Stratigraphy modeling and thermal conductivity computation in areas characterized by Quaternary sediments

Giordano Teza ^{a,*}, Antonio Galgaro ^{a,b}, Elisa Destro ^b, Eloisa Di Sipio ^b

^a *Dipartimento di Geoscienze, Università di Padova, via Gradenigo, 6, I-35131 Padova, Italy*

^b *Consiglio Nazionale delle Ricerche, Istituto di Geoscienze e Georisorse (CNR-IGG), U.O.S. Padova, Via
Gradenigo, 6, I-35131 Padova, Italy*

Abstract

A feasibility plan for the exploitation of low enthalpy geothermal resources at regional scale requires a map of the ground thermal conductivity. Available information often comes from heterogeneous well log data related to boreholes drilled for various purposes (e.g. irrigation) and whose spatial distribution is suboptimal. Therefore, a suitable interpolation technique is necessary. An approach based on local evaluation of the mode of distribution of the subsoil geological species is proposed here. It provides a stratigraphy model, the equivalent ground thermal conductivity and the corresponding statistical/geometrical significance in the nodes of a regular grid. Besides a validation with synthetic data of the MATLAB/GNU Octave implementation of the method (GRIDWELL toolbox), the results of its application to the data from Apulia's plain areas are shown here.

Keywords: Geoexchange; Borehole Heat Exchanger; Thermal Ground Characterization; Stratigraphy; Thermal Conductivity; Spatial Statistics.

* Corresponding author. *Email address:* giordano.teza@unipd.it, Tel. +39 049 8279123, fax +39 049 8279134 (G. Teza)

1. Introduction

Nowadays, the increasing cost of energy sources like oil and coal and the corresponding environmental impact suggest an extensive use of geothermal energy for heating and/or cooling of buildings. In particular, closed loop geothermal systems, often used in conjunction with ground-coupled heat pumps (GCHPs), have relatively high installation costs which are paid off by low operation and maintenance costs, leading to their use in a wide range of areas (see e.g. Banks, 2012). A GCHP consists in the following: one or more borehole heat exchangers (BHEs), with ~100 m typical depth, which ensure coupling with the ground; some circulation pumps, which circulate the carrier fluid; and a heat pump, which uses a compression-expansion cycle to boost the temperature of the heat to a level that can be practically used for heating or cooling purposes. The carrier fluid (pure water or a mixture of water and propylene glycol or ethylene glycol) exchanges heat both with the ground, through the BHE, and with the heat pump.

In order to allow an evaluation of the degree of suitability of an area for low enthalpy geothermal energy exploitation purposes, and therefore to allow a good exploitation of the geothermal resources, a general thermo-stratigraphical assessment of the subsurface that should host BHEs is useful (Blum et al., 2011). Local insight provided by Ground Response Test (GRT) and thermal modeling are needed for correct planning in the case of large BHE fields. In this way, a map of the areal distribution of thermal conductivity (TC) can be obtained. Such a map is an informative tool useful both for evaluating the feasibility of geo-exchange systems and land development planning by the competent public authorities. A detailed stratigraphical characterization of the ground is unnecessary for the design of a small GCHP system, where a successful design mainly depends on the correct evaluation of heating and cooling building demand. Nevertheless, it should be emphasized that the invariant component of the chain building-heat pump-subsoil consists in the thermo-physical properties of the subsurface to which a sound thermo-technical design should be adapted. Moreover, if

1 large GCHP systems are planned in areas where groundwater flows exist, these flows should
2 be adequately modeled (Sutton et al., 2003). For all these reasons, in some areas extensive
3 campaigns of thermo-physical characterization of the ground are carried out by means of
4 several techniques, from well logging analyses to active and/or passive geophysical techniques
5 (see e.g. Soengkono et al, 2013).

6 A large amount of information related to the ground is often available. In rare cases, data
7 have been taken specifically for geothermal energy deployment purposes. This is the standard
8 if a large BHE field is planned or for the installation of the first geothermal system, both of
9 which are typically preceded by a thermal response test (see e.g. Sharqawy et al., 2009),
10 sometimes accompanied by numerical modeling (Signorelli et al., 2007). In the majority of
11 cases, information about the soil is collected from geophysical explorations and/or boreholes
12 drilled for other purposes, e.g. irrigation, use of hydrocarbon resources, planning and/or design
13 of large civil engineering structures, waste management, analysis of presence and distribution
14 of possible pollutants in industrial or rural zones, or even scientific research purposes. The
15 quality of the data is heterogeneous, the depth can vary from few tens to several hundred
16 meters, and a certain degree of subjectivity in geological assessment descriptions can exist. A
17 possible solution to the problem was shown by Cultrera et al. (2012), where heterogeneous
18 well log data and passive seismic data were used to develop a new stratigraphic model of the
19 Venice area by means of a mode-based statistical analysis.

20 In this paper, a mode-based statistical approach to stratigraphy modeling for low
21 enthalpy geothermal energy exploitation purposes is proposed. Equivalent well logs (EWLs)
22 are obtained in the nodes of a regular grid defined in the studied area. The results are, for each
23 grid node: the lithological sequence of the modeled EWL, the effective TC related to such an
24 EWL, an estimation of the corresponding error and an evaluation of its geometrical/statistical
25 significance. The MATLAB/GNU Octave implementation of the proposed method

(GRIDWELL toolbox) was applied in the plain areas of a 19,000 km² region (Apulia, Southern Italy), leading to results that can be directly used for planning purposes.

2. The method and its MATLAB/Octave implementation

The proposed method is based on these operations: (1) possible import of data from a digital terrain model; (2) generation of a regular grid; (3) recognition of the subarea to which each available real well log (RWL) belongs; (4) mode-based EWL modeling for each subarea on the basis of the corresponding RWLs; (5) possible interpolation of water table data; (6) computation of the effective TC for each EWL; (7) visualization of the results and/or their export to a geographical information system (GIS). Although a well log generally is defined as a detailed record of the geologic formations penetrated by a borehole, in this paper the term RWL indicates all kinds of data on the subsoil stratigraphy, regardless of the data source (out-and-out well log, simple report on water well drilling, or also other sources).

The GRIDWELL toolbox, which is available as Supplementary Material to this paper, is a collection of MATLAB/GNU Octave (MO) functions (GNU Octave is the open source clone of MATLAB). If GRIDWELL runs under MATLAB, the user can either use the command line or graphical interfaces like combo boxes and input dialog boxes. If it runs under GNU Octave, the command lines, without graphical interfaces, must be used instead. In order to warrant a whole compatibility of GRIDWELL with the currently available GIS packages, the input and output data are EXCEL .xls or .xlsx files (it should be noted that the old format .xls has strong limitations in the maximum number of manageable rows with respect to the most recent format, i.e. .xlsx). The input data generally are well logs whose spatial distribution is suboptimal, i.e. whose typical distance is too high. Nevertheless, in some cases pre-processed data related to a digital terrain model (DTM) could be used. The toolbox can be utilized in both cases. The detailed flowchart is shown in Fig. 1. All the operations implemented in

GRIDWELL are now described in detail. Further information can be found in the GRIDWELL User's guide.

[Figure 1]

2.1 Use of a digital terrain model (optional)

Sometimes the user could have a digital terrain model (DTM) that represents the geometry of the study area, exported from a GIS package. The function `gw_extradtm` allows the data extraction from a DTM and the generation of files must be managed by GRIDWELL functions. It operates a planarization of the model, with ground level at depth 0. This means that if DTM elevation is z_{DTM} and the top level of a layer is z_{TL} , the corresponding values (i.e. depths) of the plane-parallel stratigraphy are 0 and $z_{DTM} - z_{TL}$ respectively. The grid used in GRIDWELL calculations can be equal to or different from the DTM one, according to the user's choice. Moreover, relatively small differences between DTM and soil can be managed by means of `gw_extradtm` function, as shown in the GRIDWELL User's Guide.

2.2 Grid generation

The subdivision of the studied area into square subareas by means of a regular grid is a straightforward operation. If the RWL having the extreme South-West position has coordinates $\mathbf{x}_{\min} = [x_{\min}, y_{\min}]^T$, where $x_{\min}$ and $y_{\min}$ are its East and North coordinate respectively, and s is the chosen grid side (horizontal step), the EWLs are modeled in the points $\mathbf{x}_{hk}^G = [x_{hk}^G, y_{hk}^G]^T$, where

$$\mathbf{x}_{hk}^G = [x_{\min} + (h+1)s/2, y_{\min} + (k+1)s/2]^T, \quad h = 1, 2, \dots, N_x - 1, \quad k = 1, 2, \dots, N_y - 1 \quad (1)$$

and N_x, N_y depending on the studied area size. The corresponding GRIDWELL function is `gw_wellgrid`.

The user can choose between two options for stratigraphy modeling: (a) using data from the RWLs belonging to the corresponding subarea for each EWL; (b) using data from the RWLs belonging to both the corresponding subarea and the eight nearest neighbor subareas for each EWL (Fig. 2). Other options are unavailable because extrapolation on high distances is not recommended as it could lead to meaningless results in soil property evaluation. In case (a), an EWL cannot be modeled if the corresponding subarea is empty. In case (b), an EWL cannot be modeled if the corresponding subarea and the nearest neighbors are empty. Therefore, the EWLs are consecutively row-wise numbered with an integer number $m = 1, 2, \dots, N$, where N is the number of effectively modeled EWLs.

The input data, which have previously been exported from a GIS environment or have been managed by means of the `gw_extradtm` function, are the identifiers, positions and depths of the RWLs. The output variables are cell arrays which can be saved in a MATLAB .mat file.

[Figure 2]

2.3 Equivalent well log modeling

Step (3), i.e. the search of the RWLs belonging to each subarea, and step (4), i.e. the mode-based statistical approach to model an EWL for each subarea center, are implemented together in the `gw_wellmod` function.

Before the data processing, the user must define the maximum depth of the EWL to be modeled, H , and the vertical step Δz . They are defined from the ground level $z = 0$. Here, positive z values are related to the subsoil. For the m^{th} EWL, let n_m be the number of depth intervals for which at least one RWL is available. Depth intervals from z_{k-1} to z_k where $z_k = z_{k-1} + \Delta z = k \Delta z$, $k = 1, 2, \dots, n_m$, are therefore considered.

The mode is the most frequent value in a distribution (strictly speaking, the point at which the probability mass function has its maximum value). If the mode of the lithology

distribution for the k^{th} depth step for the m^{th} EWL is evaluated, it corresponds to the most frequent lithology and therefore can be assumed to be the lithology for that step. In order to evaluate the mode, each lithology is labeled by means of an identifier, e.g. a string or a number. Although the lithology is defined by some quantitative parameters (e.g. mean density, granulometry, mechanical properties), the corresponding label has qualitative character only. This is not a problem because the mode, unlike mean and median, can be used to evaluate either quantitative data or data that, although related to measurements, cannot be directly measured or expressed by a number.

The recognition of the most suitable lithology for a depth step is very simple if the label distribution is unimodal. Unfortunately, in some cases the label distribution could be bimodal or also multimodal. In order to solve the problem of possible bimodality in a completely automatic way, the following preventative measure is necessary. The lithological species should be labeled by means of numerical codes in a consecutive way according to their geological properties, in particular their typical TCs. For example, in the case study 30 different labels are used. Species apparently similar but having very different TC cannot be adjacent in the label ranking. Examples of such species are sand, whose typical values are $\lambda \approx 0.6 \text{ W m}^{-1} \text{ K}^{-1}$ in anhydrous conditions and $\lambda \approx 2.6 \text{ W m}^{-1} \text{ K}^{-1}$ in saturated conditions respectively, and silty sand, with typical $\lambda \approx 0.5 \text{ W m}^{-1} \text{ K}^{-1}$ (anhydrous) and $\lambda \approx 1.7 \text{ W m}^{-1} \text{ K}^{-1}$ (saturated). On the contrary, limestone and conglomerate, which apparently are very different, could be adjacent in the label ranking because of their similar TC. Since the water table thickness along a 100-200 m column usually is prevalent on the dry material one, as typical for low enthalpy systems, two extreme values are considered and the intermediate moist / damp value between these two extremes are neglected.

For unimodality check purposes, two relative frequencies f_1 , f_2 are assumed to be equal if

$$|f_2 - f_1| \leq 0.15 \max\{f_1, f_2\}, \quad (2)$$

where the value 0.15, i.e. 15% difference, was suggested by a series of numerical simulations.

The unimodality test is passed if only a maximum of the frequency distribution exists. If the unimodality test is passed, the median and standard deviation (SD) of the distribution are computed. Strictly speaking, as numerical values these quantities are meaningless because the entries are labels, not results of measures. Nevertheless, they can be used to estimate the distance between mode and distribution center. A lithology L_{mk} is attributed to the k^{th} depth step for the m^{th} EWL if the main distribution peak (i.e. the mode) is very sharp and is far from the median less than 3 times the SD (Fig. 3.a), or if the peak is not sharp but is not as far from the median as the SD (Fig. 3.b). In the other cases, no lithology is recognized and a NaN (not-a-number) value is returned (Fig. 3.c).

If the unimodality test is not passed and the distribution is sharply bimodal, the two modes are recognized (each mode is obtained in a way like the unimodal case one), and the mode closest to the median leads the modeled value $L_{m,k}$ (Fig. 3.d). If no main peaks exist, or three or more peaks exist, a NaN value is returned (Fig. 3.e).

In order to provide insight as to the significance of the results, the flag $flag_{mk}$ is defined. In particular, $flag_{mk} = 1$ if the median is obtained from at least three RWLs belonging from the subarea, $flag_{mk} = 2/3$ if data from two RWLs are used, $flag_{mk} = 1/3$ if data from one RWL are used, $flag_{mk} = 0$ if either no data for the k^{th} depth interval exist or the m^{th} subarea is empty and L_{mk} is obtained from one or more nearest neighboring subareas (clearly, in this case under the condition that the interpolation from neighboring subareas be enabled by the user).

If the lowest modeled layer, related to the interval $k = n_m$, is a rock lithology and $H_m < H$, it is likely related to a bedrock. For this reason, another option is available to the

1 user, i.e. the depth prolongation to H in those cases where the last modeled interval
2 corresponds to a rock lithology.

3 The input variables of the `gw_wellmod` function are the following: the matrix provided
4 by `gw_wellgrid`; a matrix containing the RWL identifiers and, for each layer, the
5 corresponding top and bottom levels; the lithology and the code, provided by a GIS package;
6 the maximum depth H ; the vertical step Δz . These are the output variables: the matrix of the
7 EWL positions, containing their identifiers, depth and coordinates; the matrix of the EWL
8 data, containing the EWL identifier; the top and bottom extremes of each depth interval; the
9 corresponding modeled code; the flag and the number of used RWLs.

10 [Figure 3]

12 ***2.4 Water table data interpolation***

13 As already said, in some cases the TC of a material strongly changes if such a material is
14 anhydrous or saturated. Therefore, an interpolation of water table data on the EWL positions
15 could be necessary before the computation of the effective TC.

16 The `gw_watertable` function, included in the GRIDWELL toolbox, allows water table
17 computation with one of these techniques: nearest neighbor interpolation/extrapolation, natural
18 neighbor interpolation and linear interpolation. This function is generally enough in the
19 practical cases, where all the computations are therefore carried out in MO. Nevertheless, in
20 some cases the complex structure of the subsoil could require more sophisticated interpolation
21 techniques (e.g. kriging). In those cases, the localization output files provided by `gw_wellmod`
22 function can be imported in GIS environment and, after the data interpolation, the
23 localization/water table files can be re-imported in MO (Fig. 1).

25 ***2.5 Effective thermal conductivity computation***

The last step is the computation of the effective TC for each EWL, which is implemented in the `gw_lambdacalc` function. For the m^{th} EWL, the corresponding TC, $\lambda_{E,m}$, is computed as weighted mean of the TCs of the depth intervals. The weight for each modeled layer is $\Delta z / H_m$, and $\lambda_{E,m}$ is therefore expressed by

$$\lambda_{E,m} = \frac{\Delta z}{H_m} \sum_{k=1}^n \lambda_{mk} , \quad (3)$$

where λ_{mk} corresponds to the lithology L_{mk} recognized in the previous step. Since the TC of a non-rock material can depend on moisture, the water table can also be taken into account. If the water table corresponds to a layer k^* , for $k \leq k^*$ (i.e. $z \leq \Delta z k^*$), the anhydrous material coefficient is used, while the saturated material is used otherwise.

In order to provide information about the geometrical/statistical significance of the results, for the m^{th} EWL the flag

$$GS_m = \frac{\Delta z}{H} \sum_{k=1}^n flag_{mk} \quad (4)$$

is also computed. It is $0 \leq GS_m \leq 1$. The result $GS_m = 1$ corresponds to availability of at least three experimental data for each vertical segment that concur to $\lambda_{E,m}$ computation and $H_m = H$. The result $GS_m = 0$ means that a statistical significance is never reached in the layers of the m^{th} EWL.

If the SDs of the TCs of the involved geological species are available, for each EWL an estimate of the SD of the final result is also provided. It is estimated as

$$\sigma_{E,m} = \frac{\Delta z}{H_m} \sqrt{\sum_{k=1}^n \phi_{mk} \sigma_{mk}^2} , \quad (5)$$

where the weights ϕ_{mk} make the uncertainties worse if the statistical significance of the corresponding values is lower than 1. These weights are defined in the following way:

$$\phi_{mk} = \begin{cases} 1/flag_{mk} & \text{if } flag_{mk} \geq 0.11 \\ 3 & \text{if } flag_{mk} < 0.11 \end{cases}, \quad (6)$$

where the first row of (6) comes from the fact that if a mean $\bar{x}$ and an SD σ are obtained from N measured values $x_1, x_2, \dots, x_N$, σ is an estimate of the uncertainty related to the single measure and $\sigma/\sqrt{N}$ is an estimate of the SD related to the mean $\bar{x}$. The second row of (6) avoids divergence problems.

The gw_lambdacalc input variables are listed here: the matrix of EWL data, provided by gw_wellmod; the matrix of the EWL positions (compulsory) and water table (optional), provided by a GIS package by adding the water table to the matrix of EWL positions returned by gw_wellmod; the matrix of material specification that shows, for each lithology, the TCs in anhydrous and saturated conditions; the optional vector of the lithologies for which the depth prolongation must be performed. The output variable is a matrix whose columns are the EWL identifiers, depths, coordinates, TCs and statistical significance of the results.

Finally, the function gw_vis allows the visualization of the patterns of λ_E , GS and, if computed, σ_E (step 7).

3. Test with synthetic data

In order to evaluate the performance of the proposed approach and compare the results with the ones provided by ordinary kriging, numerical tests based on synthetic data were performed. Fig. 4.a shows the simple simulated ground system, i.e. a parallelepiped where three planar interfaces separate four different materials whose specifications are summarized in Table 1. The data in Fig. 4.a are $a = b = 1000$ m, $H = 100$ m, $h_1 = 20$ m, $h_2 = 25$ m, $h_3 = 25$ m, $h_4 = 30$ m, $h_5 = 45$ m, $h_6 = 0$ m, $h_7 = 10$ m, $h_8 = 45$ m, $h_9 = 40$ m, $h_{10} = 5$ m, $h_{11} = 30$ m, $h_{12} = 25$ m. Thirty simulated well logs (SWL), with x- and y-coordinates randomly obtained by uniform distribution, were considered (Fig. 4.b). The mean spacing was ~ 180 m.

[Table 1]

[Figure 4]

A first series of tests was aimed at evaluating the performance in correct lithology recognition with three grid sides: 50 m, 100 m and 200 m and vertical step of 5 m. In all the tests, computations were carried out using, for each EWL, data from the SWLs belonging to both the corresponding subarea and the eight nearest neighbor subareas. Each EWL stratigraphy was compared with the stratigraphy related to the same position of the simple model. The results are summarized in Table 2, where two facts can be noted: (i) as the grid side is progressively reduced with respect to the mean RWL spacing, and the theoretical number of grid nodes increases correspondingly, the fraction of obtained EWLs with respect to the possible grid nodes decreases (from 100% for 200 m grid to 45% for 50 m grid); (ii) the fraction of correctly modeled layers, i.e. of successful lithology recognition, is in the range 91-93%. These facts indicated that GRIDWELL can provide reasonable stratigraphic sequences and that the grid size should preferably be no lower than one half of the mean RLW spacing.

[Table 2]

Another test was aimed at comparing the results of GRIDWELL-based computations with the ones provided by ordinary kriging. The MATLAB implementation of ordinary kriging by Schwanghart (2010) was used. Input data for kriging were obtained by preliminary TC calculation at the SWL positions. The results in the case of GRIDWELL computation on a 5x5 grid (200 m grid side) are described in detail. In order to evaluate the effect of similar neighbors, local kriging was carried out; the maximum lag distance of experimental semi-variogram was 400 m, similar to the GRIDWELL-based search radius, and with 15-25 points in each neighborhood (distance step: 40 m). A spherical model with ~400 m range and 0.02 sill was obtained. Fig. 5.a and 5.b show the TC patterns provided by GRIDWELL and kriging respectively, whereas Fig. 5.c and 5.d show the patterns of standard deviation and statistical significance of GRIDWELL-based results. It can be seen that the results are similar. The

relative differences between the GRIDWELL and kriging results in each grid node never reached 10% and, therefore, were lower than the corresponding uncertainty. Similar results were also obtained by means of other interpolation methods. In conclusion, the performance of GRIDWELL in TC computation is like the one of typical geostatistical methods. Nevertheless, it should be noted that that GRIDWELL also provides a reliable stratigraphy model that could be used to characterize the ground properties.

[Figure 5]

4. Case study: Quaternary cover areas of Apulia region

4.1 Climate and geological setting

The proposed method was applied to the data related to the areas characterized by Quaternary deposits in the Apulia Region (Southern Italy), i.e. its plain areas. In Apulia's mountain territory, where GRIDWELL cannot provide results, the approach described in Di Sipio et al. (2014) could be used instead.

The Apulia region, whose surface is 19,000 km², stretches for 350 kilometers in the southeastern part of Italy, between the Adriatic and Ionian Seas (Fig. 6.a). More than 50% of its territory is at level ground. The elevation ranges from zero to 1150 m. Its climate is typically Mediterranean, characterized by mild winters and dry, warm summers. The mean annual temperature is ~17 °C along the Ionian coastal area and ~11-12 °C at the most elevated areas of Subappennino Dauno and Gargano. The mean annual values of rainfall are ~600 mm for the entire region; it ranges from about 1100-1200 mm at the Gargano Promontory to below 400 mm at the Tavoliere plain. The local climate is influenced mainly by the altitude and is not significantly influenced by the distance from the sea (Zito et al., 1988). The exploitation of geothermal energy in Apulia could be characterized by significant unbalancing between winter heating and summer cooling, with summer cooling considerably more prevalent.

[Figure 6]

The geological evolution of Apulia's area led to six different landscapes: Subappennino Dauno, Tavoliere plain, Gargano promontory, Murge plateau, Brindisi-Taranto plain and Salento peninsula (Fig. 6.b). Murge plateau and Gargano promontory are characterized by rock outcrops and, therefore, are not considered here. The Tavoliere plain is a wide alluvial plain stretching from the Apennine Chain front (West) down to the Manfredonia Gulf (East), where several rivers flow; the ground mainly consists of alluvial, beach and coastal dune deposits (Upper Pleistocene-Holocene), as well as Marine terraced deposits (Middle-Upper Pleistocene). The Taranto-Brindisi plain shows both Calcarenes (Middle Pliocene-Lower Pleistocene) and marine terraced deposits (Middle-Upper Pleistocene). The Salento peninsula is marked by a landscape composed of narrow and low-elevated carbonatic ridges which separate small plains shaped on Pliocene and Quaternary sediments. There karstic sinkholes exist and, along the coastal area, esorheic relict drainage are present. More information can be found in Caldara et al. (1998) and Sella et al. (1988).

4.2 Results

Fig. 4.c shows a map of the positions of Apulia's 3684 RWLs whose data are available. The depths of these RWLs range from 3 m to 832 m; mean and median of depth distribution are 110 m and 90 m respectively. The depths of several RWLs are relatively low because these wells were generally drilled for irrigation purposes. Since Apulia's deep aquifers are salty because of saltwater intrusion and, therefore, incompatible with irrigation activities, irrigation wells must draw the freshwater that floats on these aquifers. The spatial distribution of RWLs in some areas is very good; it is the case of the Tavoliere, the Brindisi Plain and, in a small degree, of the Salento peninsula, where the mean density is 0.2 RWL/km² and the mean distance between two near RWLs is 2.4 km. The RWLs are concentrated in the plain areas, which are mainly characterized by Quaternary sediments and where the most important cities

and villages rise and, on the contrary, are almost absent in hill/mountain areas (Murge plateau and Gargano promontory), where the population density is relatively low. Therefore, GRIDWELL can be applied to nearly all of Apulia's area considered significant for the market of low enthalpy geothermics. Since the typical depth of the BHE of a low enthalpy geothermal system is ~ 100 m, $H = 100$ m was used in stratigraphy modeling.

Table 3 summarizes the TC values for the most common materials recognized in the studied areas. This table was obtained through both literature review and laboratory measurements. A general review of values described in literature for rocks and sediments was carried out, leading to the recognition of ~ 90 lithologies from more than 60 authors. The obtained value ranges were compared with those obtained from laboratory measurements performed on 294 samples collected in Southern Italy (Di Sipio et al 2013a). This because, at local scale, the different climate and environmental conditions, together with the structural and geological features of the territory, lead to TC values significantly different from the ones found in literature (Di Sipio et al 2013b). The rock species are highlighted in Table 3 because the depth prolongation to H was carried out in those cases where the last recognized layer corresponded to a rock lithology and therefore was likely related to bedrock.

[Table 3]

Fig. 7 shows the effective TC, λ_E , the corresponding geometrical/statistical significance, GS , and the positions of modeled EWLs for a 3 km grid step. Although GRIDWELL allows two options for the nearest neighbors management, the high spatial variability that characterizes Apulia's coastal area suggests the use, for each EWL, only of the data from the RWLs belonging to the corresponding subarea. The SDs on TC data used in the computations were obtained from the range of data variation in Table 1. Fig. 8 shows the GS distributions for the three grid sides 10 km, 3 km and 1 km. The amounts of modeled EWLs are 126 (10 km case), 644 (3 km) and 2785 (1 km). The 10 km case leads to an insufficient spatial resolution, whereas the results for 1 grid step are characterized by an insufficient

1 statistical significance. In the 3 km case, where the grid step is similar to the mean distance
2 between the RWLs (2.4 km), the mean GS flag is 0.63 and the fractions of EWLs
3 characterized by $GS \geq 2/3$ and $GS \geq 1/3$ are ~50% and ~70% of the total respectively.
4 Moreover, it is $GS \leq 0.1$ for only one EWL (0.2% of the total). In the 1 km case, the
5 corresponding values are 0.29 (mean GS), 10% (fraction of $GS \geq 2/3$ EWLs), 26%
6 ($GS \geq 1/3$), and 33% ($GS \leq 0.1$). There therefore exists a degradation of the statistical
7 significance of the results with the reduction of grid step.

8 [Figure 7]

9 [Figure 8]

10 The GS map (Fig. 7.b) highlights that in some areas there are several RWLs, but GS is
11 low ($GS \approx 0.2$). This problem comes from the fact that several irrigation wells do not reach
12 100 m depth. In the corresponding areas GRIDWELL provides results valid for general
13 planning purposes, but the BHE design cannot be carried out without new drillings.

16 5. Discussion

17 A reliable estimate of ground TC, combined with a good evaluation of heating and cooling
18 demand, allows an evaluation of BHE size. For planning purposes, a statistical approach
19 based on available data, mainly well logs characterized by a high degree of heterogeneity and
20 non-optimal spatial positions, can provide results with reasonable quality, under the condition
21 that these are supported by the corresponding significance.

22 The proposed approach is based on statistical operator mode. This because statistical
23 operator mode, unlike mean, median and other statistical parameters, can be used to model
24 qualitative features that can be defined in a way precise enough to allow measurements but
25 they cannot be measured or expressed by a number. Typical examples of features that can be
26 defined in a precise way, but whose quantitative analysis is very complex and generally leads

1 to very dispersed results, are the lithologies of a layered subsoil. The mode-based
2 stratigraphic characterization implemented in GRIDWELL provides, for each grid node and
3 each layer, the most probable lithology. The quantitative phase of the process is left to the last
4 part, where the data on the typical species in the area are used to obtain the final equivalent
5 TC (λ_E). In this way, the analyses can be easily repeated with other TC data to evaluate
6 possible changes in the results.

7 The computational cost of the procedures implemented in GRIDWELL toolbox is
8 reasonable. The whole analysis of the case study, i.e. the Apulia region ($\sim 19,000 \text{ km}^2$ surface
9 and 3684 RWLs) was carried out in two minutes with a common notebook, including the
10 interactive operations.

11 The results of numerical tests highlighted that GRIDWELL correctly recognized the
12 lithological species in more than 90% of modeled layers and that the TC pattern is like the
13 one obtained by means of ordinary kriging.

14 The grid step can be changed, in order to change the scale of analysis. The results
15 highlight that better results, in terms of agreement between the opposite requirements on
16 adequate spatial resolution and on reasonable geometrical/statistical significance of the λ_E
17 pattern, can be obtained with a grid step similar to the mean distance between the well logs.
18 The grid step, however, cannot be lower than half of this distance (Table 1, Figs. 7 and 8).

19 The GRIDWELL-based results obtained in Apulia's plain areas are credible because the
20 main geomorphological structures can be easily correlated to the λ_E pattern. For example, in
21 the Tavoliere area (North-Western part of the region, Figs. 6 and 7), the depositional
22 morphology of the main surface water network, in particular the basins of three rivers, can be
23 easily recognized. Similarly, in the Salento peninsula (South-Eastern part of the region) the
24 obtained λ_E pattern corresponds to the known geology of the subsoil, and the borders of
25 Brindisi and Taranto plains also appear in the λ_E map. No particular features appear where

1 the geology is expected to be continuous, i.e. no artefacts due to interpolation appear in the
2 studied region. These results seem to confirm the validity of the proposed approach based on
3 mode and, therefore, its exportability to different contexts.

4 For each EWL, the geometrical/statistical significance of λ_E , i.e. GS , is also provided.
5 Fig. 7 shows that in Murge and part of Gargano significant results cannot be obtained because
6 of lack of experimental data. Instead, significant results are obtained in Tavoliere, Salento,
7 part of Brindisi and Taranto. Equation (4) indicates that a low GS flag comes up when (i) few
8 RWLs exist in the EWL neighborhood, (ii) the depth of most of the available RWLs is lower
9 than H . In the Apulia region both of these possibilities occur because several RWLs are
10 irrigation wells related to the shallow layers only. In some areas all the available RWLs do
11 not reach 100 m depth. An example of such areas is the Brindisi plain, where often $GS < 0.3$.
12 Apparently, a large part of Apulia's territory characterized by Quaternary sediments seems to
13 be suitable for exploitation of geothermal resources because it is $\lambda_E \geq 2.0 \text{ W m}^{-1} \text{ K}^{-1}$ in ~60%
14 of its territory (Fig. 7.a). Nevertheless, the significance of the results (Fig. 7.b) must be taken
15 into account.

16 The proposed method is mainly conceived for planning purposes. In principle, the
17 results could also be used for design purposes. Nevertheless, in such a case they must be used
18 with care because their accuracy at local scale could be insufficient. The fact that an area
19 shows good conditions for extensive exploitation of the geothermal resources does not imply
20 that in a single site of the area these conditions are sufficient. The availability of an estimate
21 of the result significance allows a recognition of the conditions where they can be used for
22 design purposes. If the flag defined by Equation (4) is at least ~0.6 and the grid step is no
23 more than a few km, the obtained λ_E can be used for design purposes. In other cases, the
24 results can be used by public authorities for planning purposes only and the drillings of well
25 logs aimed at characterizing the material and thermal response testing are necessary.

6. Conclusions

This paper presented a method aimed at modeling the stratigraphy and computing the equivalent thermal conductivity, λ_E , in the nodes of a regular grid, together with its MATLAB/GNU Octave implementation (GRIDWELL toolbox).

The method's performance was evaluated by means of a simulated simple system real data from Apulia's plain areas. These results led to stratigraphy and λ_E patterns compatible with the known geomorphological structures. In conclusion, the proposed approach can aid the recognition of areas suitable, or unsuitable, for an exploitation of low enthalpy geothermal resources at regional scale and, therefore, can provide significant information for planning purposes.

7. Supplementary material

The GRIDWELL toolbox is available as Supplementary Material of the online version of this paper.

Acknowledgements

This research was carried out within the framework of the VIGOR Project, which is part of the Operational Programme “Renewable Energy and Energy Efficiency”, co-funded by European Regional Development Fund (ERDF) 2007–2013, activity line “Experimental Actions in Geothermal Energy”. The authors would like to thank the management of the VIGOR Project and in particular Simonetta Piezzo (Ministry of Economic Development, Department of Energy) and Enrico Brugnoli (Director of CNR-DTA).

References

- Banks, D., 2012. An introduction to thermogeology: Ground source heating and cooling. 2nd Edition. Wiley-Blackwell, Oxford, United Kingdom, 526 pp.
- Blum, P., Campillo, G., Kölbel, T. 2011. Techno-economic and spatial analysis of vertical ground source heat pump systems in Germany. *Energy*, 36(5), 3002–3011.
- Caldara, M., Centenaro, E., Mastronuzzi, G., Sansò, P., Sergio, A., 1998. Features and present evolution of Apulian Coast (Southern Italy). *Journal of Coastal Research*, 26, 55–64.
- Cultrera, M., Antonelli, R., Teza, G., Castellaro, S., 2012. A new hydrostratigraphic model of Venice area (Italy). *Environmental Earth Sciences*, 66(4), 1021–1030.
- Di Sipio, E., Galgaro, A., Destro, E., Teza, G., Chiesa, S., Giaretta, A., Manzella, A., 2014. Subsurface thermal conductivity assessment in Calabria (southern Italy): a regional case study. *Environmental Earth Sciences*, 72(5), 1383–1401.
- Di Sipio, E., Galgaro, A., Destro, E., Giaretta, A., Chiesa, S., Manzella, A. & Vigor Team, 2013a. Thermal conductivity of rocks and regional mapping. In *Proceedings of EGC 2013 - European Geothermal Congress, Pisa (Italy), 3-7 June 2013*, EGC-SG3-04, pp. 1–6.
- Di Sipio, E., Chiesa, S., Destro, E., Galgaro, A., Giaretta, A., Gola, G., Manzella, A., 2013b. Rock thermal conductivity as key parameter for geothermal numerical models. *Energy Procedia*, 40, 87–94.
- Schwangart, W., 2010. MATLAB functions for geostatistical analysis. Available online at: <http://www.mathworks.com/matlabcentral/profile/authors/870595-wolfgang-schwangart>, accessed: 24 February 2015.
- Sella, M., Turci, C., Riva, A., 1988. Sintesi geopetrolifera della Fossa Bradanica. *Memorie della Società Geologica Italiana*, 41, 87-107.
- Signorelli, S., Bassetti, S., Pahud, D., Kohl, T., 2007. Numerical evaluation of thermal response tests. *Geothermics* 36, 141–166.

1 Sharqawy, M.H., Said, S.A., Mokheimer, E.M., Habib, M.A., Badr, H.M., Al-Shayea, N.A.,
2 2009. First in situ determination of the ground thermal conductivity for borehole heat
3 exchanger applications in Saudi Arabia. *Renewable Energy* 34, 2218–2223.

4 Soengkono, S., Bromley, C., Reeves, R., Bennie, S., Graham, D., 2013. Geophysical
5 techniques for low enthalpy geothermal exploration in New Zealand. *Exploration Geophysics*
6 44(3), 215–227.

7 Sutton, M., Nutter, D., Couvillion, R., 2003. A ground resistance for vertical bore heat
8 exchangers with groundwater flow. *Journal of Energy Resources Technology* 125, 183–189.

9 Zito, G., Cacciapaglia, G., 1993. Precipitazioni in Puglia: mappe stagionali, In *Proceedings of*
10 *5th Workshop Progetto Strategico Clima, Ambiente e Territorio nel Mezzogiorno*, Amalfi, 28-
11 30 Aprile 1993, pp. 223–253.

12

Table Captions

Table 1. List of the lithological species considered for the numerical test.

Table 2. Results of test of correct lithology recognition.

Table 3. List of lithological species recognized in Apulia region and corresponding typical thermal conductivity in the two cases of anhydrous and saturated material, provided by both literature review and laboratory measurements. The last column indicates the rock species.

Figure captions

Figure 1. Flowchart of proposed method.

Figure 2. Example of positions of EWLs for a same horizontal step: (a) for each EWL, use of data from RWLs belonging to the corresponding subarea; (b) for each EWL, use of data from RWLs belonging to both the corresponding subarea and the eight nearest neighboring subareas.

Figure 3. Mode-based choice of the most representative lithology: (a) unimodal distribution, sharp peak, lithology recognized; (b) unimodal distribution, non-sharp peak, lithology recognized; (c) unimodal distribution, lithology not recognized; (d) bimodal distribution, lithology recognized; (e) bimodal distribution, lithology not recognized; (f) multimodal distribution, lithology not recognized.

Figure 4. Synthetic data used to validate GRIDWELL results: (a) schema of 4-layer ground system, where horizontal and vertical scales are different; (b) positions of 30 simulated well logs.

Figure 5. Analysis of data related to 30 simulated well logs: (a) contour map of equivalent thermal conductivity provided by GRIDWELL; (b) contour map of equivalent thermal conductivity provided by ordinary kriging; (c) and (d) GRIDWELL-based contour maps of standard deviation and statistical significance.

Figure 6. Apulia region: (a) climatic map; (b) geological setting, main landscapes and rivers; (c) positions of the well logs whose data are available.

1

2 **Figure 7.** Plots provided by GRIDWELL for a 3 km grid step: (a) contour map of equivalent
3 thermal conductivity for 100-m depth where the statistical significance flag is at least 0.1; (b)
4 contour map of statistical significance. The x-marks are the positions of the modeled
5 equivalent well logs.

6

7 **Figure 8.** Histograms of flag on geometrical/statistical significance of computed thermal
8 conductivity for three values of grid step: 10 km, 3 km and 1 km.

9

Table 1

Code	Thermal conductivity (Wm ⁻¹ K ⁻¹)			
	anhydrous		saturated	
	value	SD	value	SD
1	0.2	0.2	0.2	0.2
2	0.9	0.2	1.5	0.4
3	1.6	0.3	2.1	0.4
4	2.8	0.3	2.8	0.3

Table 2

Test	1	2	3
Grid side (m)	50	100	200
Grid size	20x20	10x10	5x5
Number of modeled EWLs	179 on 400 (45%)	92 on 100 (92%)	25 on 25 (100%)
Total number of layers	3580	1840	500
Correct species recognition	3329 (93%)	1678 (91%)	454 (91%)

Table 3

Material	Code	Thermal conductivity (Wm ⁻¹ K ⁻¹)		Rock
		anhydrous	saturated	
Air	1	0.2	0.2	No
Peat	2	0.3-0.5	0.7-0.9	No
Tuff	3	0.5-1.20	1.40-2.40	Yes
Water	4	0.6	0.6	No
Topsoil	5	0.7-1.0	1.0-1.5	No
Gypsum	6	0.9-1.9	1.1-2.5	No
Clay	7	0.4-0.9	1.1-1.5	No
Clay and gravel	8	0.4-0.9	1.5-2.2	No
Silty clay	9	0.5-0.8	1.3-1.6	No
Gravel with clay	10	0.4-0.9	1.6-2.5	No
Basalt	11	1.3-1.7	1.3-1.7	Yes
Silt	12	0.3-1.0	1.1-1.7	No
Silty sand	13	0.5-1.0	1.1-1.7	No
Gravel	14	0.4-0.9	1.8-3.3	No
Clay and sand	15	0.4-0.9	1.0-2.2	No
Conglomerate and clay	16	0.5-1.0	1.2-2.0	No
Shale	17	1.5-2.2	1.5-2.2	Yes
Breccia	18	1.3-2.5	1.3-2.5	Yes
Sand and gravel	19	0.5-0.8	1.5-2.1	No
Marl	20	1.0-2.2	2.2-2.8	Yes
Sand	21	0.4-0.9	1.6-2.5	No
Calcarenite	22	0.9-2.8	2.1-3.5	Yes
Sandstone	23	1.2-3.2	1.7-3.2	Yes
Limestone	24	1.0-3.0	2.5-3.9	Yes
Conglomerate	25	1.2-2.8	2.4-3.2	Yes
Marble	26	2.1-3.5	2.1-3.5	Yes
Dolomite	27	2.8-5.0	2.8-5.0	Yes

Figure 1

Figure 2

Figure 3

Figure 4

Figure 5

Figure 6

Figure 7

Figure 8