PETROCK

Lithologically weighted interpolation of petrophysical data

User's guide

Version 1.1

2011

Markku Pirttijärvi

University of Oulu

Department of Physics

1. Introduction

The PETROCK program is used to interpolate irregularly sampled petrophysical data, e.g. density or magnetic susceptibility, or geochemical data (K, Zn, Au) on an even grid. The two main purposes of the PETROCK program are: a) to create improved maps of the geographical distribution of petrophysical parameters, and b) to create initial 3-D models for potential field modeling and inversion.

The program uses moving window strategy in which the gridded values are estimated from the nearest sample points using both the mean and the inverse distance weighted mean. The program also uses the digital geological (lithological) maps to provide areal weights and background weights for points inside the same lithological unit. A limiting condition can be defined to remove outliers from the petrophysical data inside the investigation area or inside each lithological unit. If the limiting condition is based on the mode (largest class) or the median the gridding can yield the base level of the lithological units better than if the mean of the samples was used. Thus, in areas where sampled data are not available the PETROCK gridding should be more reliable than normal interpolation would be. In addition, rock rejection rules can be defined that either reject some rock types from particular lithological unit or allow only certain rocks in certain lithological unit.

The petrophysical data are read from column formatted text files. The polygon information of the geological maps is read from text files that have been converted from Mapinfo MID/MIF files into a special LIT file format using the MIF2BNA program. To reduce the number of unnecessary polygon vertices the polygon data can be optimized using POLYTUNE (see Appendix). Input processing parameters are given via keyboard or using a separate script file (PETROCK.INP). Output data are saved into column formatted text files.

2. Table of contents

Keywords: Petrophysics; Interpolation; Density; Magnetic susceptibility.

3. Interpolation method

The lithological weighting of petrophysical data is made using the digital geological maps of Finland in the scale of 1:1 million (Korsman et al., 1997) and the petrophysical database of rock density and magnetic susceptibility values measured in laboratory conditions (Korhonen et al, 1997). The digital map used in PETROCK consists of 92 lithological units corresponding to geological classes of different rock type, stratigraphy, and age. Because of the high level of details of the geological map of Finland in the scale of 1:1 million, a simplified version in the scale of 1:5 million is shown in Figure 1.

The map consists of 7922 polygons of which 1959 are duplicates resulting from holes inside the actual polygons. The polygon data consists of about 450000 polygon vertices. Because the determination of the locations of the sample points inside the polygons is the most time consuming part of the PETROCK method, the polygon data has been optimized using the LITOTUNE program (see Appendix) in advance. After the tuning the reduced number of polygon vertices is about 290000.

The petrophysical database contains more than 130 000 samples. Before the actual gridding each petrophysical data point has been given a lithological code (1-92) based on its location on the geological map. Petrophysical data points outside the borders of all lithological polygons are given class number 0 (zero). The density data are limited between 2200 and 3200 kg/m^3, which approximately refers to $\pm 3.0 \times$ the standard deviation (170 kg/m^3) around the mean value (2700 kg/m^3). The main purpose is to remove outliers, erroneous data values, and other exceptional data values obtained from ore bodies, for example. The resulting density data consists of 129 252 points. The mean value and standard deviation are 2725.0 and 134.8 kg/m^3, respectively. Likewise, the magnetic susceptibility data were pre-processed by removing values smaller than 10^{-5} SI units (including negative values) and taking the 10-base logarithm of them. The resulting $\log_{10}$-normalized susceptibility data consists of about 126 000 points, the mean and standard deviation being -2.163 (=0.00687) and 0.8551, respectively.

The locations of the sample points used in this study are shown in Figure 2. Figure 2 and the results hereafter are mapped in the national rectangular KKJ coordinate system of Finland.

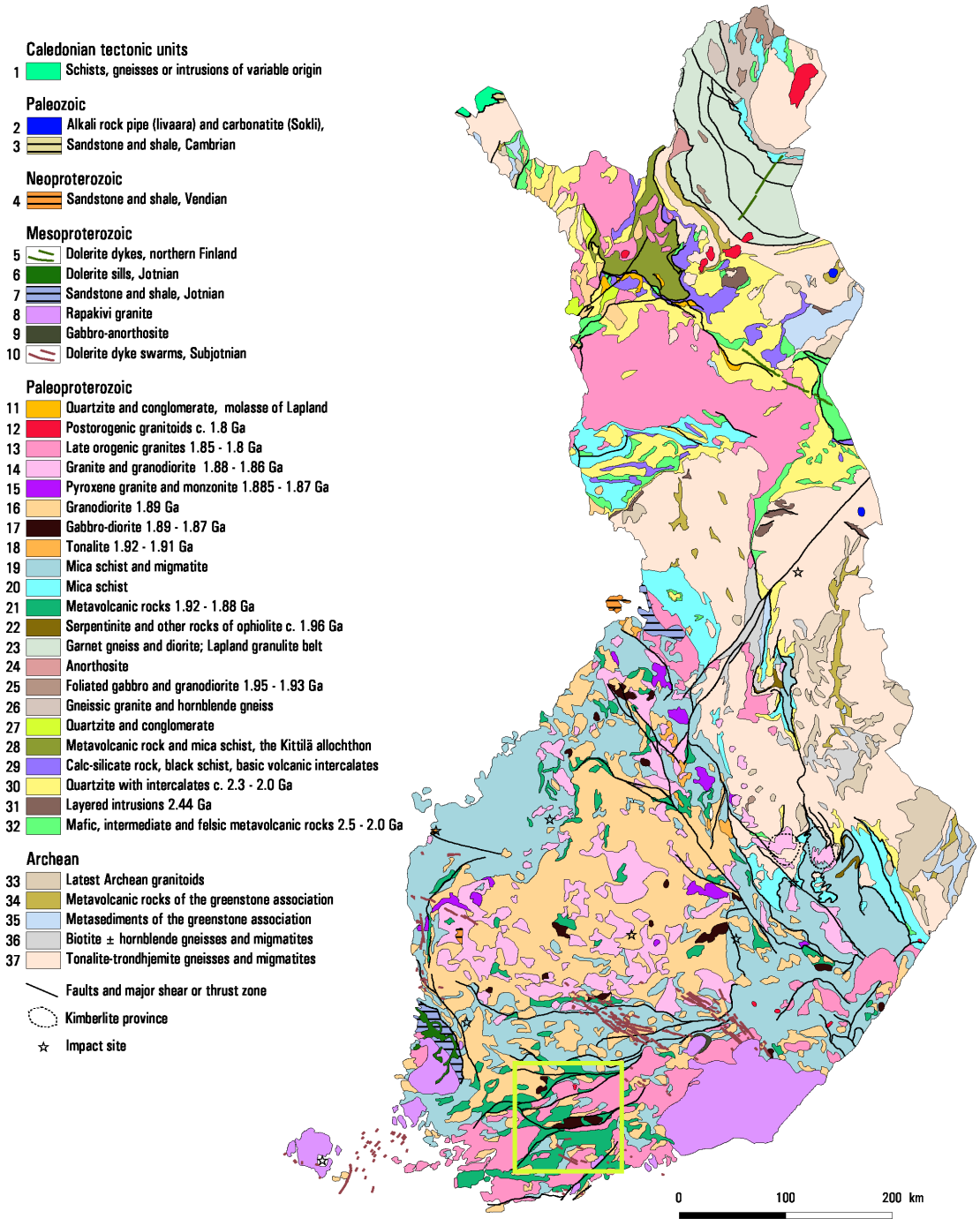

Fig. 1. Geological map of Finland (Korsman et al., 1997). The yellow square shows the location of the detailed study area discussed in the text.

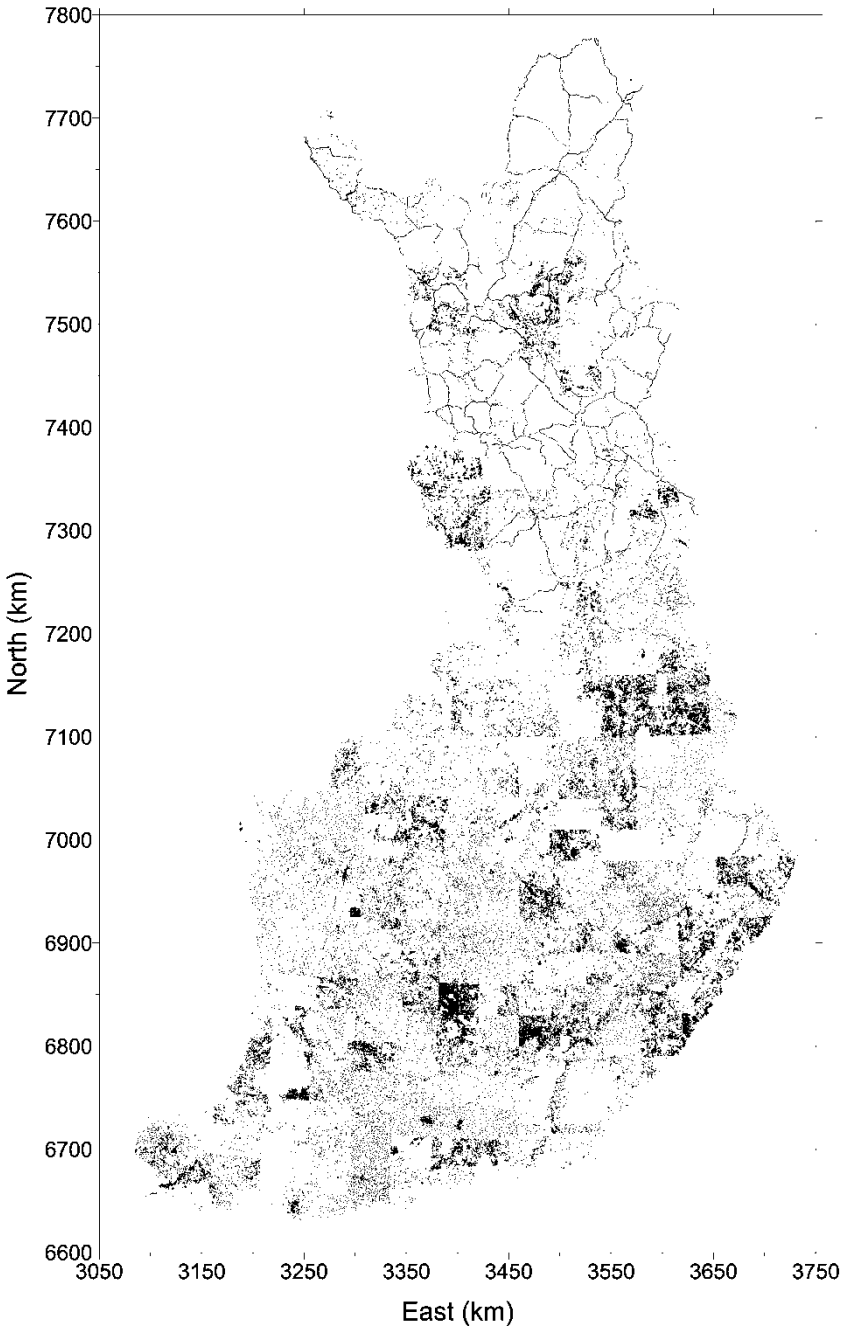

Fig. 2. Locations of the petrophysical samples of Finland.

Because of the vast size of the research area covering the whole Finland (700 km by 1200 km) and the large number of polygon vertices and petrophysical samples, the gridding is implemented using a moving window strategy. The definitions of the gridding concept are illustrated in Figure 3. A rectangular *xy* coordinate system is considered. The research area, *R*, is systematically processed by a moving window, which is the computational area, *C*, which contains the evenly spaced grid points where the mappings are to be made. The computational area is surrounded by margin area, *M*. Together *M* and *C* define the investigation area *I*. For

each *C* the petrophysical data and the lithological polygons are limited by the investigation area *I*. This makes computations faster, because it allows using only a fraction of the data and the polygons when mapping the grid points inside *C*. However, to ensure that a sufficient amount of data points locate inside *I*, the margins *M* are doubled in size as long as needed until the number of petrophysical data points inside *I* is greater than the predefined minimum value.

Before processing the individual grid points the surface areas of the lithological units below the current rectangular investigation area are computed. The lithological weights are obtained by normalizing the surface areas of each unit by the area of *I*. If *C* does not contain any lithological units, all the grid points inside *C* are given a user-defined mean value or the value reserved for missing data.

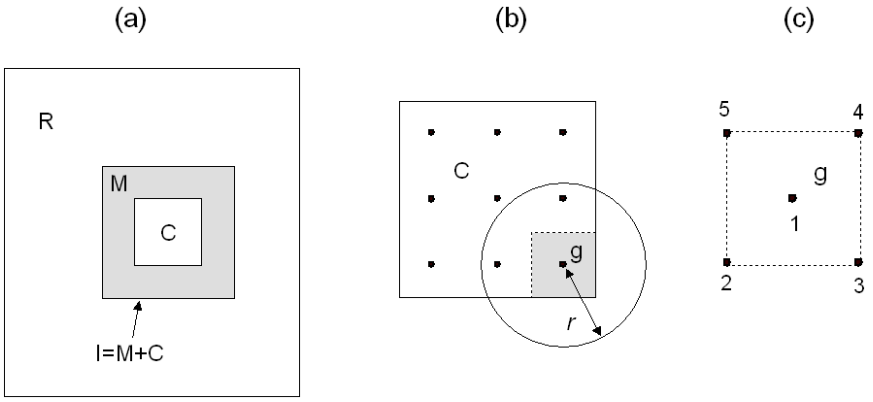

Figure 3. (a) The components of the moving window strategy: total research area (*R*), computational area (*C*), and margin area (*M*), (b) the search radius (*r*) around a grid area (*g*), and (c) the five special points of a grid point.

The grid points must also locate inside the area defined by the polygons (within national borders and islands). Thus, for each grid point the center and the four corner points are first tested (Fig. 3c). If all special points are outside the lithological polygons, the grid point is given the mean value (or missing data value). Although it would be more accurate to check if the small rectangle around the grid point (*g* in Fig. 3b) overlaps with any of the polygons, testing only the special points is faster. If at least one special point is above some polygon, the petrophysical samples within the search radius (*r*) around the grid point (Fig. 3b) are sought for and given weights based on the inverse of their distance from the grid point. If not enough sample points are found, the search radius is doubled in size as many times as needed. Once

the petrophysical data values, their inverse distance weights, and the lithological weights are known the geologically weighted mean of the petrophysical parameter can be computed at the grid point.

Before processing the data it is possible to use coordinate rounding and consequent removal of multiple data values, for example, from a drill-hole. Multiples can be replaced with the mean or the median value, or they can all be accepted or neglected. Important parameters affecting the gridding results are also the minimum number of data values per investigation area (N_{min}), the minimum number of points per grid point (n_{min}), and the minimum search radius (r_{min}) used to search for the nearest petrophysical data points. Note that the size of the computation area should be adjusted so that it contains enough data values and provides sufficiently accurate lithological weights for the computation area. PETROCK uses several weighting methods and stores all the results into the same output file. It stores the standard deviations and the number of petrophysical data points used to compute the grid values into another output file. Additional information (mean, median, maximum, minimum, standard deviation, average deviation, number of points) about each investigation area is stored into third output file.

The weighting provided by the above-mentioned method is weak when considering areas where samples are not available. To provide a better 'guess' in such locations an alternative background weighting is used. In this method the mean, median or mode (largest class) of the lithological unit is used to define a background value of each lithological unit. To improve the estimate of background value, outliers are removed using a limiting condition based on the standard deviation around the mean, median or mode of the investigation area. In addition, rejection rules based on the hierarchical rock classification of the samples can be used to reject/allow certain rocks from/in certain lithological unit. This allows, for example, removing all but sedimentary rocks from sedimentary units and, thus, greatly improves the estimate of petrophysical background value. The actual background weighting is made by using the background estimates of the lithological units below the five special points (Fig. 3c) as additional data values.

4. Program usage

Before using the program, ensure that the petrophysical data file and the file containing the lithological polygons are available and in the correct format (see chapter 6).

Immediately after the program has been started it asks whether the input parameters are read from a file (recommended) or given interactively from the keyboard. The name of the input file must be PETROCK.INP. If the input file does not exist the parameters should be given from the keyboard. In this case the PETROCK.INP file will be created automatically and it can be used in further examples thereafter. The format of PETROCK.INP and the meaning of various input parameters are discussed in the chapter 6.

Because both the petrophysical data file and the polygon file can be very large (several megabytes) it is recommended that they are stored in the program folder together with the executable file and each gridding experiment is run in a separate subfolder. This requires that each subfolder contain a copy of the PETROCK.BAT batch file. The batch file should contain only one line that defines the path to the actual executable file: "C:\PETRO\PETROCK\ PETROCK.EXE" or " ..\PETROCK.EXE", where "..\" denotes the previous directory level. The PETROCK program can then be run double-clicking the icon of the batch file (inside the Explorer program of Windows). Note that separate versions of the PETROCK.INP script file should be located inside each subfolder.

The program stores its results in column formatted text files. To visualize the gridded data as a contour or an image map a separate plotting program must be used (e.g., Golden Software Surfer or Geosoft Oasis Montaj). If the results are not satisfactory the PETROCK program can be run again using different initial processing parameters.

5. Input parameters

The following example describes the syntax of the PETROCK.INP script file. Note that all file names (text strings) must be put inside hyphens (e.g. `'here'`). Note also that if the program is run interactively, the parameters are provided in the same order as in the script.

```
row | text
-------------------------------------------------------------------------
-3  | #
-2  | # input parameters file for PETROCK4 program
-1  | #
 1  | '..\xydnkr3.dat'                    ! petrophysics file
 2  | '..\lito1b.lit'                     ! polygons file
 3  | 3575. 3750. 6925. 7000.  1.  1.     ! xmi,xma,ymi,yma,xst,yst
 4  | 200.  200.  25.   25.              ! xwin, ywin, xmar, ymar
 5  | 2                     ! sub-area cut method
 6  | 1.                    ! sub-area max dev
 7  | 100                   ! sub-area points
 8  | 5                     ! data points per grid point
 9  | 5.   1.               ! radius minimum and scale length
10  | 2.                    ! inverse power
11  | 2.                    ! lithological weight
12  | 2700.                 ! missing values
13  | 0                     ! rock type weights
14  | 2                     ! duplicate points
15  | 0                     ! coordinate rounding
16  | 1                     ! store used petrophysical data
17  | 0                     ! store used lithological polygons
18  | 'Petrock3.dat'        ! output mean parameters file
19  | 'Stdrock3.dat'        ! output std deviations file
20  | ' '                   ! output for bloxer program
21  | 9                    ! number of rejection rules
22  | 1 -4
23  | 2 -4
24  | 3 -4
25  | 4 -4
26  | 6 -4
27  | 5  4
28  | 5  1
29  | 8 -1
30  | 9 -1   ! lito-rock exclusion/inclusion pairs
34  |
35  | # 1= felsic, 2= dyke, 3= volcanic, 4= sediment, 5=metam., 6= other
```

The three lines that start with a #-character on the top are used as comments lines.

Petrophysics file: The 1.st line defines the name and path to the petrophysical data file.

Polygons file: The 2.nd line defines the name and path to the lithological polygons file.

`Xmi,xma,ymi,yma,xst,yst`: The 3.rd line defines *x* and *y* coordinates of the start (West and South) and the end (East or North) of the rectangular study area over which the data will be gridded (i.e., the research area *R* in Fig. 1a). The last two parameters define the grid spacing in *x* and *y* directions. Note that in PETROCK *x* and *y* coordinates refer to East-West and South-North directions, respectively.

`Xwin, ywin, xmar, ymar`: The 4.th line first defines the *x* and *y* size of sub-area that is the moving data window that passes through the main mapping area (*C* in Fig 1a). The third and the fourth parameter on this line define the *x* and *y* width of the margin area (*M* in Fig. 1) around the computation area. Note that the width of the investigation area (*I*) is actually $C+2\times M$. For example, the *x* width is `xwin`+2×`xmar`.

`Sub-area cut method`: The 5.th line defines the reference value(s) used to limit the petrophysical data within each sub-area. This parameter can have a value IWEI= 0, 1, 2 or 3. If IWEI= 0 the mean of all data values inside the sub-area are used as a reference point. The other IWEI values use (1) the mean, (2) the median, or (3) the mode (largest class) of each lithological unit as separately for each lithological unit. If the parameter deviations within the sub-area or lithological unit are large this method can be used to cut out outliers and samples from incorrectly classified petrophysical samples.

`Sub-area max dev`: The 6.th line defines the multiplier (*a*) for the standard deviation that sets the limiting range around the reference point discussed above. Only the values within range $p^* - dp < p < p^* + dp$, where p^* is the reference value (IWEI) and $dp=a \cdot std$, where *std* is the standard deviation of the data inside the sub-area (IWEI=0) or inside each lithological unit (IWEI= 1,2,3). IWEI and *a* are used to cut out or to reduce the effect of values that for some reason have been given incorrect lithological code. Remember that the geological boundaries do not represent actual formations of the nature. For example, if a low density sedimentary unit contains lots of dense sample points that actually belong to the surrounding area, setting IWEI=3 can remove these outliers provided that their amount is small compared to the number of samples of the main sedimentary unit.

`Sub-area points`: The 7.th line defines the minimum number of sample points inside the sub-area (MSP). If the sub-area (investigation area *I* in Fig. 1a) contains too few samples the

size of the sub-area will be increased by the size of the margins (*M*) until the number of points is large enough.

Data points per grid point: The 8.th line defines the minimum number of sample points per grid point (MGP). If the number of points is too small, the search radius (*r* in Fig. 1b) is doubled in size until the number of points is large enough.

Radius minimum and scale length: The 9.th line defines the minimum length of the search radius (*r*) and the scale length (*s*) used to normalize the distances. The initial value of the search range should be large enough that it would contain the MGP points by default. However, large search radius can flatten the gridding and enhance the effect of anomalous sample points. The scale length depends on the size of the computation area and the power of the inverse distance weighting. Normally the minimum radius is equal to the grid spacing and the scale length is equal to one. If the distances are large it may be advantageous to increase the scale length to distribute the inverse power to greater distances.

Inverse power: The 10.th line defines the power of inverse distance weighting. The higher the power the more details the grid can show. This, however, may not always be desirable since the effect of anomalous points will create spike-like dots in the map. Normally the inverse power should be 1 or 2. The inverse distance weights are computed as $w_r = s/(s+t^k)$, where *s* is the scale length and *t* is the distance of the sample point from the grid point. Thus, if the distance is equal to zero the weighting factor is equal to one. If the distance is equal to the scale length the weighting factor is equal to 0.5.

Lithological weight: The 11.th line defines the standard lithological weight (LWEI). Lithological weighting means that the mean value of the lithological unit (units) the grid point belongs to is used as an additional data value(s) (as if it were a point inside the search radius). LWEI gives additional importance to this mean value and, thus, creates a grid map, where each lithological unit stands out having a common level. However, since the mean value is computed using the points inside current sub-area, the level can change between sub-areas, which creates a kind of mosaic pattern if the size of the sub-area is small and the lithological weighting is strong.

Missing values: The 12.th line defines the parameter value used for grid points that do not fit inside any lithological polygon. Thus grid points that locate outside the national borders or over the sea area are given a constant value. Typically this value should represent the mean value of the petrophysical data. Note that points outside the lithological units are given code 0. Although these values will not be used in lithological weighting they will be added to the sub-area mean.

Rock type weights: The 13.th line defines the rock type weight (RWEI). This parameter can have values 0= no weighting, 1= weighting per sub-area, or 2= weighting per lithological unit. Rock type weighting gives additional weight for grid points and works like the lithological weighting. Instead of the proportional area of the unit the abundance of the rock type (the proportion of specific rocks to the total number of samples in the sub-area) is used as a weight. Thus, the main rock type in the sub-area (RWEI=1) or in the separate lithological units (RWEI= 2) is given more weight than others. Note that this parameter is still purely experimental and its use is not recommended.

Duplicate points: The 14.th line defines how to handle duplicate data points. This parameter can have values 0= use all data points, 1= use the mean, 2= use the median of duplicate points, or -1= discard all duplicate data points. Duplicate data points arise, for example, from petrophysical samples obtained from deep drillings. Because most samples are taken from the weathered rocks from the surface, their petrophysical properties can be quite different. Therefore, it may be advantageous to remove duplicate points totally.

Coordinate rounding: The 15.th line defines the rounding of coordinates in terms of units. Value 0.0 means that the coordinates are not rounded at all. Since the coordinates are defined in kilometers value 0.1 would round them to the accuracy of 100 meters. If duplicate data points are replaced by the mean or the median, the rounding could be used to decrease the original amount of data points. This could be advantageous in areas where there are lots of samples available.

Store used petrophysical data: The parameter on 16.th line defines whether the petrophysical data used in the gridding are stored in a separate file (1) or not (0). The name of the file is PETROSIN.DAT. When processing local areas it may be useful to plot the location of data points that were actually used to create the grid and to discard the rest of the values.

Store used lithological polygons: Likewise, the parameter on the 17.th line defines if the lithological units used to create the grid would be stored (1) or not (0). The name of the file is LITHOSIN.BNA and it uses the Atlas BNA file format. These files can be read in as overlay maps in Golden Software's Surfer and BLOXER.

Output mean parameters file: The 18.th line defines the name of the output file where the parameter values of the gridding are stored. This file contains the results from different kind of weighting methods. See the next chapter to find out the file format.

Output std deviations file: The 19.th line defines the name of an output file where the standard errors of the gridded data are stored. See the next chapter to find out the file format.

Output bloxer file: The 20.th line defines the name of an output file for BLOXER program (Pirttijärvi, 2005b). This file is used to import petrophysical data into the top elements of 3-D density models used in gravity modeling software such as GRABLOX (Pirttijärvi, 2009). The exported value corresponds to background weighting with inverse distance weighting (i.e., without basic areal weighting). A dummy column with zeros is added for elevation/height values. An additional column derived from to the search radius can be used to assign weights (fix/free status) for the data.

Number of rejection rules: The 21.st line defines the number of rejection rules (NRR) to be applied during background weighting. The maximum value is 92 rules (one for each unit).

Lito-rock exclusion/inclusion pairs: The following NRR lines (lines 22-30) define the actual rejection rules as pairs of lithological code (LIT= 1-92) and rock type (ROK= 1-6). The last line in the PETROCK.INP file is used as a comment to remind the user how the rock types are defined in the hierarchal classification system used by Geological Survey of Finland. Currently only the first hierarchy level of the rock types is used: 1= felsic, 2= dyke, 3= volcanic, 4= sediment, 5= metamorphic, and 6= other rocks, although in the petrophysical data file the rock types are saved as floating point values (two level hierarchy). More information about the classification is given in a GTK report (Korja, 1989).

The rejection can be done per rock type or per lithological unit. If rock type is defined as a positive value (ROK > 0), all samples of that rock type are rejected from the background estimate of the lithological unit. If the rock type is defined as a negative value (ROK<0), that particular lithological unit is allowed to contain samples from that rock type only (absolute value of the given value, of course). In the example script, all but sedimentary rocks were allowed in (sedimentary) lithological units 1, 2, 3, 4 and 6, and only felsic rocks were allowed in (granitic) units 8 and 9. More over sedimentary and felsic rocks were not allowed in unit 5.

6. File formats

The PETROCK program saves the results (the gridded parameter values and their standard deviations) into two files the names of which are given as the end of the PETROCK.INP input file. In the SUBAREAS.DAT file PETROCK stores information, such as the basic mean, lithologically weighted mean, standard deviation, median, mode, and the number of points inside sub-area about each sub-area. If the user chooses to, the petrophysical data and the lithological polygons actually used (Lines 16 and 17 in PETROCK.INP) are stored in the PETROSIN.DAT and LITHOSIN.BNA files.

6.1 Input petrophysical data

The columns of the petrophysical data file are: X, Y, P, IC, IR, where:

X=	the *x* coordinates (easting) of the sample points
Y=	the *y* coordinates (northing) of the sample points
P=	the value of the petrophysical sample
IC=	the lithological code of the sample (0-92)
IR=	the rock type of the sample point (1-6).

Note that in PETROCK the *x* and *y* coordinates are given in a rectangular coordinate system, where *x* and *y* axes are positive towards East and North, respectively. Geographical coordinates defined as latitude and longitude values must be converted into rectangular coordinates beforehand. Note also that normally distances are defined in kilometers [km].

The rock type is based on the hierarchical classification system of the rock samples used by the Geological Survey of Finland. In this system, plutonic rocks are coded into 85 classes as 1.x.y.z, dyke-like rocks are coded into 36 classes as 2.x.y.z, volcanic rocks are coded into 89

classes as 3.x.y.z, sedimentary rocks are coded into 33 classes as 4.x.y.z, metamorphic rocks are coded into 118 classes as 5.x.y.z, other rocks and ore bodies are coded into 36 classes as 6.x.y.z. For the PETROCK program the rock type is re-defined as a real number (floating value) using only the first two digits of the rock type code. Note that the original character codes of the rock types must have been transformed into real numbers before the data can be used in PETROCK.

To be able to handle duplicate data values correctly the petrophysical data file must be sorted manually according to increasing *x* and *y* coordinates before it can be used in PETROCK.

6.2 Input polygon data

The *.LIT file containing the vertex points of the lithological polygons of the digital geological map. The LIT file has been prepared using the MIF2BNA program. See the documentation of the MIF2BNA program (Pirttijärvi, 2005) for information how to transform the original polygon data from Mapinfo MIF/MID files into LIT format.

The following example illustrates the format of the LIT files:

```
    7922     458000
     139          1     173.8139     80   1
3536.588   7780.130
3537.366   7775.912
3537.539   7773.585
3537.701   7771.419
3538.661   7769.882
3542.303   7764.564
...
3536.587   7780.138
3536.588   7780.130
     257          2     412.8682     81   1
3529.204   7777.384
3530.652   7776.194
...
etc
```

The first line defines two important parameters: a) the number polygons (NOPA=7922), and b) the total number of polygon vertices (NOVA=458000). Note that NOVA is used merely to allocate enough memory for the tables but NOPA must be defined exactly. See chapter 5 for additional information about these parameters.

The second line is the header of the first polygon. It defines the number of vertices (NOV=139), the index number of the polygon (IP=1), the area of the polygon (A=173.8139 km^2, read from the MID file), the code number of the lithological unit (IC=80), and the host status of the polygon (IH=1). The host status defines if the polygon is an actual host polygon (IH= 1) or if it is a hole inside a host polygon (IH=0).

The following (139) lines define the *x* and *y* coordinates of the first polygon. Note that in LIT files the *x* and *y* coordinates refer to West-East and South-North directions, respectively. The last vertex point of the polygon must be equal to the first one, i.e., the polygons must be closed. The line immediately after the last vertex point of the first polygon contains the header of the second polygon, and so on.

6.3 Output files:

The columns of the output file (e.g. PETROCK.DAT) that contains the gridded values are:

X= the *x* coordinate of the center of the grid point
Y= the *y* coordinate of the center of the grid point
P1= the mean of the data points inside the search radius
P2= inverse distance weighted (IDW) mean of the data points inside the search radius
P3= areal lithologically weighted mean
P4= areal lithologically weighted mean + IDW
P5= areal + background lithological weighting
P6= areal + background lithological weighting + IDW
P7= background lithological weighting
P8= background lithological weighting + IDW
P9= the median of the data values inside search radius
P10= background value of the lithological unit below the grid point
LI= lithological code of the unit below the grid point
RD= search radius used for the grid point

Note that grid coordinates start from the SW corner of the study area. The end points of the sub-area are shared with the two adjacent sub-areas. In practice, this means that the eastern and northern sides of the sub-area are processed twice.

In the STDROCK.DAT file the columns of the file of the standard deviations are: X, Y, S1, S2, S3, S4, S5, S6, S7, S8, RD, NU, and NI, where S1-S8 are the standard deviations of the

gridded values corresponding to the weighting methods P1-P8, RD is the search radius actually used, NU is the number of petrophysical sample points used to evaluate the grid point, and NI is the number of sample points ignored due to the additional limiting condition based on the standard deviation.

The SUBAREAS.DAT file contains statistical information about each investigation area (moving window). The columns of the file are:

X=	the *x* coordinate of the center of the sub-area
Y=	the *y* coordinate of the center of the sub-area
ME1=	simple mean
ME2=	lithologically weighted mean (areal weighting)
MED=	median
MOD=	mode
MIN=	minimum value
MAX=	maximum value
SD1=	standard deviation related to ME1
SD2=	standard deviation related to ME2
AD=	average deviation related to ME1
ND=	number of data points inside the sub-area.

The LITOVALS.DAT and LITOERRS.DAT files contain the petrophysical background values and standard deviations per each investigation area. The files are formatted so that the columns are the lithological units. The order of the columns depends on the study area and therefore it is defined in the header of the file.

In the PETROSIN.DAT file the columns are the same as in the original petrophysical data file. The LITHOSIN.BNA file is stored in Atlas BNA file format, which is defined for example in the user manual of the Surfer program by Golden Software. The following example illustrates the BNA format:

```
'Pname','Sname',  -4
    3535.463     7779.234
    3537.069     7775.823
    3537.310     7771.241
    3538.019     7769.422
'Pname','Sname',  -2
    3538.646     7776.931
    3538.229     7777.292
```

The example defines two line segments, which consist of four and two points, respectively. Each line segment has a header line that contains two character variables that identify the objects and a numeric parameter (NPL). The absolute value of NPL defines the number of points of each line segment. A negative value means that the object will be a line, whereas a positive value means that the object will be interpreted as a closed area (polygon, sphere or ellipse). The polygon must be closed so that the first and the last point are the same. For a single point NPL=1 and for an ellipse (and a circle) NPL=2. To learn more about BNA file format see Surfer documentation or the web-document at <http://www.softwright.com/faq/support/boundary_file_bna_format.html>.

7. Example

Figure 4 shows various mapping results from an area around the cities of Hyvinkää and Hämeenlinna in the southern Finland. The size of the study area is 100 km by 100 km. Geologically the area is characterized by high-density gabbros in the center and the low-density granites in the surroundings. The rest are metavolcanic rocks, mica schists, granodiorites, and granitoids (cf. Fig 1). Map (a) shows the geological units (1:5 million scale) and map (b) shows the locations of the 2351 petrophysical samples. Map (c) shows the interpolation results from conventional inverse distance (power of two) interpolation method (Golden Software's Surfer 7). When comparing map (c) with the background geology the inverse distance interpolation reveals two main weaknesses. It brings up too many details from the individual samples and incorrectly interpolates the sparsely sampled area in the south. Nonetheless, it should be noted that in this specific case the inverse distance interpolation was much better than other conventional interpolation methods (kriging, minimum curvature, etc.)

Map (d) shows the mean bulk density obtained using the new mapping method without any geological or inverse distance weighting. Map (e) shows the density map obtained using the basic weighting scheme, where the areal proportions of the lithological units are used as weights. Map (f) was obtained using the alternative lithological weighting, which emphasizes the lithological boundaries. In these examples the grid size was 1×1 km^2, the size of the computational area was 5×5 km^2, and the margins were 25 km wide. The minimum number of sample points per investigation area and per grid point and were 100 and 10, respectively.

Duplicate points were replaced with their median value. Additional limiting condition was used where the reference point was the median and the limits were ±1.0 × the standard deviation of each lithological unit in each investigation area.

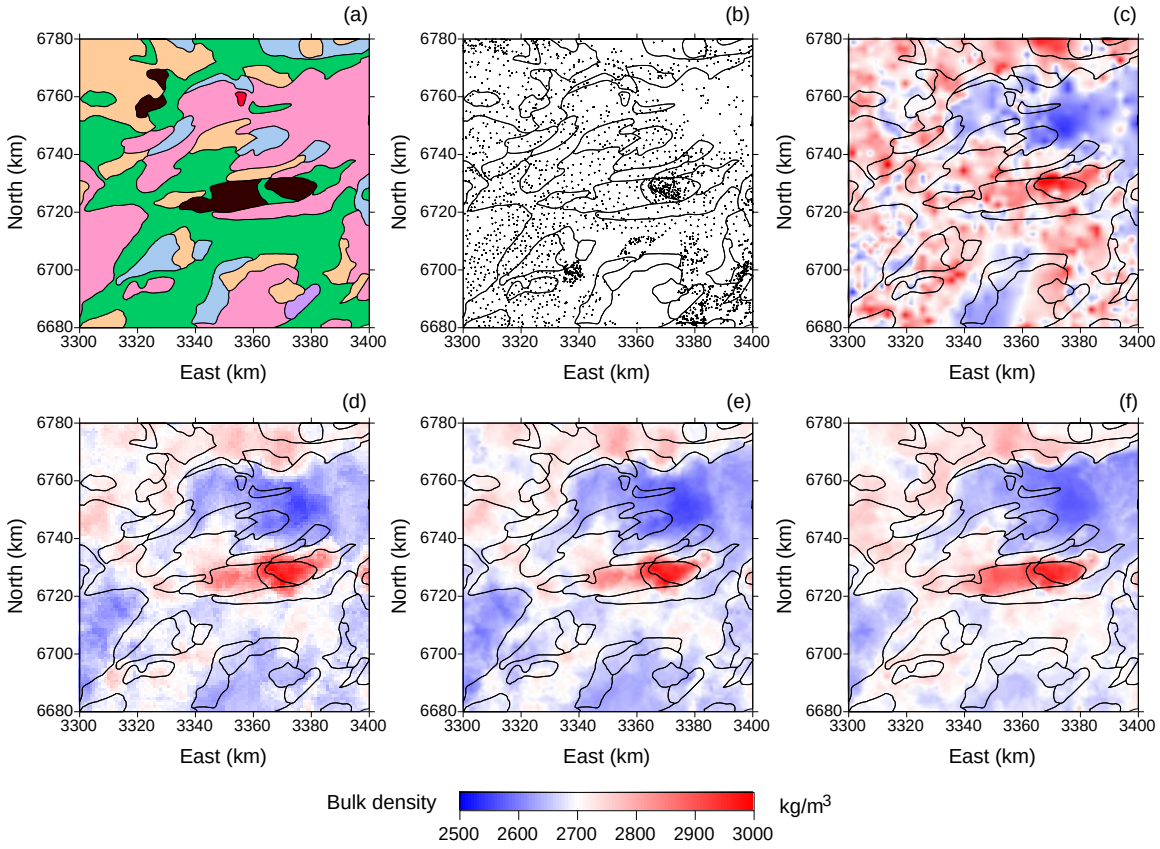

Figure 4. Mapping example: (a) geological units (cf. Fig. 1), (b) sample locations, (c) conventional inverse distance interpolation, (d) mean of the closest points, (e) mean + basic lithological weighting, (f) mean + alternative lithological weighting.

Because at least ten points were used to evaluate each gridded value, maps (d)-(f) show much smoother and flatter variations than map (c). The flattening is also caused by the additional limiting condition, which removes some (outlying) high-density samples. On the other hand, maps (d)-(f) correspond to the underlying geology much better than map (c), which was obtained using conventional interpolation. Even without any geological weighting map (d) shows good correlation with the underlying geology. Because of the additional limiting the new maps (d)-(f) reveal the base level of the more correctly, particularly above the granite areas of low density. The basic lithological weighting used in map (e) delineates the lithological units. The alternative lithological weighting further emphasizes the boundaries of

different geological units. Note that the gridding was made using the more detailed 1:1 million digital map instead of the 1:5 million map shown in Fig 1.

8. Miscellaneous

Note that PETROCK only works with rectangular coordinates (e.g., Finnish National KKJ coordinate system). Geographical coordinates defined as latitude and longitude values must be converted into rectangular coordinates beforehand.

The various input parameters have different kind of an effect on the results. In particular, the number of points per a grid point and the multiplier of the additional limiting condition affect the results. Moreover, different weighting methods give rise to different results and artifacts may appear at some locations. Because the results of the gridding are not unique, it may be necessary to run the program multiple times with different input parameters to get acceptable results.

I wrote the PETROCK program in Spring 2004 at the Geological Survey of Finland, in Espoo as a part of the 3DCM (3-D crustal model) project funded by the Academy of Finland. I added background weighting scheme and rock rejection rules in 2010 when writing a scientific paper on the subject. The program was written for the 32-bit Windows operating system using standard Fortran90 language. The program can be compiled and run on other computer platforms provided that the associated petrophysical data files and geological map files exist.

9. References

Korhonen, J.V., Säävuori, H. and Kivekäs, L., 1997. Petrophysics in the crustal model program of Finland. In: Autio S. (Ed.) Geological Survey of Finland, Current research 1995-1996. Geological Survey of Finland, Special Paper 23, 157-173.

Korja, A.,1989. The new rock classification of the petrophysical database and planning of concurrent classifications. Report Q11/27.0/89/1, Geological Survey of Finland, 47pp.

Korsman, K., Koistinen, T., Kohonen, J., Wennerström, M, Ekdahl, E., Honkamo, M, Idman H. & Pekkala Y (Eds.) 1997. Suomen kallioperäkartta -Berggrundskarta över Finland - Bedrock map of Finland 1: 1 000 000. Geological Survey of Finland, Special Maps, 37.

Pirttijärvi, M. 2005a. MIF2BNA - Conversion of polygon data of Mapinfo MIF/MID files into LIT file format. User manual. Report Q16.2/2005/1, Geological Survey of Finland.

Pirttijärvi, M., 2005b. BLOXER: visualization and maintenance of 3-D block models, version 1.5a user's guide. University of Oulu, Department of Physics. 38 pp. <https://wiki.oulu.fi/display/~mpi@oulu.fi/Block+model+maintenance >

Pirttijärvi, M., 2009. GRABLOX2: 3D modelling and interpretation of gravity and gravity gradient data, version 2.0 user's guide. University of Oulu, Department of Physics. 62 pp. < https://wiki.oulu.fi/display/~mpi@oulu.fi/Gravity+inversion+using+block+model+2>

Appendix

POLYTUNE – Fortran90 algorithm for tuning the vertex points of polygon data.

```
!-----------------------------------------------------------------------
! POLYTUNE  a subroutine that divides and marks the points of a polygon
! into important and unimportant (unnecessary) ones based on the maximum
! distance and maximum opening angle between the points.
!
! Language: Fortran-90 (Digital Visual Fortran 6.6)
!
! Input:
!
!   XP=  real array (dim=MP), x coordinates of the polygon vertices
!   YP=  real array (dim=MP), y coordinates of the polygon vertices
!   NP=  int*4, actual number of polgon vertices
!   RMA= real*4, the maximum distance between two adjacent points
!   ALF= real*4, the maximum angle the point it can have with respect to
!        the two surrounding points
!
! Output:
!
!   IP=  int*4 array (dim=MP), with  1= important & 0= unimportant point
!   N=   int*4, the total number of important points (just in case if needed)
!
! Note: the polygon must be closed so that the last point is same as the
! first point, i.e.m XP(NP)= XP(1) and YP(NP)= YP(1), this means also that
! the actual (unique) number of vertices is NP-1
!
! Note: small polygons that have less than MNP points are not processed
! Note: if angle ALF=0, then the opening angle condition is not used
!-----------------------------------------------------------------------
! M. Pirttijärvi, 2003-2004

  subroutine polytune(xp,yp,np,rma,alf,ip,n)

    implicit none
    integer, parameter :: mnp= 5
    integer :: i,i1,i2,np,n
    integer, dimension(np) :: ip
    real :: rma,alf,x0,y0,rr,ang1,ang2
    real, dimension(np) :: xp,yp

    ip(1:np)= 1
    if (np > mnp) then
      x0= xp(1)
```

```
    y0= yp(1)
    do i= 2,np-1
      rr= rr+(xp(i)-x0)**2+(yp(i)-y0)**2
      ip(i)= 0
      if (rr > rma) then
        ip(i)= 1
        x0= xp(i)
        y0= yp(i)
        rr= 0.
      else if (alf /= 0) then
        i1= i-1
        i2= i+1
        ang1= atan2(yp(i)-yp(i1),xp(i)-xp(i1))
        ang2= atan2(yp(i2)-yp(i),xp(i2)-xp(i))
        if (abs(ang1-ang2) > alf) then
          ip(i)= 1
          x0= xp(i)
          y0= yp(i)
          rr= 0.
        end if
      end if
    end do
  end if
  n= sum(ip,np)

  return
end subroutine polytune
```

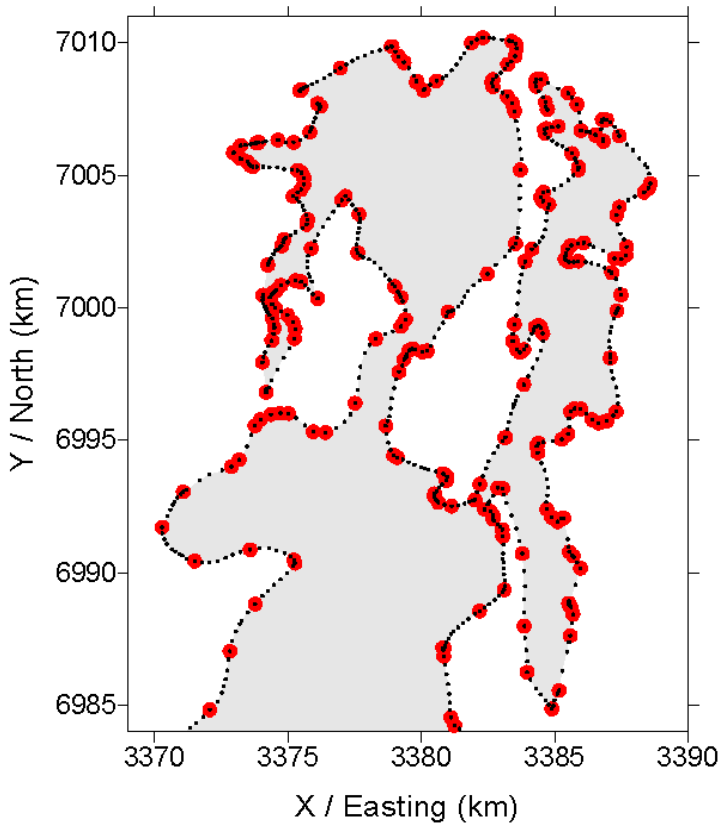

Figure A.1. Example of the tuning of polygon vertices (a detail of an actual polygon). The small black dots are the original points and the large red dots are the optimized ones. For this polygon (n:o 3317) the number of vertex points was reduced from 2633 to 570.