SPATIAL ANALYSIS OF UNCERTAIN THERMOBAROMETRIC DATA: APPLICATION TO THE SWISS CENTRAL ALPS

Tivadar M Tóth[1,2]

[1] Department of Mineralogy, Geochemistry and Petrology, University of Szeged, Hungary (mtoth@geo.u-szeged.hu)

Abstract

Several approaches exist to estimate peak pressure and temperature (PT) conditions of a single metamorphic rock sample. Because of many reasons from analytical problems to uncertainty in mineral solution models, all these calculations are rather uncertain making spatial interpretation of the data set for a whole metamorphic terrain problematic. In this study Error Kriging (de Marsily, 1984) is used to calculate PT maps for the Swiss Central Alps and the results are compared to those got by other kriging methods.

Keywords: thermobarometry, uncertainty, error kriging, Central Alps.

1. INTRODUCTION

When one seeks to interpret the results of thermobarometric studies, the spatial context is obviously important. Graphical representations, e.g. isotherms and isobars on maps or profiles, are helpful in interpreting thermobarometric data sets in a regional context. For geologists to assess the significance and implications of P-T data, it may indeed be crucial to see them together with results from geophysical, tectonic, or geochronological studies. Similarly, in modeling studies it may be most useful to compare the thermal or baric field of an orogen with predictions based on theoretical simulations. For all these reasons, we need reliable graphical representations of spatially discrete P-T data sets. As these data sets may include various forms of information, commonly containing quite variable uncertainties, the task of interpolating and extrapolating such data sets demands adequate tools.

2. PT data

Several methods exist to estimate P and T data for a given metamorphic assemblage, either from traditional Fe-Mg exchange thermometers and net-transfer type barometers (such as GASP), or from more reliable multi-equilibria techniques (Berman 1991, Gordon 1992). A

number of authors (e.g. Kohn, Spear 1991, Lieberman, Petrakakis 1991, Gordon 1992, Holland, Powell 1994) have investigated the uncertainty of various PT-estimation methods, which are due to numerous sources of error, like electron microprobe calibration, counting statistics during measurement, mineral formula calculation, mineral solution models among many others. Evaluation of these individual uncertainties may allow their propagation into a finite statistical error of P and T data. Such results may then be quoted as „crisp" PT data (e.g. 500±25°C, 2.8±0.6 kbar). Other thermobarometric approaches rely on phase diagrams (such as petrogenetic grids calculated using DOMINO/THERIAK software) and produce a "permitted" PT-window for each assemblage (e.g. 500-540°C). In this case no statistical error value is calculated; uncertainty appears as an inequality type of datum, i.e. an interval in T or P.
A different way to present uncertainty is by means of fuzzy numbers (e.g. Dubois, Prade, 2000). Fuzzy numbers are defined through "membership functions". The value of a membership function (m(*x*) in the [0,1] interval) depends on how *possible* (not how mathematically probable!) the datum (*x*) is. Defining the fuzzy numbers is usually based on qualitative or semiquantitative information. For example, if temperature is with known (without any doubt) to lie between 450 and 600 °C, and is possibly in the 500-550 °C interval, the proper fuzzy number would be a trapezoid:

$$m(T) = \begin{cases} 0, & \text{if } T < 450\ °C \\ (T-450)/50, & \text{if } 450\ °C < T < 500\ °C \\ 1, & \text{if } 500\ °C < T < 550\ °C \\ (600-T)/50, & \text{if } 550\ °C < T < 600\ °C \\ 0, & \text{if } T > 600\ °C \end{cases}$$

There are two reasons to prefer fuzzy numbers to standard error. Firstly, there is in most cases no proof that the uncertainty in P- and T-data are of probability type (e.g. Gaussian). Secondly, both error and inequality type data are easily transformed to fuzzy data, which thus offer a way to construct a data set having a uniform measure of uncertainty.
Although, there are different ways to define and calculate uncertainty in data, their spatial representation is more problematic; interpolating between data that have different uncertainty is not a simple matter. In what follows, different interpolation methods are presented and tested, which are considered promising for mapping P- and T-data.

3. Interpolation methods applied

Thermobarometry yields P and T data in spatially discrete form, hence for many questions spatial interpolation (or limited extrapolation) is required, most notably in the production of maps and profiles. Traditional interpolation techniques, such as linear interpolation (by hand or machine), trend surface analysis or inverse distance interpolation have long been found useful, yet they all have significant disadvantages. Inverse distance method, for example, tends to generate unrealistic „bull's-eye" shaped structures surrounding the position of data points. None of them take into account the real spatial structure of the data set, and none of them allow an estimation of interpolation error.

Kriging, a family of stochastic interpolation methods, has a fundamental role in geostatistics for decades. Various texts and handbooks go deep into its basic concept and are not discussed here (e.g. Matheron, 1970, Cressie, 1991), Wackernagel, 1995). In what follows ordinary kriging system (OK) with and without nugget effect will be used.

4. Kriging with uncertain data

Ordinary kriging is a good interpolator in cases where reliable data exist in sufficient number (and spatial density). However, geostatistics has spread into several areas including hydrology and soil sciences, where the conditions and requirements for OK are not always satisfied. A common problem is the insufficient quantity or quality of measurements. In order to get a useful interpolator in such cases, a number of attempts have been made to incorporate additional information into the kriging system. Bárdossy et al. (1988) present an exhaustive collection of these approaches. Some of these shall be discussed here in the context of thermobarometry, the goal being to represent regional results as maps of continuous isotherms and isobars.

In the case of mapping metamorphic P and T, difficulties arise because the data are of variable type and precision, and their spatial distribution tends to be far from uniform. Methods used to estimate statistical errors in thermobarometric data have received attention in recent years (e.g. Kohn, Spear 1991, Lieberman, Petrakakis 1991, Gordon 1992, Holland, Powell 1994). Where the data set suggests that other types of uncertainty should be incorporated in the kriging system, this is possible by two approaches called *soft kriging* (Journel, 1986) and *fuzzy kriging*. A soft kriging system works with inequality type data and/or constraint intervals. This procedure may be useful in PT mapping, where the mineral assemblages or petrogenetic net model allow only an estimation of P-T intervals (e.g. T_{min} and T_{max}). Both error and interval type uncertainties can easily be transformed to fuzzy numbers, making it possible to use data sets with mixed types of information. However, at the present stage, the fuzzy kriging estimator tends to use either only fuzzy or only crisp data for interpolating, depending on the initial conditions (Bárdossy et al., 1990a, b). For fuzzy kriging to yield reliable results, considerable computational effort is needed.

Error kriging [1] (EK) follows very simply from OK: One uses an error ε_i associated with each datum $Z(x_i)$, with the following constraints (Marsily, 1984):

- $E[\varepsilon_i] = 0, i = 1 ... n$ — ε_i is not a systematic error. (E is the mean);
- $Cov[\varepsilon_i, \varepsilon_j] = 0, \forall\, i \neq j$ — errors (ε) are not correlated with each other;
- $Cov[\varepsilon_i, Z(x_j)] = 0, \forall i, \forall j$ — errors (ε) are not correlated with data;
- σ_i^2 is known for each i.

Compared to OK, the only difference in these conditions is that the kriging system now has values of $-\sigma_i^2$ (instead of zero) in the diagonal elements of the error matrix. As a consequence, the estimator uses normal distributions with a given mean and variance for interpolation. It is probably fair to say that only the best thermobarometric data sets may come close to satisfying the above four conditions reasonably well. For example, even if multi-equilibrium techniques are used, a (minor) systematic error cannot be ruled out. Additionally, at low temperature

[1] Following a suggestion by A. Bárdossy (pers. comm.), the term "kriging with uncertain data" used by Marsily (1984) has been changed to "error kriging".

conditions T estimation becomes less accurate, hence there is likely a slight correlation between T-data and their error. However, for good thermobarometric data sets, estimation errors should come close to satisfying all four constraints fairly closely, in which case EK is the method of choice to interpolate P-T data.
The use of OK and EK kriging to generate PT maps is presented here for the example of the Swiss Central Alps. By way of example, the main steps required in any type of kriging analysis and the graphical representation of the results are documented first. In the subsequent section the results are tested by cross validation (e.g. Wackernagel, 1995), and the maps obtained by these two different approaches are compared.

5. A case study: P, T maps of the Swiss Central Alps

120 T-data and 97 P-data for the Swiss Central Alps have been selected for constructing isotherms and isobars. For a description of the geology, methods, and data sets see Engi et al. (1995) and Todd, Engi (1997). Error data in P and T were calculated using program INERSX, if at least three independent reactions existed. For all other data points T_{err} was assumed as 50°, and P_{err} = 2 kbar.

5.1. Sequence of spatial data analysis

The spatial distribution of data points in both cases (P and T) is rather uneven. In relatively large areas, geological conditions are unfavorable, and no data points exist, whereas clusters of data exist in valleys with suitable outcrops, where detailed and multiple sampling was possible. Clustered distributions of data points usually lead to an estimation error in variography (Armstrong, 1984). In order to avoid this problem, a moving window declustering method was used prior to the calculation of variograms. This process substitutes data that fall into a given rectangle (spatial window) by their arithmetic mean. Moving this window and calculating the means over the study area result in a data distribution of equal density. The size of the windows is chosen on the basis of the average spacing between locations and the size of the entire area being studied. If the window is too large, the number of points left after the process insufficient. If it is chosen too small, no reliable statistics can be obtained for most windows. For the area studied in the Central Alps (~4000 km^2), overlapping windows 5*5 km in size were found satisfactory for both T and P. For moving window statistic, calculations the program MWINDOW (Murray, Baker, 1991) was used. Variograms were calculated based on the cluster means rather then the original database. Both experimental variogram calculations and theoretical variogram fitting were performed using the program VARIOWIN (Pannatier, 1994).

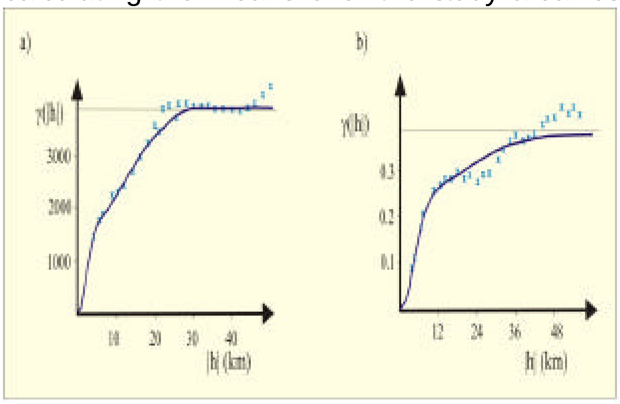

Fig.1 Estimated variograms for a) temperature, b) pressure for the Swiss Central Alps. For best-fit models see text.

On the experimental variogram that characterizes the spatial variation in temperature data, two different sills can be distinguished. The nested structure variogram (Serra, 1968) fitted to it is the sum of two Gaussian type variograms. One of them has small (5 km) range, the other significantly larger (29.5 km) range: $\gamma(h)=1520*G(5)+2520*G(29.5)$, where "G" denotes the Gaussian variogram function (***Fig. 1/a***). A sum of two Gaussian variograms was also found best for pressure (***Fig.1/b.***) ($\gamma(h)=0.234*G(10.2)+0.148*G(43.6)$), but the range of both individual variograms for P is significantly higher than in the case of T. The fit for P is weaker than for T, probably due to a hole effect (Cressie, 1991), shown by a negative correlation in a small range of the variogram (***Fig. 1/b***). Anisotropy was not calculated for either case.

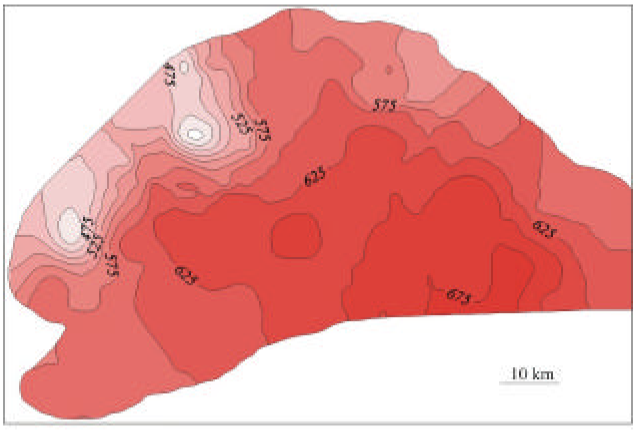

Fig.2/a Calculated PT maps for the Swiss Central Alps using various interpolation methods. EK map for temperature (T in °C)

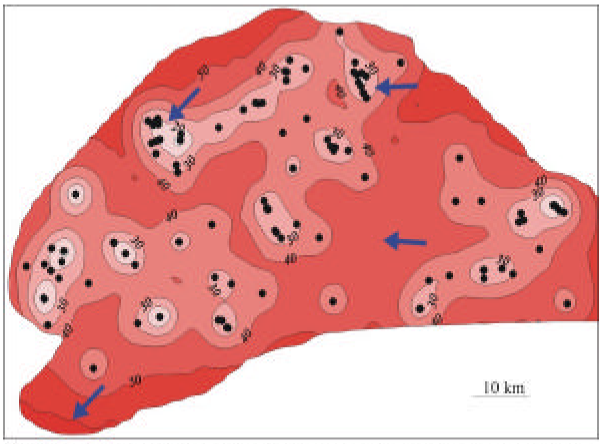

Fig.2/b Calculated PT maps for the Swiss Central Alps using various interpolation methods Kriging variance for the EK map of temperature (T in °C; for arrows see text)

Finally, theoretical variogram parameters were combined with the original data for OK and EK in a program written by Bárdossy, A. (unpublished). This code allows both data estimation and kriging error calculation to be performed simultaneously. Each kriging parameter was chosen to be the same in the two processes, but for OK the estimated errors were set zero. Contour lines were then generated using program SURFER (Golden Software Inc., 1994) (***Fig. 2/a-g***).

5.2. Results and Comparison of Different Methods

EK error maps for both P and T exhibit low kriging standard deviations (***Fig. 2/a-d***). In both cases the interpolation is best where data points form clusters, while the error is largest in areas for which no data exist (see arrows on ***Figs 2/b, 2/d***). Although the structure of the two error maps is similar, the dimension of low error areas is significantly smaller for pressure (***Fig. 2/d)***. In addition, data clusters with a large error exist, indicating that the uncertainties in metamorphic pressure estimates are larger than in T.

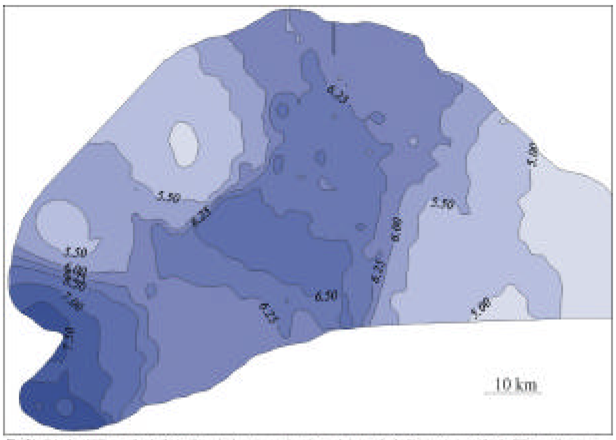

Fig.2/c Calculated PT maps for the Swiss Central Alps using various interpolation methods. Kriging variance for the EK map of pressure (P in kbar) for arrows see text

OK maps for both T and P exhibit unrealistic results, with peaks and ditches forming in many areas. Extreme values in P – up to 20 kbar or as low as 0 kbar – were locally interpolated (***Fig. 2/e***). Estimated temperature data vary

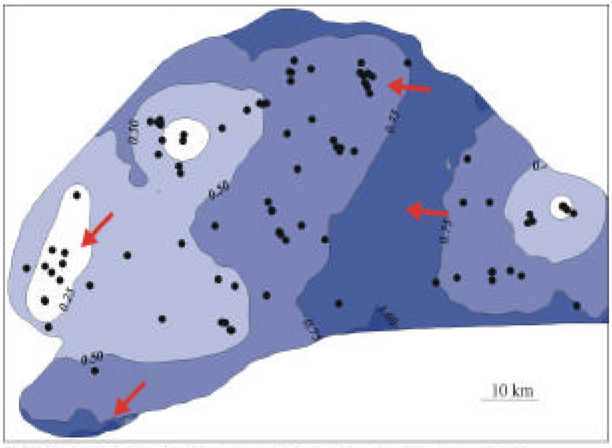

Fig.2/d Calculated PT maps for the Swiss Central Alps using various interpolation methods. EK map for pressure (P in kbar)

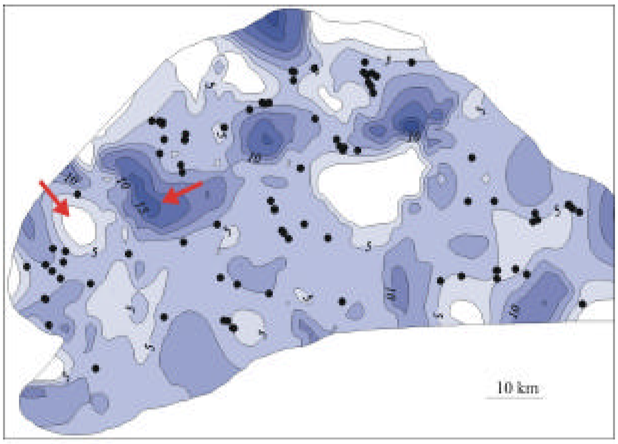

Fig.2/e Calculated PT maps for the Swiss Central Alps using various interpolation methods. OK map for temperature (T in °C)

between 0 and 1700 °C (not presented). The results emphasize the problem of using exact interpolation methods for data sets that contain uncertainty. The reason of the conspicuous (mis)estimation is likely that significantly different values, both in the initial P and T data sets, exist close to each other. Because no nugget effect was involved in the variogram used, OK is an exact interpolator, i.e. the estimated surface tends to pass through each datum point. This ambition leads to bad interpolation results in areas where significantly different values exist. If OK with a nugget effect model (smoothing interpolator) is used, a more realistic result may be obtained. By using this variogram for kriging, one can generate smoothed maps even in areas having higher uncertainty in the data (***Fig. 2/f***). The error of estimation using OK is significantly larger than for EK in areas where data points are numerous (***Fig. 2/g***). On the other hand, for extrapolation OK appears to be more reliable, as shown by the small error values towards the border of the study area as well as in domains lacking data.

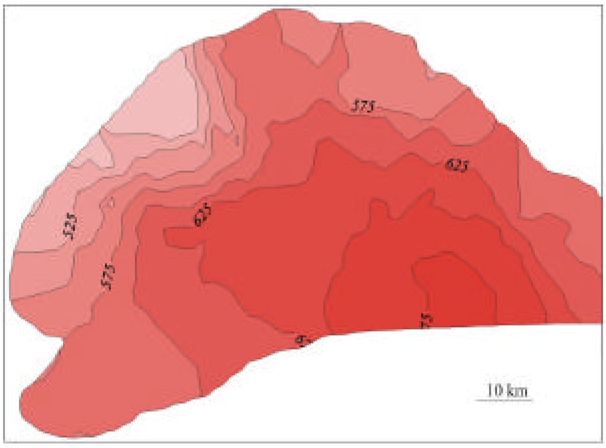

Fig.2/f Calculated PT maps for the Swiss Central Alps using various interpolation methods. OK map (with nugget effect) for temperature (T in °C)

The reliability of two maps calculated by EK and OK with nugget effect, respectively, can also be tested by cross a validation procedure. In this process each sample value at location x_0 is removed in turn from the data set and $Z^*(x_0)$ is estimated using the other samples. Comparison of the measured and the estimated data helps to identify apparent bias. For EK and OK the average of the T differences is as low as 2.5 and 0.8 °C, respectively, no systematic overestimation is seen in either case. Correlation coefficients

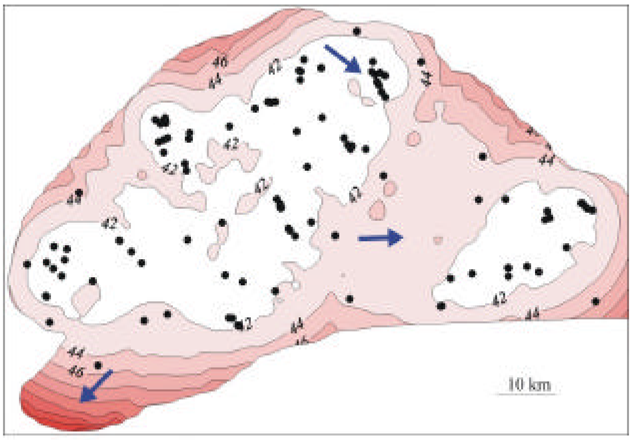

Fig.2/g Calculated PT maps for the Swiss Central Alps using various interpolation methods. Kriging variance for the OK map (with nugget effect) of temperature (T in °C); for arrows see text.

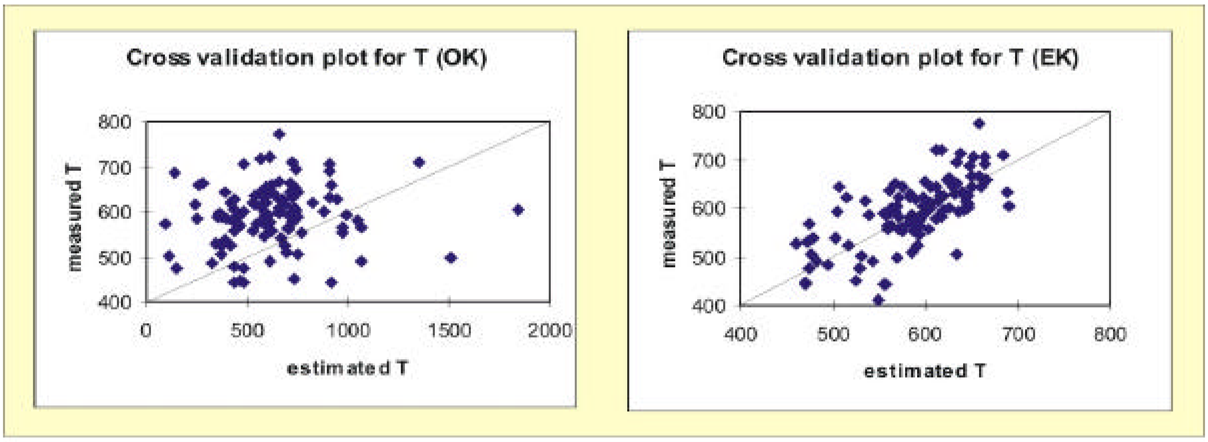

Fig.3 Cross validation plots for T (°C) calculated by a) EK and b) OK respectively. For details see text.

between estimated and measured data are similar in the two cases (r=0.67 for both), and the two estimations produce similar data (r=0.99). On ***Fig. 4*** one can see that EK reproduce the original data well, while the measured and the estimated T data are barely correlated (r=0.14) in the case of the OK map (without nugget effect). Whether the slight underestimation for high T data and the tendency to overestimate low T data (***Fig. 3***) for both EK and OK is characteristic of these interpolation methods, or whether these effects are inherited from the data set used is out of the scope of the present study.

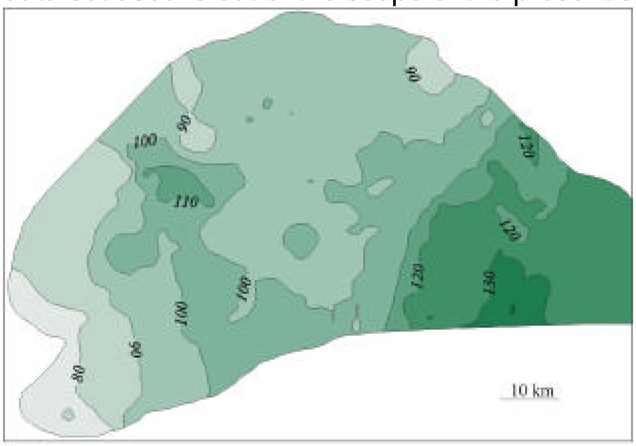

Fig.4 Metamorphic T/P (°C/kbar) map for the Swiss Central Alps

Based on the previous reasoning one can say that reliable maps can only be obtained if the uncertainty in the data is involved in the interpolation algorithm. The two approaches used (EK, OK) have different advantages and disadvantages. If a sufficient number of data points exist, with a spatial distribution close to uniform, EK is preferred because this method produces the smallest kriging error. For highly clustered data sets, as well as to get reliable maps close to the border of the area, OK is the superior method.

5.3. Geological representation of kriging results

For possible application of the PT maps in geological interpretation two examples are briefly presented. In both cases maps got by EK method are used.

1) The T/P ratio may be an informative parameter when comparing different metamorphic terrains. Based on isotherms and isobars, the construction of a T/P map is a simple calculation (***Fig. 4***). To propagate errors into the T/P map, the following expression should be used:

$$s_{T/P} = \sqrt{\sum_{P} \sum_{T} \left(\frac{\partial (T/P)}{\partial P} \right) * \left(\frac{\partial (T/P)}{\partial T} \right) * s_P * s_T * r_{PT}}$$

where r_{PT} is the correlation coefficient between P and T.

2) The exact position of the sillimanite$_{in}$ isograd in the Central Alps has been argued for a long time. By taking into account the experimentally determined P-T-location (e.g. Holdaway, 1971) of the kyanite-sillimanite reaction, the appropriate combination of isotherms and isobars results in the representation of the univariant curve in the real space. If the error maps are considered as well (by using a function equivalent to the one above), one can represent the stability limit for sillimanite with an error envelope (***Fig. 5***). This confidence interval may then be compared to the mineral zone boundary as delimited by field observation (e.g. Irouschek 1982, Todd, Engi 1997).

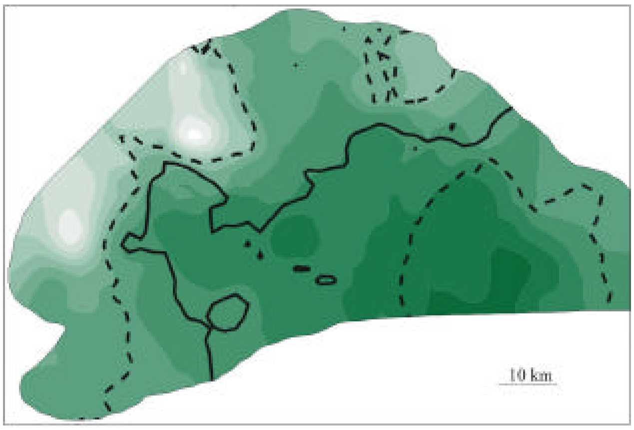

Fig.5 Calculated sillimanite-in isograd for the Swiss Central Alps. Dashed line shows the 90% probability envelope

Acknowledgments

Discussions with A. Bárdossy on methods in spatial statistics, and the availability of his computer code for kriging have helped this project. His interest and contributions are gratefully acknowledged. M. Engi and C. S. Todd are thanked for their valuable remarks and for cleaning the English. The Bolyai Research Grant supported the project.

References

Armstrong, M. (1984): Common problems seen in variography. Mathematical Geology, **16/3**, 305-313.

Bárdossy, A., Bogárdi, I., Kelly, W. E. (1988): Imprecise (fuzzy) information in geostatistics. Mathematical Geology, **20**, 287- 311.

Bárdossy, A., Bogárdi, I., Kelly, W. E. (1990): Kriging with imprecise (fuzzy) variograms. I. Theory. Mathematical Geology, **22**, 63-79.

Bárdossy, A., Bogárdi, I., Kelly, W. E. (1990): Kriging with imprecise (fuzzy) variograms. II. Application. Mathematical Geology, **22**, 81-95.

Berman, R. G. (1991): Thermobarometry using multi-equilibrium calculations: a new technique, with petrological applications. Canadian Mineralogist, **29**, 833-855.

Cressie, N. A. (1991): Statistics for spatial data. Wiley & Sons, New York, 900 p.

Dubois, D., Prade, H. (2000): Fundamentals of fuzzy sets. Kluwer Academic Publishers

Engi, M., Todd, C. S., Schmatz, D. R. (1995): Tertiary metamorphic conditions in the eastern Lepontine Alps. Schweizerische Mineralogisches Petrographisches Mitteulungen, **75**, 347-369.

Golden Software Inc. (1994): SURFER for Windows User's guide. Golden Software Inc., Golden, Colorado, variously paged

Gordon T. M. (1992): Generalized thermobarometry: Solution of the inverse chemical equilibrium problem using data for individual species. Geochimica Cosmochimica Acta, **56**, 1793-1800.

Holland T. J. B., Powell R. (1994): Optimal geothermometry and geobarometry. American Mineralogist, **79**, 120-133.

Irouschek, A. (1983): Mineralogie und Petrographie von Metapeliten der Simano-Decke unter besonderer Berücksichtigung cordieritführender Gesteine zwischen Alpe Sponda und Biasca. Unpublished PhD Dissertation, University of Basel, Switzerland.

Isaaks, E. H., Srivastava, R. M. (1989): Applied geostatistics. Oxford Univ. Press, New York, 544 p.

Journel, A. G. (1986): Constrained interpolation and qualitative information - The soft kriging approach. Mathematical Geology, **18/3**, 269-286.

Kohn M. J., Spear F. S. (1991): Error propagation for barometers: 2. Application to rocks. American Mineralogist, **76**, 138-147.