies for accounting for renewable contributions, especially in the heating and cooling sector. This section outlines and analyses the methods mandated by Commission Delegated Regulation (EU) 2022/759, amending Annex VII of Renewable Energy Directive (EU/2018/2001). A focus is given to the applicability of the methodologies to GHPs.

Historically, Annex VII of Directive (EU) 2018/2001 enabled the calculation of renewable energy contributions from heat pumps for heating purposes. This, however, fell short in addressing the cooling aspect, creating a gap in the renewable energy framework. Such a gap is particularly relevant for Member States experiencing non-negligible cooling demands, thereby limiting their potential contributions towards the EU's energy targets. Such a resolution is important, especially when considering the increasing demands for cooling driven by anthropogenic climate change and urban heat island effects (Miranda et al., 2023).

The amendment introduced by Part B of Regulation (EU) 2022/759 offers a much-needed remedy to this gap. It delineates a methodology that separately accounts for the heating and cooling modes, acknowledging the distinct operational efficiencies inherent to each. We understand that the methodology supports the fact that seasonal performance factors (SPF) are contingent upon the temperature differential between the heat source or sink and the serviced side: the smaller the differential, the greater the system's efficiency.

Furthermore, the amendment holds particular relevance for shallow energy systems, which makes them ideally suited to the separated accounting methods. This consideration means that geothermal cooling can be better represented and encouraged within the EU's renewable energy mix.

Subsequent sections will present these methodologies and offer a critical appraisal, aligning with the objectives of this report.

3.2. Heating

Annex VII of the EU's Renewable Energy Directive (EU/2018/2001) outlines the methodology for the accounting of energy harnessed from heat pumps, encompassing aerothermal, geothermal, and hydrothermal energy. The Directive considers this energy as renewable and eligible for inclusion in the renewable energy calculations for Member States. The core calculation procedure is defined by:

$$E_{RES} = Q_{usable} \times (1 - 1/SPF) \quad \textbf{Eq. 1}$$

where:

- Q_{usable} is the estimated total usable heat delivered by heat pumps;
- SPF is the average seasonal performance factor, reflecting the ratio of delivered heat to consumed electricity over a season.

To qualify, heat pumps must have an SPF greater than $1.15 \times 1/\eta$, where η signifies the EU-wide average ratio of gross electricity production to primary energy consumption, based on Eurostat data.

While Annex VII defines the basic calculation procedure, it sets out three parameters that are still needed to run the calculation. To bridge the gap between policy and practical application, the Commission's Decision 2013/114/EU elaborates on determining Q_{usable}, expected to be applied by the Member States. In this approach**,** Q_{usable} is defined as:

$$Q_{usable} = P_{rated} \times H_{HP} \quad \textbf{Eq. 2}$$

where:

- P_{rated} is the heat pump's rated capacity for heating;
- H_{HP} is the annual equivalent heat pump hours.

Note that P_{rated} denotes the capacity of the vapour compression cycle or sorption cycle of the unit at standard rating conditions. H_{HP} is the assumed annual number of hours a heat pump has to provide heat at rated capacity to deliver the total usable heat delivered by heat pumps, expressed in h. The power system efficiency η is set at 45.5%, translating to a minimum SPF of 2.5 for electrically driven heat pumps to be considered as renewable.

The Commission Decision 2013/114/EU establishes predefined values for H_{HP} and SPF, accommodating the specific characteristics of various heat pump technologies and three major climate regions. Consequently, unique values are assigned to each distinct combination of heat pump type and climate zone.

Organisations like EHPA leverages this framework to estimate the environmental benefits of the heat pump stock. But the Commission Decision 2013/114/EU does not provide specific values for P_{rated}. Consequently, EHPA's calculations use industry data and expert estimates to determine average installed capacities for various types of heat pumps across different countries (as detailed in **Table 1**).

This approach is applied to all categories of heat pumps, culminating in an aggregated determination of the total renewable energy contribution from heat pumps (E_{RES}) for each country. In short, while the procedure outlined provides a foundational framework, it seems that its application to GHPs still necessitates a more refined examination to ensure correctness, relevance, and adaptability to specific technological and environmental conditions.

Table 1. EHPA's average installed capacities in kW per type of heat pump and country.

Country	Air/ Water	Water/ Water	Brine/ Water	Dir. expansion/ water or dir. condensation	Dir. expansion/ dir. condensation	Exhaust Air	Sanitary Hot Water	Reversible
AT	12	10	14	10	10	2	3	4
BE	12	13	13	13	13	2	3	10
CH	12	10	14	10	10	2	3	4
CZ	12	26	13	0	2	2	2	5
DE	12	10	14	10	10	2	3	4
DK	8	8	10	8	8	0	3	4
EE	13	0	13	0	0	2	3	5
ES	12	18	0	0	0	2	8	6
FI	12	0	13	0	0	2	3	4
FR	11	15	14	11	10	2	3	10
HU	18	22	26	0	0	2	12	18
IE	13	14	13	0	0	2	3	0
IT	15	20	19	0	0	2	3	12
LT	14	17	12	10	0	2	3	4
NL	10	48	12	0	8	2	1	8
NO	10	0	15	0	0	2	3	5
PL	12	14	14	12	0	2	2	12
PT	15	12	12	12	12	2	3	22
SE	9	0	11	0	0	2	3	4
SK	10	13	16	10	0	2	10	12
UK	15	12	12	12	12	2	3	8

3.3. Cooling

Part B of the Delegated Regulation (EU) 2022/759 outlines the calculation methodologies for the share of renewables in cooling. This includes renewable energy from cooling systems (including district cooling) towards the EU's renewable energy target. The regulation provides definitions, sets the scope, and details the calculations required renewable energy from cooling. It mentions in Article 6 that "the cold source can be ambient energy or geothermal energy", which highlights the importance of geothermal systems among various renewable sources.

The methodology for accounting renewable energy used for cooling (Part B of the Delegated Regulation (EU) 2022/759) includes renewable energy from cooling systems, including district cooling, towards the EU renewable energy target. It specifies definitions, scope, and calculations needed to quantify renewable energy from cooling, highlighting the significance of geothermal energy systems among other renewable sources.

The methodology sets minimum efficiency requirements, expressed as the SPF, for cooling systems to be considered in producing renewable energy. It distinguishes between active and passive cooling, specifying the calculation of the renewable energy quantity for cooling (E_{RES-C}) based on the efficiency and energy input of cooling systems. The regulation aims to standardise the accounting for renewable cooling across Member States.

In order to quantify the amount of renewable energy produced by cooling systems that can be counted towards the EU's renewable energy targets, the following formula is defined:

$$E_{RES-C} = (Q_{source} - E_{input}) \times S_{SPF_p} = Q_{C_{Supply}} \times S_{SPF_p} \qquad \textbf{Eq. 3}$$

where,

- Q_{source} is the amount of heat released to the ambient environment (air, water, or ground) by the cooling system. For geothermal cooling systems, this would typically be the heat removed from a building and released into the ground (soil or groundwater);
- E_{input} is the energy consumed by the cooling system, including any auxiliary systems that support its operation. In the case of district cooling, it includes the energy used to distribute the cooling to various customers;
- $Q_{C_{Supply}}$ is cooling energy that the cooling system actually supplies. For systems that use geothermal energy, this would be the amount of cooling delivered to the end-use, such as a building's interior;
- S_{SPF_p} is a percentage defined at the cooling system level that reflects the share of the cooling supply considered renewable, in accordance with the SPF requirements. It essentially adjusts the total cooling supply $Q_{C_{Supply}}$ by the efficiency and renewable proportion to calculate the final renewable energy contribution.

To sum up, for geothermal cooling systems, Q_{source} would be the energy extracted from the ground or groundwater through the geothermal system, and E_{input} would involve the energy used to drive the heat pump. The S_{SPF_p} would then be determined to calculate the renewable proportion of this cooling supply.

By using **Eq. 3**, EU Member States can determine the amount of renewable energy provided by cooling systems, including geothermal, that contributes to the renewable energy target. **Eq. 3** makes sure that the energy accounted for is not only renewable but also delivered efficiently. This is important because it reflects the dual aims of increasing renewable energy uptake and promoting energy efficiency.

Further details on the steps and recommendations to derive $Q_{C_{Supply}}$ and S_{SPF_p} can be found in Sections 3.2 to 3.4 of the Delegated Act. A narrative that synthesises the technical complexities of the methodology into an articulated structure is given below due to the increasing need for cooling in the future and the significance of using renewable energy sources for cooling (Miranda et al., 2023).

Scope

Regarding the scope, it is defined that Member States shall count in the calculation active cooling, including district cooling, regardless of whether it is free cooling or a cooling generator is used. What should not be counted is also defined. This includes passive cooling, some pre-defined technologies or processes of cooling and energy used for cooling in industry uses (i.e., power generation plants; cement, iron and steel manufacturing; wastewater treatment plants; information technology facilities such as data centres; power transmission and distribution facilities; and transportation infrastructures)

Efficiency thresholds and seasonal performance

A first pillar of the accounting methodology is the implementation of a minimum efficiency requirement for cooling systems, delineated by the primary Seasonal Performance Factor (SPFp). Only systems greater than this efficiency baseline contribute to the renewable energy targets. SPFp is determined according to established EU regulations, reflecting a commitment to a harmonised and up-to-date energy performance assessment across Member States. Crucially, the minimum SPFp (i.e., SPFpLOW) is set at 1.4. For a cooling system's output to quality as 100% renewable, the SPFp (i.e., SPFpHIGH) should be at least 6. For values between SPFpLOW and SPFpHIGH, the share of the SPF that qualifies as renewable energy (SSPF) scales with the SPFp of the cooling system.

Standardised and measured seasonal performance factors

The calculation of the renewable energy quantity for cooling applies both standardised and measured SPF values. Standardised values, underpinned by Ecodesign requirements, are readily available for certain cooling generators up to defined capacities. In instances where these are not available or measurement is normally used, particularly for larger systems and district cooling, measured SPF values are employed. This dual approach is to bring flexibility and precision, accommodating diverse cooling system technologies and scales.

Primary energy factors and SPF calculations

The SPF values are fundamentally expressed in terms of primary energy efficiency. An update has been made to primary energy factor for electricity adjusted to 2.1. The primary energy factor for other carriers, such as heat or gas, is standardised at 1, recognising the direct utilisation of energy without transformation. SPF boundary conditions for cooling generators are defined in concordance with EU regulations, ensuring a consistent framework for SPF determination.

Simplified method for smaller cooling systems

For individual cooling systems with a capacity of less than 1.5 MW with a standard SPF, a simplified estimation method is in place, leveraging nominal cooling capacity and equivalent full load hours (EFLH) of operation. The method adopts default equations for computing EFLH, taking into account cooling-degree days and factors specific to sectors or processes.

Measured values for larger systems

Systems above 1.5 MW capacity, without standard values, or district cooling systems must base calculations on measured values for both energy input and cooling energy supply.

District cooling systems

In the context of district cooling, net cooling supply is central, with methodologies in place to account for and deduct thermal losses within distribution networks. Importantly, district cooling systems can be stratified into subsystems, facilitating granular measurement and accountability. This division allows for a careful assessment of renewable energy contributions from complex cooling system infrastructures.

Allocation of auxiliary energy

A worth mentioning aspect of the methodology is the treatment of auxiliaries within the cooling systems. The energy consumption of such components is either directly included within subsystem measurements or, when not directly allocable, distributed proportionally based on the cooling energy supplied by respective subsystems. This safeguards an ample accounting of all energy inputs, enhancing the precision of renewable energy calculations.

3.4. Critical analysis

This section scrutinises the methodologies outlined in both Commission Decision 2013/114/EU and Delegated Regulation (EU) 2022/759. The goal is to shed light on some critical aspects that we believe merit attention for the ongoing refinement of the methodologies, especially in the context of GHP systems. The critical aspects and recommendations for improvement from our analysis are the following:

a) Linear models from Equations 1 and 3

When SPF values are predefined and constant for each heat pump type and climate condition, the term $(1 - 1/SPF)$ in **Eq. 1** becomes a constant multiplier for Q_{usable}. In this case, SPF no longer acts as a variable that introduces non-linearity through its reciprocal value but rather as a fixed parameter that influences the scale of Q_{usable}.

So, given a fixed SPF, **Eq. 1** simplifies to a form where E_{RES} is directly proportional to Q_{usable}, fitting the criteria of a linear model: $y = mx + c$, where m and c are constants. The same reasoning can be applied to **Eq. 3**. Therefore, **Eqs. 1 and 3** offer simplified models for estimating the amount of renewable energy used for heating and cooling. While they serve to standardise the accounting across EU member states, both models inherently assume a linear relationship between heating or cooling output and the seasonal performance factor SPF.

Concerning heat pump systems, these systems are known to operate in a non-linear fashion. This is due to the complex interplay between the components of the system and operational conditions. While the SPF is a measure that captures the resultant efficiency across these varying conditions, including any non-linearity in performance across different operational conditions, the equations reliance on SPF as a linear proxy for system efficiency may oversimplify these dynamics. Assuming this oversimplification has some ground, it could be a factor contributing inaccuracy when estimating the true energy contribution from renewable sources.

b) Data assumptions

In the case of the calculation procedure for heating, the reliance on default values for SPF and H_{HP} may create a potential bias in the calculation of the actual renewable energy contribution. No uncertainty analyses around the average estimates is suggested. Without this, it is hard to conjecture the direction of this potential bias, both under- and overestimations could be possible.

The distinction made by the approach from Commission Decision 2013/114/EU, which provides fixed default values for H_{HP} and SPF across various types of heat pumps and broad climatic categories, should be acknowledged. Indeed, the categorisation into warmer, average, and colder climates is a step towards addressing variability, but it may still be overly simplistic. The granularity and applicability of these categories in capturing the realistic performance of GHPs warrant further examination. In the case of GHPs, the stable thermal resource accessed by these systems exhibits other sources of variability in temperature that is not solely dependent on air temperature or broad climate zones. Geothermal gradients, soil composition, and groundwater flow, among other aspects, can influence the efficiency and also the operational hours H_{HP} of installations equipped with GHPs, beyond what climate categories alone can capture. These aspects are not directly accounted for in the formulas.

Even with climate-specific default values, the inherent variability in geothermal system performance across similar climate zones due to sub-surface conditions suggests that a more specific and localised approach might be necessary for more accurate geothermal energy accounting. The default values, while practical, may not fully account for site-specific advantages or challenges faced by geothermal installations.

To address the limitations identified, it is essential, except in those situations involving standard residential energy load profiles, to systematically incorporate microclimate and geological assessments into the evaluation of GHP performance. These assessments are fundamental for gaining deeper insights into local conditions affecting GHP performance, which in turn can allow for a more precise estimation of default values.

Additionally, the integration of cooling-degree days into the calculation procedures for renewable cooling should be acknowledged. Both heating-degree days and cooling-degree days provide a measure of the demand for heating and cooling based on outside temperature fluctuations, which can be associated with the operational efficiency of GHPs.

Consequently, a *dynamic adjustment mechanism* for SPF and EFLH values based on local heating-degree days and cooling-degree day's data could in principle be more easily and quickly applied. This adjustment alone has the potential to allow for a more precise representation of GHP performance across different climates, at least in theory better incorporating the fact that the efficiency of GHPs in heating or cooling modes is influenced by the outdoor air temperatures.

c) Technology specificity

Geothermal systems encompass a range of technologies, as shown in **Sec. 1.2**. The methodology would benefit from greater specificity regarding different technologies, considering their unique operational modes and efficiencies.

For example, it is noteworthy that open systems, which directly utilise groundwater for heat exchange, typically outperform closed systems. Additionally, the potential for 'free cooling' by exploiting groundwater directly for cooling purposes, highlights a highly efficient mode of operation with minimal energy input.

Thus, more is needed to accurately capture the full potential of geothermal energy within the EU's renewable energy targets.

3.5. Final remarks

In light of the aspects discussed, it is unmistakeable that the existing methodologies provide a robust foundation for renewable energy accounting. Yet, the particularities of geothermal systems and the inherent complexity of heat pump operations would call for a more refined approach from a technical point of view.

The reliance on linear models and broad categorisations, though practical for standardisation across Member States, may not fully capture the intricacies of the various heat pump types. This is particularly true when considering the diverse operational modes and efficiencies of GHPs, which can vary depending on many factors such as local geology and microclimate conditions.

Specifically, for shallow geothermal energy systems, it is argued that changes are necessary to better account for the GHPs role in renewable heating and cooling, including:

i. **Improved data collection practices:** Focus on comprehensive collection efforts of empirical data to gather detailed information on the performance of GHPs across different geological and climatic conditions is key. This includes granular data on installation specifics, environmental, and long-term performance indicators. This information would provide a strong empirical basis for refining the calculation methodologies and setting more accurate default values for the required parameters in the accounting methodology.

ii. **Advanced reporting mechanisms:** There is an urgent need to standardise the reporting mechanisms to align advancements in geothermal technology and data generation more closely within the EU framework. Regulatory support might be needed to facilitate adherence with new reporting standards.

It should be emphasised that these two points listed above are inherently connected to the data requirements issue, which is the core aspect of the present report, and is addressed further in the next sections (**Sec. 4** and **Sec 5**).

4. Reporting situation in the partner countries

This part of the report reflects how data on GHP installations are currently generated and shared among the GeoBOOST partner countries, including **Austria**, **Germany**, **Ireland**, **Netherlands**, **Poland**, **Spain**, and **Sweden**. The analysis is focused on identifying the existing practices in data acquisition, the data parameters collected, and how this information is disseminated among stakeholders.

An additional enthusiasm behind this analysis was to formulate *high-level suggestions* for each country for continued progress. By identifying areas of challenge in current practices, it is hoped that these recommendations can provide some guidance for encouraging more refined GHP installation data management.

This analysis draws on responses from a tailored questionnaire (**Appendix 2**). The methodology underpinning the questionnaire is presented below followed by narratives for each country's analysis.

4.1. Questionnaire methodology

Structure

A structured questionnaire was designed to capture a wide range of GHP-related installation data. The questionnaire was divided into three major parts: *General Questions*, *Closed Loop Systems*, and *Open Loop Systems*. Each section was further categorised into sub-sections focusing on *Data Generation* and *Data Accessibility and Sharing*.

The questionnaire comprised both closed and open questions, enabling respondents to select predefined options for certain queries while offering detailed descriptions where necessary. This approach facilitated the collection of both quantitative and qualitative information.

Application

The questionnaire was disseminated electronically to a predetermined list of respondents (i.e., the project partners). We focused on one filled-in questionnaire by each participating country. The respondents have vast expertise in their countries' situation. Additionally, the selection of respondents covers a range of experiences, thereby providing valuable perspectives on the sector's data generation and sharing practices. The applied procedure prioritised efficiency and minimal constraints in reaching the intended audience. Follow-up interactions with respondents were conducted to clarify some responses and gather additional insights, with the goal of enhancing the quality of the information collected.

4.2. Austria

Overview of the current situation

In Austria, the management of data on both closed and open systems is significantly influenced by the regulatory framework, with the local and regional governmental authorities playing a central role. Data generation for the GHP installations occurs at the Federal States level. While all open systems require authorisation, warranting their inclusion in the Water Registries throughout Austria, closed systems are not uniformly documented due to different permit requirements. Both system types present inconsistencies in data completeness and availability, with geographical coordinates and commissioning year generally being more reliably documented. The absence of a unified national database complicates public access to data across the country.

Analysis of closed systems

Framework for data generation

Austria's regulatory framework plays a crucial role in the data generation practices for both closed and open systems. The primary GHP installation data is generated by local and regional governmental authorities. This structure aids in standardisation but also results in regional discrepancies.

For new installations requiring a license or permit, designated as a "water right," district authorities are typically responsible for issuing these and forwarding the details to a Water Registry (*Wasserbuch*) organised at the federal state level. However, some federal states gather this information directly. In regions where BHEs are exempt from permit requirements, they do not qualify as 'water rights.' As a result, BHEs are not listed in the Water Registry, leading to a lack of systematic documentation. Closed systems not assigned a water right are often recorded in the *Baugrundkataster*, an ample drilling registry, though not uniformly across all Federal States.

The Water Registries of the Austrian Federal States can be accessed here: https://info.bml.gv.at/themen/wasser/wisa/datenverbund/wis-bl.html

Data parameters

Data available from the Water Registries typically include technical specifications and operational metrics, such as geographical coordinates, year of construction, installed capacity, annual thermal work, and specifics regarding BHEs. However, completeness varies significantly; while location and installation year are more consistently recorded, other parameters are less reliably documented. Data completeness and availability also vary by state. Additionally, the data on horizontal collectors is incomplete and sparsely collected. It should be noted that the capacity and annual thermal work data can be found in official licensing documents (*Bescheide*), however not typically digitalised and transferred into the online Water Registries.

Variations in data availability suggest that different Federal States may have different priorities and capabilities for gathering and maintaining data on such installations. This regional variability has important implications for conducting comprehensive national-level analyses. A lack of standard templates or guidelines to standardise data collection has been noted, which could help explain the issues with overall consistency and completeness of the data parameters collected.

Data accessibility and sharing

Data accessibility is considered moderately easy because it is fragmented across Austria's nine Federal States, thereby complicating uniformity and ease of use. While the general public can access primary generated data, the process may involve specific requests to authorities, which can be considered a barrier to straightforward access. This level of accessibility, combined with regular updates of the Water Registries (weekly to monthly), supports transparency. But it remains cumbersome for everyone seeking comprehensive and direct access.

Analysis of open systems

Framework for data generation

Local and regional authorities are responsible for primary data generation of open systems. On the contrast to closed systems, all open systems require a "water rights" license, warranting their inclusion in the Water Registries. It is notable however that the generated data relates to groundwater abstraction, not specifically groundwater heat pump systems.

Data parameters

Due to the mandatory licensing requirement, data collection is better for open systems than for closed systems. This shows the importance of the licensing procedure. Data parameters collected for open systems include geographical coordinates, year of construction, installed capacity, annual thermal work, flow rates, water temperatures, and specifics about the wells. Note that the capacity and annual thermal work data are stated in official licensing documents but have largely not been digitised for the online Water Registries.

Geographical coordinates and year of construction are the parameters which are more consistently recorded. There is considerable regional variability in how completely other parameters are documented, which poses challenges for maintaining uniformity across the national level. Furthermore, it is important to stress that the licensed flow rate data represent the upper limits (maximum permitted values). The actual operational values can evidently differ to lower values.

Another challenge within the available flow rate data is distinguishing the specific uses of pumped water. Firstly, it is often difficult to accurately separate the share of extracted water used for heating, cooling, or both. Secondly, where pumped water is additionally used for other purposes such as irrigation or as process water puts further complications in differentiating the flow rates according to the purpose. Capturing these data parameters with sufficient accuracy is important due to possible thermal or hydraulic interference with new or neighbouring installations.

Data accessibility and sharing

The accessibility of data for open systems is moderate, however better than for closed systems. The general public can access the primary data, but it often involves submitting specific requests to multiple authorities. This fragmentation highlights the need for more streamlined processes and platforms to facilitate easier access and use of data. Additionally, the absence of standardised templates or guidelines indicates broader issue within the sector regarding the lack of unified methodologies for data generation and sharing.

Recommendations

To address the identified gaps, the following suggestions are proposed:

a) **National data templates for both systems:** Develop and implement uniform data templates or guidelines for registering and generating primary GHP installation data. These templates outline the procedures, methods, and best practices for gathering data, including a definition of parameters and quality control measures. Templates can simplify the integration of the generated data from potentially different sources.

b) **Centralised national database for both systems:** Start a centralised national database where all data pertaining to GHP installations can be stored and freely accessed. This repository should be designed with user-friendly interfaces to accommodate various stakeholders. This approach could help overcome the challenges posed by the current decentralised data generation model.

c) **Retrospective documentation for closed systems:** In those Federal States without mapped BHE installations, a similar tactic as done in Sweden could be applied. That is, property owners are required to report their GHP installations to local authorities in Federal States lacking current data. This could backfill data gaps and support the integrity of a centralised national database.

4.3. Germany

Overview of the current situation

In Germany, the generation of primary data for closed and open systems primarily involves local and regional governmental authorities, supplemented by research institutions, drilling companies, installers and operators. This decentralised system, largely driven by the specific water laws of each Federal State, leads to high variability in GHP data management practices. While the diversity of data sources can be considered a strength, in the case of Germany poses significant issues for data harmonisation. Both system types have limitations in data completeness and availability. Only the coordinates of the installations is usually more reliably documented. The absence of a national database complicates access to comprehensive data on GHP installations throughout Germany.

Analysis of closed systems

Framework for data generation

The primary data on closed systems in Germany are generated mainly by local and regional governmental authorities, in addition to research institutions. This represents a layered governmental involvement in data generation, exposing a structured yet decentralised collection framework. The data generation procedure cover entries by authorities into databases, academic research contributions, and submissions by drilling companies based on licensing or permits. Although the diversity of data sources can be a strength, it poses challenges in standardisation and data harmonisation across different bodies.

While an arranged data generation process is apparent, it is based on different regulatory influences. The presence of templates or guidelines, which are specific to certain regions or vary by organisation, mirror the decentralised nature of regulatory compliance. This variability reflects the regional responsibility for Water Law, leading to a lack of a unified national framework. Consequently, the regulatory framework for closed systems is considered to be fragmented across Germany. This is likely a major factor impacting the uniformity and integration of data generation practices across Germany.

Data parameters

Technical and operational parameters collected for closed systems include geographical coordinates, year of construction, probe types, operation modes, and specifics about boreholes, along with antifreeze types and Thermal Response Tests (TRTs). However, it is noteworthy that certain parameters, such as the total and average depth of boreholes, are inconsistently documented. Likewise, in general, there are no specific quality criteria, and the available data is heterogeneous, affecting its comprehensiveness.

Moreover, complications with missing historical data were noted. It seems that older installations are not documented as thoroughly as desired. While some data is available at regional geological surveys or local water administrations, it might not consistently cover all older installations.

Data accessibility and sharing

Data on closed systems is accessible to the public, reflecting a commitment to transparency. For example, Bavaria offers access through its environmental atlas (www.umweltatlas.bayern.de). However, obtaining detailed data typically requires specific requests to authorities (for all further data besides the location), with operational data being even more challenging to obtain. This scenario indicates moderate data accessibility. It also points towards an area for improvement, aiming for more efficient and comprehensive access to GHP installation data.

Analysis of open systems

Framework for data generation

Data on open systems, similar to closed systems, is produced by local and regional authorities, as well as installers and operators. This collaboration among various stakeholders shows the involvement of governmental bodies at multiple levels, due to the decentralised and layered regulatory oversight, and the industry's direct involvement in documenting installation details. This process of data generation, influenced by regional and local regulations, emphasises the significance of regional control due to water law being a regional responsibility.

Data parameters

Data for open systems includes well details and thermal characteristics of the water used, with a focus on flow rates and inlet and outlet temperatures. However, the generation and availability of data on open systems are markedly influenced by the size of the system and regional practices. For systems with an abstraction rate equal to or above 100,000 m^3 $year^{-1}$, more detailed data are reported on the basis of annual reports. This data encompasses average flow rates, peak load flow rates, and temperature levels for both average use and peak load conditions. Conversely, for smaller systems, only permitted values are documented, indicating a significant reduction in data detail for these installations. Temperature data, including average inlet and outlet temperatures, is variably recorded, with specific temperature levels noted for average and peak load uses. However, this information is not consistently available across all installations, particularly for average outlet water temperatures, which are only sometimes documented.

Data accessibility and sharing

Open access to the primary data generated is beneficial as it indicates transparency. However, there are still challenges in accessing operational data, as previously noted. The requirement for specific authority requests to obtain detailed information shows the need for more effective data sharing mechanisms. Furthermore, the variation in templates and guidelines across different organizations and regions points to a fragmented approach. This, along with the absence of a national framework, highlights the potential difficulties in standardising data collection and sharing procedures.

Recommendations

It is argued that addressing the challenges of data standardisation, harmonisation, and accessibility in Germany requires a multi-faceted approach. Yet, the following recommendations are suggested:

a) **National data templates for both systems:** Develop and implement national data templates for both systems to standardise data generation and improve integration across different sources.

b) **Centralised national database for both systems:** Establish a centralised national database to store all data related to GHP installations, increasing accessibility and usability for all users.

4.4. Ireland

Overview of the current situation

In Ireland, the current generation of primary data for both closed and open systems lacks a proper regulatory framework, leading to reliance on unofficial data from industry associations and installers/operators. This situation results in varied and incomplete data collection across both systems. Closed systems particularly suffer from a lack of official data generation, with parameters such as construction year, installed capacity, probe type, and operation mode made available from some entities. Open systems, although subject to a more regulated environment due to their interaction with groundwater, still rely on unofficial data for further data (e.g., capacity, annual thermal output). Official records concentrate on coordinates, license application dates, water abstraction rates, the type of abstraction, and how thermally used water is disposed of. The expected regulatory changes aim to improve data standardisation and accessibility, suggesting a future shift towards more formalised data generation and reporting practices for GHP installations in Ireland.

Analysis of closed systems

Framework for data generation

The current practice in Ireland involves no official data generation by governmental authorities for closed systems. In the absence of comprehensive regulatory mandates, there is a reliance on unofficial data sourced from industry associations and installers/operators.

The current state, characterised by a lack of standardised data generation protocols, stresses the significant gap in Ireland's GHP data management. The anticipation of regulatory changes suggests an evolving context that could address these gaps. As a consequence, a move towards more formalised data collection and reporting practices is expected.

At this point, a brief overview of the current state and anticipated changes in the data reporting framework for geothermal energy systems in Ireland is helpful. The landscape is shaped by contributions from several key entities. Notably, the Sustainable Energy Authority of Ireland, a vital national body, has been documenting installations of heat pumps since the introduction of grant funding in 2007, even though without examining subsurface details. Geological Survey Ireland have been developing data and information on shallow geothermal energy resources through a series of public web viewers and guides for end users since 2015.

In parallel, Irish Water, the Local Authorities alongside the Environmental Protection Agency oversee the documentation of substantial groundwater abstraction and re-injection licensing systems, through which larger scale heat pump systems are permitted. Furthermore, the Geothermal Association of Ireland maintains a selective record of commercial closed systems, offering insights into their installed capacity, location, and system type since its inception in 1998. However, this repository lacks detailed operational data, relying instead on generalised estimates of energy output. Additionally, the Heat Pump Association of Ireland contributes to this dataset by tracking heat pump sales, with these figures subsequently integrated into the broader EHPA database from circa 2018 onwards. The anticipated regulatory updates aim to enforce comprehensive registration and detailed annual monitoring for selected systems of a defined size. This evolution signals a shift towards standardised data reporting and a more complete understanding of Ireland's geothermal energy infrastructure.

Data parameters

As authorities do not generate official data, only unofficial data related to specific parameters such as construction year, building type, installed capacity, probe type, and operation mode are available. Thus, these parameters are collected in a setting where official databases and standardised reporting are missing. Yet, it shows the importance of industry associations in compiling and disseminating data at least about some aspects of closed systems.

Data accessibility and sharing

The current practice of updating available data annually, coupled with its poor accessibility, indicates major challenges in sharing and using this information. Even though industry associations play an important role in data generation and sharing, the restricted access and reliance on technical reports limit the availability and usability of data for stakeholders across the sector.

Analysis of open systems

Framework for data generation

Open systems operate under a slightly more regulated environment, particularly concerning groundwater abstraction. However, there is a similar reliance on installers and operators for data beyond the scope of current regulations. This dependence on unofficial sources for additional data implies a regulatory shortcoming in capturing the broad scope of information relevant to open systems. Hence, the need for regulatory advancements to integrate both data sources is clear.

Official data is generated by local and national government authorities and independent consultants, primarily to meet regulatory compliance. It should be stressed that the official data generated is not specific to groundwater heat pump systems, but to groundwater abstraction licenses. Meanwhile, installers and operators contribute unofficial data, which adds operational insights.

Data parameters

Unofficial data detailing operational aspects of open systems include installed capacity, annual thermal output and operation mode. Official records are focused on geographical coordinates, year of construction, water abstraction rates, type of abstraction, disposal of the thermally utilised water and building construction year (license application date). The available official records on water abstraction are limited, with no comprehensive data on the specifics of open systems. The dual-path process to data generation suggests a situation where unofficial data seems to target the gaps left by official sources.

Data accessibility and sharing

High accessibility and user-friendly nature for the official data was noted, as supported by online databases and publicly accessible websites. This is a strength in data dissemination practices. The register of groundwater abstraction licenses is maintained by the Environmental Protection Agency. The update frequency is annually, which gives a relatively current dataset. Unofficial data is maintained by the Geothermal Association of Ireland is also reported on an annual basis. So, the dual-path data generation process impacts data accessibility; official data might be constrained by regulatory dissemination channels, while unofficial data could provide broader, though potentially less regulated access.

The presence of national templates or guidelines for data generation in open systems suggests a more structured approach compared to closed systems. This approach, mandated by national regulations for water abstraction, indicates a higher level of regulatory oversight and potential for data quality assurance in open systems. However, the limited scope of official records, focusing mainly on water abstraction, confirms the gaps in capturing the diverse operational parameters of these systems.

Note that the regulatory focus on water use is structured around a three-tiered registration and licensing system established by the 2018 EU regulations. Important, official groundwater abstraction records available on the internet through a web viewer are limited to 2018, indicating a historical gap in the generation of primary data for open systems. Older records exist but are less easily accessible. Upcoming changes in licensing and regulation are expected to bridge the data gaps.

Recommendations

To address the identified gaps and enhance data generation and sharing practices in Ireland, the following measures are suggested to address foundational challenges in both closed and open system:

a) **Unified data management and reporting framework:** The need for a coherent approach to data collection across both closed and open systems is apparent. Combining improvements in data generation processes, accessibility, sharing practices, and the integration of unofficial data sources into a unified data management and reporting framework would address many of the identified issues. This comprehensive framework could include the development of a dedicated national database for providing data on GHP installations.

b) **Enhanced regulatory framework:** The anticipation of a new regulatory framework for the deployment of GHPs highlights an opportunity to align and integrate this with a broader regulatory structure for geothermal resources. This should not only focus on the deployment aspects but also encompass data generation, quality assurance, impact assessments, and operational monitoring. It is argued that new rules, if developed alongside better data generation and reporting practices, will have a greater impact.

4.5. Netherlands

Overview of the current situation

In the Netherlands, data management for both closed and open systems is characterised by a well-structured yet evolving framework. Both types of systems benefit from a cooperation between authorities and industry stakeholders like installers and operators. This many-sided approach to data generation, leveraging regulatory compliance and active stakeholder participation, is standardised through national templates and guidelines. For closed systems, data generation is systematic, involving licensing requirements, with smaller systems up to 70 kW requiring only a notification, whereas larger systems need a full license. Many parameters are collected from geographical coordinates to technical and operational specifics. However, data collection obligations for closed systems began in 2013, indicating a gap in historical data coverage. Although open systems have long been subject to licensing requirements, obtaining comprehensive data on these installations is more challenging due to decentralised data management. Data on open systems is not centrally stored or readily accessible in databases, demanding engagement with various authorities to gather detailed information.

Analysis of closed systems

Framework for data generation

Local governmental authorities and installers/operators are identified as the primary generators of data for closed systems in the Netherlands. The data generation process is facilitated through a combination of direct entries by authorities into databases and submissions by drilling companies based on licensing requirements. This is complemented by front-end user interfaces for installers to submit data, highlighting a multifaceted strategy to data collection that leverages both regulatory compliance and stakeholder participation. National templates and guidelines standardise this process across the country, ensuring uniformity in data capture. This standardisation can address potential disparities in data quality and reliability, providing a more solid foundation for accurate and consistent data generation.

The regulatory framework is shaped by a combination of national, regional, and local regulations, illustrating a comprehensive and layered approach to oversight. Data collection obligations that began in 2013 point towards an evolving regulatory environment, with a focus on enhancing the scope and depth of data over time. Smaller closed systems up to 70 kW require notification, while closed systems larger than this size must apply for a license. This however should not affect the specific data parameters generated for closed systems. However, considering that the historical coverage remains a gap, a retrospective documentation would be beneficial if one it is interested in deriving insights into long-term market trends.

Data parameters

A wide array of parameters is collected. These comprise geographical coordinates, year of construction, number of boreholes per installation, operational mode, operational capacity, and other technical specifications like antifreeze types and concentrations. The dual approach with involvement of different stakeholders from the outset implies a collaborative effort in data generation and a potentially robust mechanism for data accuracy and comprehensiveness.

Data accessibility and sharing

Data accessibility and sharing practices highlight an open-access approach, with the general public granted access to the data. The principal providers of the primary data generated are local and regional governmental authorities by online databases with user query capabilities or by specific requests to authorities, reportedly in a highly accessible and user-friendly manner. Yet, the frequency of data updates is not specified, suggesting potential variability in data currency and availability. It should be noted that the main database for installations for both closed and open systems (www.wkotool.nl) also provides information on regulations such as zones with drilling-depth limitations or ground-source energy plans.

Analysis of open systems

Framework for data generation

Open systems exhibit a parallel structure in data management, involving regional governmental authorities and installers/operators as the primary data generators. This indicates a more regionally based approach to data generation compared to closed systems.

Similar to closed systems, the data generation process is supported by regulatory compliance mechanisms and stakeholder submissions. However, regional specificities influence the adoption of templates and guidelines, which may reflect particular considerations about potential environmental impacts of these systems across different locales. Implementing uniform national templates and guidelines, particularly tailored to address the unique challenges of this system, can help to reduce variability and enhance data quality.

The regulatory framework for open systems is defined by a combination of national, regional, and local regulations, with a notable emphasis on environmental impact studies. This underscores the stringent regulatory compliance requirements specific to open systems, aimed at safeguarding environmental integrity.

Data parameters

The scope of parameters collected is broad, covering not only installation and operation specifics but also specific technical metrics like flow rates at various units and inlet/outlet temperatures.

Data accessibility and sharing

Data accessibility for open systems, while still committed to transparency, encounters some challenges. Due to their direct interaction with groundwater resources, open systems have long been subject to licensing requirements. However, the management of installation data reveals a layered scenario; although licenses are mandatory, comprehensive data on these installations is not centrally available or easily accessible in any database. This often requires stakeholders to engage directly with authorities for insights into regional or neighbourhood-specific installations, pointing to a decentralised approach to data management.

Recent efforts by many Dutch authorities to update online databases with historical data, at least concerning the location of systems, suggest a move towards improving accessibility. However, the variation in practices across different authorities and the partial availability of historical data suggest a "*grey zone*" in open systems data management, with implications for the understanding and evaluation of the sector's historical development.

In general, the accessibility for open systems has been identified as moderate, pointing to the need for improvements in user-friendly access and data sharing practices. The irregular or as-needed basis for data updates further complicates the timely availability of information, which stresses a need for more steady data management practices.

Recommendations

Specific areas have been identified where targeted improvements could enhance the already advanced Dutch data management of both closed and open systems. These suggestions have been tailored to address the unique challenges and operational specifics of each system type.

a) **Specificity in SPF reporting for closed systems:** The current practice of providing the Seasonal Performance Factor (SPF) as an overall average across systems is respectable but there is room for enhancement. Specifically, for closed systems, SPF could be distinguished by different operating conditions, such as space heating, domestic hot water production, and space cooling. This adjustment would allow for an improved understanding of system efficiency across various conditions, with implications for example for renewable energy share accountability.

a) **Permitted volume vs. actual use in open systems:** A noted discrepancy between the permitted volume (e.g., m^3 $year^{-1}$) and actual use in open systems presents a chance for advance. For example, this issue has connections with the assessment of thermal influence areas. Enhancing the faithfulness of data reporting in this regard can support environmental management, regulatory adherence, and the optimisation of resource use.

b) **Data accessibility and update frequency for open systems:** Efforts to enhance the user-friendliness of online databases and interfaces has been specifically noted for open systems. Establishing a consistent schedule for data updates is also recommended.

c) **Expanding historical data coverage for both systems:** Initiatives to document pre-2013 installations could be beneficial for both closed and open systems. Filling the gaps in historical data could provide valuable insights for policy development, system optimisation, and studies looking at long-term market trends.

d) **Addressing human resource constraints for both systems:** The identified limitation in human resources affects both closed and open systems, suggesting that targeted investment in training and capacity building could be done. A well-trained workforce is fundamental to the effective generation, management, and analysis of GHP installation data.

4.6. Poland

Overview of the current situation

In Poland, data management for both closed and open systems is primarily driven by regional and national governmental authorities, supplemented by industry associations. However, there are notable challenges in regulatory compliance, especially concerning documentation requirements for closed systems. While standards and guidelines are employed nationally, there are significant gaps in regulatory compliance and data completeness, as indicated by the limited submissions to the National Geological Archive. Similarly, for open systems, data generation is under governmental oversight, with a reliance on hydrogeological documentation submitted to local Water Authorities. But there are also important information gaps, particularly regarding the geothermal usage of wells. Substantial data gaps in the documentation within the National Geological Archive of Poland highlights a critical area for improvement.

Analysis of closed systems

Framework for data generation

Regional and national governmental authorities in Poland, supplemented by industry associations, predominantly generate the primary data on closed systems. This indicates a structured but potentially fragmented approach to data collection.

There are challenges in regulatory compliance, especially regarding the residual documentation in the National Geological Archive. This points to a need for enhanced enforcement and incentives for comprehensive reporting. More in detail, in Poland boreholes used for ground heat extraction in closed systems are legally classified as geological activities. Thus, those extending beyond 30 meters are mandated to provide detailed as-built documentation in accordance with regulatory requirements. Regrettably, this mandate is not universally applied, leading to a situation where only a select few submissions (such as those associated with governmental, credit, or grant requirements) are actually made. Awareness of this requirement is growing, albeit slowly and incompletely. Research conducted by the Polish Geological Institute indicates that the proportion of documented installations, out of all ground source heat pump systems sold up to 2020, stands at approximately 2.6%. This figure has modestly increased to 5% in 2022. Often, even when as-built documentation is provided to regional authorities, it is not consistently forwarded to the national level (National Geological Archive).

Furthermore, the geological administration’s role in ensuring the completeness and correctness of documentation is strictly limited to instances where the provided geological documentation fails to meet legal standards. There lacks a proactive approach towards enhancing the quality and reliability of this documentation; the focus remains solely on legal compliance.

Standards or guidelines are employed nationally, offering a uniform approach to data documentation. However, the mention of residual documentation and the limited submission to the National Geological Archive reveal significant gaps in regulatory compliance and data completeness.

Data parameters

The available data encompasses different parameters, from geographical coordinates to system-specific details like the type of antifreeze solution used. As noted before for other countries, the focus on both operational and technical parameters suggests a dual interest in understanding the geographic distribution of systems and their technical specifications.

Data accessibility and sharing

Data accessibility and sharing are marked as strengths, with general public open access provided through various platforms. Yet, the overall reliability of the primary data generated on closed systems is rated as poor, reflecting significant room for improvement in both the collection and sharing of data.

Data is broadly accessible to the public through open data portals and websites, a strength that promotes transparency and stakeholder engagement. Nonetheless, the poor overall reliability rating of the data generated raises concerns about the accuracy and utility of the information available, suggesting areas for improvement in data validation and update frequency.

Analysis of open systems

Framework for data generation

Similar to closed systems, the primary data for open systems is generated by regional and national governmental authorities. This reflects a common approach across different system types in Poland, which suggests that data collection remains under governmental oversight.

The drilling of a well and the water permit are important, codified and legally required documents. Accordingly, there is an identified reliance on hydrogeological documentation for primary data generation. Such hydrogeological documentation is done by drilling companies or geologists on behalf of the end-user and submitted further to the local Water Authority. This highlights a similar third-party dependency as seen with closed systems. This process is beneficial for leveraging specialised expertise, but may suffer from the same challenges related to completeness and compliance.

Data parameters

The parameters for open systems include both operational metrics like flow rates and technical specifications such as well depths, number of abstraction and injection wells, and type of abstraction (where is water taken from, e.g. surface water or groundwater).

The adherence to national templates or guidelines for open systems mirrors the approach taken with closed systems. However, there is a major information gap regarding the geothermal usage of wells.

Data accessibility and sharing

Although the data is available to the public, moderate accessibility issues indicate that improvements are needed in both the ease of access to data portals and the completeness of the data provided. The irregular update frequency further impairs these challenges, affecting the timeliness of the data available.

Recommendations

The suggested solutions to address the identified issues in both closed and open systems in Poland are more of a fundamental nature. While additional measures like a national database could be advocated, it is understood that preliminary steps are first necessary:

a) **Enhanced compliance and documentation submission:** The residual nature of the documentation in the National Geological Archive of Poland for both system types is a major area for progress. Increasing compliance through stricter enforcement of submission requirements and incentivising widespread reporting could address this issue. Providing clear, accessible guidelines and support for stakeholders involved in documentation processes may also be beneficial.

b) **Improvement in data generation processes:** Developing mechanisms to directly collect detailed installation data from operators could mitigate the reliance on third-party submissions, closing at least some of the gaps in data coverage and the reliability of the data collected. This could involve digital platforms that simplify the submission process, coupled with verification processes to ensure data accuracy.

c) **Standards or guidelines re-evaluation:** While national templates or guidelines are in place, the noted challenges in data quality and regulatory compliance suggest a need for their re-evaluation. Updating these guidelines to address current challenges and ensure they are effectively communicated and implemented across all relevant entities would be important.

4.7. Spain

Overview of the current situation

The present narrative shows Spain's use of a statistical, census-based approach for deriving total GHP installation counts with a focus on renewable energy accountability. While this approach is clever and serves its purpose, Spain could enhance its geothermal energy data management practices, drawing inspiration from the peer countries. A strategic shift towards a "first principles" approach, akin to approaches observed in the other investigated countries, could streamline data generation and sharing procedures. Without access to comprehensive GHP installation data, conclusions cannot be drawn on the temporal development of the systems, including potential changes in well and borehole depths and in system size, and the spatial spread of the installations.

The current narrative for Spain is based on a different arrangement compared to the other analysed countries. The narrative reflects Spain's focused effort to know the role of GHP into the broader national energy mix. It seems that the Spanish approach is mainly driven by obligations with renewable energy accountability. This approach is detailed in Section 2.2 of the document: "*IDAE notebooks. Statistics: Methodology used in heat pump statistics*". Here, we only give an outline of Spain's approach in light of our objectives.

The initial stage involves gathering raw data on GHP installations across Spain, using a census-based method. This method has been chosen due to the sector's low market penetration and the fragmented nature of the installations. The reference period of the information starts in 1992. The research unit to which the data refer is the heat pump installation. The reporting units comprised several stakeholders (e.g., Renewable Energy Departments of Autonomous Communities, regulatory bodies, industry experts, and manufacturers). It has been reasoned that this broad engagement helped to achieve a thorough aggregation of data, although with challenges in information standardisation and duplication avoidance.

Key variables included the type of heat pump (geothermal or hydrothermal), commissioning date, installation characteristics, and energy output metrics. This data collection necessitated data cleaning and inference methods to address informational gaps and inconsistencies.

Following data collection, a cleaning process is undertaken to attempt the integrity of the information. This involves eliminating duplicates, verifying the accuracy of installation details, and rectifying inconsistencies. The document also outlines a process for cross-referencing data against official records and industry reports. This highlights the challenge of maintaining data quality in the face of diverse information sources. For data points that are missing or incomplete, the methodology employs inference techniques to fill gaps. This involves for example the use of technical catalogues and assumptions based on similar installations.

At one hand, one of the distinctive aspects of Spain's approach was the rigorous methodology applied to infer data for missing fields, relying on a mix of official sources, technical catalogues, and statistical assumptions. This methodology thus relies on a combination of empirical data and educated guesses to realise the status of the GHP situation. At the other hand, despite these efforts, the challenge of ensuring data completeness and accuracy remains a difficulty, making the case for improved data generation and standardisation practices.

Despite its own challenges, the recording of installations from "*first principles*" as typically done in other investigated countries has the potential to provide a more foundational and accurate basis for understanding the market.

a) **Adoption of a "first principles" approach to primary data generation:** The sector in Spain could benefit from the adoption of a "first principles" approach to GHP installation data generation, like the practices observed in other investigated countries. This involves collecting data at the most fundamental level, starting with the mapping of the installations and their year of construction. It also implicates distinguishing between closed and open systems and capturing a range of operational and technical parameters for each installation. By obtaining detailed information from the outset, a more solid foundation for understanding the market's dynamics can be achieved.

4.8. Sweden

Overview of the current situation

In Sweden, data management for both closed and open systems is shaped by a robust regulatory framework dating back to 1976, mandating compulsory reporting from drillers. Namely, primary data generation for both systems is anchored in this regulatory mandate. The Geological Survey of Sweden (SGU) manages a national database (*Brunnsarkivet*) to store all information. Many parameters are consistently collected for both closed and open systems. However, there are chances for extending the data collection of open systems to include more operational and technical details, similar to closed systems. Systematic differentiation regarding whether an installation is operating for heating, cooling, or under both modes, could also be improved. Despite this structured approach, challenges persist particularly regarding coordinate precision and historical documentation.

Analysis of closed systems

Framework for data generation

In Sweden, the infrastructure for closed systems exhibits an ample approach towards data generation. There is an active involvement of installers and operators as the primary data generators. This reflects a decentralised yet coordinated mechanism for capturing system information at the point of installation and operation. Mandatory reporting from drillers occur since 1976. The generated data lands in the SGU's database, known as *Brunnsarkivet* (https://apps.sgu.se/kartvisare/kartvisare-brunnar.html).

Regulatory compliance is shaped by a synergy of national, regional, and local regulations. This synergy may act to ensure that data generation practices are aligned with legislative requirements across different governance levels. Also, this multi-tiered regulatory approach may contribute to increase the consistency of the data collected. The process of data generation is homogenous by means of national templates and guidelines, which serve as a basis for bringing uniformity in the information captured across installations.

SGU is the responsible authority in Sweden providing open access geological data of rock, soil and groundwater for the public. As mentioned, all boreholes or wells that are drilled must be registered in the SGU's database. Meanwhile, municipalities provide supplementary data, especially for closed systems. Importantly, property owners must notify local authorities of their geothermal installations.

The SGU's database faces challenges with location accuracy (100 m accuracy, or higher), particularly for older installations. The data of the SGU mainly feature an estimated indication of location, typically positioned at the property's centroid. In instances where a property contains several wells or boreholes, these are frequently recorded under a singular location coordinate. The municipalities' records often complement the records in the SGU's database by offering more precise location details, which are important for planned installations. The practice of cross-referencing data between the SGU and municipal databases (when there is information from both) mitigates some of the accuracy issues, ensuring a more reliable dataset for planning. But this is usually done for small regions where a single project takes place. It is has yet to be realised at the national level.

Moreover, despite the obligation for drillers and project consultants to report completed installations, there's an acknowledgment of an estimated 20% of installations, predominantly older ones, missing from the database (Juhlin, 2016). This gap underscores the ongoing need for enhancing the completeness of the registry, a challenge compounded by the sometimes reduced location accuracy in the SGU's records.

Data parameters

The data parameters collected are varied and tailored to encapsulate the technical and geographical specifics of the installations, including geographical coordinates, year of construction, installed capacity, probe type, area of horizontal collectors, number of boreholes, borehole diameter, total and average depth of boreholes, antifreeze type, and antifreeze solution concentration.

Data accessibility and sharing

Data accessibility is relatively straightforward, with the SGU database being a valuable resource openly accessible to the public. The frequency of data updates on a weekly and monthly basis indicates a mechanism for keeping the database current.

Analysis of open systems

Framework for data generation

Parallel to closed systems, open systems in Sweden benefit from a well-arranged framework for data generation, involving local and regional governmental authorities alongside installers/operators. The approach behind data generation for open systems is anchored in national templates or guidelines, mirroring the structured approach observed in closed systems. There is thus apparently a consistency in standards across system types, with implications for more high-quality and reliable data. The regulatory framework is predominantly dictated by national regulations.

Data parameters

The scope of parameters collected for open systems includes geographical coordinates, year of construction, well depth, flow rates (in $l\ h^{-1}$), groundwater level, and soil depth. However, there are opportunities for broadening the data collection to include more operational and technical details, akin to closed systems. For example, currently there seems to be no clear differentiation whether an installation is operating under heating, cooling or both modes.

Data accessibility and sharing

Data on open systems is also made openly accessible in the SGU database (*Brunnsarkivet*). Finally, it should be noted that the SGU database serves as a resource for evaluating prospective drilling sites for new installations. Insights into soil depth, rock type, and groundwater levels from nearby installations contribute essential country-wide data for selecting well/borehole design and execution methods. In addition, this information gives a more reliable estimation of drilling depth and costs.

Recommendations

Potential areas for improvement that have emerged from Sweden’s analysis are given below for both closed and open systems.

a) **Data completeness and accuracy:** Initiatives aimed at closing the gap in unreported installations, combined with efforts to improve the geographical accuracy of the SGU's database, could be prioritised.

b) **Data integration:** Developing a more coordinated mechanism for integrating SGU and the available data from the municipalities could streamline access and enhance the utility of the information available to a great extent.

5. Data template sheets

5.1. Objectives of the data sheets

The data sheets for the GHP systems provided herein serve as templates showing:

i. Parameters that we recommend to be generated or collected in at least state-wide or better country-wide databases.
ii. How the generated data could be shared with third-party users.

The purpose of the tables in these datasheets is to ensure that the shared data is consistent and clearly structured. Once the data is made available in this format, it is almost ready for further analysis, reducing the time and effort required for initial data processing. It is important to note that this data should be made publicly available without any associated costs.

It should be noted that the datasheets are not intended to be an exhaustive list of all possible parameters that could be collected. The parameters in these datasheets should be seen as recommendations, which may be expanded upon according to particular needs or interests.

5.2. Documentation separated according to main systems

The individual sheets are adapted to the main GHP systems:

- **BHE:** Borehole heat exchangers.
- **GWHP:** Groundwater heat pumps.
- **HOR:** Horizontal collectors (e.g., normal horizontal collectors, slinky collectors, earth baskets).
- **TAG:** Thermoactive geostructures (e.g., energy piles, thermally activated underfloor elements and diaphragm walls).

5.3. Structure of the data sheets

Tables can be seen as a two-dimensional format to store data. There are two main table structures which are widely used: the wide and long data formats. These offer distinct advantages and limitations depending on the nature of the data being processed and the specific requirements of the analysis at hand.

In the wide format, data is spread horizontally, with separate columns representing different variables and each row typically representing a unique entity. The long format arranges data vertically, stacking observations under a unified set of variable columns, which can include multiple entries for each entity.

The wide format is particularly useful for displaying all relevant information about one entity in a single row. This makes it easier for readers to understand the context of each entity's data without needing to scroll down across multiple rows. Therefore, the presented data sheets are mainly, but not only, based on the wide format.

Presently, a majority of statistical and data processing software packages can be used to convert between wide and long data formats. Additionally, given that the data template sheets should incorporate geographical coordinates and their corresponding reference systems, the transformation of data from tabular representations (e.g., CSV) to geospatial vector data formats (e.g., SHP) is relatively uncomplicated.

In the context of the previous paragraphs, the "***entity***" represents a geothermal installation as a whole or a system component. Entity as used here should not however be confused with the same term used in **Appendix 3** for another meaning. So, depending on the level of detail being addressed:

- A geothermal installation as a whole refers to the entire setup of a BHE or a GWHP system, encompassing all its components and the site it occupies.
- A system component refers to a part within that geothermal installation. These parts are separate wells and boreholes that are described in more detail for example by stating the depth and diameter of a particular borehole, the material of the probes, and so forth.

We have included an Excel file as an **Annex 1** to this report to demonstrate how the different catalogued parameters can be reported to end users. This file features two sheet formats for the most common system types, BHE and GWHP, which end with *_1* & *_2*.

5.4. Format_1

The tables belonging to ***Format_1*** employ elements of both wide and long formats to capture comprehensive data about geothermal installations and their individual components within a single structure. This format is uniquely suited for datasets where both installation and component information are important.

Also, this format offers detailed component tracking without losing sight of the installation-wide perspective. Such a format is suitable for studies that require granular details of geothermal installations, e.g. on the location of the individual system components.

- *Column labelling:* To distinguish between the levels of data, column names are prefixed with 'installation' or 'component'.
- *Installation-level data:* Columns prefixed with "installation"are populated only once per installation instance (i.e., in the row corresponding to the first borehole of an installation) and is left as blank or marked as `NA` in subsequent rows pertaining to the same installation. This prevents the redundancy of installation-wide data across multiple component entries.

- *Component-level data:* Columns with the "component" prefix include physical characteristics, materials used, and other technical features that could vary from one system to another.
- *Notes column:* An additional column is included for inserting comments regarding either the installation or specific system components. It gives some sort of flexibility for the inclusion of extra information such as geological conditions or peculiarities currently not covered by the catalogued parameters. However, from the general idea given, it is not demanding to expand the tables to include more parameters if desired.

5.5. Format_2

The tables belonging to ***Format_2*** are built in a "pure" wide format. In this case, each row represents a unique entity (in this case, an installation identified by *installation_id*), and different variables related to each entity are spread across multiple columns. This includes identifiers, measurements, and attributes specific to each installation, such as *installation_name*, *x_coordinate*, *y_coordinate*, *commissioning_date*, *status*, and various technical specifications related to the heat pumps and their operation.

Thus, this format offers a summarised outline of the geothermal installation. It simplifies the documentation and sharing process because detailed information on each constituent part of the system is not present. This aggregated data presentation is particularly advantageous for overviews or comparative analyses of geothermal installations focusing on general system characteristics rather than the minutiae of component-specific data.

5.6. Recommendation

For reasons of completeness, the detailed compilation of the individual system parts (***Format_1***) is recommended. The aggregated format (***Format_2***) can be derived from this if desired.

5.7. Final remarks

Format_1 and ***Format_2*** show the benefits of data structuring principles, such as those embodied by the wide and long formats, to the requirements of reporting and analysing GHP installation data. The aim of these sheets is not to attempt to replicate the complexity of a relational database, but hopefully provide a clearer and more efficient format for data delivery from the data providers. For an insight into how the data could be structured for more complex database systems, we prepared the document titled "*Data Models in SQL Databases: The Basics*", to be found in **Appendix 3**.

5.8. BHE_1

Within the present and following sub-sections, the developed data template sheets are shown. This is done by presenting "sample tables" for the borehole heat exchangers (BHE_1 and BHE_2). For the other GHP systems, these sample tables follow the same idea and are therefore not repeated herein.

Sample tables rather than the entire developed tables are presented because they offer a tangible glimpse into the many parameters we have catalogued, while bringing conciseness within the confines of this report. Given the expansive nature of the information, directly incorporating the entire tables into the report would exceed the spatial limitations of the pages.

In the sample table for ***Format_1***, we use abbreviations for both installations ("i" for installation) and their components ("c" for component) as a deliberate choice to save space. In both sample tables for ***Format_1*** and ***Format_2***, we use the ellipsis (...) in the header serves as a placeholder to indicate that additional columns are involved but not explicitly displayed. For readers desiring a deeper dive into the complete data template sheets, the accompanying Excel file is made available (**Annex 1**). This file contains the complete tables.

Subsequently, all catalogued parameters are given in the metadata tables.

Finally, we provide entity-relationship diagrams illustrating the structure and relationships of variables related to each system.

Table 2. Sample table for the BHE_1 template sheet where "i" means installation and "c" means component.

i_id	i_name	i_commissioning_date	i_status	i_n_heatpumps	...	c_borehole_id	c_borehole_x_coordinate	c_borehole_y_coordinate	c_borehole_depth	c_borehole_diameter	...
bhe_001	a	01-12-10	decommissioned	1	...	bhe_001_01	503505	235002	100	150	...
bhe_001	NA	NA	NA	NA	...	bhe_001_02	503512	235019	110	150	...
bhe_002	b	02-12-10	operational	1	...	bhe_002_01	523701	239801	90	120	...
bhe_003	c	03-12-10	operational	2	...	bhe_003_01	520001	219010	115	150	...
bhe_003	NA	NA	NA	NA	...	bhe_003_02	520002	219011	115	150	...
bhe_003	NA	NA	NA	NA	...	bhe_003_03	520003	219012	110	150	...
bhe_003	NA	NA	NA	NA	...	bhe_003_04	520004	219013	90	150	...
...	**...**	**...**	**...**	**...**	**...**	**...**	**...**	**...**	**...**	**...**	**...**

Table 3. Metadata table for the BHE_1 template sheet.

Column	Description	Data level	Unit	Data type
installation_id	Unique identifier for each geothermal installation	Installation	-	Text
installation_name	Name or label for the geothermal installation	Installation	-	Text
installation_commissioning_date	Date when the geothermal installation was commissioned (optionally approved or notified). Format: Day/Month/Year	Installation	-	Date
installation_status	Current operational status of the installation (e.g., planned, operational, decommissioned)	Installation	-	Text
installation_n_heatpumps	Number of heat pumps associated with the installation	Installation	-	Numeric
installation_thermal_capacity_heating	Total installed thermal capacity of the geothermal installation for heating in kilowatts. It represents the amount of energy the system can generate and it is referred to the theoretical maximum installed thermal capacity of the system	Installation	kW	Numeric
installation_thermal_capacity_cooling	Total installed thermal capacity of the geothermal installation for cooling in kilowatts. It represents the amount of energy the system can generate and it is referred to the theoretical maximum installed thermal capacity of the system	Installation	kW	Numeric
installation_thermal_energy_delivered_heating	Total amount of thermal energy delivered by the geothermal system for heating. It is measured in megawatt-hours per annum (MWh/a) and reflects the energy provided by the heat pump	Installation	MWh/a	Numeric
installation_thermal_energy_delivered_cooling	Total amount of thermal energy delivered by the geothermal system for cooling. It is measured in megawatt-hours per annum (MWh/a) and reflects the energy provided by the heat pump	Installation	MWh/a	Numeric
installation_eflh_heating	Equivalent full load hours of operation for heating	Installation	h	Numeric
installation_eflh_cooling	Equivalent full load hours of operation for cooling	Installation	h	Numeric
installation_operation_mode	Operational mode of the installation (e.g., heating, cooling, heating and cooling, passive/active cooling/heating)	Installation	-	Text
installation_fluid	Type of fluid used inside the probes (e.g., water, water/ethanolglykol)	Installation	-	Text
installation_TRT	Indicates whether a Thermal Response Test was conducted (e.g., conducted, not done, planned)	Installation	-	Text
installation_num_sim	Indicates if a hydraulic numeric simulation has been done (e.g., yes, no)	Installation	-	Text
component_borehole_id	Unique identifier for each borehole within an installation	Component	-	Text
component_borehole_x_coordinate	The X spatial coordinate of the borehole. Used to represent longitude in a projected coordinate reference system. The	Component	m	Numeric

	coordinate system must be consistently used across all entries in the database			
component_borehole_y_coordinate	The Y spatial coordinate of the borehole. Used to represent latitude in a projected coordinate reference system. The coordinate system must be consistently used across all entries in the database	Component	m	Numeric
component_borehole_depth	Depth of the borehole in meters	Component	m	Numeric
component_borehole_diameter	Diameter of the borehole in millimetres	Component	mm	Numeric
component_borehole_target_dist	Minimum target distance between boreholes during the design stage in meters	Component	m	Numeric
component_borehole_filling_thermal_cond	Thermal conductivity of the material used to fill the borehole around the heat exchanger pipes in Watts per meter-Kelvin	Component	W/(m·K)	Numeric
component_borehole_backfilling_material	Type of material used to backfill the borehole (bentonite, cement-bentonite, sand and gravel mixtures, additive-based blends, etc)	Component	-	Text
component_probe_type	Type of probe used in the borehole (e.g., double-U, single-U, coaxial)	Component	-	Text
component_probe_diameter	Nominal outside diameter of the probe used in the borehole in millimetres	Component	mm	Numeric
component_probe_material	Material of the probe used in the borehole (e.g., PE100, PEX, PE-RT, PE-RC)	Component	-	Text
notes	Additional notes or comments related to the installation or specific boreholes	-	-	Text

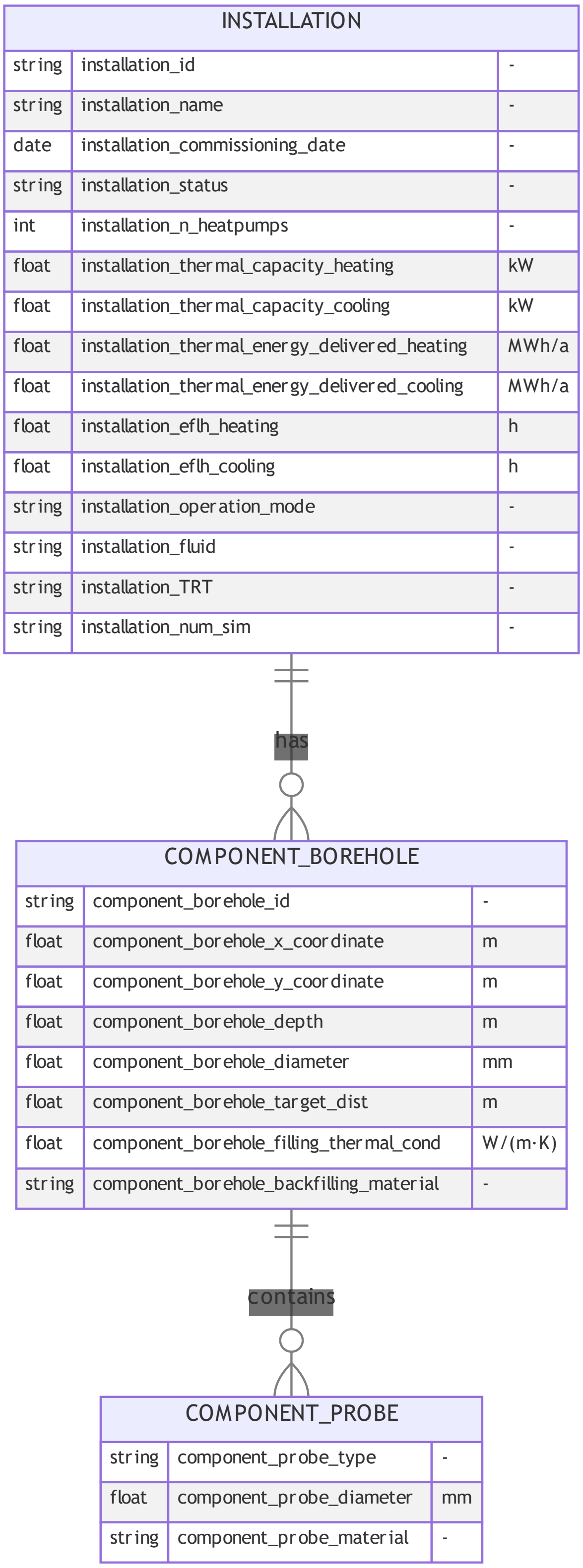

Fig. 3. *Entity-relationship diagram showing the structure and relationships of the catalogued parameters for BHE_1.*

5.9. BHE_2

Table 4. Sample table for the BHE_2 template sheet where "i" means installation and "c" means component.

installation_id	installation_name	x_coordinate	y_coordinate	commissioning_date	status	n_heatpumps	thermal_capacity_heating	thermal_capacity_cooling	thermal_energy_delivered_heating	thermal_energy_delivered_cooling	...
bhe_001	a	503500	235000	01/12/2010	decommissioned	1	10	NA	28	NA	...
bhe_002	b	470300	235000	02/12/2010	operational	1	5	NA	14	NA	...
bhe_003	c	520002	219000	03/12/2010	operational	2	20	NA	56	NA	...
bhe_004	d	553503	234803	04/12/2010	operational	1	7	NA	19.6	NA	...
bhe_005	e	513504	194804	05/12/2010	operational	1	NA	10	NA	28	...
bhe_006	f	503505	234805	06/12/2010	operational	1	10	8	25	14	...
bhe_007	g	483506	214806	07/12/2010	planned	1	5	NA	14	NA	...
...	...	...	...	...	...	...	...	...	...	...	...

Table 5. Metadata table for the BHE_2 template sheet.

Column	Description	Unit	Data type
installation_id	Unique identifier for each geothermal installation	-	Text
installation_name	Name or label for the geothermal installation	-	Text
x_coordinate	The X spatial coordinate of the installation (e.g., the centroid). Used to represent longitude in a projected coordinate reference system. The coordinate system must be consistently used across all entries in the database	-	Numeric
y_coordinate	The Y spatial coordinate of the installation (e.g., the centroid). Used to represent latitude in a projected coordinate reference system. The coordinate system must be consistently used across all entries in the database	-	Numeric
commissioning_date	Date when the geothermal installation was commissioned (optionally approved or notified). Format: Day/Month/Year	-	Date
status	Current operational status of the installation (e.g., planned, operational, decommissioned)	-	Text
n_heatpumps	Number of heat pumps associated with the installation	-	Numeric
thermal_capacity_heating	Total installed thermal capacity of the geothermal installation for heating in kilowatts. It represents the amount of energy the system can generate and it is referred to the theoretical maximum installed thermal capacity of the system	kW	Numeric
thermal_capacity_cooling	Total installed thermal capacity of the geothermal installation for cooling in kilowatts. It represents the amount of energy the system can generate and it is referred to the theoretical maximum installed thermal capacity of the system	kW	Numeric
thermal_energy_delivered_heating	Total amount of thermal energy delivered by the geothermal system for heating. It is measured in megawatt-hours per annum (MWh/a) and reflects the energy provided by the heat pump	MWh/a	Numeric
thermal_energy_delivered_cooling	Total amount of thermal energy delivered by the geothermal system for cooling. It is measured in megawatt-hours per annum (MWh/a) and reflects the energy provided by the heat pump	MWh/a	Numeric
eflh_heating	Equivalent full load hours of operation for heating	h	Text
eflh_cooling	Equivalent full load hours of operation for cooling	h	Text
operation_mode	Operational mode of the installation (e.g., heating, cooling, heating and cooling, passive/active cooling/heating)	-	Text
fluid	Type of fluid used inside the probes (e.g., water, water/ethanolglykol)	-	Text
TRT	Indicates whether a Thermal Response Test was conducted (e.g., conducted, not done, planned)	-	Text
num_sim	Indicates if a hydraulic numeric simulation has been done	-	Text
n_boreholes	Total number of boreholes of the installation	-	Numeric
avg_depth	Average depth of all boreholes within an installation in meters	m	Numeric
borehole_diameter	Average diameter of all boreholes within an installation in millimetres	mm	Numeric

borehole_target_dist	Minimum target distance between boreholes during the design stage in meters	m	Numeric
borehole_filling_thermal_cond	Thermal conductivity of the material used to fill the borehole around the heat exchanger pipes in Watts per meter-Kelvin	W/(m·K)	Numeric
borehole_backfilling_material	Type of material used to backfill the borehole (bentonite, cement-bentonite, sand and gravel mixtures, additive-based blends, etc)	-	Text
probe_type	Probe type, e.g.: Double-U, single-U, coaxial. Most common probe type, in the case of different types	-	Text
probe_diameter	Average nominal outside diameter of the probes used in the boreholes in millimetres	mm	Numeric
probe_material	Material of the probe used in the borehole (e.g., PE100, PEX, PE-RT, PE-RC)	-	Text
notes	Additional notes or comments related to the installation or specific boreholes	-	Text

INSTALLATION		
string	installation_id	-
string	installation_name	-
float	x_coordinate	m
float	y_coordinate	m
date	commissioning_date	-
string	status	-
int	n_heatpumps	-
float	thermal_capacity_heating	kW
float	thermal_capacity_cooling	kW
float	thermal_energy_delivered_heating	MWh/a
float	thermal_energy_delivered_cooling	MWh/a
float	eflh_heating	h
float	eflh_cooling	h
string	operation_mode	-
string	fluid	-
string	TRT	-
string	num_sim	-
int	n_boreholes	-
float	avg_depth	m
float	borehole_diameter	mm
float	borehole_target_dist	m
float	borehole_filling_thermal_cond	W/(m·K)
string	borehole_backfilling_material	-
string	probe_type	-
float	probe_diameter	mm
string	probe_material	-

Fig. 4. *Single-entity diagram illustrating the catalogued parameters for BHE_2. In this diagram, all the listed attributes are given under a single entity called INSTALLATION without showing relationships to any other entities to denote the aggregated format.*

5.10. GWHP_1

Table 6 and **Fig. 5** present the data sheets and diagram for the Groundwater heat pump systems (GWHP), respectively, related to ***Format_1.***

Table 6. Metadata table for the GWHP_1 template sheet.

Column	Description	Data level	Unit	Data type
`installation_id`	Unique identifier for each geothermal installation	Installation	-	Text
`installation_name`	Name or label for the geothermal installation	Installation	-	Text
`installation_commissioning_date`	Date when the geothermal installation was commissioned (optionally approved or notified). Format: Day/Month/Year	Installation	-	Date
`installation_status`	Current operational status of the installation (e.g., planned, operational, decommissioned)	Installation	-	Text
`installation_n_heat pumps`	Number of heat pumps associated with the installation	Installation	-	Numeric
`installation_thermal_capacity_heating`	Total installed thermal capacity of the geothermal installation for heating in kilowatts. It represents the amount of energy the system can generate and it is referred to the theoretical maximum installed thermal capacity of the system	Installation	kW	Numeric
`installation_thermal_capacity_cooling`	Total installed thermal capacity of the geothermal installation for cooling in kilowatts. It represents the amount of energy the system can generate and it is referred to the theoretical maximum installed thermal capacity of the system	Installation	kW	Numeric
`installation_thermal_energy_delivered_heating`	Total amount of thermal energy delivered by the geothermal system for heating. It is measured in megawatt-hours per annum (MWh/a) and reflects the energy provided by the heat pump	Installation	MWh/a	Numeric
`installation_thermal_energy_delivered_cooling`	Total amount of thermal energy delivered by the geothermal system for cooling. It is measured in megawatt-hours per annum (MWh/a) and reflects the energy provided by the heat pump	Installation	MWh/a	Numeric
`installation_eflh_heating`	Equivalent full load hours of operation for heating	Installation	h	Numeric
`installation_eflh_cooling`	Equivalent full load hours of operation for cooling	Installation	h	Numeric
`installation_operation_mode`	Operational mode of the installation (e.g., heating, cooling, heating and cooling, passive/active cooling/heating)	Installation	-	Text
`installation_pumping_test`	Indicates whether a pumping test has been conducted (e.g., conducted, not done, planned).	Installation	-	Text
`installation_aquifer_type`	Describes the type of aquifer (primary or secondary porosity, confined or unconfined)	Installation	-	Text
`component_id`	Unique identifier for each well within an installation	Component	-	Text

component_type	Type of well referring to its purpose (extraction or injection). This field can also be used to identify a standing column well	Component	-	Text
component_waters	Indicates which body of water is used (e.g. groundwater, running waters, sewer, old mines, caves)	Component	-	Text
component_x_coordinate	The X spatial coordinate of the installation component (e.g., well). Used to represent longitude in a projected coordinate reference system. The coordinate system must be consistently used across all entries in the database	Component	m	Numeric
component_y_coordinate	The Y spatial coordinate of the installation component (e.g., well). Used to represent latitude in a projected coordinate reference system. The coordinate system must be consistently used across all entries in the database	Component	m	Numeric
component_well_depth	Total depth of the well from ground level in meters	Component	m	Numeric
component_well_top	Depth to the top of the installed screen from ground level in meters	Component	m	Numeric
component_well_bottom	Depth to the bottom of the installed screen from ground level in meters	Component	m	Numeric
component_well_diameter	Diameter of the well in millimetres	Component	mm	Numeric
component_flow_rate_l_s	Maximum extraction or injection flow rate in litres per second.	Component	l/s	Numeric
component_flow_rate_m3_d	Maximum extraction or injection flow rate in cubic meters per day.	Component	m^3/d	Numeric
component_flow_rate_m3_a	Maximum extraction or injection flow rate in cubic meters per year.	Component	m^3/a	Numeric
component_avg_water_temp	Average temperature of the water in degrees Celsius. This usually will be the natural groundwater temperature at the extraction well	Component	°C	Numeric
component_temp_inj_heating	Average temperature of the injection water in heating mode in degrees Celsius	Component	°C	Numeric
component_temp_inj_cooling	Average temperature of the injection water in cooling mode in degrees Celsius	Component	°C	Numeric
notes	Additional notes or comments related to the installation or specific wells	-	-	Text

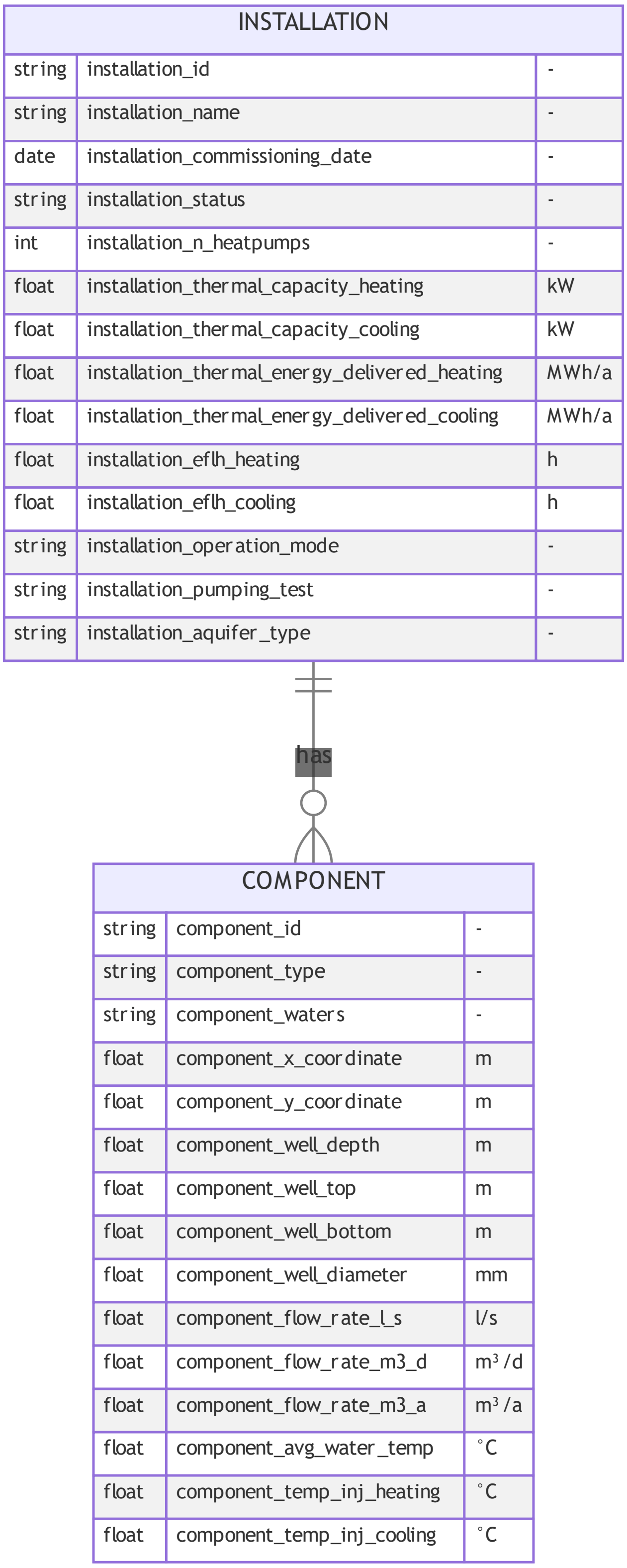

Fig. 5. *Entity-relationship diagram showing the structure and relationships of the catalogued parameters for GWHP_1.*

5.11. GWHP_2

Table 7 and **Fig. 6** present the data sheets and diagram for the Groundwater heat pump systems (GWHP), respectively, related to ***Format_2.***

Table 7. Metadata table for the GWHP_2 template sheet.

Column	Description	Unit	Data type
installation_id	Unique identifier for each geothermal installation	-	Text
installation_name	Name or label for the geothermal installation	-	Text
x_coordinate	The X spatial coordinate of the installation (e.g., the centroid). Used to represent longitude in a projected coordinate reference system. The coordinate system must be consistently used across all entries in the database	m	Numeric
y_coordinate	The Y spatial coordinate of the installation (e.g., the centroid). Used to represent latitude in a projected coordinate reference system. The coordinate system must be consistently used across all entries in the database	m	Numeric
commissioning_date	Date when the geothermal installation was commissioned (optionally approved or notified). Format: Day/Month/Year	-	Date
status	Current operational status of the installation (e.g., planned, operational, decommissioned)	-	Text
n_heatpumps	Number of heat pumps associated with the installation	-	Numeric
thermal_capacity_heating	Total installed thermal capacity of the geothermal installation for heating in kilowatts. It represents the amount of energy the system can generate and it is referred to the theoretical maximum installed thermal capacity of the system	kW	Numeric
thermal_capacity_cooling	Total installed thermal capacity of the geothermal installation for cooling in kilowatts. It represents the amount of energy the system can generate and it is referred to the theoretical maximum installed thermal capacity of the system	kW	Numeric
thermal_energy_delivered_heating	Total amount of thermal energy delivered by the geothermal system for heating. It is measured in megawatt-hours per annum (MWh/a) and reflects the energy provided by the heat pump	MWh/a	Numeric
thermal_energy_delivered_cooling	Total amount of thermal energy delivered by the geothermal system for cooling. It is measured in megawatt-hours per annum (MWh/a) and reflects the energy provided by the heat pump	MWh/a	Numeric
eflh_heating	Equivalent full load hours of operation for heating	h	Numeric
eflh_cooling	Equivalent full load hours of operation for cooling	h	Numeric
operation_mode	Operational mode of the installation (e.g., heating, cooling, heating and cooling, passive/active cooling/heating)	-	Text
pumping_test	Indicates whether a pumping test has been conducted (e.g., conducted, not done, planned)	-	Text
extraction_type	Indicates where the water to be used is pumped from (e.g. groundwater, running water, old mines, caves,	-	Text

	groundwater-SWC). Note that 'groundwater-SCW' can be used to identify standing column wells		
injection_type	Indicates where the thermally utilised water is returned (e.g. groundwater, running water, old mines, caves, groundwater-SWC)	-	Text
n_ext_well	Number of extraction wells in the installation	-	Numeric
n_inj_well	Number of injection wells in the installation	-	Numeric
depth_ext_well	Average depth from ground level of all extraction wells within an installation in meters	m	Numeric
depth_inj_well	Average depth from ground level of all injection wells within an installation in meters	m	Numeric
ext_well_top	Average depth to the top of the installed screens in all extraction wells, measured from ground level, in meters	m	Numeric
ext_well_bottom	Average depth to the bottom of the installed screens in all extraction wells, measured from ground level, in meters	m	Numeric
inj_well_top	Average depth to the top of the installed screens in all injection wells, measured from ground level, in meters	m	Numeric
inj_well_bottom	Average depth to the bottom of the installed screens in all injection wells, measured from ground level, in meters	m	Numeric
diameter_ext_well	Average diameter of all extraction wells within an installationin millimeters	m	Numeric
diameter_inj_well	Average diameter of all injection wells within an installationin millimeters	m	Numeric
extraction_l_s	Maximum extraction flow rate in liters per second	l/s	Numeric
extraction_m3_d	Maximum extraction flow rate in cubic meters per day	m^3/d	Numeric
extraction_m3_a	Maximum extraction flow rate in cubic meters per year	m^3/a	Numeric
injection_l_s	Maximum injection flow rate in liters per second	l/s	Numeric
injection_m3_d	Maximum injection flow rate in cubic meters per day	m^3/d	Numeric
injection_m3_a	Maximum flow rate in cubic meters per year	m^3/a	Numeric
avg_water_temp	The average temperature of the water in degrees Celsius. This usually will be the natural groundwater temperature at the extraction well	°C	Numeric
temp_inj_heating	Average temperature of the injection water in heating mode in degrees Celsius	°C	Numeric
temp_inj_cooling	Average temperature of the injection water in cooling mode in degrees Celsius	°C	Numeric
notes	Additional notes or comments related to the installation or specific wells	-	Text

INSTALLATION		
string	installation_id	-
string	installation_name	-
float	x_coordinate	m
float	y_coordinate	m
date	commissioning_date	-
string	status	-
int	n_heatpumps	-
float	thermal_capacity_heating	kW
float	thermal_capacity_cooling	kW
float	thermal_energy_delivered_heating	MWh/a
float	thermal_energy_delivered_cooling	MWh/a
float	eflh_heating	h
float	eflh_cooling	h
string	operation_mode	-
string	pumping_test	-
string	extraction_type	-
string	injection_type	-
int	n_ext_well	-
int	n_inj_well	-
float	depth_ext_well	m
float	depth_inj_well	m
float	ext_well_top	m
float	ext_well_bottom	m
float	inj_well_top	m
float	inj_well_bottom	m
float	diameter_ext_well	mm
float	diameter_inj_well	mm
float	extraction_l_s	l/s
float	extraction_m3_d	m^3/d
float	extraction_m3_a	m^3/a
float	injection_l_s	l/s
float	injection_m3_d	m^3/d
float	injection_m3_a	m^3/a
float	avg_water_temp	°C
float	temp_inj_heating	°C
float	temp_inj_cooling	°C

Fig. 6. *Single-entity diagram illustrating the catalogued parameters for GWHP_2. In this diagram, all the listed attributes are given under a single entity called INSTALLATION without showing relationships to any other entities to denote the aggregated format.*

5.12. HOR_2

Table 8 and **Fig. 7** present the data sheets and diagram for the horizontal collector systems (HOR), respectively, related to ***Format_2.***

Table 8. Metadata table for the HOR_2 template sheet.

Column	Description	Unit	Data type
installation_id	Unique identifier for each geothermal installation	-	Text
installation_name	Name or label for the geothermal installation	-	Text
x_coordinate	The X spatial coordinate of the installation (e.g., the centroid). Used to represent longitude in a projected coordinate reference system. The coordinate system must be consistently used across all entries in the database	m	Numeric
y_coordinate	The Y spatial coordinate of the installation (e.g., the centroid). Used to represent latitude in a projected coordinate reference system. The coordinate system must be consistently used across all entries in the database	m	Numeric
commissioning_date	Date when the geothermal installation was commissioned (optionally approved or notified). Format: Day/Month/Year	-	Date
status	Current operational status of the installation (e.g., planned, operational, decommissioned)	-	Text
n_heatpumps	Number of heat pumps associated with the installation	-	Numeric
thermal_capacity_heating	Total installed thermal capacity of the geothermal installation for heating in kilowatts. It represents the amount of energy the system can generate and it is referred to the theoretical maximum installed thermal capacity of the system	kW	Numeric
thermal_capacity_cooling	Total installed thermal capacity of the geothermal installation for cooling in kilowatts. It represents the amount of energy the system can generate and it is referred to the theoretical maximum installed thermal capacity of the system	kW	Numeric
thermal_energy_delivered_heating	Total amount of thermal energy delivered by the geothermal system for heating. It is measured in megawatt-hours per annum (MWh/a) and reflects the energy provided by the heat pump	MWh/a	Numeric
thermal_energy_delivered_cooling	Total amount of thermal energy delivered by the geothermal system for cooling. It is measured in megawatt-hours per annum (MWh/a) and reflects the energy provided by the heat pump	MWh/a	Numeric
eflh_heating	Equivalent full load hours of operation for heating	h	Numeric
eflh_cooling	Equivalent full load hours of operation for cooling	h	Numeric
operation_mode	Operational mode of the installation (e.g., heating, cooling, heating and cooling, passive/active cooling/heating)	-	Text
fluid	Type of fluid used inside the probes (e.g., water, water/ethanolglykol)	-	Text
sub_type	Specific type of horizontal collector system (e.g. normal horizontal collectors, slinky collectors, earth baskets)	-	Text

total_area	Total surface area covered by the horizontal collector in square meters	m^2	Numeric
avg_depth	Average depth at which the horizontal collector system is installed in meters	m	Numeric
total_pipe_length	Total length of all piping used in the horizontal collector system in meters	m	Numeric
pipe_diameter	Average diameter of the pipes used in the system in millimetres	mm	Numeric
n_exc_points	Total number of excavation points (trenches or pits) used in the system	-	Numeric
n_circuits	Total number of circuits in the trenches or pits used in the system	-	Numeric
avg_distance_exc_points	Average center-to-center distance between each excavation point (trenches or pits) in meters	m	Numeric
avg_length_exc_points	Average length of the excavation points used in the system, encompassing both trenches and pits, measured in meters. In the case of pits for earth baskets, the term 'length' is more related to the width of pits	m	Numeric
collector_position	Indicates the position of the collector: reclined (parallel to the ground surface) or standing (perpendicular to the ground surface). More relevant to slinky collectors	-	Text
loop_pitch	Loop pitch (i.e., loop spacing) in meters. More relevant to slinky collectors	m	Numeric
ring_diameter	Ring diameter in millimetres. More relevant to slinky collectors	mm	Numeric
notes	Additional notes or comments related to the installation	-	Text

INSTALLATION		
string	installation_id	-
string	installation_name	-
float	x_coordinate	m
float	y_coordinate	m
date	commissioning_date	-
string	status	-
int	n_heatpumps	-
float	thermal_capacity_heating	kW
float	thermal_capacity_cooling	kW
float	thermal_energy_delivered_heating	MWh/a
float	thermal_energy_delivered_cooling	MWh/a
float	eflh_heating	h
float	eflh_cooling	h
string	operation_mode	-
string	fluid	-
string	sub_type	-
float	total_area	m²
float	avg_depth	m
float	total_pipe_length	m
float	pipe_diameter	mm
int	n_exc_points	-
int	n_circuits	-
float	avg_distance_exc_points	m
float	avg_length_exc_points	m
string	collector_position	-
float	loop_pitch	m
float	ring_diameter	mm

Fig. 7. *Single-entity diagram illustrating the catalogued parameters for HOR_2. In this diagram, all the listed attributes are given under a single entity called INSTALLATION without showing relationships to any other entities to denote the aggregated format.*

5.13. TAG_2

Table 9 and **Fig. 8** present the data sheets and diagram for the thermoactive geostructure systems (TAG), respectively, related to ***Format_2.***

Table 9. Metadata table for the TEG_2 template sheet.

Column	Description	Unit	Data type
installation_id	Unique identifier for each geothermal installation	-	Text
installation_name	Name or label for the geothermal installation	-	Text
x_coordinate	The X spatial coordinate of the installation (e.g., the centroid). Used to represent longitude in a projected coordinate reference system. The coordinate system must be consistently used across all entries in the database	m	Numeric
y_coordinate	The Y spatial coordinate of the installation (e.g., the centroid). Used to represent latitude in a projected coordinate reference system. The coordinate system must be consistently used across all entries in the database	m	Numeric
commissioning_date	Date when the geothermal installation was commissioned (optionally approved or notified). Format: Day/Month/Year	-	Date
status	Current operational status of the installation (e.g., planned, operational, decommissioned)	-	Text
n_heatpumps	Number of heat pumps associated with the installation	-	Numeric
thermal_capacity_heating	Total installed thermal capacity of the geothermal installation for heating in kilowatts. It represents the amount of energy the system can generate and it is referred to the theoretical maximum installed thermal capacity of the system	kW	Numeric
thermal_capacity_cooling	Total installed thermal capacity of the geothermal installation for cooling in kilowatts. It represents the amount of energy the system can generate and it is referred to the theoretical maximum installed thermal capacity of the system	kW	Numeric
thermal_energy_delivered_heating	Total amount of thermal energy delivered by the geothermal system for heating. It is measured in megawatt-hours per annum (MWh/a) and reflects the energy provided by the heat pump	MWh/a	Numeric
thermal_energy_delivered_cooling	Total amount of thermal energy delivered by the geothermal system for cooling. It is measured in megawatt-hours per annum (MWh/a) and reflects the energy provided by the heat pump	MWh/a	Numeric
eflh_heating	Equivalent full load hours of operation for heating	h	Numeric
eflh_cooling	Equivalent full load hours of operation for cooling	h	Numeric
operation_mode	Operational mode of the installation (e.g., heating, cooling, heating and cooling, passive/active cooling/heating)	-	Text
fluid	Type of fluid used inside the probes (e.g., water, water/ethanolglykol)	-	Text
sub_type	Specific type of component activation used (e.g., energy pile, diaphragm wall, underfloor), multiple entries possible	-	Text

n_elements	Total number of thermal energy geostructure elements, including for example energy piles and diaphragm walls, used within the installation	-	Numeric
avg_depth_elements	Average depth of the thermal energy geostructure elements, including for example energy piles and diaphragm walls, used within the installation	m	Numeric
avg_distance_elements	Average distance between adjacent thermal energy geostructure elements, including energy piles and diaphragm walls, within the installation, in meters.	m	Numeric
total_area	Total area of the surface thermally activated element used (e.g., for underfloor slabs and diaphragm walls)	m^2	Numeric
notes	Additional notes or comments related to the installation	-	Text

INSTALLATION		
string	installation_id	-
string	installation_name	-
float	x_coordinate	m
float	y_coordinate	m
date	commissioning_date	-
string	status	-
int	n_heatpumps	-
float	thermal_capacity_heating	kW
float	thermal_capacity_cooling	kW
float	thermal_energy_delivered_heating	MWh/a
float	thermal_energy_delivered_cooling	MWh/a
float	eflh_heating	h
float	eflh_cooling	h
string	operation_mode	-
string	fluid	-
string	sub_type	-
int	n_elements	-
float	avg_depth_elements	m
float	avg_distance_elements	m
float	total_area	m^2
float	pipe_diameter	mm
string	notes	-

Fig. 8. *Single-entity diagram illustrating the catalogued parameters for TEG_2. In this diagram, all the listed attributes are given under a single entity called INSTALLATION without showing relationships to any other entities to denote the aggregated format.*

6. Conclusion

This report underscores the importance of augmenting data collection, monitoring, and standardisation for geothermal heat pumps (GHPs) across Europe. The current mechanisms for reporting and monitoring of market indicators related GHPs were explored on European level and on national level, within the GeoBOOST partner countries (Austria, Germany, Ireland, Netherlands, Poland, Spain, and Sweden).

The main market indicator at the European level to show progress in GHPs, is the number of heat pump units sold. The EHPA Market Report is based on heat pump sales at the national scale across European countries. However, it is not workable to isolate GHP system types when sales data is given by energy source and distribution medium. More precise categorisation of GHPs according to their primary energy sources (e.g., groundwater, boreholes, topmost soil) would be beneficial to get a better understanding of market dynamics. Another European market report is issued by EGEC and focuses on GHP sales data. Refining its methodologies to ensure more consistent data collection due to the reliance on national coordinators could be beneficial. Further detailing to distinguish between specific GHP systems would provide more insights into segments of the market.

The analysis of the reporting practices has revealed significant gaps in data standardisation, completeness, and accessibility across all investigated partner countries. In countries such as Germany and Austria, decentralised reporting creates high variability in the available data across federal states. This shows the need for uniform data templates and the establishment of centralised national databases. Cohesive regulatory frameworks that mandate the countrywide registration of installations are fundamental for supporting comprehensive data generation practices. Accordingly, Sweden and the Netherlands have implemented more effective strategies by fostering collaborative data generation coupled with national standardisation and are seen as best practice examples. These practices appear to address the more fragmented approaches to GHP data management observed in the other GeoBOOST partner countries, which lack uniform regulations underpinning strategies for national registration of installations.

Importantly, this work has revealed inconsistencies in how separate metrics for heating and cooling are captured in GHP installation data, a crucial aspect not fully considered in all investigated countries. To help overcome the encountered limitations, the annex of this report provides a template with a wide range of parameters that we recommend to be collected in comprehensive state-wide or, ideally, country-wide databases. The suggested parameters focus on the underground and include, amongst others, the depth of the installation and annual thermal energy delivered and installed capacity. The data template sheets are tailored to the main GHP systems (borehole heat exchangers, groundwater heat pumps, horizontal collectors and thermoactive geostructures) in the participating countries. Lastly the report suggests how the generated data can best be shared with third-party users (e.g., geological survey organisations, energy planners and policymakers).

Providing this comprehensive information on regional, national, but also European level about installations is vital ***(i)*** for managing potential thermal or hydraulic interference among systems and ***(ii)*** for determining the share of renewable energy in the heating and cooling sector from GHPs. Hence, reliable GHP installation data is not just beneficial but necessary. Improved data management of the installations will provide a robust foundation for research, policymaking, and the broader adoption of GHP systems.

References:

Bayer, P., Attard, G., Blum, P., Menberg, K., 2019. The geothermal potential of cities. Renewable and Sustainable Energy Reviews 106, 17–30. https://doi.org/10.1016/j.rser.2019.02.019

EGEC, 2023. 2022 EGEC GEOTHERMAL MARKET REPORT (No. 12th Edition, July 2023). EGEC - European Geothermal Energy Council.

EHPA, 2023. European Heat Pump Market and Statistics Report 2023 (European Heat Pump Association).

García-Gil, A., Goetzl, G., Kłonowski, M.R., Borovic, S., Boon, D.P., Abesser, C., Janza, M., Herms, I., Petitclerc, E., Erlström, M., Holecek, J., Hunter, T., Vandeweijer, V.P., Cernak, R., Mejías Moreno, M., Epting, J., 2020. Governance of shallow geothermal energy resources. Energy Policy 138. https://doi.org/10.1016/j.enpol.2020.111283

Gazzetta Ufficiale, 2012. DECRETO 14 gennaio 2012 [WWW Document]. MINISTERO DELLO SVILUPPO ECONOMICO. URL https://www.gazzettaufficiale.it/eli/id/2012/02/14/12A01451/sg (accessed 5.2.24).

Juhlin, K., 2016. Kartering och sammanställning av större geoenergisystem i Sverige – En studie baserad på SGUs brunnsarkiv (No. Uppsala universitet).

Liu, X., Ho, J., Winick, J., Porse, S., Lian, J., Wang, X., Liu, W., Malhotra, M., Li, Y., Anand, J., 2023. Grid Cost and Total Emissions Reductions Through Mass Deployment of Geothermal Heat Pumps for Building Heating and Cooling Electrification in the United States (No. ORNL/TM-2023/2966). Oak Ridge National Laboratory (ORNL), Oak Ridge, TN (United States). https://doi.org/10.2172/2224191

Lund, J.W., Huttrer, G.W., Toth, A.N., 2022. Characteristics and trends in geothermal development and use, 1995 to 2020. Geothermics 105, 102522. https://doi.org/10.1016/j.geothermics.2022.102522

Miranda, N.D., Lizana, J., Sparrow, S.N., Zachau-Walker, M., Watson, P.A.G., Wallom, D.C.H., Khosla, R., McCulloch, M., 2023. Change in cooling degree days with global mean temperature rise increasing from 1.5 °C to 2.0 °C. Nat Sustain 6, 1326–1330. https://doi.org/10.1038/s41893-023-01155-z

Pieve, M., Trinchieri, R., 2019. The heat-pump market in Italy: an in-depth economic study about the reasons for a still unexpressed potential. Clean Energy 3, 126–143. https://doi.org/10.1093/ce/zkz002

Sarsentis, D., Orso, L., 2023. How the UK compares to the rest of Europe on heat pump uptake. Electricity, gas and other fuel prices across Europe [WWW Document]. URL https://www.nesta.org.uk/report/how-the-uk-compares-to-the-rest-of-europe-on-heat-pump-uptake/electricity-gas-and-other-fuel-prices-across-europe/ (accessed 5.3.24).

Somogyi, V., Sebestyén, V., Nagy, G., 2017. Scientific achievements and regulation of shallow geothermal systems in six European countries – A review. Renewable and Sustainable Energy Reviews 68, 934–952. https://doi.org/10.1016/j.rser.2016.02.014

Tsagarakis, K.P., Efthymiou, L., Michopoulos, A., Mavragani, A., Anđelković, A.S., Antolini, F., Bacic, M., Bajare, D., Baralis, M., Bogusz, W., Burlon, S., Figueira, J., Genç, M.S., Javed, S., Jurelionis, A., Koca, K., Ryżyński, G., Urchueguia, J.F., Žlender, B., 2020. A review of the legal framework in shallow geothermal energy in selected European countries: Need for guidelines. Renewable Energy 147, 2556–2571. https://doi.org/10.1016/j.renene.2018.10.007

Witte, H.J.L., 2023. Catalogue of existing and promising technical solutions. Deliverable D.5.1 of the GeoBOOST project (Boosting geothermal heat pumps to mainstream cost-effective and efficient renewable heating and cooling in buildings).

Annex 1. Accompanying Excel File

This Annex encloses an Excel file that demonstrates how the various catalogued parameters can be reported to end users for the most common system types: **BHE** (borehole heat exchangers), **GWHP** (groundwater heat pumps), **HOR** (horizontal collectors such as normal horizontal collectors, slinky collectors and earth baskets) and **TAG** (thermoactive geostructures such as energy piles, thermally activated underfloor elements and diaphragm walls).

Appendix 1. Main differences between ATES and BTES

Introduction

The most prevalent forms of underground thermal energy storage (UTES) systems are those utilising boreholes, known as borehole thermal energy storage (BTES), and those employing aquifers, referred to as aquifer thermal energy storage (ATES).

These systems are predominantly used for the purpose of storing energy over different seasons. ATES systems function by installing at least two wells in a suitable aquifer. These systems operate by extracting groundwater from one well, using its thermal energy, and subsequently reinjecting the water back into the aquifer through a separate well. On the other hand, BTES systems offer an alternative method of underground thermal energy storage. These systems are versatile and can be adapted to a wide range of ground conditions. They generally comprise one or multiple boreholes which are used to store thermal energy beneath the Earth's surface. This stored energy is typically used on a seasonal basis (Javadi et al., 2022).

This appendix aims to shed some light on the basic distinctions that set ATES systems apart from their BTES counterparts, offering a concise and informative overview of their unique characteristics.

Main differences

Table A1-1 presents the main differences between ATES and BTES systems (Daniilidis et al., 2022; Fleuchaus et al., 2018; Goetzl et al., 2022; Jiang et al., 2022).

Summary

ATES and BTES serve the purpose of underground thermal energy storage. However, they differ in geological prerequisites, scale, thermal efficiency, regulatory constraints, and versatility. ATES is more reliant on specific geological formations and may face stricter regulations, but it is often more suitable for large-scale applications. BTES, on the other hand, offers more versatility in location and scale but usually demands a greater surface area and careful design to achieve high efficiency.

Table A1-1. Comparison of ATES and BTES systems.

Aspect	ATES	BTES
Geological formation	Relies on natural aquifers, which are water-bearing permeable layers in the earth.	Does not require natural aquifers; it uses instead the thermal properties of the ground itself (soil or rock).
Drilling requirements	Typically requires fewer but deeper wells aimed at reaching specific aquifer layers.	Often involves more numerous but generally shallower boreholes, without the need to target water-bearing layers.
Scale and surface area needs	Usually suited for larger-scale applications because a single well can cover a significant area, depending on aquifer properties.	More flexible in scale but may require a greater surface area due to the number of boreholes needed.
Thermal efficiency	Generally high thermal efficiency, aided by the natural insulating properties of the surrounding geological formations.	Efficiency can be very high but is more dependent on the design, including the material used to backfill the boreholes.
Regulatory constraints	May face stricter regulations due to the use of natural aquifers, with potential issues related to water rights and possible aquifer contamination.	Usually less encumbered by water rights issues but still subject to land use and environmental regulations.
Geological versatility	Highly dependent on the presence and suitability of aquifers, thereby geographically limited.	More versatile, as it can be implemented in a wider range of geological settings.
Heat transfer medium	Uses native groundwater as the heat transfer medium.	Uses a heat transfer fluid, commonly water or a water-antifreeze mix, circulated through pipes.

Bibliography

Daniilidis, A., Mindel, J.E., De Oliveira Filho, F., Guglielmetti, L., 2022. Techno-economic assessment and operational CO_2 emissions of High-Temperature Aquifer Thermal Energy Storage (HT-ATES) using demand-driven and subsurface-constrained dimensioning. Energy 249, 123682. https://doi.org/10.1016/j.energy.2022.123682

Fleuchaus, P., Godschalk, B., Stober, I., Blum, P., 2018. Worldwide application of aquifer thermal energy storage – a review. Renewable and Sustainable Energy Reviews 94, 861–876. https://doi.org/10.1016/j.rser.2018.06.057

Goetzl, G., Chicco, J., Schifflechner, C., Figueira, J., Tsironis, G., Zajacs, A., 2022. Pathways to Better Integrate Geothermal Energy at Its Full Technological Scale in European Heating and Cooling Networks. European Geologist 54, 44–50.

Javadi, H., Urchueguía, J.F., Badenes, B., Mateo, M., Nejad Ghafar, A., Chaudhari, O.A., Zirgulis, G., Lemus, L.G., 2022. Laboratory and numerical study on innovative grouting materials applicable to borehole heat exchangers (BHE) and borehole thermal energy storage (BTES) systems. Renewable Energy 194, 788–804. https://doi.org/10.1016/j.renene.2022.05.152

Jiang, M., Rindt, C., Smeulders, D.M.J., 2022. Optimal Planning of Future District Heating Systems—A Review. Energies 15, 7160. https://doi.org/10.3390/en15197160

Appendix 2. Questionnaire

Your information			
First name		Last name	
First name		Last name	
Organisation			
Country			
City			
Email			
Date	*DD/MM/YYYY*		

Goal of the questionnaire

The objective of this questionnaire is to assess how data on shallow geothermal energy systems are primarily generated and shared in the project partner countries.

The underlying aim is to identify and compare current practices in data acquisition, types of data gathered, and their sharing with stakeholders. Establishing this baseline will enable us to propose enhanced approaches for market monitoring.

Instructions

The questionnaire is composed of closed and open questions:

- ***For closed questions***: Please select all options that best apply to your situation, unless specifically instructed or implied.
- ***For open questions:*** You are strongly encouraged to provide as much detail as possible.

Questionnaire structure:

1. **General questions** (related to all types of shallow geothermal energy systems).
2. **Closed loop systems.**
3. **Open loop systems.**

Sections 2 and 3 are further organised into sub-sections with the following headings: "***Data Generation***" and "***Data Accessibility and Sharing***". This serves as a guide for the main content of the questions within each sub-section.

Important definitions

- **Primary generation of data** refers to the earliest step of recording information from shallow geothermal systems. Currently this step is mostly done by authorities.
- **National authorities** and **national regulations** refer to the country level.
- **Regional authorities** and **regional regulations** refer to a sub-national level within defined geographical areas like states, provinces, or regions.
- **Local authorities** and **local regulations** refer to the smallest units of government, like a city, town, or district.

Contacts

If you require assistance or have any queries regarding the questionnaire, please contact us at:

Marlon Brancher , marlon.brancher@geosphere.at
Cornelia Steiner , cornelia.steiner@geospher.at

Part 1: General Questions

Q1.1. What main categories of shallow geothermal heat exchanger systems are subject to primary generation of data in your country? *Select all that apply.* If you select "**No primary data generated**", you may skip all subsequent questions of Parts 1, 2 and 3.

☐ Closed loop ☐ Open loop ☐ No primary data generated

Q1.2. Which specific types of heat exchangers are subject to primary generation of data in your country? *Select all that apply.*

☐ Vertical/inclined borehole
☐ Horizontal collector
☐ Energy pile
☐ Groundwater
☐ Groundwater ATES
☐ Surface water
☐ All types
☐ Other (please specify):

Q1.3. For which specific types of heat exchangers is primary data generated on a more widespread basis across your country, rather than being confined to very specific regions? *Select all that apply.*

☐ Vertical/inclined borehole
☐ Horizontal collector
☐ Energy pile
☐ Groundwater
☐ Groundwater ATES
☐ Surface water
☐ All types
☐ Other (please specify):

Q1.4. Are there any major challenges in the primary data generation process in your country? *Select all that apply.*

☐ **Inadequate infrastructure:** Lack of proper infrastructure (e.g., physical tools used to generate data, systems in place for storing and maintaining it) for data generation
☐ **Limited human resources:** Insufficient skilled personnel or technical expertise for effective data generation
☐ **Regulatory or legal hurdles:** Regulatory issues or legal barriers affecting data generation
☐ **Data privacy and security concerns:** Concerns regarding the privacy and security of generated data
☐ **Citizen science initiatives:** Concerns regarding the privacy and security of generated data
☐ Other (please specify):
☐ **Inconsistent data standards:** Lack of uniform data generation entries/standards to databases
☐ **Funding or budget constraints:** Limited funding or budgetary support for data generation efforts
☐ **Public awareness and government action:** Insufficient public awareness, resulting in limited government initiative or effort in data generation processes
☐ **Data integration and compatibility issues:** Challenges in integrating primary data generated from different sources
☐ **No significant challenges**
☐ **Not sure**

Further elaboration (if applicable): *You can elaborate herein on your reply if needed.*

Part 2: Closed loop systems

Data generation

Q2.1. Who generates the primary data on closed loop systems? *Select all that apply. If you select "**No primary data generated**", you may skip all subsequent questions of Part 2.*

- ☐ Local governmental authorities
- ☐ Regional governmental authorities
- ☐ National governmental authorities
- ☐ Installers/Operators
- ☐ Not-for-profit organisations
- ☐ Independent consultants
- ☐ Research institutions
- ☐ Industry associations
- ☐ No primary data generated
- ☐ Other (please specify):

Further elaboration (if applicable): *If you have selected any of the governmental authority options (local, regional, national), please provide further details about the specific authority.*

Q2.2. What specific parameters of the primary generated data are available for installations based on closed loops?

- ☐ Geographical coordinates
- ☐ Year of construction
- ☐ Installed capacity (kW)
- ☐ Annual thermal work (kWh/year)
- ☐ Operation mode (heating/cooling)
- ☐ Number of boreholes
- ☐ Borehole diameter
- ☐ Total depth of the boreholes (in case of multiple probes)
- ☐ Average depth of the boreholes (in case of multiple probes)
- ☐ Average spacing between boreholes
- ☐ Number of heat pumps
- ☐ Type of building
- ☐ Year of building construction
- ☐ Probe type
- ☐ Area (m^2) of horizontal collectors
- ☐ Antifreeze type
- ☐ Antifreeze solution concentration
- ☐ Indication if TRT was conducted
- ☐ Other (please specify):

Q2.3. Are the primary generated data from different types of closed loop systems combined or summed together?

- ☐ Yes
- ☐ No

If "Yes", please indicate for which types of closed loop systems the data is combined:

Q2.4. If available, can you please provide references or links to the sources of the primary generated data in your country?

Q2.5. How is the primary data generated? *Select all that apply.*

- ☐ Authorities enter installations in their database
- ☐ Drilling companies submit their records based on licenses/permits (e.g. water rights)
- ☐ Front-end user interfaces for drillers or installers to submit data
- ☐ Initiatives led by industry associations
- ☐ Other (please specify):
- ☐ Data collection through academic research, studies, or projects
- ☐ Front-end user interfaces for the user of a geothermal system to submit data
- ☐ Citizen science initiatives
- ☐ Not sure

Further elaboration (if applicable): *You can elaborate herein on your reply if needed.*

Q2.6. Are there standard templates or guidelines used for the primary generation of data on closed loop systems in your country? *Please select the option that best describes the situation in your country.*

- ☐ Yes, there are national templates or guidelines used throughout the country
- ☐ Yes, there are templates or guidelines, but they are specific to certain regions within the country
- ☐ Yes, templates or guidelines exist, but they vary by organisation
- ☐ No, there are no standard templates or guidelines
- ☐ Not sure/Do not know

Further elaboration (if applicable): *If standard templates or guidelines are used, please briefly describe them or share them with us if possible.*

Note: 'Standard templates or guidelines' refer to any formalised documents, procedures, or protocols that are eventually used within your country. These could include official forms, recommended practices, data recording standards, or any other structured approach used to guide the primary generation of data.

Q2.7. How does the spatial coverage of the primary data relate to the regulatory compliance? *Please select the option that best describes the situation in your country.*

- ☐ National regulations prescribe national-level data generation
- ☐ Regional regulations prescribe regional-level data generation
- ☐ Local regulations prescribe local-level data generation
- ☐ Data collection is unregulated and varies without specific spatial coverage guidelines
- ☐ A combination of national, regional, and local regulations affects data generation
- ☐ Other (please specify):

Further elaboration (if applicable): *You can elaborate herein on how the spatial coverage is influenced by the regulatory framework, including any unique aspects or practices in your country.*

Q2.8. Are there noticeable gaps in the generation of primary data for closed loop systems in your country, particularly in terms of historical coverage? For example, there are installations dating back to 2002, but data generation only started from 2010 onwards.

Q2.9. Are you aware of measures in place to ensure the quality and reliability of the primary data generated? If yes, could you briefly specify which measures are used?

Q2.10. Based on your knowledge and experience, how would you rate the overall reliability of the primary data generated on closed loop systems in your country?

- ☐ Poor
- ☐ Fair
- ☐ Good
- ☐ Very Good
- ☐ Excellent

Tip: *For "Poor" to "Excellent" options, you may consider factors such as accuracy, consistency, and comprehensiveness of the different types of data to arrive at your response.*

Data accessibility and sharing

Q2.11. Who can access the primary data generated? *Select all that apply. Selecting 'General public (open access)' implies that the data is also accessible to all other listed entities.*

- ☐ General public (open access)
- ☐ Industry associations
- ☐ Research institutions
- ☐ Government agencies
- ☐ International organisations
- ☐ Not sure
- ☐ Restricted access
- ☐ Other (please specify):

Q2.12. Who are the principal providers of the primary data generated on closed loop systems? *Select all that apply.*

- ☐ Local governmental authorities
- ☐ Regional governmental authorities
- ☐ National governmental authorities
- ☐ Installers/Operators
- ☐ Other (please specify):
- ☐ Not-for-profit organisations
- ☐ Independent consultants
- ☐ Research institutions
- ☐ Industry associations

Note: *This questions deals with the entities responsible for making the primary data accessible. The entity responsible for generating the data may not always be the same as the one sharing it.*

Q2.13. How is the primary data generated made available in your country? *Select all that apply.*

- ☐ Online databases with user query capabilities (e.g., web interface, API)
- ☐ Open data portals
- ☐ Publicly accessible websites
- ☐ Dashboards
- ☐ Through reports
- ☐ Specific requests to authorities (e.g., by forms or email)
- ☐ Fee-based access
- ☐ Other (please specify):

Q2.14. What is the approximate frequency that the primary data generated is updated?

- ☐ As soon as it is generated (almost real-time)
- ☐ Daily
- ☐ Weekly
- ☐ Monthly
- ☐ Quarterly
- ☐ Other (please specify):
- ☐ Annually
- ☐ Irregular / As-needed basis
- ☐ No data availability
- ☐ Not sure

Q2.15. How accessible is the data?

- ☐ Highly accessible and user-friendly
- ☐ Moderately accessible but with some accessibility issues
- ☐ Poorly accessible and difficult to access

Q2.16. Are there specific parameters that are currently not being generated or made accessible, but which you believe would be beneficial? If so, please specify.

Q2.17. Regarding the extent of the implementation of current data generation and accessibility practices in your country, how satisfied are you?

- ☐ Poor
- ☐ Fair
- ☐ Good
- ☐ Very Good
- ☐ Excellent

Tip: *For "Poor" to "Excellent" options, you may consider factors such as accuracy, consistency, and comprehensiveness of the different types of data to arrive at your response.*

Part 3: Open loop systems

Data generation

Q3.1. Who generates the primary data on open loop systems? *Select all that apply. If you select "**No primary data generated**", you may skip all subsequent questions of Part 3.*

☐ Local governmental authorities	☐ Independent consultants
☐ Regional governmental authorities	☐ Research institutions
☐ National governmental authorities	☐ Industry associations
☐ Installers/Operators	☐ No primary data generated
☐ Not-for-profit organisations	☐ Other (please specify):

Further elaboration (if applicable): *If you have selected any of the governmental authority options (local, regional, national), please provide further details about the specific authority.*

Q3.2. What specific parameters of the primary generated data are available for installations based on open loops?

☐ Geographical coordinates	☐ Flow rate (l/s)
☐ Year of construction	☐ Flow rate (m^3/h)
☐ Installed capacity (kW)	☐ Flow rate (m^3/h)
☐ Annual thermal work (kWh/year)	☐ Flow rate (m^3/y)
☐ Operation mode (heating/cooling)	☐ Flow rate (l/s)
☐ Number of abstraction wells	☐ Average inlet water temperature (°C)
☐ Number of re-injection wells	☐ Average outlet water temperature (°C)
☐ Type of abstraction (where is water taken from, e.g. surface water or groundwater)	☐ Number of heat pumps
☐ Type of disposal of the abstracted water	☐ Design system heating temperature (°C)
☐ Well diameter	☐ Design system cooling temperature (°C)
☐ Total depth of the wells	☐ Type of building
☐ Average depth of the wells (in case of multiple wells)	☐ Year of building construction
☐ (Average) spacing between wells	☐ Other (please specify):

Q3.3. Are the primary generated data from different types of open loop systems combined or summed together?

☐ Yes ☐ No

If "Yes", please indicate for which types of open loop systems the data is combined:

Q3.4. If available, can you please provide references or links to the sources of the primary generated data in your country?

Q3.5. How is the primary data generated? *Select all that apply.*

☐ Authorities enter installations in their database	☐ Data collection through academic research, studies, or projects
☐ Drilling companies submit their records based on licenses/permits (e.g. water rights)	☐ Front-end user interfaces for the user of a geothermal system to submit data
☐ Front-end user interfaces for drillers or installers to submit data	☐ Citizen science initiatives
☐ Initiatives led by industry associations	☐ Not sure
☐ Other (please specify):	

Further elaboration (if applicable): *You can elaborate herein on your reply if needed.*

Q3.6. Are there standard templates or guidelines used for the primary generation of data on open loop systems in your country? *Please select the option that best describes the situation in your country.*

☐ Yes, there are national templates or guidelines used throughout the country
☐ Yes, there are templates or guidelines, but they are specific to certain regions within the country
☐ Yes, templates or guidelines exist, but they vary by organisation
☐ No, there are no standard templates or guidelines
☐ Not sure/Do not know

Further elaboration (if applicable): *If standard templates or guidelines are used, please briefly describe them or share them with us if possible.*

Note: *'Standard templates or guidelines' refer to any formalised documents, procedures, or protocols that are eventually used within your country. These could include official forms, sheets, recommended practices, recording standards, or any other structured approach used to guide the primary generation of data.*

Q3.7. How does the spatial coverage of the primary data relate to the regulatory compliance? *Please select the option that best describes the situation in your country.*

☐ National regulations prescribe national-level data generation
☐ Regional regulations prescribe regional-level data generation
☐ Local regulations prescribe local-level data generation
☐ Data collection is unregulated and varies without specific spatial coverage guidelines
☐ A combination of national, regional, and local regulations affects data generation
☐ Other (please specify):

Further elaboration (if applicable): *You can elaborate herein on how the spatial coverage is influenced by the regulatory framework, including any unique aspects or practices in your country.*

Q3.8. Are there noticeable gaps in the generation of primary data for open loop systems in your country, particularly in terms of historical coverage? For example, there are installations dating back to 2002, but data generation only started from 2010 onwards.

Q3.9. Are you aware of measures in place to ensure the quality and reliability of the primary data generated? If yes, could you briefly specify which measures are used?

Q3.10. Based on your knowledge and experience, how would you rate the overall reliability of the primary data generated on open loop systems in your country?

☐ Poor
☐ Fair
☐ Good
☐ Very Good
☐ Excellent

Tip: *For "Poor" to "Excellent" options, you may consider factors such as accuracy, consistency, and comprehensiveness of the different types of data to arrive at your response.*

Data accessibility and sharing

Q3.11. Who can access the primary data generated? *Select all that apply. Selecting 'General public (open access)' implies that the data is also accessible to all other listed entities.*

☐ General public (open access)
☐ Industry associations
☐ International organisations
☐ Not sure

☐ Research institutions
☐ Government agencies
☐ Restricted access
☐ Other (please specify):

Q3.12. Who are the principal providers of the primary data generated on open loop systems? *Select all that apply.*

☐ Local governmental authorities
☐ Regional governmental authorities
☐ National governmental authorities
☐ Installers/Operators
☐ Other (please specify):
☐ Not-for-profit organisations
☐ Independent consultants
☐ Research institutions
☐ Industry associations

Note: *This questions deals with the entities responsible for making the primary data accessible. The entity responsible for generating the data may not always be the same as the one sharing it.*

Q3.13. How is the primary data generated made available in your country? *Select all that apply.*

☐ Online databases with user query capabilities (e.g., web interface, API)
☐ Open data portals
☐ Publicly accessible websites
☐ Dashboards
☐ Through reports
☐ Specific requests to authorities (e.g., by forms or email)
☐ Fee-based access
☐ Other (please specify):

Q3.14. What is the approximate frequency that the primary data generated is updated?

☐ As soon as it is generated (almost real-time)
☐ Daily
☐ Weekly
☐ Monthly
☐ Quarterly
☐ Other (please specify):
☐ Annually
☐ Irregular / As-needed basis
☐ No data availability
☐ Not sure

Q3.15. How accessible is the data?

☐ Highly accessible and user-friendly
☐ Moderately accessible but with some accessibility issues
☐ Poorly accessible and difficult to access

Q3.16. Are there specific parameters that are currently not being generated or made accessible, but which you believe would be beneficial? If so, please specify.

Q3.17. Regarding the extent of the implementation of current data generation and accessibility practices in your country, how satisfied are you?

☐ Poor
☐ Fair
☐ Good
☐ Very Good
☐ Excellent

Tip: *For "Poor" to "Excellent" options, you may consider factors such as accuracy, consistency, and comprehensiveness of the different types of data to arrive at your response.*

Appendix 3. Data Models in SQL Databases: The Basics

Introduction

A data model, often referred to as a schema in the context of the structured query language (SQL) environment, is a conceptual representation of the data structures and relationships within a database.

The current appendix aims to provide a general overview of data models within SQL databases, a subject of significance in database development and management. The focus is on delineating the basic structure and function of data models, which are essential for organising and manipulating data in relational databases.

A data model typically reflects a wide range of processes. Oftentimes, collaboration occurs between a data professional and a process expert to grasp how the data interrelates within the process under consideration. Simultaneously, the business expert can gain insights from the data professional, increasing their understanding of the process-related actions by potentially better observing the interactions and connections within the data.

The target audience of this appendix includes individuals seeking a fundamental understanding of data models in relational databases.

Main concepts of data models

In what follows, the main concepts of data models are presented (Kleppmann, 2017):

Tables: The fundamental building blocks

In SQL databases, tables are the primary structure for storing data. Each table is analogous to a spreadsheet, comprised of rows and columns, and represents a specific entity or concept within the database, such as `customers`, `orders`, or `products`. The manner in which tables are structured and interlinked forms the heart of the data model. The real effectiveness of a database often depends on the logical organisation of these tables.

Columns and rows: Organising data

The concept of *tidy data* is important for structuring datasets. In the context of SQL databases, this principle is applied in the organisation of columns and rows within tables. Columns in this structure represent variables, with each column dedicated to data of the same type. Rows, on the other hand, correspond to individual observations. Each row contains a complete record of data across these variables. Such an arrangement ensures uniformity where each column holds values of the same data type, each row presents data for a single record, and every cell in the table holds a single value. This methodical organisation of data promotes data integrity and simplifies querying and future analyses.

Primary keys and foreign keys: Unique identifiers and relational links

A primary key uniquely identifies each record in a table, while foreign keys establish links between tables by referencing primary keys in other tables. Namely, foreign keys are used to establish and enforce links between tables. They reference primary keys in other tables, creating a web of relationships that is crucial to the relational database model. These keys enable the database to maintain data integrity and support complex queries involving multiple tables.

Relationships: Connecting data

Relationships between tables are the essence of the relational database concept. The most common types are one-to-one, one-to-many, and many-to-many. Each type serves a distinct purpose in linking data across tables. For example, a one-to-many relationship might link a single customer to multiple orders. Understanding and effectively implementing these relationships is crucial for designing a functional and efficient data model.

Creating a data model in SQL databases

Designing a data model involves planning and consideration of how the data is interrelated and will be accessed. The following steps are useful for creating an effective data model for SQL databases (Kleppmann, 2017):

- *Identification of entities and relationships*: The initial step involves a detailed analysis of the domain **to identify the entities** that the database needs to represent. These entities typically correspond to the real-world things that the database is intended to store, like customers, products, or orders. Concurrently, it is very important to **establish** the relationships among these entities, determining how they interact with one another within the context of the database.

- *Definition of tables and columns*: Following the identification of entities, the next step is to define tables that represent these entities. Each table should include columns that encapsulate the attributes of the entity. These attributes should be carefully selected and defined with appropriate data types. This step attempts to ensure that the data model correctly mirrors the structure of the underlying data.

- *Establishment of key constraints*: Key constraints are instrumental in maintaining the integrity and uniqueness of the data. Primary keys are designated for each table to uniquely identify each record. Foreign keys are defined to create associations between tables, thereby establishing the relational aspect of the database. These constraints are used to enforce the integrity of data and in defining the navigational pathways through the relational model.

- *Application of normalisation principles*: Normalisation, a process of organising data to minimise redundancy and dependency, is an essential consideration in data model design. The data model can be refined to eliminate redundant data and ensure that each piece of information is stored in only one place. However, strategic replication of variables can be a deliberate and beneficial design choice to enhance performance and usability in certain scenarios.

- *Planning for scalability and performance*: An effective data model should be designed with future growth in mind. It involves considering how the addition of new data and the complexity of queries might impact the performance of the database.

Practical applications

In practice, the application of a well-designed data model is evident in the efficiency and scalability of a database. For instance, in a retail database, tables might include `customers`, `products`, and `orders`. Each of these tables would have columns suited to the data they store, and relationships would be established to link customers to their orders and the products within those orders. Such a structure allows for efficient data retrieval, such as quickly finding all orders placed by a particular customer or listing all customers who purchased a specific product.

As a real-world example, the data model used in the `smonitor` framework is a nice case in point of efficiently organising environmental monitoring data. It structures time-series data across seven core tables that cover different aspects of air quality monitoring, such as site details, measurement processes, and observation records (**Figure A3-1**). This model exemplifies how distinct types of data can be systematically categorised and related, thereby enabling effective data handling and analysis.

In the case of shallow geothermal energy systems, issues could arise from trying to accommodate diverse data types and system-specific attributes within a single table. Given these challenges, it is perhaps more effective to use separate tables for significantly different system types, like groundwater heat pump (GWHP) systems and borehole heat exchangers (BHE). Then, these tables can be linked through relational database techniques.

This approach allows for more straightforward, organised data storage and can make further analysis easier and faster. The following is an outline of this possibility:

- **Common attributes table**:
  - Holds data common to both GWHP and BHE systems.
  - Columns might include: **Main_ID** (primary key), **Location**, **Installed_Year**.
- **GWHP specific table**
  - Columns include: **GWHP_ID** (primary key), **Main_ID** (foreign key), and other system-specific attributes.
- **BHE specific table:**
  - Columns include: **BHE_ID** (primary key), **Main_ID** (foreign key), and other system-specific attributes.

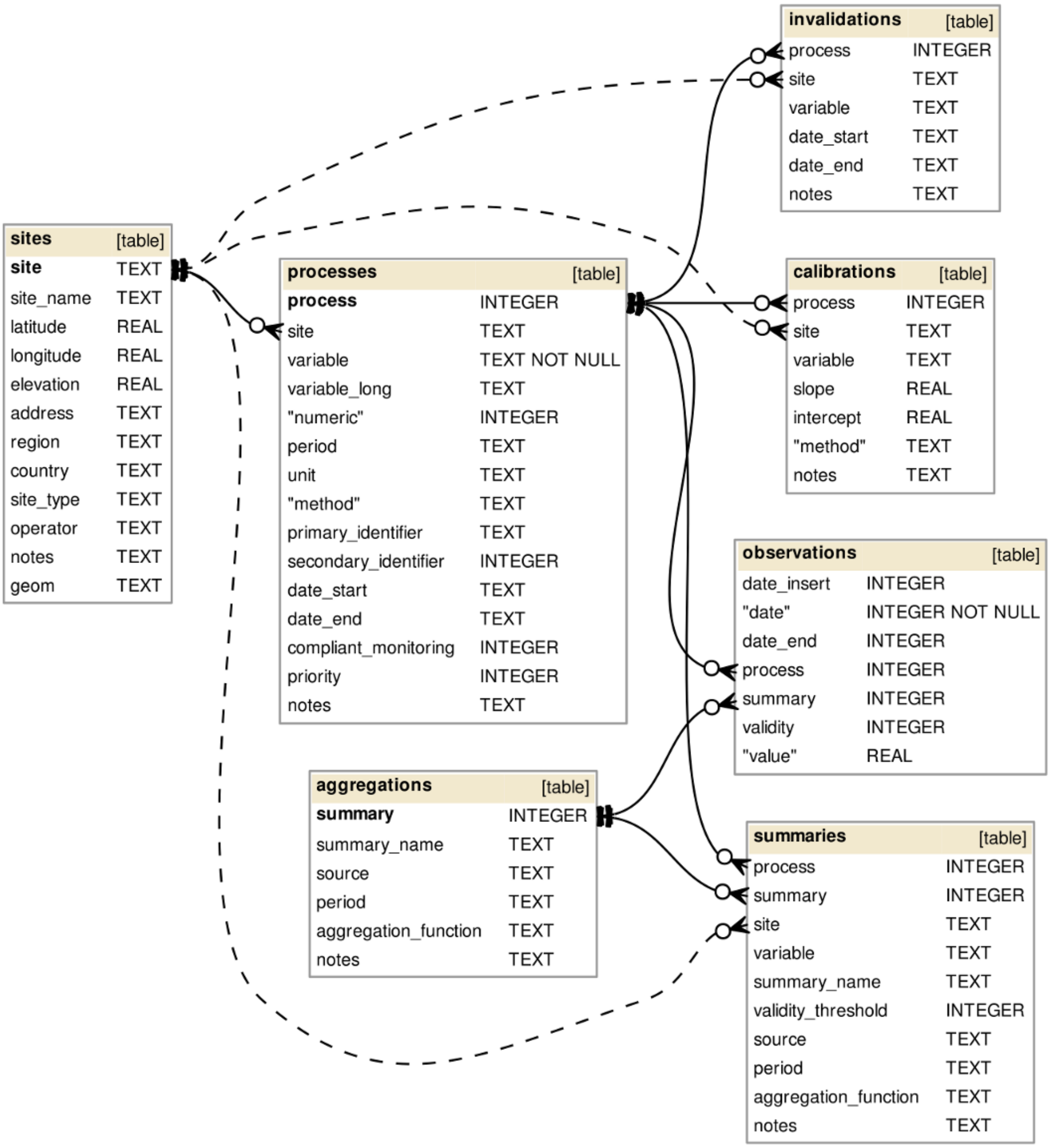

Fig. A3-1. Entity-relationship diagram of the `smonitor` core data model. Source: smonitor.

Concluding remarks

The preceding text on data models in SQL databases aligns primarily with the conceptual data model, the highest level of abstraction in data modelling. This sets the groundwork for subsequent, more detailed data modelling phases.

To sum up, data models are fundamental to the effective use of SQL databases. They provide a structured approach to primary data storage and subsequent retrieval. This facilitates data integrity and future analyses. The principles of tidy data, when applied to the design of tables, rows, and columns, is expected to further enhance the usability and integrity of a relational database.

Bibliography

Kleppmann, Martin. 2017. Designing Data-Intensive Applications: The Big Ideas Behind Reliable, Scalable, and Maintainable Systems. O'Reilly Media, Inc. https://www.oreilly.com/library/view/designing-data-intensive-applications/9781491903063/.

Co-fu
the E

Deliverable D.2.2: Reporting and monitoring geothe
Date: April, 2024

Objectives of the data sheets

The data sheets for shallow geothermal systems provided herein serve as templates showing (***i***) v should be generated or collected, and (***ii***) how the generated data could be shared with third-par the tables in these sheets is to ensure that the shared data is consistent and clearly structured. Or available in this format, it is almost ready for further analysis, reducing the time and effort require processing.

It should be noted that the sheets are not intended to be an exhaustive list of all possible paramet collected. The parameters in these sheets should be seen as recommendations, which may be exp to particular needs or interests.

Documentation separated according to main geothermal systems

The individual sheets are adapted to the main shallow geothermal systems:
BHE: Borehole heat exchangers
GWHP: Groundwater heat pumps
HOR: Horizontal collectors (e.g., normal horizontal collectors, slinky collectors, earth baskets)
TAG: Thermoactive geostructures (e.g., energy piles, thermally activated underfloor elements anc

Structure of the data sheets

To demonstrate how the different catalogued parameters can be reported, we present two sheet
Format_1: Contains individual entries for each borehole or well within a geothermal installation.
Here, specific parameters at well or borehole level and additional information relating to the enti
This format is suitable for studies that requires granular details of geothermal installations, e.g. o

Format_2: Contains a summarised overview of the geothermal installation.
This aggregated data format enables simple documentation of the entire geothermal installation

Recommendation

For reasons of completeness, we recommend the detailed compilation of the individual system p
The aggregated format (Format_2) can be derived from this if desired.

Final remarks

The aim of these sheets is ***not*** to attempt to replicate the complexity of a relational database but
For an insight into how the data could be structured for more complex database systems, we prep
This document can be found in Appendix 3 of the Technical Report (Deliverable D.2.2).

nded by
ıropean Union

ɜrmal heat pumps in Europe

vhich data we believe
ty users. The purpose of
ıce the data is made
ɜd for initial data

ɪers that could be
ɔanded upon according

ɪ diaphragm walls)

formats for the most common system types, BHE and GWHP. These sheets end with *_1* & *_2* :

ɾe system are captured.
ı the location of the individual system components.

without specific information for each part of the system.

arts (Format_1).

perhaps to provide a clearer and more efficient format for data delivery from the data providers (e
›ared the document titled "**Data Models in SQL Databases: The Basics**".

e.g., local authorities).