A machine learning approach for mapping the very shallow theoretical geothermal potential

Dan Assouline[1]*, Nahid Mohajeri[2], Agust Gudmundsson[3] and Jean-Louis Scartezzini[1]

*Correspondence:
dan.assouline@epfl.ch
[1] Solar Energy and Building
Physics Laboratory (LESO-PB),
Ecole Polytechnique Federale
de Lausanne, 1015 Lausanne,
Switzerland
Full list of author information
is available at the end of the
article

Abstract

The very shallow geothermal potential (vSGP) is increasingly recognized as a viable resource for providing clean thermal energy in urban and rural areas. This is primarily due to its reliability, low-cost installation, easy maintenance, and little constraints regarding ground-related laws and policies. We propose a methodology to extract the theoretical vSGP (installed in the uppermost 10 m of the ground, and mostly at depths of 1–2 m) at the national scale for Switzerland, based on a combination of Geographic Information Systems, traditional modelling, and machine learning (ML). The theoretical vSGP is based on the estimation of three thermal characteristics of the ground that impact significantly the geothermal potential, namely the monthly temperature at various depths in the surface layer, the thermal conductivity, and the thermal diffusivity. Each of the three variables is estimated separately, to a depth of 1 m below the surface, using the following general strategy: (1) collect significant data related to the variable, (2) if not existing, extract values for the variable at available locations with the help of traditional models and part of the data as input for these models, (3) train a ML model (with the Random Forests algorithm) using the extracted variable values as examples (training output labels) and related information contained in the data as features (training input samples), (4) use the trained ML model to estimate the variable in unknown locations, (5) estimate the uncertainty attached to the estimations. The methodology estimates values for (200×200) (m^2) pixels forming a grid over Switzerland. The strategy, however, can be generalized to any country with significant data (topographic, weather, and surface layer/soil data) available. The results indicate a very non-negligible potential for very shallow geothermal systems in Switzerland.

Keywords: Geothermal potential, Very shallow system, Geographic Information Systems, Machine learning, Switzerland

Introduction

In Switzerland there is currently great effort to increase the use of renewable energy. This effort is motivated by the Swiss Energy Strategy 2050, which sets as a goal to cease the use of nuclear power as a part of the energy mix by 2035, and reduce the CO_2 emissions by factor of 70 % by 2050. One promising solution is the large-scale deployment of ground-source heat pumps. While for many pumps the heat drawn from the surface layer is primarily from the sun, other pumps, particularly those using vertical boreholes, draw their heat from geothermal resources. Here we refer to the heat sources

as geothermal in the sense of being related to heat inside or from the earth. Geothermal energy is currently a commonly used renewable energy resource in Switzerland (Kemmler et al. 2018). Its main asset is that a ground connected heat pump can theoretically be installed almost anywhere (subject to regulations) and extract energy from the ground in an efficient way. Its use has been particularly common for single houses. Yet, a real impact over the whole country can be only reached with a large-scale deployment of ground-source heat pump. To do so, a thorough analysis of the thermal proprieties of the surface layers all the over the country is needed. Since, however, many different factors affect the thermal behavior of the ground, mapping the large-scale ground-source heat pump potential is a challenging task.

There have been several studies proposing large-scale (regional or national scale) methodologies for shallow (50 to 200 m depth) geothermal potential estimation. In particular, the potential for vertical Borehole Heat Exchangers (BHEs) has been studied often, since their high coefficient of performance (usually between 3 and 6) make them one of the most attractive means of exploiting shallow geothermal energy (Lund et al. 2004; Sanner et al. 2003).

Some studies extract the theoretical potential for shallow geothermal energy estimating a thermal ground-related variable at a large scale. In most studies, the estimated variable is the thermal conductivity, which is understandable given its high impact on the potential. Beamish (2013) made a GIS study on the thermal conductivity all over the UK using statistical sampling with airborne electromagnetic data together with an available geological database for the country. Di Sipio et al. (2014) made a GIS study based on sampling to extract thermal conductivity values for the Calabria region in Italy. Kalogirou et al. (2015) explored the use of machine learning, more specifically Neural Networks, in order to estimate a thermal conductivity contour map for Cyprus, based on a measurement training data of 41 points at different locations in the island. The features used as inputs for the models for each point were as follows: the lithology class, the elevation, air temperature statistics, rainfall, and the x and y coordinates. These studies, however, lack the estimation of other variables (e.g., thermal heat capacity, ground temperature gradient, etc.) to provide a full geothermal potential.

Many studies have developed large-scale methodologies for extracting the complete technical potential for shallow geothermal systems. Ondreka et al. (2007) proposes a methodology using GIS together with geological, hydrogeological, and lithological ground information, based on the German Verein Deutscher Ingenieure (VDI) guideline 4640, for extracting the technical geothermal potential for two study areas in Germany. The VDI guideline is a table that provides heat extraction values depending on the type of soil or surface rocks/sediments of a particular location, the number of operating hours of the pump, and other factors (Verein Deutscher Ingenieure. VDI Richtlinie 4640 Thermische Nutzung des Untergrundes 2001). Despite its being comparatively crude, the VDI guideline is very practical to obtain potential estimations and is particularly suitable for large-scale studies, where only very general geological information may be available. VDI has been used in many later studies, including a study by García-Gil et al. (2015), focusing on the groundwater flow to extract the technical potential, and in another recent study by Schiel et al. (2016), using the VDI guideline and demand values for extracting the very shallow geothermal potential in the urban area of a city in

Assouline *et al. Geotherm Energy*   (2019) 7:19

Page 3 of 50

Germany. Another strategy to extract the technical potential consists in first estimating significant ground thermal variables (commonly the thermal conductivity, ground temperature, and heat capacity, forming a theoretical geothermal potential) using suitable methods, and combining these variables with modelling in order to extract the technical potential. Galgaro et al. (2015) use multiple models to estimate the annual air and ground temperature, thermal conductivity, and monthly energy loads, to finally extract technical potential values using empirical modelling. Casasso and Sethi (2016) propose a quantitative method called G.POT to map the shallow geothermal potential. The thermal heat capacity, thermal conductivity, and ground temperature are first extracted from typical values (based on the type of rock/sediment) and empirical models. Then, the G.POT method uses analytical models to simulate the heat transfer in the ground and in the borehole with varying thermal properties and operational and design parameters of the system. It results in the technical potential estimation in the form of heat extraction values per year. In a later study (Casasso and Sethi 2017), the G.POT method was revised so as to extract the potential for open-loop installations such as groundwater heat pumps, in addition to BHEs. Furthermore, many studies have considered the geothermal potential of shallow aquifers (Casasso and Sethi 2017; García-Gil et al. 2015). A notable recent study estimates the geothermal potential of shallow aquifers in Finland (Arola et al. 2014). In this study, the heating capabilities of groundwater for buildings were extracted from the entire country based on the heat flux, temperature, thermal heat capacity of groundwater, and design of buildings.

Since very shallow geothermal systems are particularly suitable for cities, several studies attempted to account for specific urban conditions, and most notably the impact of urban heat islands. In particular, studies by Allen et al. (2003), Zhu et al. (2010), and more recently Arola and Korkka-Niemi (2014) and Rivera et al. (2017) show that the urban heat island effect has a very significant positive impact on the geothermal potential for BHEs installed in urban areas. These results are, in theory, also valid for shallower geothermal installations.

While traditional shallow geothermal installations (50 to 200 m depth) such as vertical Borehole Heat Exchangers can be of great use for urban settings and single family houses given their very high efficiency (Sanner et al. 2003), very shallow geothermal systems (VSGs) ($\leq$ 10 m depth) can often offer a viable alternative solution. Despite their efficiency being lower than that of the BHEs, VSGs offer many advantages, including (1) easy maintenance, (2) low-cost installation, (3) fewer legal constraints than for deeper installations, (4) possibility for technical improvements, and (5) a potential for installation almost everywhere. For the heat to be easily replenished around the collector, however, steady groundwater flow in the surface layer (normally soil or sediment) is the ideal condition. Thus, water-saturated surface layers offer much better heat sources for the heat pump collectors than dry and non-cohesive soils and sediments such as sand (Arola and Korkka-Niemi 2014; Lee and Hahn 2006; Milenić et al. 2010).

There have been recent efforts to develop, in addition to traditional horizontal loops or heat collectors, new VSGs using less space and with increased efficiency such as Slinky loops, helical heat exchangers, stacked tubes, and heat baskets. Several studies have focused on the main factors influencing the latter VSGs (Di Sipio and Bertermann 2017). Some other studies provide design guidelines and models for multiple VSGs

(Boughanmi et al. 2015; Dehghan et al. 2016; Xiong et al. 2015; Zarrella et al. 2013a, b; Zarrella and De Carli 2013). Thus, when the heat potential exists at a very shallow depth, it is now possible to model and size appropriate VSGs at any location. It is necessary, however, to conduct a detailed study of the thermal characteristics of the ground/surface layers at such depth before deciding on the best location of such systems.

Few studies have been made on large-scale vSGP estimation. Some recent articles test multiple methods to compute the apparent (very shallow) thermal diffusivity from ground temperature time series, using various analytical methods and numerical models based on the 1D heat equation. Such studies include those of Busby (2015), Rajeev and Kodikara (2016), and Andújar Márquez et al. (2016). These studies, however, aim at providing estimations for specific study areas rather than large regions. The main large-scale vSGP study is the one initiated by the ThermoMap project (Bertermann et al. 2013, 2014, 2015). This project aims at extracting the very shallow (top 10 m) geothermal potential in Europe as a whole, including specific case studies. Values are estimated for heat capacity and thermal conductivity based on near surface geology and hydrogeology information, the United States Department of Agriculture (USDA) soil texture classification, and equations from Kersten (1949) and Dehner (2007). For the ThermoMap project, the continental scale is used, which naturally reduces the resolution of the study. Eventually, a reliable methodology for vSGP at large (regional/national) scale is still to be developed.

Regarding the estimation of energy values and environmental modelling in general, machine learning (ML) methods have recently become widely used. In particular, many different ML algorithms have been explored for geospatial modelling of multiple environmental variables, including solar radiation and wind speed (Kanevski and Maignan 2004; Kanevski et al. 2009), forecasting of solar radiation over horizontal and tilted surfaces (Alessandrini et al. 2015; Hassan et al. 2017; Hussain and AlAlili 2017; Lou et al. 2016; Yadav and Chandel 2014), and short-term forecasting of wind speed and wind power prediction (Heinermann and Kramer 2016; Najeebullah et al. 2015; Treiber et al. 2016). Also, Joshi et al. (2014) used ML algorithms to perform a rooftop classification and provide a solar potential estimation over rooftops, and Assouline et al. (2017, 2018) used a combination of GIS and ML methods (Support Vector Machines and Random Forests) to map the technical solar rooftop potential in Switzerland. It has been very rarely used, however, for ground-related variable estimations. One of the main related studies is the work of Kalogirou et al. (2012), which used neural networks for extracting ground temperature maps at various depths in Cyprus, based on measurement data from 41 boreholes. In addition, Beardsmore et al. (2016) and Beardsmore (2014) developed a Bayesian inference (which can be seen as an ML sub-family of methods) software tool for geophysical joint inversions, in the framework of the National ICT Australia, helping for the detection of promising locations for geothermal energy exploration. Nevertheless, ML has never, to the best of our knowledge, been used for a geothermal potential mapping study.

In the present study, we propose a methodology to map the theoretical geothermal potential at very shallow depths (the top 1 meter of the ground/surface layer) and at the national scale for Switzerland. To do so, we use a combination of Geographic Information Systems, traditional modelling, and machine learning methods. The theoretical potential in this study includes an estimation of three significant thermal

properties of the ground/surface layer: thermal conductivity, thermal diffusivity, and temperature. Each of the three variables is separately treated and estimated, using multiple sources of data for processing and estimation, as detailed in the different sections of the paper. In the remaining sections, the paper focuses on the following: (1) the methods, (2) the data sources used in the study along with their characteristics and the associated pre-processing steps, (3) the presentation of the results for the three ground variable mentioned above, and (4) the discussion of the results and their implications for Switzerland.

Methods

Elements of soil structure and texture

The soil structure consists of the arrangement of solid, liquid, and void parts within the soil. It is defined by the following variables: the volume of air, water, solid soil, void, and total volume, which are, respectively, noted as V_a, V_w, V_s, V_v, and V_T. The void volume can be filled with air and water, so that $V_v = V_a + V_w$, and the total volume is given by $V_T = V_v + V_s = V_a + V_w + V_s$. The masses of the different parts are also of use and are noted as M_s and M_w, respectively, for the mass of solid soil and the mass of water.

Various soil texture quantities are often used to describe the amount of water or air within the soil:

- Volumetric Water Content (VWC) $:= V_w/V_T$
- Gravimetric Water Content (GWC or w) $:= M_w/M_s$
- Porosity $(n_p) := V_v/V_T = e/(1+e)$
- Void ratio $(e) := V_v/V_s = n_p/(1-n_p)$
- Saturation degree $(S_r) := V_w/V_v$
- Particle density $(\gamma_s) := M_s/V_s$
- Dry (bulk) density $(\gamma_d) := M_s/V_T$
- Water density $(\gamma_w) := M_w/V_w \approx 1\,\text{g/cm}^3$

Formulas can be derived to link some of these quantities, using their respective definitions. One of these formulas, expressing VWC as a function of the GWC, will be used in the present study. It states that

$$\text{VWC} = \text{GWC}\frac{\gamma_d}{\gamma_w}. \tag{1}$$

The soil texture differentiates soil types based on the arrangement of particles or grains and their sizes, defined by their diameter $\varnothing$. Often, the very coarse grains with $\varnothing > 2$ mm (block, rocks, and gravels) are isolated, and the soil texture is defined by the respective percentage of the three fine soil elements, namely sands (50 μm $< \varnothing < 2$ mm), silts (2 μm $< \varnothing < 50$ μm), and clays ($\varnothing < 2$ μm). Note that the percentages of sand, silt, and clays are given independently of the coarse minerals, meaning that the sum of the three percentages is 100%. In order to create a finite set of typical soil textures, soil texture classes can be extracted. The most common classification is the one created by the USDA, followed by that of the American Society of Agronomy, and used in the United States (Gee and Bauder 1986).

Interpretation of vertical electrical sounding data

Vertical Electrical Sounding (VES) is a classical geophysical method for estimating the electrical resistivity or conductivity of a medium. As it is one of the oldest and most trusted methods for extracting resistivity values and one of the least expensive to perform per unit depth, it is a very commonly conducted type of study. It is based on the measurement of the voltage between grounded electrodes that are installed at multiple distances from each other in order to reach various depths of the ground. The apparent electrical resistivity/conductivity is then given by Ohm's law as a function of the induced current, the measured voltage, and the geometry of the installation setup.

Although there exist multiple configurations for the installation of a VES study, the geometry of the electrode array can be generalized in order to extract generic formulas. A generalized form of array is defined by two main points A and B, in between which a current is induced. In the line joining A and B are two other points M and N, between which the difference of potential δV is measured. [For an illustration of such an array, see for example (Reynolds 2011)]. The value of δV and therefore the resistivity depend on the distance between A and B and the larger this distance the deeper in the ground the current flows between the two electrodes. As a result, the AB distance is gradually increased in practice in order to extract the apparent resistivity at increasing depths of the ground. The list of AB distance (or AB/2) with the corresponding measured apparent resistivities is what is often called a VES curve, or in the present paper VES data.

VES curves are said to be interpreted (or inversed) when the depth and corresponding resistivity of the different ground strata are extracted from the AB and apparent resistivity measurements. There are multiple methods to interpret VES curves, including fitting simple curve shapes, graphical modelling, or numerical modelling. The latter is, however, the most rigorous and up to date general method. Many algorithms and methods have been developed for the automatic inversion of VES curves (Friedel 2003; Haber 2004; Loke and Barker 1996, 1995; Reynolds 2011; Zohdy 1989). In the present study, a 1D inversion function from a C++/python library [pyGIMLi (Rücker et al. 2017)] is used. pyGIMLi generally uses regularization methods to perform inversion, with different schemes, including the popular Marquardt scheme. For more details on pyGIMLi and details about the inversion algorithms, please see (Rücker et al. 2017).

For more details on Vertical Electrical Sounding studies and geophysical exploration methods in general, the reader is invited to see (Reynolds 2011).

Fourier modelling for thermal diffusivity estimation

The behavior of the daily average soil or sediment temperature can be described at any location by the 1D heat conduction equation. The equation depends on the volumetric heat capacity c_v and the apparent thermal conductivity λ, which are generally functions of time and depth. In the commonly made assumption of uniform physical properties of the soil, however, both variables can be considered constant and the heat equation becomes

$$\frac{\partial T}{\partial t} = \alpha \frac{\partial^2 T}{\partial z^2}, \tag{2}$$

where T is the soil temperature at depth z, t the time, and $\alpha = \frac{\lambda}{c_v}$ the apparent thermal diffusivity, also considered constant by the uniformity assumption. The apparent

thermal diffusivity is therefore representative of the ground at one location, independently of time and space.

At very shallow depths (meaning < 10 m), the daily average soil/sediment temperature is characterized by a periodic behavior over the years, due to the seasonal changes of weather. Note that, despite this behavior, there exists a critical depth under which the temperature varies little throughout the year. This critical depth depends on the location and properties of the soil but is commonly 2 to 3 m. Following the assumptions that (i) the ground surface temperature (boundary condition at $z = 0$) is sinusoidal, (ii) the ground temperature is constant at infinite depth and equal to the average ground temperature, and (iii) the apparent thermal diffusivity is, as mentioned, constant with depth and throughout the year, the solution of Eq. 2 can be given by a Fourier series, defining an infinite sum of harmonic functions which can be fit to experimental data. The boundary condition [ground surface temperature $T(0, t)$], in particular, can be captured by the following Eq. 3 (Hurley and Wiltshire 1993):

$$T(0, t) = T_0 + \sum_{n=1}^{\infty} T_{Sn} \sin(n\omega t + C_n), \tag{3}$$

where T_0 is the average ground surface temperature over a year (the period of the Fourier series), T_{Sn} and C_n are the amplitude and phase of the harmonics defined by n, and $\omega = \frac{2\pi}{P}$ is the angular frequency of one period P (a year or 365.24 days).

The general Fourier solution of Eq. 2 is then given by (Hurley and Wiltshire 1993):

$$T(z, t) = T_{0,z} + \sum_{n=1}^{\infty} R_n(z) \sin(n\omega t + \phi_n(z)), \tag{4}$$

where

$$R_n(z) = T_{Sn} \exp\left(\frac{-z\sqrt{n}}{D}\right) \text{ and } \phi_n = \frac{-z\sqrt{n}}{D} + C_n \tag{5}$$

and $T_{0,z}$ is the average value of $T(z, t)$ over 1 year period (and also the constant of the Fourier series, usually noted as c_0) and D is damping depth, which traduces the decrease of the temperature amplitude when the depth increases (Carslaw and Jaeger 1959; Wijk 1963). D is given by

$$D = \sqrt{\frac{2\alpha}{\omega}}. \tag{6}$$

The coefficients $\phi_n(z)$ and $R_n(z)$ are the phase and amplitude of the harmonics of the solution given by Eq. 4 and can therefore be computed using Fourier analysis based on a set of temperature measurements at various depths and times (Hurley and Wiltshire 1993).

In order to use discrete temperature data, $R_n(z)$ and $\phi_n(z)$ are computed in practice with the Fast Fourier Transform (FFT). The FFT, in this study used from the NumPy Python library, approximates the Fourier series by computing N frequencies rather than an infinity of frequencies. The Nyquist sampling theorem gives the value for N as

$N = \left\lfloor \frac{N_s}{2} \right\rfloor$, where $N_s = 365$ (days) is the number of sampling points (one sample temperature value for each day).

The damping depth D, often assumed to be 1–3 m as indicated above, and therefore the apparent thermal resistivity can be computed using Eq. 5 together with the previously estimated values for $R_n(z)$ and $\phi_n(z)$. An exact computation, however, requires another Fourier analysis to extract T_{Sn} and C_n from Eq. 3. In practice, it has been shown that the slope of $\ln(R_n)$ vs. $z\sqrt{n}$ provides a reliable estimate of the damping depth (Carson 1963; Wijk 1963), which allows for an easier estimation of the apparent thermal diffusivity α.

Machine learning/Random Forests

Machine learning (ML) aims at building models that develop predicting abilities by gathering information from examples (training data) (Hastie et al. 2001; Jordan and Mitchell 2015). Generally, the larger the number of examples, the better the performance of the model. Formally, let Y denote the output variable (or target) that requires prediction and $X_1, ..., X_d$ the input variables (or features) of interest that were gathered for their expected impact on the behavior of Y. Let us also use small letters to denote a realization (meaning one point value) of the corresponding variables denoted by the capital letters, i.e., y for a realization of Y, and $x_1, ..., x_d$ for a realization $X_1, ..., X_d$. A regression task, in an ML framework, can be described as follows: given a training dataset $(\mathbf{x}_i, y_i)$ $(i = 1, ..., N)$ of N points (training samples), where $\mathbf{x}_i \in \mathbb{R}^d$ is the input feature vector and $y_i \in \mathbb{R}$ is the output value (or label) for point i, we seek to learn a function $\varphi : \mathbb{R}^d \to \mathbb{R}$ so that a prediction $\varphi(\mathbf{x})$ is as close as possible to the observed output y. While there exists a great variety of ML algorithms [see (Hastie et al. 2001) for reference], Random Forests was chosen for its multiple qualities, as discussed further later in the section.

Random Forests (RFs) (Breiman 2001) is an ML algorithm part of the Ensemble Learning family of methods. Ensemble Learning aims at training multiple weak learners (models with poor accuracy) and aggregating their prediction in order to obtain one model with a better performance. In the case of RFs, the weak learners are decision trees (Breiman 2017).

Decision trees are models that are built from the training data using a series of binary splits performed at each node of the tree. Each binary split consists in separating the data in two children nodes, defined by an inequality query on one of the variable (e.g., $X_3 < 2$). A tree is grown until either there is exactly one data point in each leaf (node that does not have children), or until a certain condition is fulfilled. A prediction can then be performed by passing a new point through the built tree, following the different queries imposed by the nodes, the estimated output value being the value of the data point contained in the leaf the new point ends in, or the average of the different data point values of the leaf. The heart of the algorithm resides in the optimization performed at each node of the tree in order to pick a variable and a threshold to form the query. For more details on decision trees, please see (Breiman 2017).

A Random Forest is a gathering of a certain number B of decision trees, trained using B bootstrapped versions (sampled with replacement) of the training data. In the case of regression, the RF prediction is given as the average of all the trees' predictions. The

Assouline *et al. Geotherm Energy*   (2019) 7:19

Page 9 of 50

key aspect of RF which differentiates it from other Ensemble Learning methods using decision trees is an additional layer of randomness at each node split: m variables are randomly picked out of all d variables and the best split using these m variables is used. This extra step increases significantly the performance of the model by decorrelating the trees, while increasing the speed of the training step.

RFs has multiple practical advantages, notably:

(i) It is mainly sensitive to only two hyperparameters: B the number of decision trees, and m the number of randomly chosen variables at each node split.

(ii) It is very fast to train.

(iii) It includes an embedded measure of error, called the Out-Of-Bag score (OOB), giving information about the generalization capability of the regressor in comparison with a simple average estimator (the closer the OOB score is to 1, the better the regressor can perform predictions for unobserved points).

(iv) It indirectly performs a feature selection process when building the trees by choosing the best feature (together with the best threshold) at each splitting. Thus, all intuition/knowledge-driven features can be used regardless of their respective significance without damaging the overall performance of the model.

(v) It includes an embedded measure of feature importance, namely the Variable Importance (VI). The VI evaluates the impact of each feature while training an RF model.

In addition, B does not require fine tuning as the performance can only increase with the number of trees, and a high number of trees cannot result in overfitting (Breiman 2001; Hastie et al. 2001). Therefore, it is advised to try increasing values of B until a performance plateau is reached and the optimal value is found (Hastie et al. 2001). Lastly, a modified version of RFs, namely Quantile Regression Forests (QRFs), allows the computation of Prediction Intervals (PIs) with a certain confidence level (90%, 95%, 99%, etc.) when using a trained RF to estimate a variable at new points (Meinshausen 2006), which provides a very useful estimation of the uncertainty of the prediction. For more details about the theory of QRFs please see Meinshausen (2006) or Assouline et al. (2018) for a quicker introduction in the framework of energy estimation. For more details about the Random Forests in general or its numerous embedded measures, please see Breiman (2001).

In the present study, each Random Forest model, using the RF implementation from the Scikit-Learn library, a python library for machine learning (Pedregosa et al. 2011), is trained following the same procedure. Each data set is split into a training data and a test data (75% for training, 25% for test). The number of variables chosen at each node split m is tuned using a sixfold Cross-Validation (CV) (Stone 1974). The choice of $K = 6$ in the K-fold CV was motivated by a rule of thumb suggested by Hastie et al. (2001), which consists in choosing a value for K that offers a CV error within one standard error of the minimum mean CV error. The best value for m is picked from a list of values that proved efficient in practice: $1, \left\lfloor \frac{d}{6} \right\rfloor, \left\lfloor \frac{d}{3} \right\rfloor, \left\lfloor \frac{2d}{3} \right\rfloor$, where d is the total number of feature variables available in the training data (Liaw and Wiener 2002). As previously mentioned, an increasing number of decision trees, B, can

only improve the performance of the model. Within the present study, $B = 1000$ trees for each model was enough to reach the performance plateau and therefore optimal accuracy for each variable prediction. Note that the depth of the trees trained within the RF model can have an impact on the performance, although less important than the previously mentioned parameters, as deep trees may lead to the overfitting of the training data (Hastie et al. 2001; Segal 2004). The RF depth is therefore controlled through the tuning of the minimum number of samples in a leaf by sixfold CV, using grid-search to find value offering the best results out of a list of possible values (1, 2, 3, 4, 5) together with *m*. The accuracy of the RF models is measured through the use of a testing error (error between predictions and actual output values in the test set). In the present study, two forms of errors are used, along with the OOB score: (i) the Root Mean Square Error (RMSE, also noted as E_R), given the same unit as the variable of interest, using Eq. (7), and (ii) the Normalized Root Mean Square Error (NRMSE, also denoted by E_{NR}), using Eq. 8, given in percentage. They are expressed as (Willmott and Matsuura 2005):

$$\text{RMSE} = E_R = \sqrt{\frac{\sum_{i=1}^{N_{test}} \left(y_i - \hat{y}_i\right)^2}{N_{test}}} \tag{7}$$

$$\text{NRMSE} = E_{NR} = \frac{E_R}{\bar{y}}, \tag{8}$$

where N_{test} is the number of test data points, y_i is the observed (real) output value, $\hat{y}_i$ is the value predicted by the trained RF, and $\bar{y}$ is the average of the observed values in the test set. Lastly, PIs with 95% confidence were computed both in the test set and for unobserved new points for suitable variables. While the PIs in the test set make it possible to cross-validate the PI computation with the known data, the PIs computed for new points allows to provide a measure of uncertainty along with predictions. The PIs in the test set are accompanied by a test confidence, that is, the fraction of observed points that occur within the boundaries of the computed PI. For a PI to be validated, the test confidence has to be around (ideally higher than) the confidence of the PI, for example 95% (0.95 in decimal form).

Finally, it should be noted that the spatial extrapolation capabilities of RFs have been questioned for some tasks in the past, particularly when the training points are located far from the points to be estimated (Rezvanbehbahani et al. 2017). Given that RF model building is highly dependent on the data itself, however, its performance must be assessed in each particular study. In the present study, and for each extrapolation task, experiments were conducted to verify such capabilities: in addition to the RF models trained/tested on the regular training/testing sets (randomly extracted from the labeled data, and as presented in the following sections), additional models we trained and tested, respectively, on clearly separated (spatially) sets. It showed good performances, as the RMSE and NRMSE appeared only very slightly higher than in the randomly extracted sets.

Data

Data sources and scale of study

Switzerland is a Central European country located at a latitude of 46°57′N and a longitude of 7°27′N, and has an area of 41,285 km^2, divided between the mountain chains of the Alps in the south and the Jura in the north. In between the mountain chains is the Swiss Plateau, where most of the urban areas are located. The population is around 8.5 million (2018). Switzerland is divided into 26 cantons, which are further divided into communes, the smallest administrative divisions in the country. The communes could therefore provide a natural division of Switzerland for a national-scale study of the potential for shallow geothermal energy or heat pumps. One issue, however, besides the relatively large size of the communes, is that the boundaries of communes have often changed over the years. To avoid that issue and provide a higher precision to the study, a grid of (200×200) (m^2) cells spanning over the entire country is used. The whole grid covering Switzerland consists of 1140×1925 pixels. The location of the cells is based on a Swiss digital elevation model (DHM25), presented later in the section. Throughout the study, the grid cells will be called *pixels* in order to avoid the possible confusion with raster data cells.

Several datasets, provided by federal offices, companies or particular studies were gathered in order to extract significant features impacting on the variables of interest and train machine learning models to predict the behavior of these variables. Table 1 gives a summary of all data sources used in the study, along with references and sources of data, including links to access the data, when available. The use of the dataset within the study is also specified: *features* means that the dataset is used for extracting one or more features to train an RF model; *labels* means that the dataset is used for extracting examples (training values) of a variable of interest to be further extrapolated over Switzerland using an RF model; *modelling* means that the dataset is used to collect inputs for a model (inversion, FFT) aiming at extracting examples for a variable of interest. Note that several data sources have multiple uses. For example, some datasets are used for two different estimations (the ground temperature data are used for both the ground temperature extrapolation and diffusivity estimation) or in two consecutive estimations steps (e.g., the precipitation data are used first as features for precipitation extrapolation, and then the obtained precipitation maps are used as features for the ground temperature estimation). Also note that more details on the processing and use of these datasets are given in the following parts of the article. In particular, the *features* dataset details are given in the following data section, while the *labels* and *modelling* dataset details are given in their respective sections (e.g., the VES processing is detailed in a dedicated subsection further in the article).

Data processing—feature selection

Feature selection is one of the most important steps when building a machine learning model. It consists in selecting significant features (variables) that impact on the quantity one would like to predict. Automated procedures exist to select good, or ideally the "best" features possible out of all the features gathered (Chandrashekar and Sahin 2014; Siedlecki and Sklansky 1993). Some algorithms (e.g., RFs) include an embedded feature

Assouline *et al. Geotherm Energy*   (2019) 7:19

Page 12 of 50

Table 1 Data characteristics used for estimating the theoretical potential

Data	Region	Inputs	Period	Time Res.	Spatial Res.	Error	Source	Use
Sunshine duration	Switzerland	Meteo stations	1981–2010	Monthly means	66 stations	Negligible	MeteoSwiss (IDAWEB 2018; Meteoswiss 2018)	Features, labels
Precipitation	Switzerland	Meteo stations	1981–2010	Monthly means	417 stations	± 4–40%	MeteoSwiss (IDAWEB 2018; Meteoswiss 2018)	Features, labels
Temperature	Switzerland	Meteo stations	1981–2010	Monthly means	91 stations	Negligible	MeteoSwiss (IDAWEB 2018; Meteoswiss 2018)	Features, labels
Cloud cover	Switzerland	Human observations	1981–2010	Monthly means	23 stations	NA	MeteoSwiss (IDAWEB 2018; Meteoswiss 2018)	Features, labels
Snow depth	Switzerland	Meteo stations + human observations	> 2000	Daily sum	291 stations	Negligible	MeteoSwiss (IDAWEB 2018; Meteoswiss 2018)	Features, labels
Ground temperature (various depths)	Switzerland	Measurement stations	> 2000	Hourly	47 stations	± 0.1–0.3 °C	MeteoSwiss (IDAWEB 2018; Meteoswiss 2018)	Labels, modelling
Soil moisture (various depths)	World	Satellite images	> 01/04/2015	30 s	3 km × 3 km	± 0.05 m^3/m^3	NASA SMAP (Das et al. 2018)	Features, labels
Digital elevation model (DHM25)	Switzerland	Elevation maps + aerial images	NA	NA	25 m × 25 m	± 2 m	Swisstopo (geodata4edu 2018; Swisstopo 2018)	Features
Geology cover polygons (GK500)	Switzerland	Geological Atlas of Switz. + other maps	NA	NA	13,320 polygons	± 0.02 m	Swisstopo (geodata4edu 2018; Swisstopo 2018)	Features
Soil structure (NABODAT)	Switzerland	Measurements (Samples)	NA	NA	6212 points	NA	FOAG (NABODAT 2018)	Features
Vertical Electrical Soundings	Switzerland	Measurements	NA	NA	4144 points	Coords: ± 10–30 m	SGPK (Dumont and Chapellier 2003)	Modelling, labels
Electrical/Thermal resistivity data	India	Experimental Measurements ("electrical res. box")	NA	NA	118 points (in lab.)	± 4%	various studies (Erzin et al. 2010; Sreedeep et al. 2005)	Features, labels

SGPK Swiss Geophysical Commission, *FOAG* Federal Office for Agriculture

selection in their core, which prevents them from seeing their predictive power drop in the presence of low-impact or redundant features (Breiman 2001). It is crucial, however, to gather significant features that were built from either intuition or expert knowledge in the domain of interest, given all the data available at hand. This procedure is called *feature engineering* (Dong and Liu 2018). In the present study, we look for features that have an impact on electrical and thermal properties of ground rocks and soils.

The main source of information on Swiss geology at the national level is the GK500 (or GeoCover500) dataset, provided by Swisstopo, which gives information on surface geological formations and materials for the whole of Switzerland. It is used to extract the first family of features that will be used within this study, the *geology features*. The data are available in a GIS vector polygon format. Each polygon represents the boundaries of a surface geological formation and includes various pieces of information about the formation. Some features were therefore naturally extracted from this dataset. Furthermore, as it is reasonable to assume the ground properties to generally remain similar within one formation polygon, all additional geological features are aggregated within the GK500 polygons (if not specified otherwise).

The features available from the GK500 data are all categorical (class-based) features, and include

- geological period [code: PERIOD] (classes include quaternary, tertiary, etc.),
- main types/classes of rocks [code: TYPE ROCHE] (classes are sedimentary, igneous, and metamorphic),
- detailed rock-type classification [code: LITH PET] (rock type classes include sand, silt, clay, limestone, gneiss, gabbro, basalt, andesite, etc.),
- yydrogeological characteristics [code: HYDRO] (classes include surface water, presence or absence of aquifers, etc.), and
- productivity of aquifers [code: PRODUCTIV] (classes include saturated from 2 to 10 m, saturated from 10 to 20 m, etc.).

The previously mentioned features need to be converted into real values. A "one hot encoding" approach is used in order to obtain real numbers from the categorical features: for each original variable, we create as many binary features as there are classes, defining the binary features as "variable=class1," "variable=class2," etc. and label with 0 or 1 each feature (1 if the point belongs to the class for this variable, 0 if it does not). All possible classes of each GK500 variable are listed in Appendix 1. It results in 107 geology features, as shown in Tables 4, 5, 6, 7, and 8.

Additional *soil/sediment texture features* are gathered for the study. Soil texture information (in the first meter of the ground) was extracted from the NABODAT (NAtionale BOdenDATenbank) dataset from the Swiss Federal Office for the Environment. The dataset contains various soil textural information for an array of 6212 measurements at various locations in Switzerland (mostly in the Swiss plateau), including the sand, silt and clay content, as well as gravel and stones content, in fractional values. The percentage sum of the sand and gravel fractions in the soil, denoted by *F*, is also computed at each measurement location. As the measurements are often available at multiple depths, all the content values were aggregated at each location,

using a weighted average with the depths as the weights. Therefore, the aggregated content value (for sand or silt, etc.) at a location l is computed as $\bar{q}_l$ defined as follows:
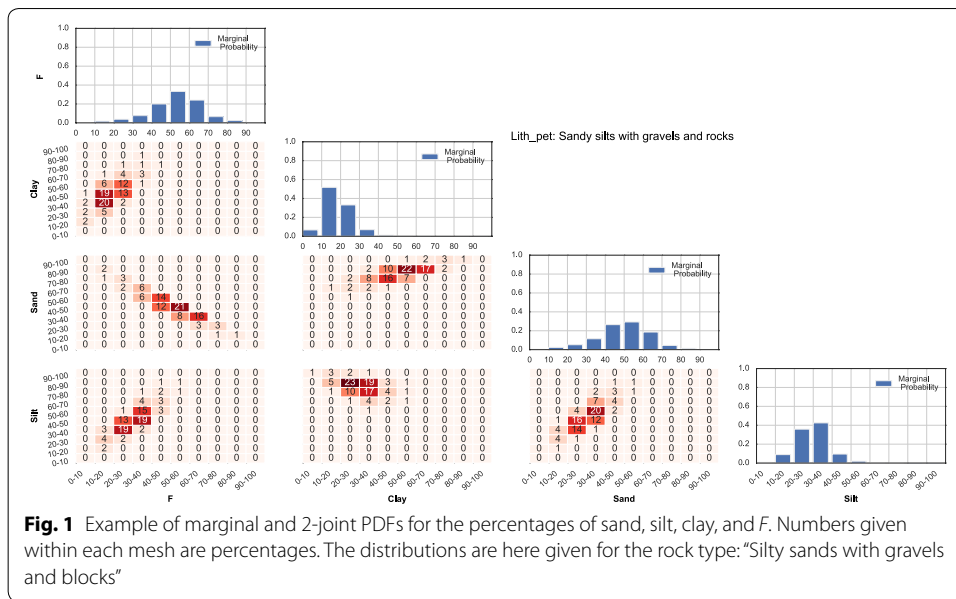
$$\bar{q}_l = \frac{\sum_{i=1}^{n_l} q_{i,l} h_{i,l}}{\sum_{i=1}^{n_l} h_{i,l}}, \tag{9}$$

where $q_{i,l}$, $h_{i,l}$, and n_l are, respectively, the measured content in the rock of interest in stratum i, the thickness (depth) of stratum i, and the number of strata considered in location l. Given that the soil texture depends more on the type of rock (LITH_PET) rather than on spatial location of the measurement, it was decided to aggregate the soil texture information by type of rock rather than within each GK500 geology polygon. For example, for the rock type "silts a sables avec graviers et blocs" ("Silty sands with gravels and blocks"), 1614 NABODAT points spanning over 1962 different GK500 polygons were recorded and considered for extracting statistical information. This information is then considered correct for each of these 1962 polygons containing this kind of rock. The number of polygons of each rock type and the number of NABODAT samples for these rock types are specified in Appendix 1. Note that, for statistical sampling validity, rock types covered by only 3 or less than 3 NABODAT sample points were not considered in the study, independently of the number of GK500 polygons typical for these types of rock. For example, the 291 GK500 polygons of one type of gneiss (more specifically "Gneiss with sericitic and chloritic schists") could not be considered, as only 2 NABODAT samples are of this rock type. Eventually, 10,812 polygons are characterized by soil texture information, out of the 13,320 GK500 polygons.

The statistical information for the soil texture is extracted in two forms: (i) classical summary statistics, and (ii) Probability Density Functions (PDFs). The first form of statistical information is extracted to serve directly as features. Statistics for sand, silt and clay content, as well as F, are computed for each rock type, using the ArcGIS Joint Spatial tool: the minimum, maximum, range, standard deviation, median, 25th and 75th percentiles. It results in 28 soil texture statistical features. The second form of statistical information is extracted to serve as weights (when converting electrical to thermal resistivity). PDFs for percentages of sand, silt, clay, and F are computed for each rock type using 10% wide bins (for example, we compute the probability that the percentage of sand is within $0-10\%$, etc.). We denote the random variables for sand, silt, and clay content, respectively, as Sd, St, and Cl, and the possible intervals as $I_1, I_2, \ldots, I_{10}$, for $[0-10\%]$, $[10-20\%]$, ..., $[90-100\%]$. In order to store all possible marginal and joint PDFs for the four structure variables efficiently, we compute the full joint PDF $\mathbb{P}\big(\text{Sd} \in I_i, \text{St} \in I_j, \text{Cl} \in I_k, F \in I_l\big)$, where $(i, j, k, l) \in \{1, 2, ..., 10\}^4$. Since the four structure variables are considered non-independent random variables, the full joint PDF is expressed as a function of conditional probabilities using a chain rule:

$$\mathbb{P}(\text{Sd}, \text{St}, \text{Cl}, F) = \mathbb{P}(\text{Sd} \mid \text{St}, \text{Cl}, F)\mathbb{P}(\text{St} \mid \text{Cl}, F)\mathbb{P}(\text{Cl} \mid F)\mathbb{P}(F), \tag{10}$$

where we omitted the I intervals to lighten the notation. Any marginal or joint PDF (with 2 or 3 variables out of the 4) can easily be extracted from the full joint by summing the probabilities over the unconsidered variables. As an example, Fig. 1 shows the marginals and the 2-joint PDFs extracted from the NABODAT samples for one example of rock type ("silty sands with gravels and blocks"). Note that the conditional probabilities were

Fig. 1 Example of marginal and 2-joint PDFs for the percentages of sand, silt, clay, and *F*. Numbers given within each mesh are percentages. The distributions are here given for the rock type: "Silty sands with gravels and blocks"

computed within each rock type in a frequentist fashion, counting the number of samples with particular sand, silt, clay, and *F* percentages.

Further terrain and weather features are aggregated within the GK500 polygons to bring more additional general information to the machine learning models. These features include the following:

- *Space features* Latitude, longitude, and altitude. Originally available for the whole Switzerland at a 25×25 (m²) resolution from the DHM25 digital elevation model (DEM), they are resampled at a 200×200 (m²) resolution. The latitude and longitude of each GK500 polygon are computed by extracting its centroid using the Graphics and Shapes toolbox (Jenness 2011). Statistics are computed for the altitude within each GK500 polygons (minimum, maximum, range, mean, standard deviation, and sum). It results in 8 space features.
- *Terrain features* Ground surface slope and aspect. They are computed from the resampled 200×200 (m²) DHM25 DEM using the Spatial Analyst toolbox from ArcGIS. To condense the terrain information, slope and aspect classes are created, as it is often done in the literature. We consider 9 aspect classes: 1 = flat, 2 = North, 3 = North-East, 4 = East, 5 = South-East, 6 = South, 7 = South-West, 8 = West, 9 = North-West; and 12 slope classes: 1 = [0–5°], 2 = [5−10°], ..., 9 = [40−45°], 10 = [45−50°], 11 = [50−60°] and 12 = [60−70°]. Statistics are computed for both features within each GK500 polygon based on classes: variety (number of different classes), majority (most frequent class), minority (least frequent class), "mean" class (the mean value of all registered classes using the class labels, even though it does not have any physical meaning), and median class. It results in 10 terrain features.
- *Weather features* Monthly mean air temperature, mean sunshine duration, mean precipitation, and cumulative snow depth. For each of the previously mentioned

Assouline *et al. Geotherm Energy* (2019) 7:19

Page 16 of 50

weather variables, except snow depth, monthly rasters derived in Assouline et al. (2018) are used. The rasters are built based on MeteoSwiss measurement data and the DHM25 DEM, using Random Forests. The monthly cumulative snow depth rasters are estimated in a similar fashion, using snow depth measurement data from MeteoSwiss. Statistics are computed for each of the weather variables within each GK500 polygons (minimum, maximum, range, mean, standard deviation, and sum). It results in 24 weather features.

Lastly, *soil moisture features* are extracted for the study. Soil moisture (in the top 5 cm of the ground) was extracted from the recently available Soil Moisture Active Passive (SMAP) satellite data from NASA, in the form of Volumetric Water Content (WVC). As the SMAP data were very recently collected by NASA (from 2015), it should be noted that our estimation of the soil moisture in Switzerland will only reflect its behavior during the past 3 years. Furthermore, the specificity of the SMAP mission is that the data are collected during 6:00 a.m. descending or 6:00 p.m. ascending half orbits (see https://smap.jpl.nasa.gov/data/ for more information). It therefore does not reflect the fluctuations of the soil moisture during the entire day. Lastly, the data were only available for 6 months in Switzerland during the last 3 years (January, February, March, October, November, December), and yearly average values (considering the six mentioned months) were therefore used. The data do not, as a result, allow for precise information about the monthly moisture fluctuations during summer. The spatial variations of the moisture, however, are nonetheless captured by the data. Originally with a resolution of 3 km × 3 km, and resampled to a resolution of 1 km × 1 km by NASA, the data were further resampled to follow the 200 m × 200 m pixel grid. A grid of SMAP polygon cells was built in ArcGIS based on the original SMAP points defining the centroids of the cells, using Thiessen polygons (polygons generated based on the centroids, so that any point within one polygon is closer to its centroid than to the other centroids). Then, the ArcGIS Joint Spatial tool was used in order to associate each 200 m × 200 m pixel with the moisture value of the SMAP 1 km × 1 km cell that contains the centroid of the pixel (the option `HAVE THEIR CENTER IN` is used when performing the Joint Spatial). While covering a large portion of the territory, the SMAP data do not span over the whole country (the spatial coverage is different depending on the month as well). Therefore, an RF model is trained for each available month using the soil moisture values in pixels as labels, and the previously mentioned space, weather, and terrain features. Testing errors, that is, the RMSE and the NRMSE are shown for each model in Table 2. The obtained yearly soil moisture (VWC) map of Switzerland is shown in Fig. 2.

Statistics of VWC are computed within each GK500 polygon to form the soil moisture features: VWC minimum, maximum, range, mean, standard deviation and sum. It results in 6 soil moisture features.

Then, a total of 183 features are computed for each GK500 polygon. Note that not all features will be systematically used to estimate each of the three ground thermal variables. Depending on the variable, an adequate subset of the mentioned features will be used.

Assouline *et al. Geotherm Energy*   (2019) 7:19

Page 17 of 50

Table 2 Testing RMSE (E_R) and NRMSE (E_{NR}, in percentage) for Random Forest models trained for monthly ground soil moisture (Volumetric Water Contents)

Month	Soil moisture		
	E_R (cm³/cm³)	E_{NR} (%)	OOB (−)
Jan.	0.08	19.61	0.82
Feb.	0.08	20.82	0.82
Mar.	0.07	22.75	0.85
Oct.	0.07	26.65	0.87
Nov.	0.07	23.88	0.87
Dec.	0.07	19.78	0.84

Fig. 2 Yearly soil moisture map estimated from the SMAP data in cm^3/cm^3

Results

In order to extract the very shallow geothermal theoretical potential (vSGP) for the surface layers or ground of entire Switzerland, three main variables must be estimated for the uppermost meters (in the present study, 1 m): the monthly ground temperature, the ground thermal conductivity, and the ground thermal diffusivity. Note that while these three variables are among the most important variables, they are not the only ones that affect the theoretical potential (Di Sipio and Bertermann 2017). The local groundwater behavior is notably another important factor, which is not studied here because of a lack of data (but could be added in a future study to complete the potential estimation). Note also that the thermal heat capacity can be derived from the conductivity and the diffusivity. The three variables are ultimately estimated at the aggregated level of 200×200 (m^2) pixels. In the following sections, we explain the details of the multiple steps leading
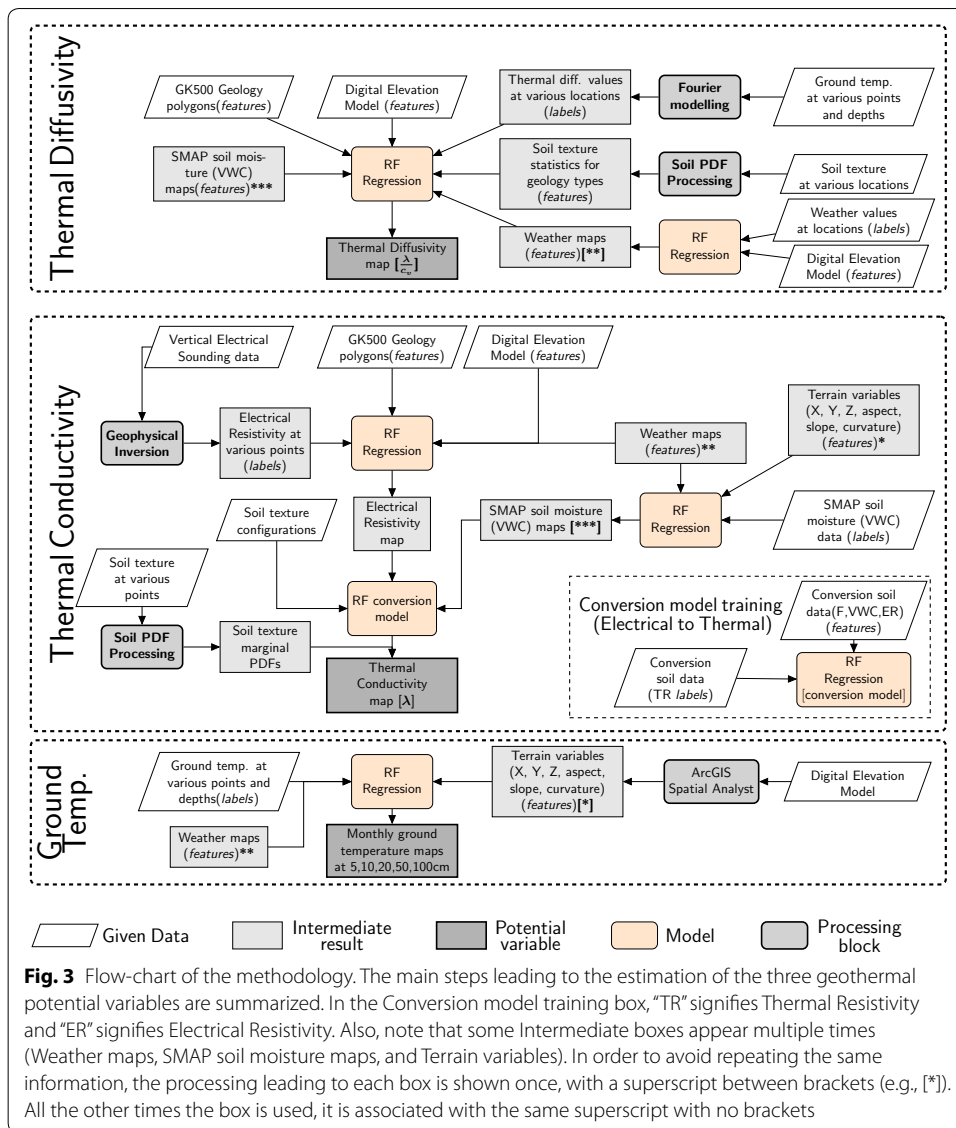
to the estimation of each of the three potential variables. These steps are summarized in a flow-chart in Fig. 3.

Ground temperature estimation

The first step in the potential study is the estimation of ground temperature maps in Switzerland at different shallow depths in order to assess the shallow thermal gradient. Given that we consider very shallow horizontal ground loop collectors or systems (loops), the seasonal weather variations may have a significant impact on the ground temperature, and yearly temperature values would not bring sufficient information. As indicated above, below the depth of 2−3 m, the soil temperature changes little throughout the year. However, most horizontal ground loop collectors are at the depth of 1−2 m, where there are significant changes in temperature throughout the year. The reasons why the collectors are placed at this shallow depth relate to the lack of thicker soil/sediment cover in the particular location, the higher costs of deeper trenches, and their suitability to provide an economic space heating.
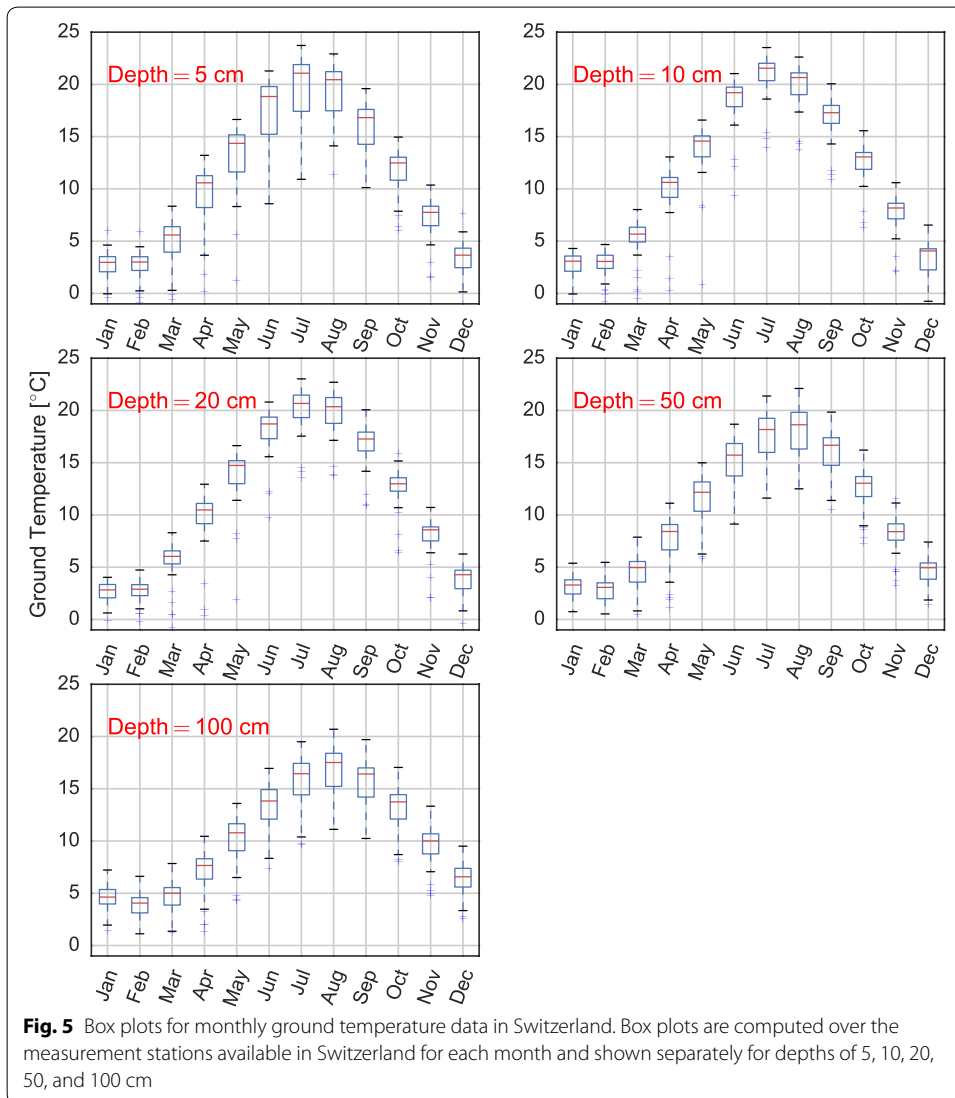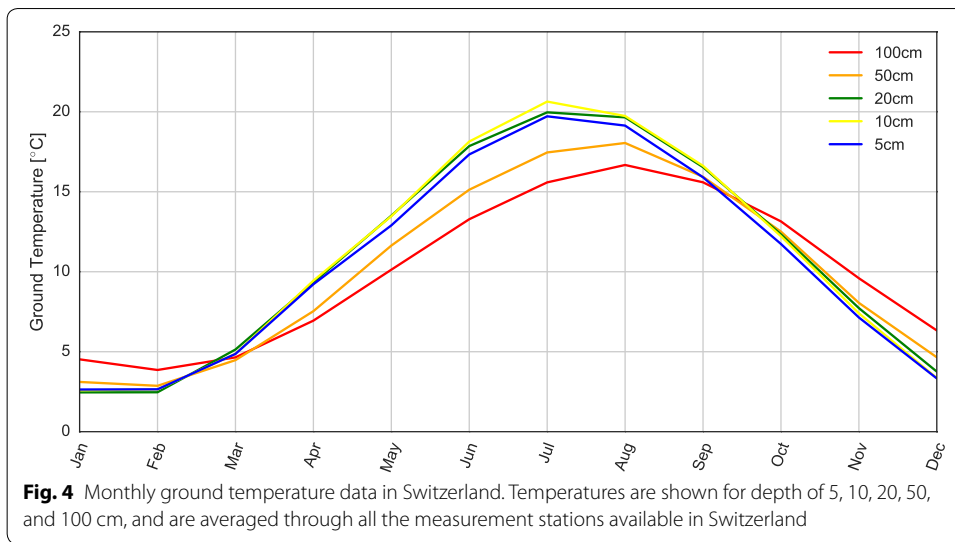
As a result, monthly ground temperature maps have been computed. The estimation is based on an hourly ground temperature time series data, available from MeteoSwiss (see Table 1) for multiple locations and at multiple depths, namely 5, 10, 20, 50, and 100 cm. The data are not always available for all the depths at the same locations for the same years. Therefore, the five datasets for the five different depths are treated separately. Each data is aggregated monthly through the available years, allowing for 12 typical average monthly values for each location. Intuitively, the seasonal variations should be attenuated with larger depths, as it is shown in Fig. 4, showing the monthly ground temperature at the five different depths, averaged through all locations available for each depth. Box plots are also computed for each month and depth (Fig. 5), showing the variability of temperature values across the multiple measurement stations. To allow for the estimation to be at the resolution of the 200×200 (m^2) pixels, the values of the available measurement stations are assigned to their nearest pixels. Each measurement corresponds to the pixel whose centroid is the closest to the location of the considered station. In the case where multiple stations are located within one pixel, the multiple measured values are averaged through the stations in order to provide one ground temperature value for the pixel.

RF models are trained using the pixel ground temperature values as training labels and weather and terrain variables as features: latitude, longitude, altitude, ground aspect and slope, and monthly precipitation, sunshine duration, snow depth, air temperature as defined in the data section. One RF model is built for each depth and for each month, leading to 60 different models. The model are used to extrapolate the training temperature data and build a ground temperature map for each month and each depth. Testing errors, that is, the RMSE and the NRMSE are shown for each model in Table 3. The resulting monthly maps for a depth of 100 cm as well as yearly maps for 5, 10, 20, 50, and 100 cm are shown in Figs. 6 and 7. Furthermore, 95% Prediction Intervals (PIs) for ground temperature have been computed at all depths and for all month, both in the test set and for new predicted points. A visualization of the PIs for 2 months (January and June) is shown in Fig. 8.
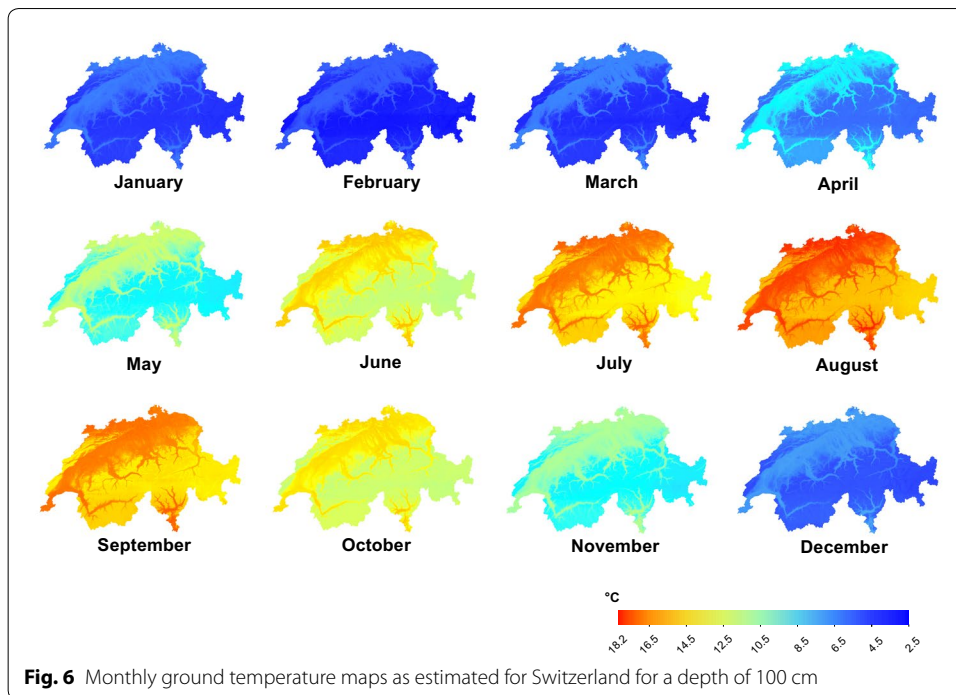
Fig. 3 Flow-chart of the methodology. The main steps leading to the estimation of the three geothermal potential variables are summarized. In the Conversion model training box, "TR" signifies Thermal Resistivity and "ER" signifies Electrical Resistivity. Also, note that some Intermediate boxes appear multiple times (Weather maps, SMAP soil moisture maps, and Terrain variables). In order to avoid repeating the same information, the processing leading to each box is shown once, with a superscript between brackets (e.g., [*]). All the other times the box is used, it is associated with the same superscript with no brackets

Thermal conductivity estimation

Due to a lack of data, the estimation of the ground thermal conductivity is less straightforward than that for the ground temperature. It includes multiple steps that aim at taking maximum advantage of the current data available at national level. The steps are as follows: (i) geophysical inversion of Vertical Electrical Sounding (VES) data, (ii) estimation of electrical resistivity values, (iii) spatial extrapolation of electrical resistivity in Switzerland, (iv) conversion of electrical resistivity into thermal conductivity. VES data are perhaps the most common resistivity data as it involves standard equipment and is a practical non-invasive geophysical study often performed to extract basic ground properties. The methodology presented here can therefore be re-used in other locations, should this sort of data be available. Also note that the conductivity could have been simply assumed from typical rock values. The focus of the study, however, is precisely to

Fig. 4 Monthly ground temperature data in Switzerland. Temperatures are shown for depth of 5, 10, 20, 50, and 100 cm, and are averaged through all the measurement stations available in Switzerland

Fig. 5 Box plots for monthly ground temperature data in Switzerland. Box plots are computed over the measurement stations available in Switzerland for each month and shown separately for depths of 5, 10, 20, 50, and 100 cm

Table 3 Testing RMSE (E_R), NRMSE (E_{NR}, in percentage) and OOB score for Random Forest models trained for monthly ground temperature at multiple depths

Month	5 cm			10 cm			20 cm			50 cm			100 cm		
	E_R (°C)	E_{NR} (%)	OOB (−)	E_R (°C)	E_{NR} (%)	OOB (−)	E_R (°C)	E_{NR} (%)	OOB (−)	E_R (°C)	E_{NR} (%)	OOB (−)	E_R (°C)	E_{NR} (%)	OOB (−)
Jan.	0.81	25.40	0.17	0.69	20.81	0.46	0.57	19.18	0.43	0.63	23.17	0.25	0.87	17.55	0.26
Feb.	0.57	18.69	0.15	0.65	20.20	0.48	0.69	22.24	0.34	0.66	26.88	0.27	0.86	20.27	0.24
Mar.	0.78	13.74	0.35	0.94	16.24	0.46	1.34	20.44	0.40	1.32	35.38	0.40	1.04	20.58	0.44
Apr.	0.88	08.49	0.59	1.12	10.62	0.46	1.35	12.30	0.46	2.00	32.11	0.44	1.11	15.36	0.52
May.	1.48	10.38	0.52	0.97	06.62	0.34	1.35	08.86	0.34	1.80	17.21	0.33	1.19	11.58	0.47
Jun.	2.34	12.48	0.56	1.22	06.41	0.42	1.22	06.35	0.42	2.02	14.62	0.31	1.34	09.95	0.47
Jul.	2.22	10.55	0.46	1.08	05.06	0.48	1.18	05.59	0.40	2.21	13.77	0.25	1.54	09.74	0.40
Aug.	1.68	08.28	0.53	1.22	05.95	0.48	0.96	04.62	0.46	2.09	12.47	0.29	1.51	08.88	0.37
Sep.	1.10	06.57	0.46	1.02	05.86	0.45	1.05	05.99	0.38	1.90	12.90	0.23	1.44	09.03	0.42
Oct.	0.83	06.61	0.37	0.87	06.66	0.34	0.71	05.43	0.28	1.57	13.65	0.33	1.12	08.25	0.46
Nov.	0.93	11.90	0.24	0.86	10.44	0.32	0.60	07.18	0.31	1.19	16.35	0.30	1.14	11.25	0.38
Dec.	1.10	29.08	0.06	0.78	18.89	0.35	0.60	14.27	0.43	0.80	19.53	0.32	1.05	15.31	0.39

Fig. 6 Monthly ground temperature maps as estimated for Switzerland for a depth of 100 cm
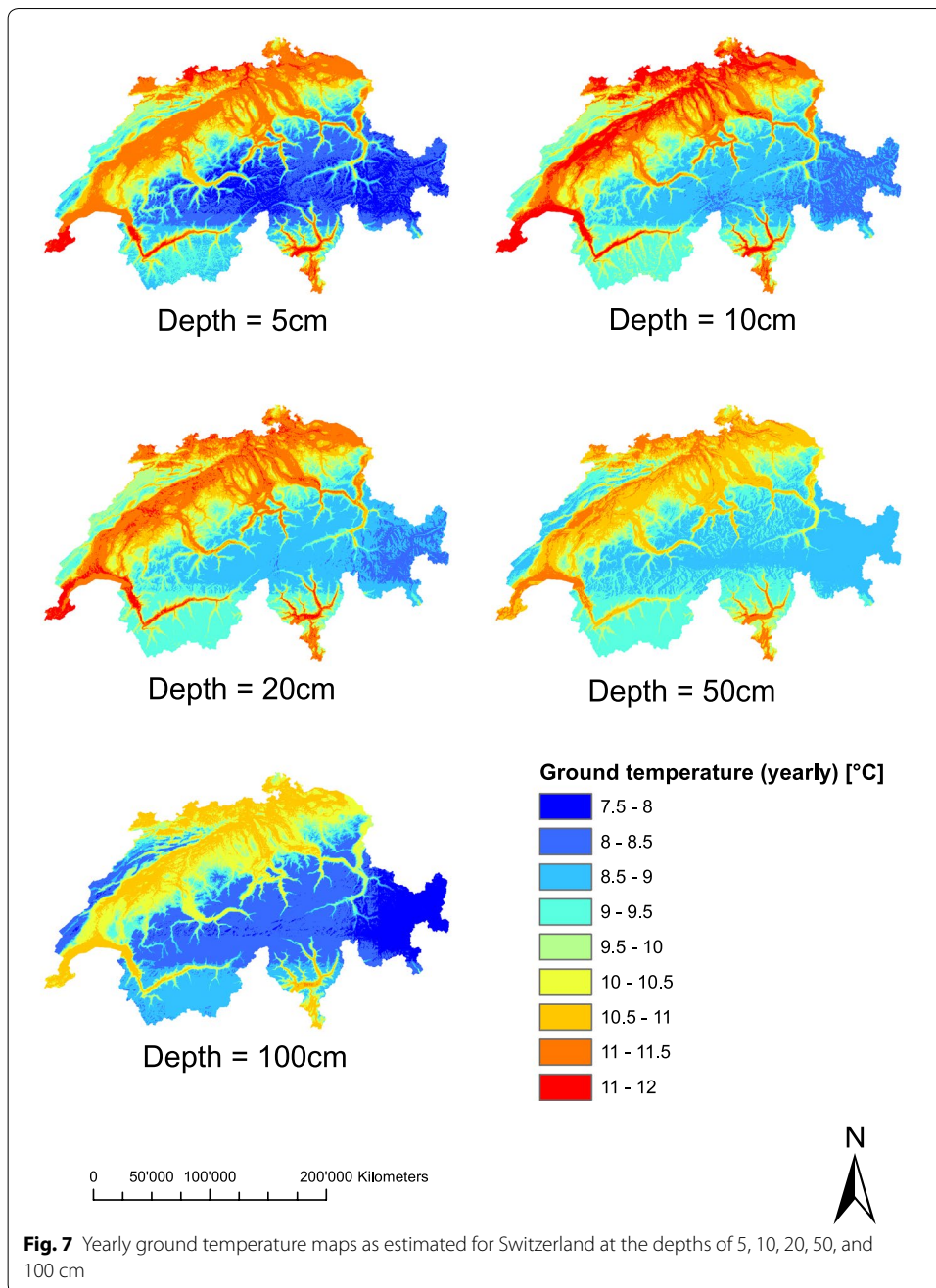
attempt to extract more accurate values from real data. The remaining parts of the section aim at explaining the details of the four previously mentioned steps.
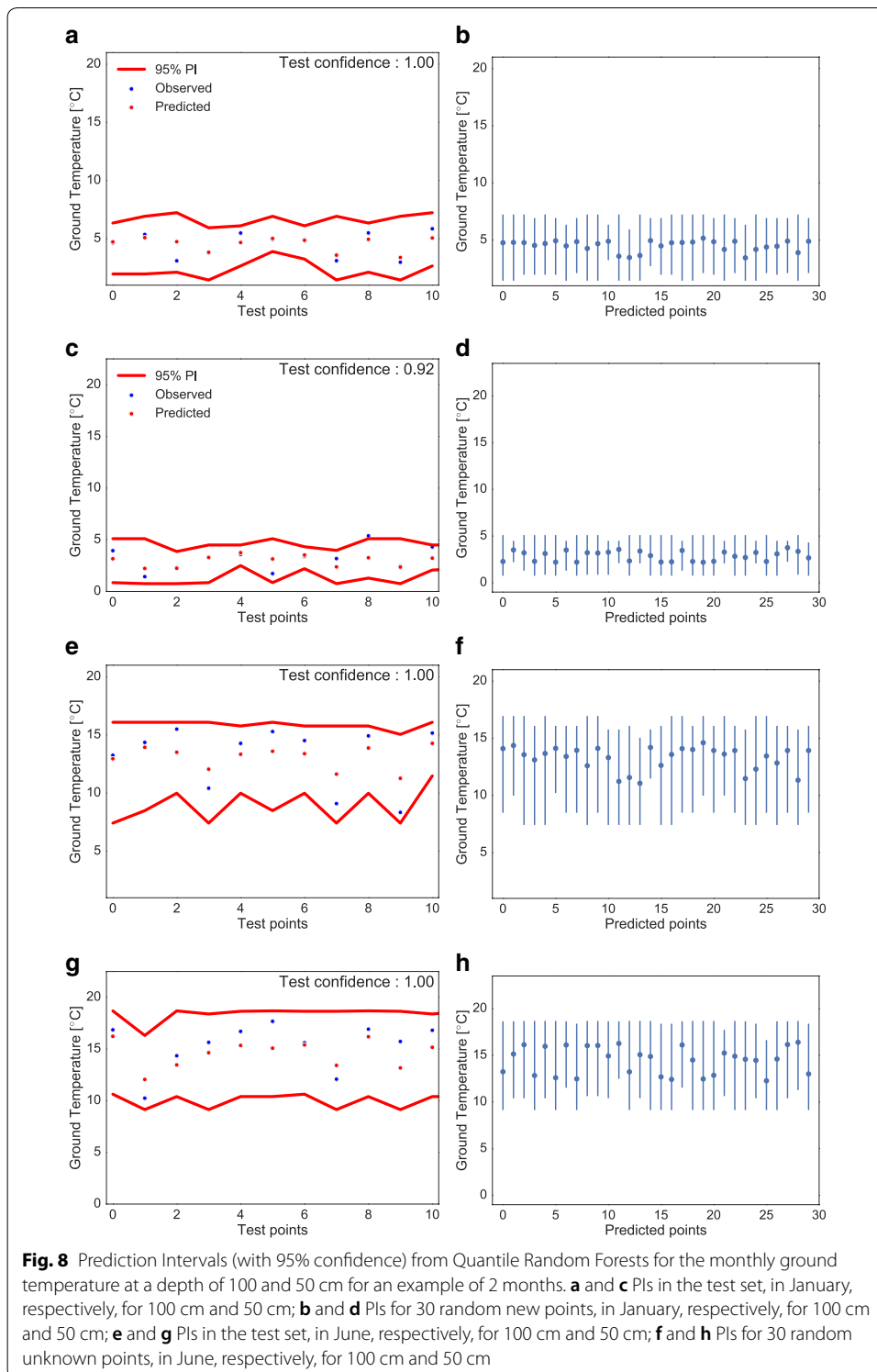
Processing and interpretation of Vertical Electrical Sounding data

The estimation of electrical resistivity across Switzerland is based on a Vertical Electrical Sounding (VES) dataset created by the Swiss Geophysical Commission (see Table 1). The data were built in an effort to gather multiple measurement studies performed over the last few years by multiple Swiss laboratories and universities (Dumont and Chapellier 2003). The dataset is split into two parts: (i) the raw electrical measurements which require interpretation (inversion), and (ii) the already inverted/interpreted electrical measurements. As a result, the two parts of the dataset were processed individually. The locations of all the points are shown in Fig. 9.

The first part of the dataset, which requires interpretation, includes 4144 points. For each point, the 1D inversion set of functions from the pyGIMLi library (Rücker et al. 2017) is used in order to provide an interpretation of the sounding data. Note that the number of different resistivity layers of the soil n_s at the measurement location is a parameter of the inversion algorithm. Therefore, a simple tuning strategy is performed: for each point, (i) the point is inversed separately with multiple values of n_s from 2 to 10, (ii) the resulting forward model is used to compute the apparent resistivities corresponding to the multiple distances between the electrodes, (iii) the n_s minimizing the RMSE between the original measurements and the forward modelled values is picked. Once n_s is picked, we obtain the resistivity in each different layer of the ground at the measurement point. Figure 10 shows an example of the inversion results for one point. Two additional constraints are considered in order to filter noisy data points: we do not

Fig. 7 Yearly ground temperature maps as estimated for Switzerland at the depths of 5, 10, 20, 50, and 100 cm
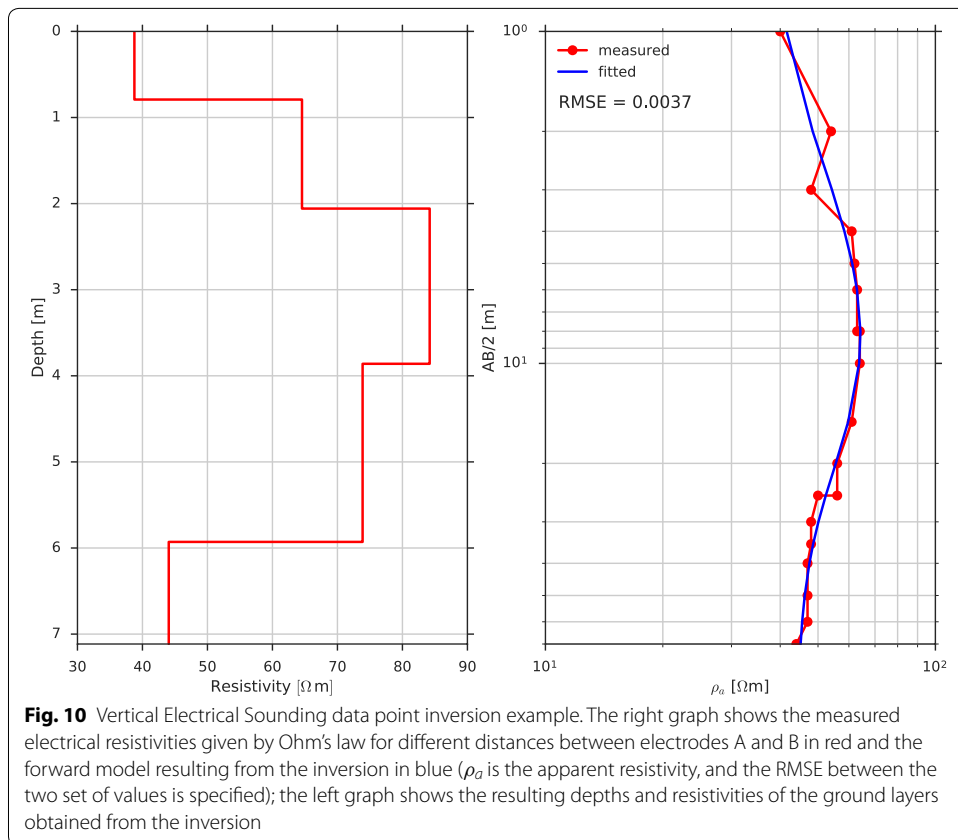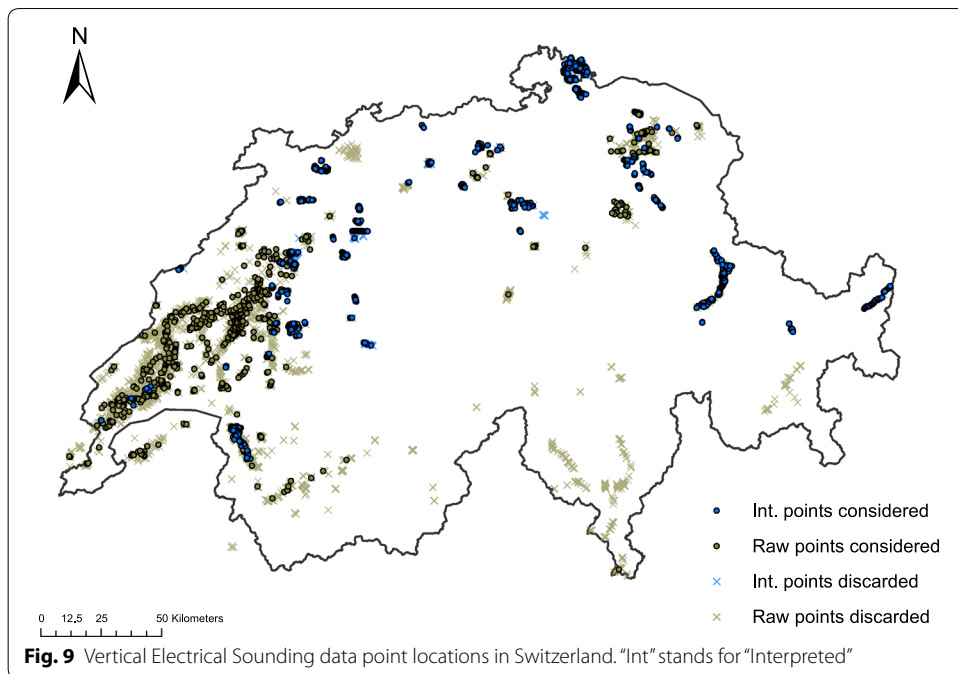
consider a point for which (i) the depths of the layers are non-strictly-increasing, as it has no physical meaning, (ii) the RMSE is greater than 30%. Note that the choice of 30% as a maximum threshold for the RMSE is motivated by a tradeoff between accuracy and number of points considered. While a higher threshold than 30% would signify poor lead to a poor accuracy, a lower one would result in a very low number of points considered (around 5% of the original 4144 points data). It results in 694 interpreted points in the first part of the data.

Fig. 8 Prediction Intervals (with 95% confidence) from Quantile Random Forests for the monthly ground temperature at a depth of 100 and 50 cm for an example of 2 months. **a** and **c** PIs in the test set, in January, respectively, for 100 cm and 50 cm; **b** and **d** PIs for 30 random new points, in January, respectively, for 100 cm and 50 cm; **e** and **g** PIs in the test set, in June, respectively, for 100 cm and 50 cm; **f** and **h** PIs for 30 random unknown points, in June, respectively, for 100 cm and 50 cm

The second part of the dataset, already interpreted, includes 1915 points. This part of the data specifies the layers' thickness and apparent resistivity at each measurement location. Note that the method or algorithm used for inversion is unknown, as the

Fig. 9 Vertical Electrical Sounding data point locations in Switzerland. "Int" stands for "Interpreted"

Fig. 10 Vertical Electrical Sounding data point inversion example. The right graph shows the measured electrical resistivities given by Ohm's law for different distances between electrodes A and B in red and the forward model resulting from the inversion in blue (ρ_a is the apparent resistivity, and the RMSE between the two set of values is specified); the left graph shows the resulting depths and resistivities of the ground layers obtained from the inversion

studies leading to these data were often performed many years ago and did not specify their strategy. Therefore, after verifying that the values are realistic (between 0 and 100,000 Ω m), the values have to be trusted. Furthermore, information (depth and/or resistivity) is missing in some points, which are excluded from the study. It results in 1521 interpreted points in the second part of the data.

The two parts of the data are merged together once both interpreted. Finally, the two data gather 2215 points across Switzerland.

Estimation and extrapolation of electrical resistivity

The whole VES data are processed to extract resistivity values at shallow depths for all 2215 filtered points. In order to obtain one shallow electrical resistivity value for each point, the resistivities at different depths are averaged through the first meter and weighted by the width of the corresponding layer. At each measured location, the average shallow resistivity ρ_{moy} is computed as follows:

$$\rho_{\mathrm{moy}} = \frac{1}{h_{n_{s<1\mathrm{m}}} - h_0} \sum_{i=1}^{n_{s<1\mathrm{m}}} \rho_i(h_i - h_{i-1}), \tag{11}$$

where ρ_i is the interpreted resistivity in layer i, h_i is the width of layer i, and $n_{s<1\mathrm{m}}$ is the number of layers in the first meter of the ground (the total width of all layers in the first two meters is therefore $h_{n_{s<1\mathrm{m}}} - h_0$).

The estimated shallow electrical resistivity values are aggregated within the GK500 geology polygons (see Table 1). Using the Joint Spatial function from ArcGIS to track the GK500 polygons covered by the VES interpreted points, the resistivity values from these points are averaged in each polygon. Note that in the case of a GK500 polygon containing only one interpreted point, this sole point defines the apparent electrical resistivity of the polygon. Eventually, 317 polygons are covered by at least one interpreted VES point and are therefore attached with an estimated shallow resistivity value.
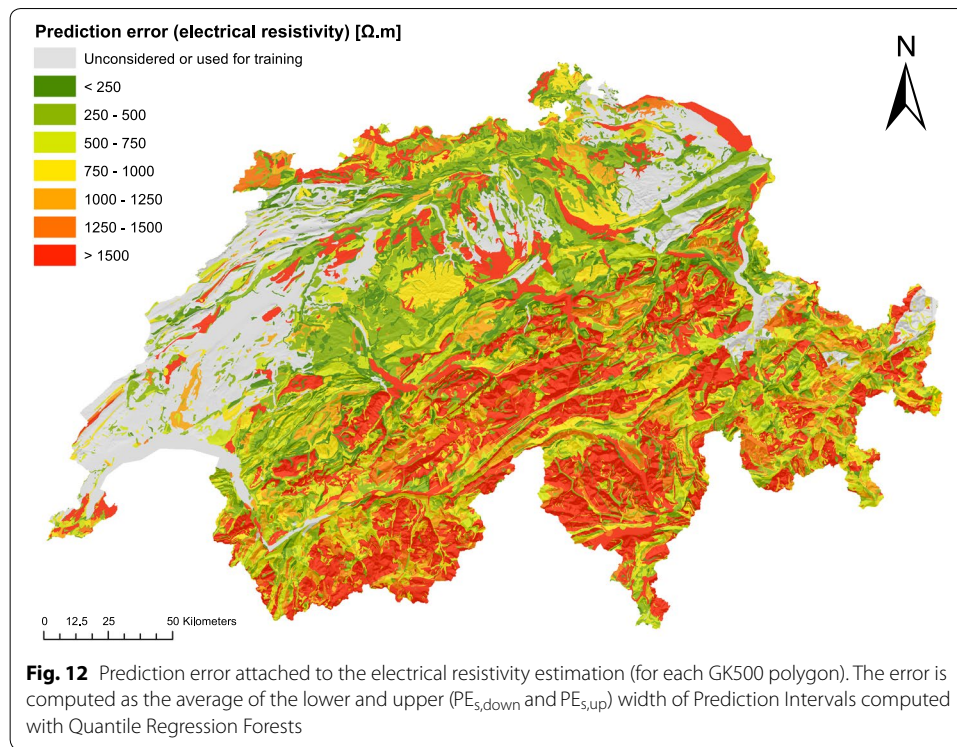
In order to further extrapolate the electrical resistivity over the whole Switzerland, an RF model is trained. The 317 resistivity values are used to build the training labels, and the space, weather, and all geological features (defined in the data processing section) sampled at the label points locations are used as training features. The label variable, however, is not the resistivity itself but $\ln\left(5 + \frac{\rho}{C}\right)$ where ρ is the resistivity within the polygon and C is the number of 200×200 (m^2) pixels contained in the polygon. This small modification allows to take advantage of the information of the polygon size while reducing the order of magnitude of the output. The test RMSE, NRMSE, and OOB score of the RF are, respectively, 0.25, 13 %, and 0.37. The RF model is finally used in order to estimate the electrical resistivity in each GK500 polygon over the whole Swiss territory. The obtained electrical resistivity map of Switzerland is shown in Fig. 11. Also, 95% Prediction Intervals (PIs) for electrical resistivity predicted values have been computed, both in the test set and for new predicted points. A visualization of the PIs for 30 test and new points is shown in Fig. 11. Note that the values plotted on the Y-axis are not the electrical resistivity ρ but the modified output used to train the RF ($\log(5 + \frac{\rho}{C})$ with C being the number of pixels in the polygon of interest). Finally, in order to show the distribution of the uncertainty attached to the estimation, PIs have been computed for all

Fig. 11 Electrical resistivity map as estimated in the study with visualization of PIs (with 95% confidence) both in the test set and for new points. **a** Electrical resistivity (ρ) map; **b** PIs in the test set for the label variable used ($\log(5 + \frac{\rho}{C})$ and not ρ) while training the RF model; **c** PIs for 30 random unknown points for the same label variable

GK500 polygons over Switzerland (besides the ones that were not considered from the start because of lack of data, or the ones used for training). The lower and upper width of the PIs can be seen as lower and upper prediction errors (respectively noted as $PE_{s,down}$ and $PE_{s,up}$), and the overall prediction error attached to the estimation of the electrical resistivity in each GK500 polygon is the average $(PE_{s,down} + PE_{s,up})/2$. The resulting error map is shown in Fig. 12.

From electrical resistivity to thermal conductivity

The conversion of electrical into thermal resistivity values requires information on the structure/texture of the soil. In particular, previous studies have shown that the most important parameters allowing this conversion are the particle size, moisture content, dry density, and saturation (Erzin et al. 2010; Singh et al. 2001; Sreedeep et al. 2005). These studies also developed models for the conversion: Sreedeep et al. (2005) suggested

Fig. 12 Prediction error attached to the electrical resistivity estimation (for each GK500 polygon). The error is computed as the average of the lower and upper ($PE_{s,down}$ and $PE_{s,up}$) width of Prediction Intervals computed with Quantile Regression Forests

a parametric model based on electrical resistivity, saturation degree S_r, and percentage sum of the sand and gravel fractions F, while Erzin et al. (2010) trained an artificial neural network using the same variables. The two previous models, however, could not be used in the present study as the saturation degree is not available over the Swiss territory. We use, instead, the data collected by both studies in order to train a conversion model to predict the thermal resistivity from the electrical resistivity. It is then straightforward to extract the thermal conductivity as the inverse of the thermal resistivity. Both data gather 135 points with experimental values of electrical and thermal resistivity, along with other soil characteristics, for different types of soils with various textures and structures. All 135 points offer, in particular, the dry density γ_d, the gravimetric water content (GWC), the saturation degree S_r, and the percentage sum of the sand and gravel fractions F. As saturation values are very challenging to gather at the scale of a country, we rather use the VWC to express the soil water content. The VWC can be obtained from γ_d and the GWC using Eq. 1.

The conversion from electrical resistivity to thermal conductivity consists of the following steps:

1. Import the combined data from (Sreedeep et al. 2005; Erzin et al. 2010).
2. Train an RF conversion model, with the following features: (i) the experimental values from the combined data for percentage sum of the sand and gravel fractions F, (ii) Volumetric Water Content (VWC), and (iii) electrical resistivity. The outputs are the thermal resistivity values from the combined data. The RMSE, NRMSE, and OOB score of the conversion RF trained in (2) are, respectively, 0.16, 17.6%, and 0.94%.

3. For each GK500 polygon in Switzerland, extract the previously estimated electrical resistivity and the VWC mean (extracted from NASA SMAP data, as presented in the data processing section), and consider the 10 possible values for F (5%, 15%, …, 95%) corresponding to the center of the 10 possible intervals (I_1,…, I_{10}) (data processing section).

4. For each GK500 polygon in Switzerland, use the trained conversion RF model together with the 10 possible configurations of VWC, electrical resistivity, and F in order to estimate the thermal resistivity for all 10 possible F values.

5. Consider the full joint PDFs extracted for each GK500 polygon (extracted from the NABODAT data, as presented in the data processing section) in order to extract the marginal PDF for F ($\mathbb{P}(F \in I_i)$). Use it to weigh the 10 possible thermal resistivity values with their respective probabilities and finally obtain the final thermal resistivity value for each GK500 polygon.

6. The thermal conductivity is then computed as the inverse of the thermal resistivity over the whole Switzerland.

Note that the joint PDF for soil texture is not used fully and only the probabilities for F are used. This is caused by the lack of sand, silt, and clay content information within the experimental data used for thermal and electrical resistivity values (Erzin et al. 2010; Sreedeep et al. 2005). We computed the full joint PDF nonetheless as it is a valuable piece of information that can be useful for later research. The obtained thermal conductivity map of Switzerland is shown in Fig. 13. In order to have an estimation of the RF prediction, 95% Prediction Intervals (PIs) were again computed for the conversion from electrical resistivity to thermal resistivity, before applying the F PDFs and the computation of the conductivity, both in the test set and for new predicted points. A visualization of the PIs for 30 test and new points is shown in Fig. 13. Note that the values plotted on the Y-axis are not the thermal conductivity λ but the modified output used to train the RF ($\log(\rho_t)$ where ρ_t is the thermal resistivity).
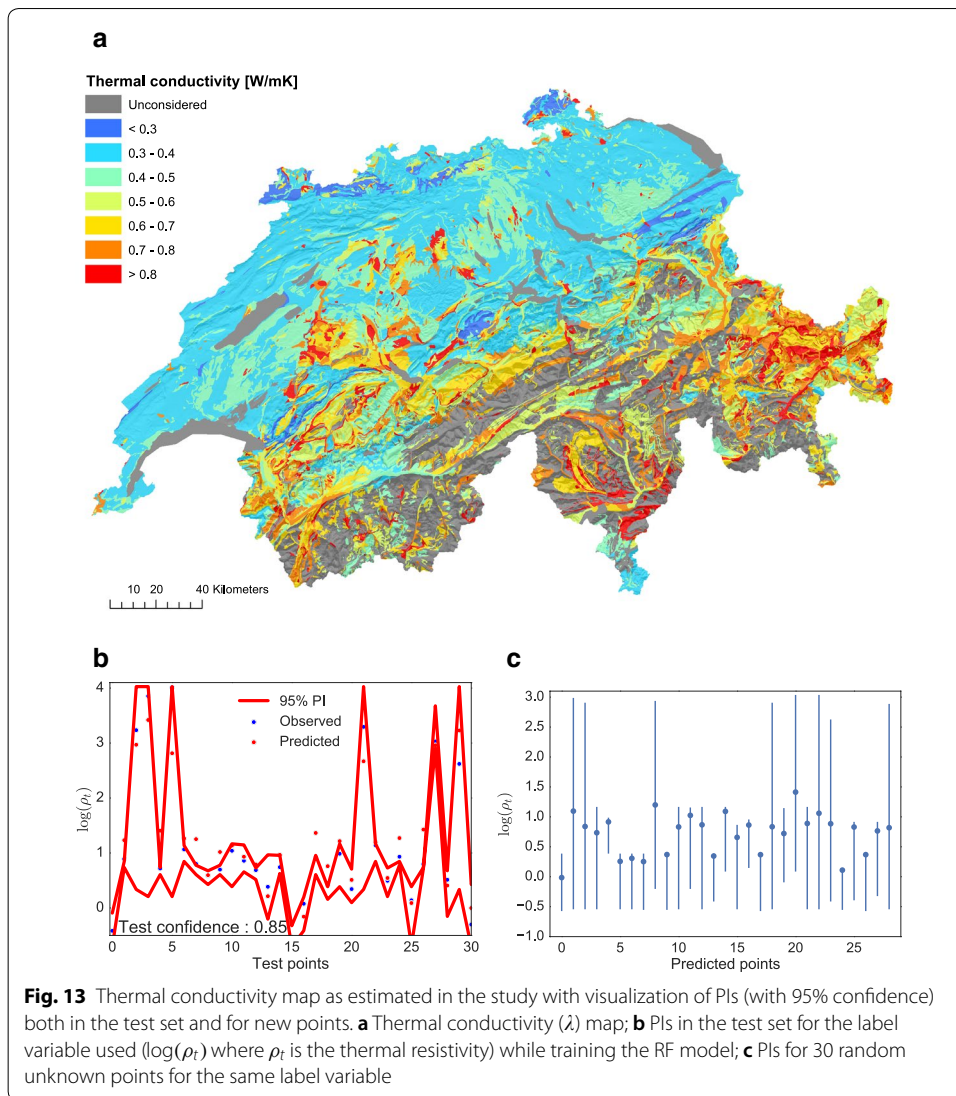
Thermal diffusivity estimation

The third and final thermal variable to estimate is the shallow thermal diffusivity, for each GK500 polygon in Switzerland. The estimation consists of two main steps: (i) estimation of shallow thermal diffusivity in locations where temperature data are available at various depths, using Fourier modelling of the 1D heat equation, and (ii) extrapolation of the diffusivity to the whole Switzerland.

Fourier modelling for thermal diffusivity estimation

The Fourier modelling strategy (presented in an earlier section) to estimate the apparent thermal diffusivity is applied to several locations in Switzerland where shallow ground temperature data (also presented earlier) is available from MetwoSwiss at 5, 10, 20, 50, and 100 cm.
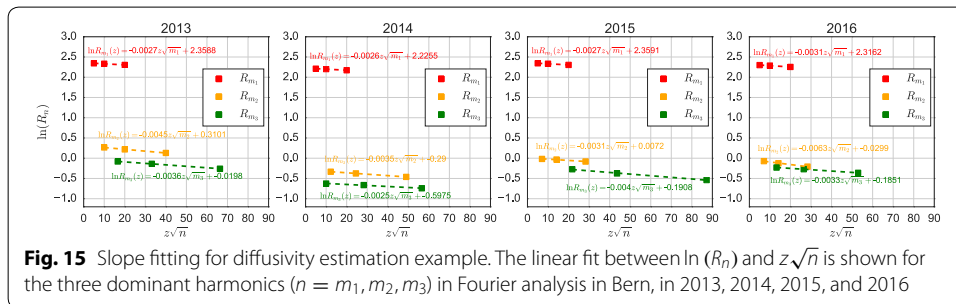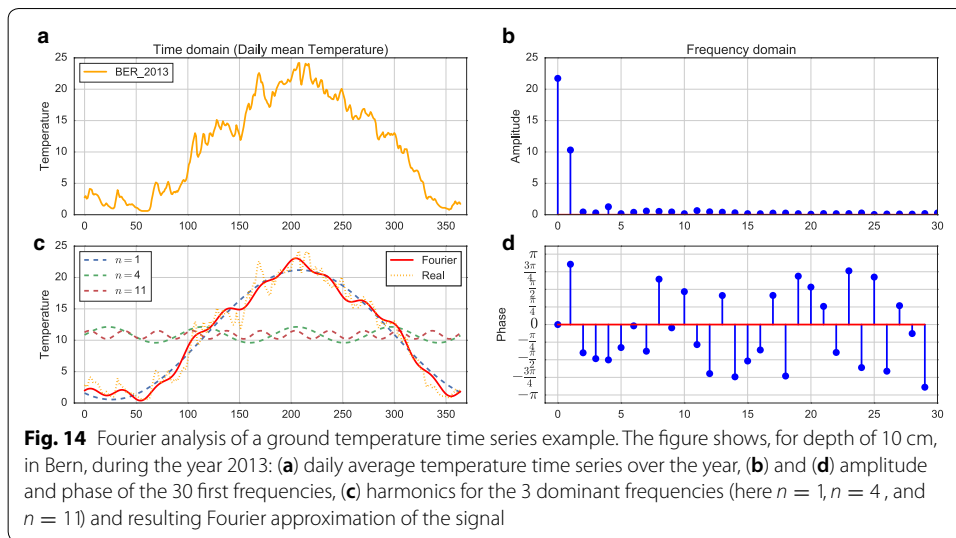
The Fourier series for the daily ground temperature is estimated at each available station, at each depth and for each available year using the FFT algorithm. The constant term of the series, which is by definition the average yearly temperature

Fig. 13 Thermal conductivity map as estimated in the study with visualization of PIs (with 95% confidence) both in the test set and for new points. **a** Thermal conductivity (λ) map; **b** PIs in the test set for the label variable used ($\log(\rho_t)$ where ρ_t is the thermal resistivity) while training the RF model; **c** PIs for 30 random unknown points for the same label variable

($T_{0,z}$), is first computed. Then, the amplitudes and phases (R_n and ϕ_n) of the frequencies for the multiple harmonics of the Fourier series are computed over a period of 1 year, meaning $P = 365.24$ days and $\omega = \frac{2\pi}{365.24}$. Although the first three harmonics ($n = 1, 2, 3$) are often enough to reproduce the signal with a good approximation (Rajeev and Kodikara 2016), we consider three dominant harmonics, namely the ones with the highest amplitudes. It allows for a better estimation of the signal. The reconstructed signal from Fourier analysis for one station (Bern), at one depth (20 cm), and for 1 year (2013) is shown in Fig. 14, together with the 30 first computed amplitude and phase values.

Following Fourier strategy, the slope of $\ln(R_n)$ vs. $z\sqrt{n}$ is then estimated, separately for the (5, 10, 20) cm and the (50, 100) cm time series. The estimated slopes in one station (Bern) for the (5, 10, 20) cm time series are shown in Fig. 15 for 2013, 2014, 2015, and 2016, where three dominant harmonics are defined by $n = m_1, m_2, m_3$. The linear fit for each of the three dominant harmonics is shown for each year. In Fig. 15,

Fig. 14 Fourier analysis of a ground temperature time series example. The figure shows, for depth of 10 cm, in Bern, during the year 2013: (**a**) daily average temperature time series over the year, (**b**) and (**d**) amplitude and phase of the 30 first frequencies, (**c**) harmonics for the 3 dominant frequencies (here $n = 1, n = 4$, and $n = 11$) and resulting Fourier approximation of the signal

Fig. 15 Slope fitting for diffusivity estimation example. The linear fit between $\ln(R_n)$ and $z\sqrt{n}$ is shown for the three dominant harmonics ($n = m_1, m_2, m_3$) in Fourier analysis in Bern, in 2013, 2014, 2015, and 2016

the slope does not vary greatly from one harmonic to the other, which validates the uniformity assumption. The mean slope is then computed by averaging all three harmonic slopes, which gives the damping depth and finally the apparent thermal diffusivity (using Eq. 6) for each of the 49 locations and each available year.

For each station, the estimated yearly thermal diffusivity estimations are cross-validated with typical values for various common rocks and soils given for two saturation states (Pahud 2002), given the type of rock from the GeoCover500 polygon data (see data section). The value is validated if it is within the typical minimum and maximum values ± 0.5 (10^{-6} m^2/s) of the corresponding rock for each year; otherwise it is discarded. The final diffusivity values for each of the 49 available stations are computed as the average year diffusivity value for that station.

Extrapolation of diffusivity

In order to further extrapolate the estimated thermal diffusivity over the whole Switzerland, an RF model is trained. The training data are represented by the GK500 polygons in which we estimated the diffusivity. In the case of multiple stations within one GK500 polygon, the considered diffusivity in the polygon is the average of the diffusivities at the included stations; otherwise the sole estimated diffusivity value defines the diffusivity over the whole polygon. It results in 47 training (polygon) points.

For each point, the estimated diffusivity is to be used as label for the RF model, and the considered features are the following: the space, weather, and GK500 geological features, as well as the soil moisture and the soil texture features (all presented in data processing section). The soil texture ones, however, are not available for all polygons throughout the country as some of the possible rock types are not represented by a NABODAT soil texture measurement point (as explained in data processing section). As a result, a second RF model is trained over the 47 training points without the soil texture features. This latter model can later be used to obtain diffusivity estimations in polygons that lack that texture information.

To improve the performance of the prediction, the label is slightly modified: instead of the diffusivity α, we consider $\ln(\alpha \times C)$, where C is the number of 200×200 m^2 pixels contained in each GK500 polygon. The resistivity estimation makes it possible to take advantage of the information of the polygon size, while reducing the order of magnitude of the output. The testing RMSE, NRMSE, and OOB score are, respectively, 0.69 (10^{-6} m^2/s), 13.7%, and 0.87, for the RF model considering the soil texture information, and 0.82 (10^{-6} m^2/s), 16.4%, and 0.78 for the RF model not considering the soil texture information. While both RF models show good performances in the test set, the model taking soil texture into account has a better accuracy, which aligns with intuition. The apparent thermal diffusivity is then estimated in all GK500 polygons in Switzerland, using the previously trained RF models (taking $e^{(\cdot)/C}$ of the prediction to recover the diffusivity from the predicted modified label). The obtained thermal diffusivity map is shown in Fig. 16. Finally, 95% Prediction Intervals (PIs) have been computed for the estimated diffusivity values, for both RF models (with and without soil texture information), both in the test set and for new predicted points. A visualization of the PIs for the test points and 30 new points is shown in Fig. 16. In addition, the prediction error estimated from the PIs computed for all GK500 polygons show the spatial variation of the uncertainty attached to the thermal diffusivity estimation (Fig. 17).
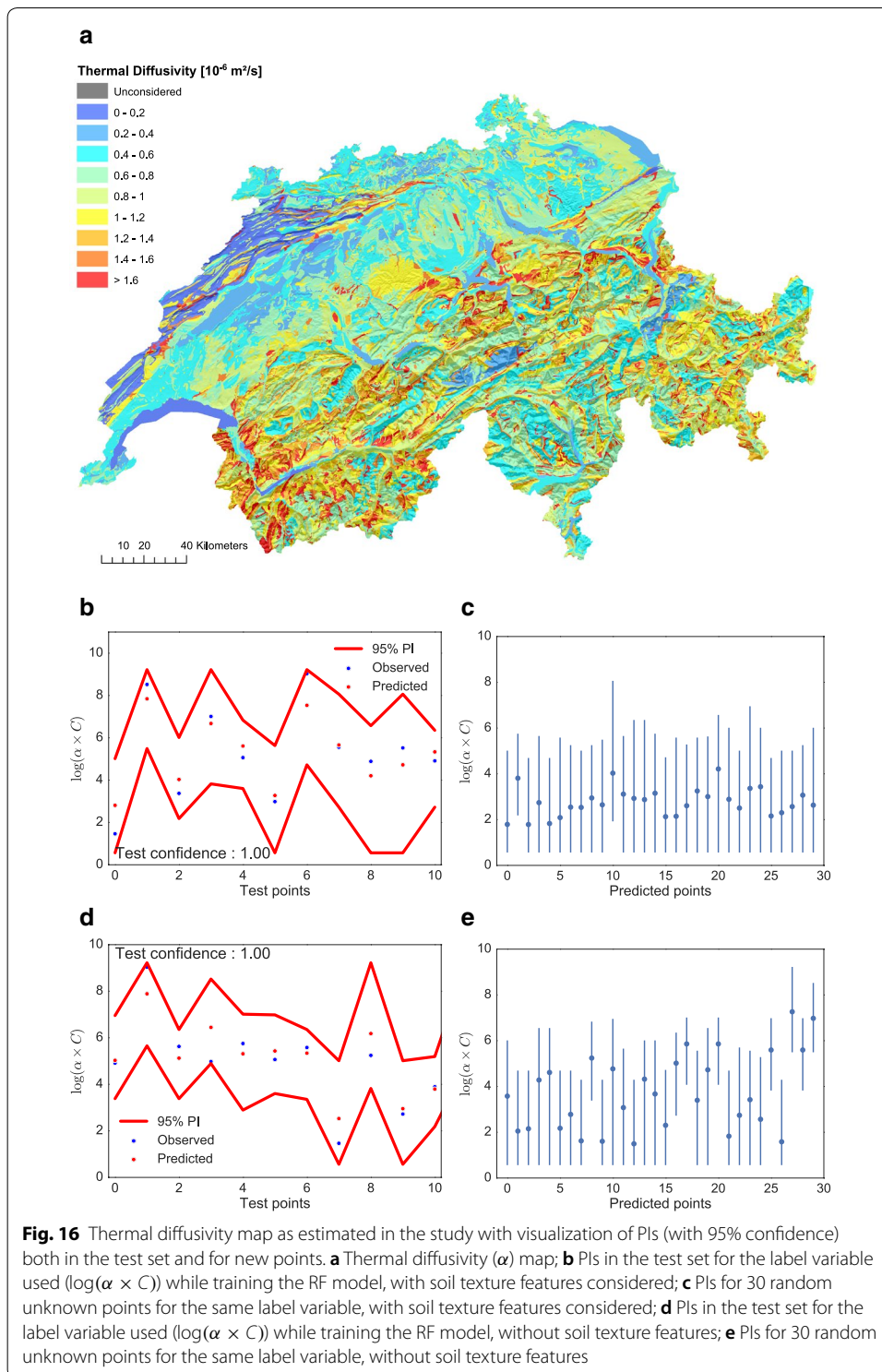
Re-aggregation in 200 × 200 (m²)

The estimated maps for thermal conductivity and thermal diffusivity were re-aggregated from the GK500 vector polygon resolution into a 200×200 (m^2) pixel raster in order to match the resolution of the ground temperature maps. For both variables, the value attributed to each pixel was computed as follows: (a) if the pixel is fully included within a GK500 polygon, then the value of the pixel is the value estimated for the polygon, (b) if the pixel is located at the boundaries of multiple polygons, then the value of the pixel is the weighted average of the values estimated for the multiple polygons, the respective polygon weights being the percentage of area they cover within the pixel. Note that for pixels in case (b), the polygon area percentages were computed using the Tabulate Intersection tool from the Statistics toolbox within ArcGIS.
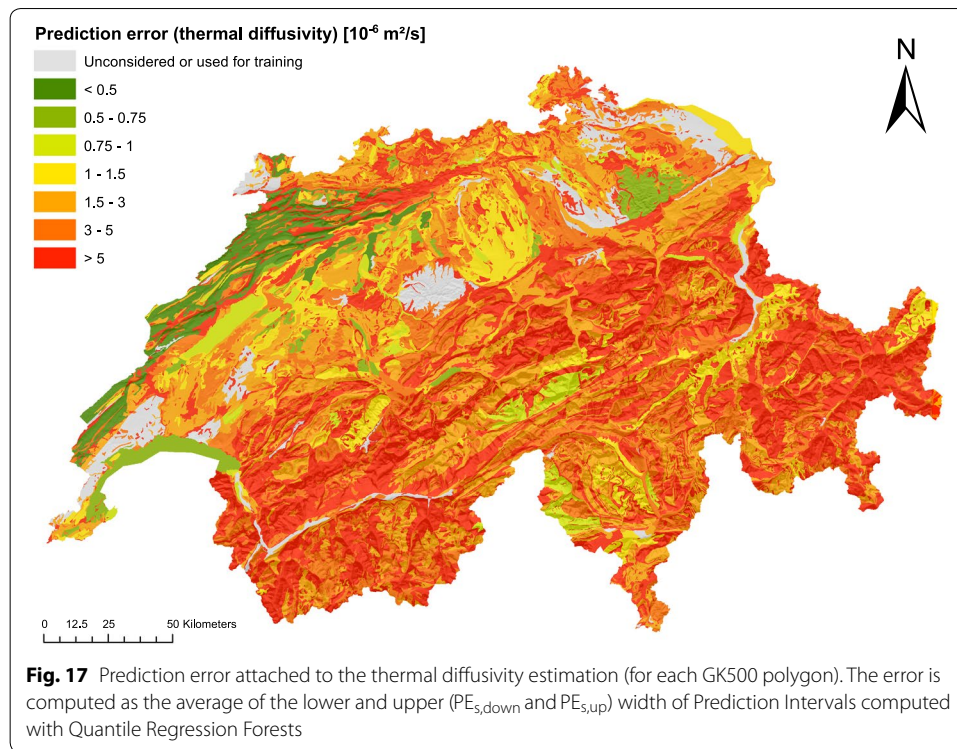
Discussion

Three main variables affecting shallow geothermal potential, namely the monthly ground temperature, the ground thermal conductivity, and the ground thermal diffusivity, have been estimated in Switzerland, at a shallow depth of 1 m over the entire Swiss territory,
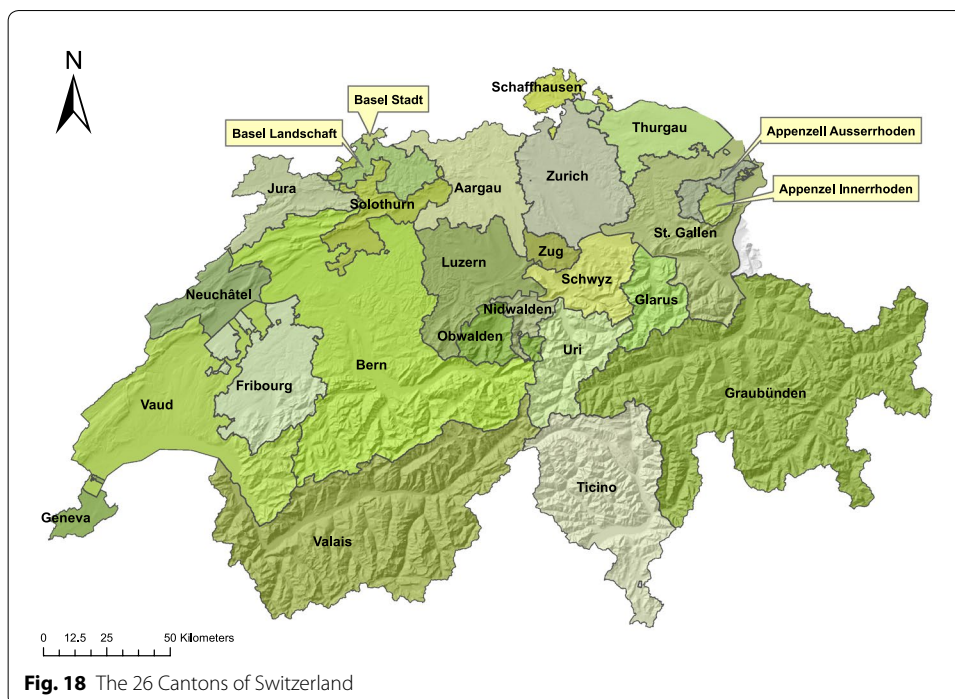
Fig. 16 Thermal diffusivity map as estimated in the study with visualization of PIs (with 95% confidence) both in the test set and for new points. **a** Thermal diffusivity (α) map; **b** PIs in the test set for the label variable used ($\log(\alpha \times C)$) while training the RF model, with soil texture features considered; **c** PIs for 30 random unknown points for the same label variable, with soil texture features considered; **d** PIs in the test set for the label variable used ($\log(\alpha \times C)$) while training the RF model, without soil texture features; **e** PIs for 30 random unknown points for the same label variable, without soil texture features

at a resolution of 200×200 (m^2) pixels. The results for the three variables may be used and combined for a future complete potential study. It is desirable, however, that each of the maps obtained for the three variables are discussed and validated in some fashion,

Fig. 17 Prediction error attached to the thermal diffusivity estimation (for each GK500 polygon). The error is computed as the average of the lower and upper (PE$_{s,down}$ and PE$_{s,up}$) width of Prediction Intervals computed with Quantile Regression Forests
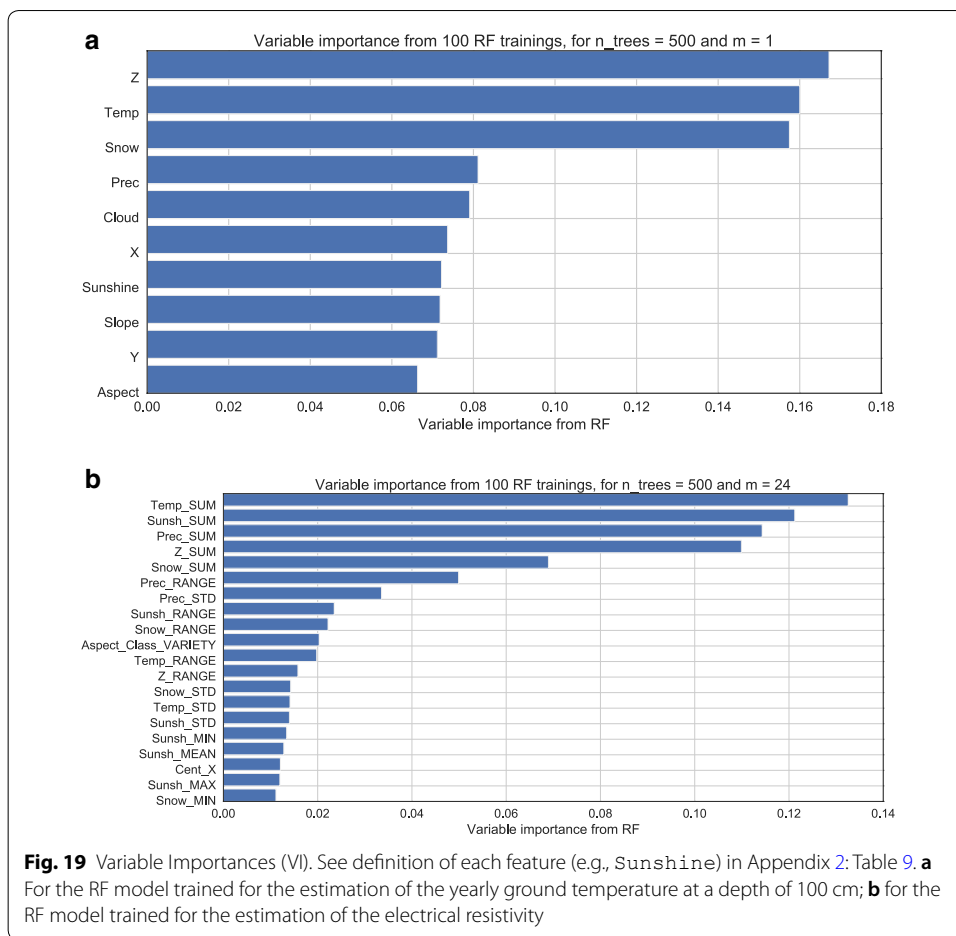
with the current available information at hand, before proceeding with the further combination of the estimations. In order to follow the discussion on the spatial variability of the multiple estimated values, a map of the Swiss cantons can be seen in Fig. 18.

Figure 7 presents the obtained yearly ground temperature maps for four different depths (5, 10, 20, 50, and 100 cm). The figure shows an instinctive pattern: the smaller the depth, the closer the ground temperature is to air temperature. Conversely, the deeper the depth, the less the ground temperature is affected by spatio-temporal variations. In particular, while the ground temperature is always higher in the Swiss plateau (where the altitude is lower) than in the Alps (where the altitude is higher), one can observe on the yearly maps that the general temperature difference between the two regions is smaller at a depth of 100 cm (difference of around 3 °C) than at a depth 5 cm (difference of around 4 to 5 °C). This latter depth of 100 cm is highly relevant for very shallow ground-source heat pump installations, the most common ones (horizontal collectors) which are mostly at 1–2 m depth, and hence is the one requiring more attention. In Fig. 6, one can observe the monthly variations of the ground temperature at 100 cm. Although the seasonal variations are significant, they are not so drastic as in the uppermost tens of centimeters, where the temperature very much resembles the air temperature. In particular, the coldest and warmest periods are characterized by less extreme temperatures. The depths below the uppermost tens of centimeters allow the use of ground-source heat pumps in the winter, as 3 to 4 °C is sufficient for heating purposes. In order to assess the impact of each feature used for the ground temperature estimation, Variable Importances (VI) (presented in section) are computed. In particular, the VI attached to a yearly RF model used for the depth of 100 cm can be observed in Fig. 19a (the features are displayed in decreasing order of importance from top to bottom) and show that altitude,

Fig. 18 The 26 Cantons of Switzerland

yearly air temperature, and yearly snow cover are the most important predictors. One can also observe on the Prediction Intervals (PIs) in Fig. 8 that the test confidence is higher than 95% for all four PIs shown in the test set. While the PI in Fig. 8c has a confidence of 92%, it is the result of one observed point being very slightly outside of the PI boundary. Also, even though the test set is very small, the respective impacts of ground depth and seasonality in Switzerland are shown in the PIs extracted for new predicted points: (i) the width of a PI is generally larger at 100 cm than 50 cm, as the temperature at 100 cm is more stable throughout the year and with a slightly broader range of possible values than at 50 cm, where a specific month dictates a narrower range of values reflecting the air temperature, (ii) the width of a PI is generally larger in spring/summer than in winter, as the temperature shows often more fluctuations in spring/summer at higher values (between 10 and 20 °C) than in the winter (between 0 and 6—7 °C), where it is naturally cold but rarely gets very cold.

The electrical resistivity map (Fig. 11) gives a first intuition of suitable locations for ground-source heat pump installations. Naturally, the most suitable locations are the ones characterized by smaller resistivity values, meaning the blue regions in the map. A significant number of areas seem to be suitable, including a medium to large portion of the plateau and the Geneva and Vaud cantons. Also, a large amount of small polygons spread all over the country show low resistivity values, notably in the plateau, in the Valais canton in the south, and the eastern cantons (Uri, Glarus, Graubunden). Variable Importances are also computed, and the twenty features with the highest VI are showed in Fig. 19b. It can be observed that yearly meteorological (statistical) features here have the largest impact on the estimation. In particular, their impact on the average resistivity value seem to be larger than the one of space variables and interestingly also higher than the one of soil/rock features. The very large

Fig. 19 Variable Importances (VI). See definition of each feature (e.g., `Sunshine`) in Appendix 2: Table 9. **a** For the RF model trained for the estimation of the yearly ground temperature at a depth of 100 cm; **b** for the RF model trained for the estimation of the electrical resistivity
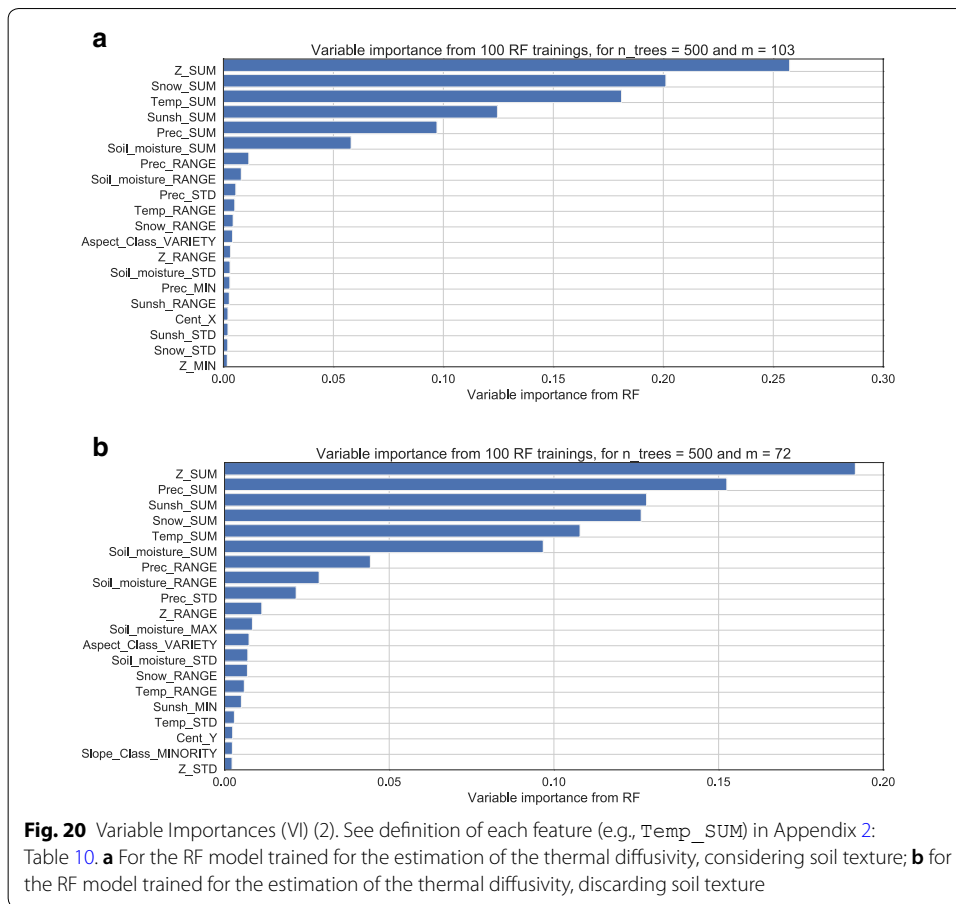
range of possible resistivity values for one type of rock/soil could partly explain this observation. When looking at the PIs in Fig. 11, one can observe that the test confidence is, even though acceptable, not very high, and more particularly lower than 95%, which indicates that the PIs for this variable are not always reliable, and must therefore be considered with caution for new predicted points. The PIs in Fig. 11c, however, give a good indication of the typical uncertainty attached to the prediction of electrical resistivity. In the case of the 30 unobserved random polygons shown, it is rather high since it is on average around ± 1 (between 2 and 4) for the modified output, corresponding to around an error of $50 \times C$ (Ω m), which can be significantly high depending on C (with C being the number of pixels in the polygon). This is partly caused by the extremely high range of possible values for electrical resistivity. In the error map (Fig. 12), the distribution of uncertainty can be seen across the Swiss territory. While the uncertainty is generally higher in the Alps region, its distribution over Switzerland is rather heterogeneous. It depends on the range of the output values, as mentioned previously, but also the density of training data available in different regions (which explains notably the high error in the Alps, where very few polygons are known), and the similarity (with respect to the features) between the unobserved polygons and the training polygons. Note, however, that this resistivity map is an

intermediate step to the estimation of the thermal conductivity and that the electrical properties of the ground may differ significantly from the thermal properties. Therefore, the thermal conductivity map must be analyzed to extract valid conclusions.

Figure 13 shows the estimated thermal conductivity values in Switzerland. A first natural observation is the range of the estimated values, which seems relatively small. These values, however (ranging from 0.25 to more than 1 (W/mK)), are perfectly normal for unsaturated clay, silt, and sand (Pahud 2002), which are the most frequent type of soil in the surface (mostly quaternary) layer of the ground (<1 m), the depth considered in the present study. The thermal conductivity seems to be higher in the center-west of the plateau, the south (Ticino and Valais cantons), and the east part of Switzerland (St. Gallen canton), which are therefore suitable locations for very shallow geothermal installations. The north part of the country, on the contrary, including some of the most dense urban areas of the plateau are not characterized by a high conductivity; notably Zurich, Vaud, Geneva, and Neuchatel cantons have also small potential for the very shallow installations. Note, however, that the systems considered in the present study are characterized by a very shallow depth, which are systems that are nowadays more adapted to rural areas, due to the lack of horizontal space in the ground (even though some interesting technologies, such as heat basket, are being developed to avoid that issue). For dense cities, Borehole Heat Exchangers (BHE) are usually preferred as they need less horizontal space and have a very high efficiency.

Ground-source heat pumps therefore still remain a very good option for the densely populated areas of Switzerland, but require a potential study at larger depths, which is not the focus of the present study. Also, note that the lack of measurement data for soil texture results in a significant number of unconsidered (gray) polygons, which are the polygons for which the type of texture was unobserved. Fortunately, the great majority of these unconsidered polygons are either lakes (Neuchatel, Bienne, Leman, Brienz, Zurich, etc.), either located in the alps, where the population is very low. The potential information is then naturally less significant in these areas. When looking at the PIs in Fig. 13, one can notice than similarly to the first estimation of electrical resistivity, the test confidence is not very high, showing again a difficulty to extract a PI for conductivity/resistivity variables. Furthermore, for the 30 random samples the width of the PI is on average $[-0.5, 1.5]$ in terms of $\log(\rho_t)$, which corresponds to an error of ± 0.13 (W/mK) in terms of thermal conductivity. This uncertainty is significant but not surprising, given the limited information used to perform the conversion from electrical to thermal resistivity. Note, however, that the uncertainty from the electrical resistivity estimation naturally propagates into the conversion step, which is not taken into account by the QRFs when computing the PIs.

Figure 16 shows the estimated thermal diffusivity in Switzerland. Although the impact of the diffusivity is less important than the conductivity, it is nonetheless of great importance when a potential study is being conducted, as it gives information on how fast the heat is being conducted through the ground and is needed in the modelling of the heat conduction between the ground and tubes of the installation (for further geographical and technical potential estimations). The regions with the highest diffusivity seem to be the Valais Canton in the south and the east cantons: Schwyz, Obwalden, Nidwalden, and Glarus. Variable Importances are computed and the twenty most important features

Fig. 20 Variable Importances (VI) (2). See definition of each feature (e.g., `Temp_SUM`) in Appendix 2: Table 10. **a** For the RF model trained for the estimation of the thermal diffusivity, considering soil texture; **b** for the RF model trained for the estimation of the thermal diffusivity, discarding soil texture
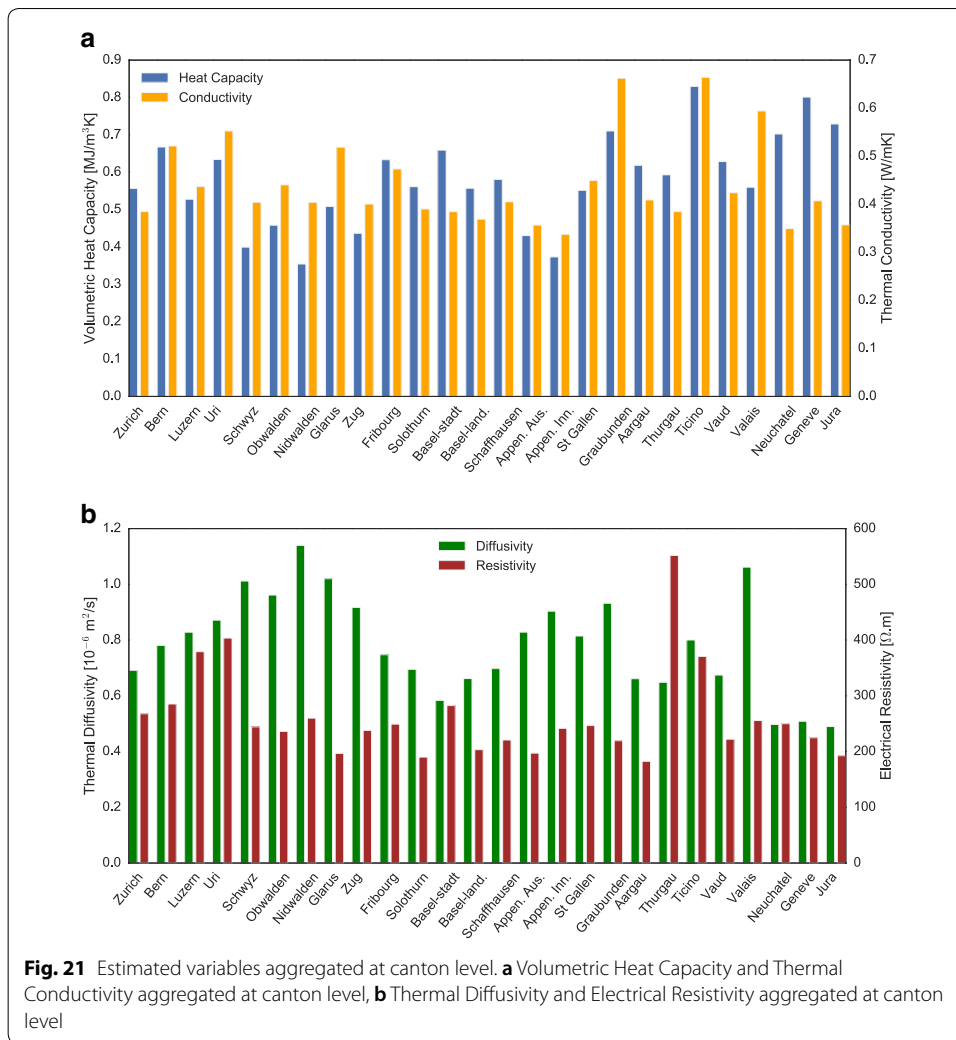
are illustrated with their respective VI for both RF models (with and without soil texture features) in Fig. 20a, b. Similarly to the electrical resistivity case, it shows that the meteorological features have a bigger impact on the estimation than the geological and soil texture features. The change of temperature and related variables therefore seem to be more important predictors than the actual types of rocks and soils, at the scale of the geological formation polygons which we are using. The PIs shown in Fig. 16 are notably more reliable than for the thermal conductivity estimation as the test confidence is better than 95% in both feature cases (with soil texture and without soil texture). The uncertainty in the prediction, as shown by the width of the PI for unobserved points, however, is still significant. In the random sample, the width of the PI is on average 1.5 in terms of $\log(\alpha \times C)$, which corresponds to $\pm 45/C$ (10^{-6} m^2/s), with C being the number of pixels in the polygon. By construction of the label, the uncertainty decreases with an increasing size of the polygon. Also, the uncertainty is slightly lower when the soil texture is not considered (Fig. 16d, e), specially for points with high predicted values. It seems that the smaller quantity of information entails a larger confidence for the model, which has fewer possibilities to consider and ultimately finds easier to predict a variable with fewer features used during the training process. It seems that the multiple trees in the forest agree more with each other when the number of features is small, which leads to a lower uncertainty. Note, however, that it does not mean that the accuracy of each prediction

is necessarily higher, but that the variance of the tree predictions is on average slightly smaller. In addition, the changes in uncertainty with space can be observed from Fig. 17. As discussed the resistivity, the uncertainty is again higher within the Alps, which is partly explained by the very small number of training points within this region. The uncertainty is also higher for small polygons, as previously explained. The general distribution is, like in the case of resistivity, quite heterogeneous, with values of uncertainty ranging from 0.2–0.5 to 5 (10^{-6} m^2/s) and even more in the most uncertain polygons. Finally, note that even though the southern and eastern regions do show a higher potential, they are characterized by high uncertainty given their low number of training samples, and are therefore possibly subject to overestimation.

The estimated variables were aggregated (averaged) at the canton level in Switzerland in order to have a higher level view of the very shallow geothermal potential. Volumetric Heat Capacity was also included in the aggregation, as it is commonly paired with thermal conductivity in geothermal potential studies. Plots for thermal diffusivity and electrical resistivity, and heat capacity and thermal conductivity can be seen in Fig. 21. Notably, it can be observed that the cantons with the highest mean thermal conductivity values are Graubunden, Ticino, and Valais, followed by Glarus, Uri, and Bern.

Very few geothermal potential studies have been conducted at very shallow depth, particularly in Switzerland, which makes comparison with other studies difficult. While the ThermoMap created by the European very shallow geothermal project (Bertermann et al. 2015) offers values for Switzerland, the resolution of these values is low, and more importantly, the depth does not match as it considers a depth of 10 m. In Switzerland, a few local organizations may have values for various thermal variables, but at larger depths. The "Systeme dÍnformation du Territoire a Geneve (SITG)," for example, offers heat capacity and thermal conductivity estimations for the Quaternary layer. The thickness of the Quaternary, however, may vary. It is therefore not possible to perform a thorough comparison with the results of the present study.

The present study suggests new methodologies for the estimation of thermal ground characteristics at a large scale and valuable information on their range of values in Switzerland. Several limitations, however, are to be noted, and possibly improved in further studies. These include (1) the lack of labeled training data. Ideally, the training data should include a few hundreds of points in order to build very reliable models. Particularly in the case of the thermal diffusivity and the conversion from electrical to thermal resistivity, the size of the data was around 50 to 100 points, which is commonly considered the minimum size to perform supervised learning. Consequently, the results are still valuable but would greatly benefit from additional data. Note that the test errors are based on the test set and are therefore only validated on this set. Low test errors unfortunately do not guarantee a good generalization outside of this test set. The bigger the test, the more reliable are the test errors. More data may be added in the future to improve the models. (2) The lack of validation data. As no dataset is available for the shallow ground characteristics of Switzerland over the entire territory, it is impossible to validate the final obtained results in unobserved points. Should another study be available, a comparison between the results would be valuable to provide some validation to the present results. (3) Using a combination of consecutive RF models together with more conventional signal processing methods (e.g., FFT) and numerical models (e.g.,

Fig. 21 Estimated variables aggregated at canton level. **a** Volumetric Heat Capacity and Thermal Conductivity aggregated at canton level, **b** Thermal Diffusivity and Electrical Resistivity aggregated at canton level

iterative inversion schemes) brings confidence, as the conventional methods have been tested and validated through the years, but also additional uncertainty, as the uncertainty of each estimation propagates through the next step, and therefore increases at each step. In that sense, a single step machine learning strategy can be considered to avoid the propagation issue. Labeled data, however, are naturally required for the final variable to estimate. There might be, however, an increasing availability for data in the future, as there is currently an effort to digitalize geological and geophysical information, which is still often stored in the form of paper maps or written information. It is particularly the case in Switzerland, where a very significant amount of geo-studies have been performed through the years.

Conclusion

We present a methodology combining GIS data processing, machine learning, and physical models in order to estimate the very shallow theoretical geothermal potential in Switzerland. This theoretical shallow geothermal potential consists of the estimation

of three significant thermal variables at shallow depth: (i) ground temperature gradient, (ii) ground thermal conductivity, and (iii) ground thermal diffusivity. The ground thermal heat capacity can then be recovered from the conductivity and the diffusivity. The estimation of the three variables is proposed at a spatial resolution of (200×200) (m^2), all over the Swiss territory, and at a monthly mean temporal resolution for the ground temperature, and a yearly mean temporal resolution for the conductivity and the diffusivity. Besides the methodological contributions of the present study, it eventually shows that, while traditional shallow geothermal systems (100−200-m-deep Borehole Heat Exchangers mainly) have been used for a very long time in Switzerland, there is also a significant potential for very shallow (first meter of the ground) geothermal energy systems, which can be a viable low-cost solution in adequate locations. There is notably a high potential for such systems in the Valais, Ticino, and St. Gallen cantons, where the highest thermal conductivity values were found.

The obtained information on ground thermal characteristics can be of great use for municipalities, stakeholders, and private holders who are considering small- to large-scale very shallow geothermal installations. With the current development of new efficient and cost-effective geothermal systems at shallow depths (including Slinky systems, helicoidal systems, and heat baskets), the estimated results could serve as a useful help to identify the optimal locations for geothermal energy and for energy-related decision making in general in Switzerland. Also, note that the methodology is mainly based on various sources of data (geological, weather, and terrain data) that are currently being digitalized in more and more countries and methods/algorithms that are already implemented in various libraries. Therefore, should similar data be available, the methodology is generalizable to any other location.

Future work to complement and further develop the methods presented here include (1) the use of newly available geological data to improve the accuracy of the different machine learning models in the methodology; (2) the extraction of the geographical potential, constraining the theoretical potential to areas where the shallow geothermal potential is efficiently extractable, and the technical potential, accounting for the losses and the practical sizing and use conditions of the installation, (3) the assessment of the uncertainty propagation through the consecutive steps of the methodology by means of statistical calculus or Monte Carlo methods.

Authors' contributions
DA collected the data, performed the estimations, and wrote the first draft of the manuscript. NM and AG participated in the evaluation and assessment of the main assumptions and research directions used in the manuscript. AG also provided important input as to shallow geothermal energy concepts and use in practice. NM, AG, and JLS reviewed and contributed to the final version of the manuscript. All authors read and approved the final manuscript.

Author details
[1] Solar Energy and Building Physics Laboratory (LESO-PB), Ecole Polytechnique Federale de Lausanne, 1015 Lausanne, Switzerland. [2] Sustainable Urban Development Programme, Department for Continuing Education, University of Oxford, Oxford, UK. [3] Department of Earth Sciences, Royal Holloway University of London, Egham, UK.

Appendices

Appendix 1—Geology classes from GeoCover (GK500) data

See Tables 4, 5, 6, 7, 8

Table 4 Hydrogeology [HYDRO] categories in GK500 dataset

Description	N_{GK500}	Area (km^2)
No information	1	69.8
Areas without productive aquifer reservoirs	2930	10,816.72
Areas without productive reservoirs	250	508.32
Surface water	250	1377.6
Glacier, Neve	377	866.28
Aquifer tanks in karstifiable coherent rocks	1796	6144.08
Low productive aquifers	3218	7758.64
Low productive aquifers in unquantifiable, cracked and porous coherent rocks	2945	10,608.32
Productive aquifers partly out of valley bottoms	1112	2061.04
Highly productive aquifer reservoirs of valley bottoms	441	1243.24

N_{GK500} is the number of polygons for each category and Area is the total area spanned by all polygons of the category

Table 5 Geological period [PERIOD] categories in GK500 dataset

Description	N_{GK500}	Area (km^2)
No information	1229	4976.52
Carboniferous	220	1040.04
Cretaceous	1264	4331.08
Devonian	86	762.6
Jurassic	1772	6826.72
Lower Paleozoic	126	642.8
Permian	342	770.64
Permian-Cretaceous	100	100.84
Quaternary	5541	13,733.56
Tertiary	1712	6815.68
Triassic	928	1453.56

N_{GK500} is the number of polygons for each category and Area is the total area spanned by all polygons of the category

Assouline *et al. Geotherm Energy*   (2019) 7:19

Page 43 of 50

Table 6 Aquifer productivity [PRODUCTIV] categories in GK500 dataset

Description	N_{GK500}	Area (km^2)
Barely exploitable, usually in fine sands	250	508.32
Surface water	250	1377.6
Glacier, Neve	377	866.28
Not locally or barely exploitable	2930	10,816.72
Not very productive	2461	9197.0
Not very productive, in the moraines	2263	6507.36
Productive, variable or low productivity	1796	6144.08
Variable productivity	484	1411.32
Variable productivity in loamy gravels	956	1321.08
Usable saturated area for a depth of 10 to 20 m	284	771.68
Usable saturated area for a depth of 2 to 10 m	1112	2061.04
Usable saturated area for a depth of more than 20 m	157	471.56

N_{GK500} is the number of polygons for each category and Area is the total area spanned by all polygons of the category

Table 7 Rock formation types [TYPE ROCHE] categories in GK500 dataset

Description	N_{GK500}	Area (km^2)
Watercourses, lakes	250	1377.6
Glaciers, snowfields	377	866.28
Magmatic rocks	323	1530.64
Unconsolidated rocks	5023	11,630.76
Metamorphic rocks	2078	7511.0
Sedimentary rocks	5269	18,537.76

N_{GK500} is the number of polygons for each category and Area is the total area spanned by all polygons of the category

Table 8 Rock/soil types [LITH PET] categories in GK500 dataset

Description	N_{GK500}	$N_{soil\ samples}$	Area (km^2)
Amphibolites with diorite passages or a hornblende gneiss	226	5	287.48
Slates with intercalations of dolomites, cornieules, gypsum, limestone, and sandstone	6	0	27.0
Phyllite slates with intercalations of sandstone and pudding stones	82	0	309.64
Ferruginous argillites	49	7	15.12
Breccia or limestone pudding stones	7	8	22.96
Breccia and pudding stones	4	0	30.04
Sandy limestones with siliceous limestones with marly shale levels	224	8	761.64
Silicious limestones	163	4	840.8
Limestones, sometimes marbled	112	4	103.68
Conglomerates and breccia rich in sericite	38	0	105.68
Rivers, lakes	250	0	1377.6
Diorites and gabbros	23	1	64.52
Dolostones and Cornieules	175	0	205.44
Dolostones with gypsum levels	47	4	42.44
Landslide and scree deposits	783	89	903.68
Glaciers	377	0	866.28
Gneiss and micaschists rich in biotite and muscovite	350	33	1525.88
Gneiss rich in biotite or muscovite, sometimes chloritic, sometimes with calc-silicates rocks or quartzitic horizons (hornfels)	7	0	124.96
Gneiss rich in feldspar	203	13	992.32
Gneiss rich in feldspar, schistous with sericite, epidote, and chlorite formation	88	3	443.12
Gneiss with two micas or biotite rich in feldspar and varied structures	98	0	679.4
Gneiss with two micas or biotite rich in feldspar, platelets	5	2	268.28
Gneiss with two micas or biotite, rich in feldspar, mainly homogeneous	28	0	252.56
Gneiss with two micas or biotite, with white, often green (phengite)	22	0	137.8
Gneiss with sericitic and chloritic schists	291	2	1278.0
Granite with passages of quartzitic diorite or quartzitic syenite	111	3	823.84
Granite with sericite, epidote, and chlorite	98	6	526.52
Gravel and sand	754	740	1500.08
Gravel and sand, sometimes clayey or silty	738	636	1597.96
Hard and compact sandstone with marly shale and limestone phyllites	189	43	1096.04
Sandstone and marls with low to moderately consolidated pudding stone levels	207	70	945.68
Quartz sandstone with sandy slates	15	0	7.2
Glauconite siliceous sandstone and echinoderm debris	52	0	126.2
Sandstone mainly calcareous and porous, with marl levels	362	396	1622.92
Sandstone with marl levels	84	20	482.8
Dolomitic marble	23	0	19.04
Marly and shale clay with limestone bench, dolomite, and sandstone	162	145	226.32
Marls with breach levels with hard and sandy shells	1	0	34.96
Well-consolidated sandstone slabs	51	0	166.76
Marbles with sandstone levels with shell-rich breccia	12	23	35.84
Marls with sandstone levels, conglomerates, or poorly consolidated pebbles	376	698	1146.72
Quartz Phyllites	1	0	0.48
Phyllites with limestone micaschists	206	0	765.84
Phyllites with limestone micaschists with limestone levels and marbled dolomites	17	0	36.68
Phyllites with limestone micaschists with greenstone levels	1	0	1.04
Quartz porphyry	59	0	73.36

Assouline *et al. Geotherm Energy*   (2019) 7:19

Page 45 of 50

Table 8 (continued)

Description	N_{GK500}	$N_{soil\ samples}$	Area (km^2)
Porphyrites and porphyry tuffs	15	0	19.6
Breccia puddings with arkoses and sandstone	80	0	174.04
Pudding stones with breccia with arkoses and sandstone	66	22	1406.8
Peridotites and olivine rocks	18	0	8.48
Quartzites	70	0	112.12
Radiolarites	11	0	3.72
Dolomitic rock, sometimes with limestone levels	480	19	817.04
Limestone rocks in general, often with marly intercalations	1175	560	4636.2
Limestone rocks with dolomitic levels	48	64	184.8
Limestone rocks with important levels of marl, shale, and marly limestone	628	159	1370.96
Limestone rocks, often marly	89	0	480.68
Volcanic and pyroclastic rocks	5	0	9.64
Sand, gravel, pebbles, and blocks	454	45	457.28
Marly shale with calcareous phyllites and interbedded sandstone	382	79	1226.52
Marly shale with limestone phyllites with tuffitic sandstone levels	42	0	90.48
Green shale with passages of eruptive rocks or eclogites	173	5	252.76
Serpentinites	78	13	84.68
Clayey silts, with clay with sand levels	264	499	686.24
Sand silts with gravel and blocks	1962	1614	6331.44
Silts with silty sands, often clayey, mainly calcareous	68	170	154.08
Syenite	12	0	13.16
Amphibolite and gneiss mixing zone	23	0	30.72

N_{GK500} is the number of polygons for each category and Area is the total area spanned by all polygons of the category.
$N_{soil\ samples}$ is the number of NABODAT soil texture points available in each category

Appendix 2—Features in variable importance figures
Tables 9, 10

Table 9 Detailed definition of features showed in VI Fig. 19a

Name	Description
Aspect	Terrain aspect (normal direction to the slope) in a pixel
Cloud	Cloud cover index in a pixel
Prec	Precipitation in a pixel
Slope	Terrain slope in a pixel
Snow	Snow cover in a pixel
Sunshine	Sunshine duration in a pixel
Temp	Air temperature in a pixel
X	*X* coordinate in the LV03 coordinate system (linear transformation of longitude in Switzerland)
Y	*Y* coordinate in the LV03 coordinate system (linear transformation of latitude in Switzerland)

All meteorological variables are yearly averaged

Table 10 Detailed definition of features showed in VI Figs. 19b, 20a, b

Name	Description
Cent_X	*X* coordinate (LV03 coordinate system) of a GK500 polygon centroid
Aspect_Class_VARIETY	Number of observed classes of pixel terrain aspect values over a GK500 polygon
Slope_Class_MINORITY	Least frequent class of pixel terrain slope over a GK500 polygon
Prec_MIN	Minimum of pixel precipitation values over a GK500 polygon
Snow_MIN	Minimum of pixel snow cover values over a GK500 polygon
Sunshine_MIN	Minimum of pixel sunshine duration values over a GK500 polygon
Z_MIN	Minimum of pixel elevation values over a GK500 polygon
Soil_moisture_MAX	Maximum of pixel soil moisture values over a GK500 polygon
Sunshine_MAX	Maximum of pixel sunshine duration values over a GK500 polygon
Prec_RANGE	Range of pixel precipitation values over a GK500 polygon
Snow_RANGE	Range of pixel snow cover values over a GK500 polygon
Soil_moisture_RANGE	Range of pixel soil moisture values over a GK500 polygon
Sunshine_RANGE	Range of pixel sunshine duration values over a GK500 polygon
Temp_RANGE	Range of pixel air temperature values over a GK500 polygon
Z_RANGE	Range of pixel elevation values over a GK500 polygon
Prec_STD	Standard deviation of pixel precipitation values over a GK500 polygon
Snow_STD	Standard deviation of pixel snow cover values over a GK500 polygon
Soil_moisture_STD	Standard deviation of pixel soil moisture values over a GK500 polygon
Sunshine_STD	Standard deviation of pixel sunshine duration values over a GK500 polygon
Temp_STD	Standard deviation of pixel air temperature values over a GK500 polygon
Z_STD	Standard deviation of pixel elevation values over a GK500 polygon
Prec_SUM	Sum of pixel precipitation values over a GK500 polygon
Snow_SUM	Sum of pixel snow cover values over a GK500 polygon
Soil_moisture_SUM	Sum of pixel soil moisture values over a GK500 polygon
Sunshine_SUM	Sum of pixel sunshine duration values over a GK500 polygon
Temp_SUM	Sum of pixel air temperature values over a GK500 polygon
Z_SUM	Sum of pixel elevation values over a GK500 polygon

All meteorological variables are yearly averaged

Appendix 3—Detailed tables of used variables

Tables 11, 12

Table 11 Detailed nomenclature for geothermal and soil-related variables

Symbol	Unit	Description
α	$(m^2\,s^{-1})$	Thermal diffusivity
γ_d	(g/cm^3)	Dry (bulk) density
γ_s	(g/cm^3)	Particle density
γ_w	(g/cm^3)	Water density
λ	$(W\,K^{-1}\,m^{-1})$	Thermal conductivity
ρ	$(\Omega\,m)$	Electrical resistivity
ω	(s^{-1})	Angular frequency of one period in Fourier series
a_n	(–)	First Fourier coefficient for the nth harmonics
b_n	(–)	Second Fourier coefficient for the nth harmonics
c_v	$(J\,m^{-3}\,K^{-1})$	Volumetric heat capacity
D	(m)	Damping depth
e	(–)	Void ratio
F	(%)	Percentage sum of sand and gravel fractions in the soil
h_i	(m)	Width of soil strata i
$I_1, ..., I_{10}$	(–)	Possible fraction intervals for soil texture variables
M_s	(g)	Mass of solid soil in ground
M_w	(g)	Mass of water in ground
n	(–)	Harmonics index in Fourier series
n_p	(–)	Porosity
R_n	(–)	Amplitude of nth harmonic of Fourier series solution for T
Sd, St, Cl	(%)	Sand, silt, and clay fraction percentage in soil
S_r	(%)	Saturation degree
t	(s)	Time
T	(°C)	Shallow ground temperature
T_0	(°C)	Average ground surface temperature over a year
V_a	(m^3)	Volume of air in ground
V_s	(m^3)	Volume of solid soil in ground
V_T	(m^3)	Total volume of ground
V_v	(m^3)	Volume of void in ground
V_w	(m^3)	Volume of water in ground
w	(%)	Gravimetric water content
z	(m)	Ground depth

(–) signifies no unit

Table 12 Detailed nomenclature for machine learning-related variables

Symbol	Unit	Description
φ	(–)	Generic predictive function learned from training process
B	(–)	Number of trees in a Random Forest
d	(–)	Number of features (predictors) in the training data for an ML task
E_R	(–)	Root Mean Square Error
E_{NR}	(–)	Normalized Root Mean Square Error
K	(–)	Number of folds consider in fold cross-validation
m	(–)	Number of variables to consider to split a node v in a decision tree
N	(–)	Number of training points (size of training data)
N_{test}	(–)	Number of testing point (size of testing data)
$\mathbb{R}$	(–)	Set of real numbers
x	(var. unit)	Generic input vector (realization of $X_1, ..., X_d$)
$X_1, ..., X_d$	(var. unit)	Generic input variables
y	(var. unit)	Generic output value (label, realization of Y)
Y	(var. unit)	Generic output variable

(–) signifies no unit. (var. unit) signifies that the unit is defined by the quantity of interest represented by the variable

References

Alessandrini S, Delle Monache L, Sperati S, Cervone G. An analog ensemble for short-term probabilistic solar power forecast. Appl Energy. 2015;157:95–110.

Allen A, Milenic D, Sikora P. Shallow gravel aquifers and the urban 'heat island' effect: a source of low enthalpy geothermal energy. Geothermics. 2003;32(4–6):569–78.

Andújar Márquez JM, Martínez Bohórquez MÁ, Gómez Melgar S. Ground thermal diffusivity calculation by direct soil temperature measurement. Application to very low enthalpy geothermal energy systems. Sensors. 2016;16(3):306.

Arola T, Korkka-Niemi K. The effect of urban heat islands on geothermal potential: examples from quaternary aquifers in finland. Hydrogeol J. 2014;22(8):1953–67.

Arola T, Eskola L, Hellen J, Korkka-Niemi K. Mapping the low enthalpy geothermal potential of shallow quaternary aquifers in finland. Geotherm Energy. 2014;2(1):9.

Assouline D, Mohajeri N, Scartezzini J-L. Quantifying rooftop photovoltaic solar energy potential: a machine learning approach. Solar Energy. 2017;141:278–96.

Assouline D, Mohajeri N, Scartezzini J-L. Large-scale rooftop solar photovoltaic technical potential estimation using random forests. Appl Energy. 2018;217:189–211.

Beamish D. The bedrock electrical conductivity map of the UK. J Appl Geophys. 2013;96:87–97.

Beardsmore G. Data fusion and machine learning for geothermal target exploration and characterization. Australia: Technical report, Technical report, National ICT Australia Limited (NICTA); 2014.

Beardsmore G, Durrant-Whyte H, McCalman L, O'Callaghan S, Reid A. A bayesian inference tool for geophysical joint inversions. ASEG Ext Abstr. 2016;2016(1):1–10.

Bertermann D, Bialas C, Morper-Busch L, Klug H, Rohn J, Stollhofen M, Psyk M, Jaudin F, Maragna C, Einarsson GM, et al. Thermomap-an open-source web mapping application for illustrating the very shallow geothermal potential in europe and selected case study areas. In: European Geothermal Congress, Pisa; 2013. p. 1–7.

Bertermann D, Klug H, Morper-Busch L, Bialas C. Modelling vsgps (very shallow geothermal potentials) in selected csas (case study areas). Energy. 2014;71:226–44.

Bertermann D, Klug H, Morper-Busch L. A pan-European planning basis for estimating the very shallow geothermal energy potentials. Renew Energy. 2015;75:335–47.

Boughanmi H, Lazaar M, Bouadila S, Farhat A. Thermal performance of a conic basket heat exchanger coupled to a geothermal heat pump for greenhouse cooling under tunisian climate. Energy Build. 2015;104:87–96.

Breiman L. Random forests. Mach Learn. 2001;45(1):5–32.

Breiman L. Classification and regression trees. New York: Routledge; 2017.

Busby J. Determination of thermal properties for horizontal ground collector loops. In: Proceedings of the world geothermal congress 2015, Melbourne, Australia, 19–25 April 2015; 2015.

Carslaw HS, Jaeger JC. Conduction of heat in solids. Oxford: Clarendon Press, Oxford; 1959.

Carson JE. Analysis of soil and air temperatures by Fourier techniques. J Geophys Res. 1963;68(8):2217–32.

Casasso A, Sethi R. G. pot: a quantitative method for the assessment and mapping of the shallow geothermal potential. Energy. 2016;106:765–73.

Casasso A, Sethi R. Assessment and mapping of the shallow geothermal potential in the province of cuneo (Piedmont, NW italy). Renew Energy. 2017;102:306–15.

Chandrashekar G, Sahin F. A survey on feature selection methods. Comput Electric Eng. 2014;40(1):16–28.

Das N, Entekhabi D, Kim S, Yueh S, Dunbar RS, Colliander A. Smap (soil moisture active passive)/sentinel-1 l2 radiometer/radar 30-second scene 3 km ease-grid soil moisture, version 1. 2018. https://doi.org/10.5067/9UWR1WTHW1WN. Accessed 23 May 2018.

Dehghan B, Sisman A, Aydin M. Parametric investigation of helical ground heat exchangers for heat pump applications. EnergyBuild. 2016;127:999–1007.

Dehner U. Bestimmung der thermischen eigenschaften von böden als grundlage für die erdwärmenutzung. Mainzer geowissenschaftliche Mitteilungen. 2007;35:159–86.

Di Sipio E, Bertermann D. Factors influencing the thermal efficiency of horizontal ground heat exchangers. Energies. 2017;10(11):1897.

Di Sipio E, Galgaro A, Destro E, Teza G, Chiesa S, Giaretta A, Manzella A. Subsurface thermal conductivity assessment in calabria (southern italy): a regional case study. Environ Earth Sci. 2014;72(5):1383–401.

Dong G, Liu H. Feature engineering for machine learning and data analytics. Boca Raton: CRC Press; 2018.

Dumont B, Chapellier D. Inventaire des sondages electriques de suisse (Publication nr. 42); 2003. http://www.sgpk.ethz.ch/opencms/opencms/publications/pub_contrib_ch/pubcon_40_49/pubcon_42/. Accessed 23 May 2018.

Erzin Y, Rao BH, Patel A, Gumaste S, Singh D. Artificial neural network models for predicting electrical resistivity of soils from their thermal resistivity. Int J Therm Sci. 2010;49(1):118–30.

Friedel S. Resolution, stability and efficiency of resistivity tomography estimated from a generalized inverse approach. Geophys J Int. 2003;153(2):305–16.

Galgaro A, Di Sipio E, Teza G, Destro E, De Carli M, Chiesa S, Zarrella A, Emmi G, Manzella A. Empirical modeling of maps of geo-exchange potential for shallow geothermal energy at regional scale. Geothermics. 2015;57:173–84.

García-Gil A, Vázquez-Suñe E, Alcaraz MM, Juan AS, Sánchez-Navarro JÁ, Montlleó M, Rodríguez G, Lao J. Gis-supported mapping of low-temperature geothermal potential taking groundwater flow into account. Renew Energy. 2015;77:268–78.

Gee GW, Bauder JW. Particle-size analysis 1 vol. Methods of soil an1. Madison: Soil Science Society of America, American Society of Agronomy; 1986.

Geodata4edu portal; 2018. http://geodata4edu.ethz.ch. Accessed 29 Jan 2018.

Haber E. Quasi-Newton methods for large-scale electromagnetic inverse problems. Inverse Probl. 2004;21(1):305.

Hassan MA, Khalil A, Kaseb S, Kassem M. Exploring the potential of tree-based ensemble methods in solar radiation modeling. Appl Energy. 2017;203:897–916.

Hastie T, Tibshirani R, Friedman J. The elements of statistical learning. Springer series in statistics. New York: Springer; 2001.

Heinermann J, Kramer O. Machine learning ensembles for wind power prediction. Renew Energy. 2016;89:671–9. https://doi.org/10.1016/j.renene.2015.11.073.

Hurley S, Wiltshire RJ. Computing thermal diffusivity from soil temperature measurements. Comput Geosci. 1993;19(3):475–7.

Hussain S, AlAlili A. A hybrid solar radiation modeling approach using wavelet multiresolution analysis and artificial neural networks. Appl Energy. 2017;208:540–50.

IDAWEB database; 2018. https://gate.meteoswiss.ch/idaweb. Accessed 29 Jan 2018.

Jenness J. Tools for graphics and shapes: extension for arcgis. Flagstaff: Jenness Enterprises; 2011.

Jordan MI, Mitchell TM. Machine learning: trends, perspectives, and prospects. Science. 2015;349(6245):255–60.

Joshi B, Hayk B, Al-Hinai A, Woon WL. Rooftop detection for planning of solar pv deployment: a case study in Abu Dhabi. In: International workshop on data analytics for renewable energy integration. Berlin: Springer; 2014. p. 137–49.

Kalogirou SA, Florides GA, Pouloupatis PD, Panayides I, Joseph-Stylianou J, Zomeni Z. Artificial neural networks for the generation of geothermal maps of ground temperature at various depths by considering land configuration. Energy. 2012;48(1):233–40.

Kalogirou SA, Florides GA, Pouloupatis PD, Christodoulides P, Joseph-Stylianou J. Artificial neural networks for the generation of a conductivity map of the ground. Renew Energy. 2015;77:400–7.

Kanevski M, Maignan M. Analysis and modelling of spatial environmental data, vol. 6501. Lausanne: EPFL press; 2004.

Kanevski M, Pozdnoukhov A, Timonin V. Machine learning for spatial environmental data: theory, applications, and software. Lausanne: EPFL press; 2009.

Kemmler A, Spillmann T, Koziel S, Piegsa A, Notter B, Laderach A, Jakob M, Catenazzi G. Analyse des schweizerischen energieverbrauchs 2000–2017 nach verwendungszwecken. Bern: Bundesamt für Energie; 2018.

Kersten MS. Thermal properties of soils; 1949.

Lee J-Y, Hahn J-S. Characterization of groundwater temperature obtained from the Krean national groundwater monitoring stations: implications for heat pumps. J Hydrol. 2006;329(3–4):514–26.

Liaw A, Wiener M. Classification and regression by randomforest. R News. 2002;2(3):18–22.

Loke M, Barker R. Improvements to the zohdy method for the inversion of resistivity sounding and pseudosection data. Comput Geosci. 1995;21(2):321–32.

Loke MH, Barker R. Rapid least-squares inversion of apparent resistivity pseudosections by a Quasi-Newton method 1. Geophys Prospect. 1996;44(1):131–52.

Lou S, Li DH, Lam JC, Chan WW. Prediction of diffuse solar irradiance using machine learning and multivariable regression. Appl Energy. 2016;181:367–74.

Lund J, Sanner B, Rybach L, Curtis R, Hellstrom G. Geothermal (ground-source) heat pumps: a world overview. Geo-Heat Center Q Bull. 2004;25(3):1–10.

Meinshausen N. Quantile regression forests. J Mach Learn Res. 2006;7(Jun):983–99.

Meteoswiss (Federal Office of Meteorology and Climatology) portal; 2018. http://www.meteosuisse.admin.ch/home.html?tab=overview. Accessed 29 Jan 2018.

Milenić D, Vasiljević P, Vranješ A. Criteria for use of groundwater as renewable energy source in geothermal heat pump systems for building heating/cooling purposes. Energy Build. 2010;42(5):649–57.

NABODAT database; 2018. https://www.nabodat.ch. Accessed 23 May 2018.

Najeebullah Zameer A, Khan A, Javed SG. Machine learning based short term wind power prediction using a hybrid learning model. Comput Electric Eng. 2015;45:122–33. https://doi.org/10.1016/j.compeleceng.2014.07.009.

Ondreka J, Rüsgen MI, Stober I, Czurda K. Gis-supported mapping of shallow geothermal potential of representative areas in South-Western Germany—possibilities and limitations. Renew Energy. 2007;32(13):2186–200.

Pahud D. Geothermal energy and heat storage. Manno: Scuola Universitaria Professionale della Svizzera Italiana; 2002.

Pedregosa F, Varoquaux G, Gramfort A, Michel V, Thirion B, Grisel O, Blondel M, Prettenhofer P, Weiss R, Dubourg V, et al. Scikit-learn: machine learning in python. J Mach Learn Res. 2011;12(Oct):2825–30.

Rajeev P, Kodikara J. Estimating apparent thermal diffusivity of soil using field temperature time series. Geomech Geoeng. 2016;11(1):28–46.

Reynolds JM. An introduction to applied and environmental geophysics. Oxford: Wiley; 2011.

Rezvanbehbahani S, Stearns LA, Kadivar A, Walker JD, van der Veen CJ. Predicting the geothermal heat flux in greenland: a machine learning approach. Geophys Res Lett. 2017;44(24):12–271.

Rivera JA, Blum P, Bayer P. Increased ground temperatures in urban areas: estimation of the technical geothermal potential. Renew Energy. 2017;103:388–400.

Rücker C, Günther T, Wagner FM. pyGIMLi: an open-source library for modelling and inversion in geophysics. Comput Geosci. 2017;109:106–23. https://doi.org/10.1016/j.cageo.2017.07.011.

Sanner B, Karytsas C, Mendrinos D, Rybach L. Current status of ground source heat pumps and underground thermal energy storage in europe. Geothermics. 2003;32(4–6):579–88.

Schiel K, Baume O, Caruso G, Leopold U. Gis-based modelling of shallow geothermal energy potential for co2 emission mitigation in urban areas. Renew Energy. 2016;86:1023–36.

Segal MR. Machine learning benchmarks and random forest regression. UCSF: Center for Bioinformatics and Molecular Biostatistics. 2004.

Siedlecki W, Sklansky J. On automatic feature selection; 1993 p. 63–87. https://doi.org/10.1142/9789814343138_0004.

Singh DN, Kuriyan SJ, Manthena KC. A generalised relationship between soil electrical and thermal resistivities. Exp Therm Fluid Sci. 2001;25(3–4):175–81.

Sreedeep S, Reshma A, Singh D. Generalized relationship for determining soil electrical resistivity from its thermal resistivity. Exp Therm Fluid Sci. 2005;29(2):217–26.

Stone M. Cross-validatory choice and assessment of statistical predictions. J R Stat Soc Ser B (Methodological). 1974;36:111–47.

Swisstopo (Federal Office of Topography) portal; 2018. https://www.swisstopo.admin.ch. Accessed 29 Jan 2018.

Treiber NA, Heinermann J, Kramer O. Wind power prediction with machine learning. In: Computational sustainability. Springer, Cham; 2016. p. 13–29.

Verein Deutscher Ingenieure. VDI—Richtlinie 4640—Thermische Nutzung des Untergrundes, B..E.W.; 2001.

Wijk WRv. Physics of plant environment. Eindhoven: Wiley; 1963.

Willmott CJ, Matsuura K. Advantages of the mean absolute error (mae) over the root mean square error (rmse) in assessing average model performance. Clim Res. 2005;30(1):79–82.

Xiong Z, Fisher DE, Spitler JD. Development and validation of a slinky^{TM} ground heat exchanger model. Appl Energy. 2015;141:57–69.

Yadav AK, Chandel S. Solar radiation prediction using artificial neural network techniques: a review. Renew Sustain Energy Rev. 2014;33:772–81.

Zarrella A, De Carli M. Heat transfer analysis of short helical borehole heat exchangers. Appl Energy. 2013;102:1477–91.

Zarrella A, Capozza A, De Carli M. Analysis of short helical and double u-tube borehole heat exchangers: a simulation-based comparison. Appl Energy. 2013a;112:358–70.

Zarrella A, Capozza A, De Carli M. Performance analysis of short helical borehole heat exchangers via integrated modelling of a borefield and a heat pump: a case study. Appl Therm Eng. 2013b;61(2):36–47.

Zhu K, Blum P, Ferguson G, Balke K-D, Bayer P. The geothermal potential of urban heat islands. Environ Res Lett. 2010;5(4):044002.

Zohdy AA. A new method for the automatic interpretation of schlumberger and wenner sounding curves. Geophysics. 1989;54(2):245–53.

Publisher's Note