Geophysica (1998), 34(1-2), 23-37

Removal of Ambiguities in Meteorological Fields Through Stochastic Minimization of the Spatial Variability

Alexander Jann

Zentralanstalt für Meteorologie und Geodynamik (ZAMG)

Hohe Warte 38, A-1190 Vienna, Austria

(Received: January 1997; Accepted: July 1997)

Abstract

Sometimes there are a finite number of options for the value of a meteorological variable at a grid point. When no physical selection criterion is available, values are likely to be chosen in such a manner that the spatial variability of the resulting field is as small as possible. An algorithm based on simulated annealing is constructed to accomplish the optimization of smoothness in case of a large number of ambiguities.

Two concrete applications are discussed, where the first one is the construction of an equivalent-potential temperature surface. There exists a rule of thumb to position the surface at the greatest possible altitude, and a couple of arguments are given why the surface resulting therefrom should be close to the smoothest possible. The simulated annealing procedure by and large confirms this presumption, but does in special situations show a distinct capability to yield considerably better results. This potential benefit of simulated annealing compared to less sophisticated approaches becomes even more apparent for the second application which deals with the selection of Meteosat infra-red cloud motion vectors from several vector fields achieved through varying the size of the targets to be tracked. Despite a certain slowness of convergence towards a final solution, simulated annealing proves to be a promising approach to the considered type of meteorological optimization problems.

The cloud motion vector experiments represent the first step of an unconventional attempt of cloud/no-cloud discrimination which is based on the hypothesis that tracking of features in infra-red imagery should yield results of different character for cloudy resp. cloud free regions.

Key words: Simulated annealing, maximum smoothness of a meteorological field, equivalent-potential temperature surfaces, cloud motion vectors

1. Introduction

It is a frequent problem in meteorology to assign geometric height or a pressure value p to a specific temperature value using a given temperature profile. Since the diverse temperatures are not necessarily strictly monotone functions of altitude, ambiguities in the resulting height may occur. A particular problem to be dealt with in the following is the construction of an equivalent-potential temperature (Θ_e) surface. Analyses on such surfaces have appeared in meteorological literature on several occasions, apparently always in connection with relative stream analysis (e.g. *Carlson*,

1980; *Bader et al*., 1995). In order to select one level out of a number of different possibilities, a rule of thumb is in use at the Zentralanstalt für Meteorologie und Geodynamik (ZAMG) (e.g. for the analyses in *Mahringer* and *Zwatz-Meise*, 1993): the greatest possible height is chosen (henceforth referred to as GPH rule). This is one of the only two reasonable options to draw the Θ_e=312 K isoline in Fig. 1. There, one would be inclined to choose either the lower levels (with the isoline disappearing below the bottom level at the right end) or the higher alternatives throughout. Any mixing would appear odd to the meteorologist albeit there is in principle little physical support to such a point of view. The subjective reluctance to draw surfaces with larger oscillations than necessary may finally lead to the definition of a criterion such as

$$\iint\limits_{\text{Area}} \left|\nabla^2 p\right| \,\mathrm{d}x\,\mathrm{d}y \rightarrow \min! \tag{1}$$

in order to establish an objective selection of height (see the final section for a motivation of the use of the norm in this smoothness criterion). The minimization problem is in general of discrete nature as there exist only a few (typically not more than five) levels p with a specified Θ_e at any individual grid point. Moist-neutral stratification with its infinite number of possible levels for some value of Θ_e has not been an issue of great concern here since the scope of the investigation allowed the restriction to equivalent-potential temperature surfaces without such an extra complication. Where this proceeding is not suitable, the applicability of algorithms used below can easily be preserved through high-resolution discretization of the range of admissible pressure values.

Fig. 1. Illustration of the philosophy behind criterion (1). Shown is a selected meridional section of Θ_e data taken from a limited area model; crosses mark heights with an equivalent-potential temperature value of 312 K. Solid lines indicate the two locations of a 312 K-surface which are subjectively acceptable. One could in principle arbitrarily connect the values as with the dashed line but this would be rejected as unjustifiably oscillatory.

Finding the minimum of (1) for limited area model (LAM) data through an exhaustive combinatorial search is inhibited by the large number of configurations to be checked. Examples comprising more than 3000 grid points with ambiguities were investigated during this study. In these cases, the discrete analogon of the integral in (1) for a regular grid point field,

$$\sum_{i=1}^{k}\sum_{j=1}^{l}\left|4p_{i,j}-(p_{i-1,j}+p_{i+1,j}+p_{i,j-1}+p_{i,j+1})\right|, \tag{2}$$

would have to be computed for a number of at least 2^{3000} configurations. For such large combinatorial minimization problems, the sophisticated method of simulated annealing (hereafter: SA) is available. Previous applications of SA to geophysical problems are described e.g. in papers by *Rothman* (1986), *Billings* (1994), and *Dittmer* and *Szymanski* (1995). More popular discussions on SA were given by *Kirkpatrick et al.* (1983) and *Press et al.* (1992), tracing the SA algorithm back to the article of *Metropolis et al.* (1953). In passing, the related method of genetic algorithms should be mentioned which also found its way into geophysical literature (e.g. *Wilson* and *Vasudevan*, 1991; *Sambridge* and *Drijkoningen*, 1992; *Tarvainen* and *Tiira*, 1995). Belonging formally to the same class of algorithms, there is moreover the Monte Carlo approach which, however, cannot compete with SA or genetic algorithms in terms of efficiency (*Sambridge* and *Drijkoningen*, 1992). Only SA was investigated in the present study. Details on the algorithm used here to solve problem (1) are given in section 2, and application to actual data is dealt with in section 3.

Section 4 covers an application of essentially the same algorithm to the production of cloud motion vector (CMV) fields. By tracking rectangular areas (called the targets) in successive Meteosat images using a cross-correlation method, displacement vectors can be computed for the three Meteosat spectral channels infrared (IR), water vapour and visible (*Schmetz et al.*, 1993). The raw IR product frequently shows some undesired peculiarities caused by too rapid variations of the tracked features. In those instances where a displaced air mass exhibits a quite different cloud pattern one image - i.e. half an hour - later, remote areas may suddenly look more similar to the rectangular target from the earlier image. The displacement with the highest cross-correlation then becomes an incorrect estimate of the true displacement. It turns out that problems in certain areas can be alleviated by altering the size of the target to be tracked. However, a clear guideline how to find the optimum target size cannot be provided. Smaller targets may be affected strongly by mesoscale cloud pattern changes, but there is also an upper limit to the target size since with too large targets, significant mesoscale features of the stream field become blurred. An idea to use „optimum“ IR CMV fields for cloud/no-cloud discrimination is outlined below, and there too large targets also cause undesired effects. Apparently, the ideal target size varies between different parts of the image resp. for stream configurations of different

scales. In order to achieve a field with the optimum number of reasonable IR vectors, it is perhaps most appropriate to try several target sizes and select from the resulting vectors those which, when assembled, form a vector field which satisfies some prescribed demands. Choosing maximum smoothness as criterion, the problem becomes one of the same type as the Θ_e-surface construction. Selection of vectors from different CMV fields is, in fact, already subject of an operational application: The cloud motion vector fields issued by EUMETSAT comprise vectors from the IR, water vapour and visible channel where for any considered segment of the image, an individual decision is made which of the three types of vectors is transmitted. The method used therefor only faintly resembles the one described in the following: Each vector is checked for its consistency with adjacent cloud motion vectors of the same type resp. with forecast wind fields, and the vector with the best mark is selected.

2. *Identification of the smoothest Θ_e-surface: The SA algorithm*

Simulated annealing finds optimum (or at least near-optimum) configurations through stochastic variations of the unknown parameters. *Dittmer* and *Szymanski* (1995) list four specifications which are necessary to define an SA algorithm. Those are:

1) *A set of parameters to alter.* Here, these are the pressure values being assignable to a specified Θ_e at the individual grid points. The values of $p(\Theta_e)$ have been achieved through linear vertical interpolation of gridded model data available at several pressure levels.
2) *The cost function E(p).* The discretized Laplacian in expression (2) describes, in fact, a comparison between a considered value $p_{i,j}$ and an average of its closest four neighbours. In order to be able to handle grid points at the boundary of the domain, it is straightforward and consistent to define $E(p)$ for a $k \times l$-element field formally as

$$E(p) = \sum_{i=1}^{k} \sum_{j=1}^{l} \left| p_{i,j} - \overline{p_N} \right|, \tag{3}$$

where p_N stands for the sub-set of $\{ p_{i-1,j},\ p_{i+1,j},\ p_{i,j-1},\ p_{i,j+1} \}$ comprising the actually existing neighbours.

3) *A method of randomly altering the parameters.* The indices of the grid points with ambiguities are ordered randomly which yields the sequence in which the points are visited. The procedure is repeated with a new random sequence after a run is completed (i.e. after each relevant grid point was visited once). When arriving at a grid point (i, j), the cost function E is evaluated for the m possible pressure values $\pi_{i,j}^{(n)}$, n=1,...,m; the individual cost functions be denoted as $E(\pi_{i,j}^{(n)})$. These are fed into the so-called heat-bath method which is more economic than the early Metropolis algorithm when - like in our case - there are a large number of unknown

variables (according to *Billings*, 1994). The heat-bath scheme establishes a probability distribution through

$$P(p_{i,j} = \pi_{i,j}^{(n)}) = \frac{\exp\left[-E(\pi_{i,j}^{(n)}) / T\right]}{\sum_{q=1}^{m} \exp\left[-E(\pi_{i,j}^{(q)}) / T\right]}, \tag{4}$$

where T is a constant explained under item 4). The sum of P over all n is, of course, 1. A random number is then drawn from a uniform distribution on the interval [0,1), and the new hypothesis about the value of $p_{i,j}$ is chosen according to the value of that random number. For example, if we have $P(p_{i,j} = \pi_{i,j}^{(1)}) = 0.2$ and $P(p_{i,j} = \pi_{i,j}^{(2)}) = 0.8$, the interval [0,1) is split into [0,0.2) and [0.2,1). If the random number is less than 0.2, it is a member of the first interval; consequently $p_{i,j}=\pi_{i,j}^{(1)}$. For a larger random number one sets $p_{i,j}=\pi_{i,j}^{(2)}$. Assuming that the current hypothesis about $p_{i,j}$ is always stored in the element with superscript 1, it has to be noted that actually $E(\pi_{i,j}^{(n)})$-$E(\pi_{i,j}^{(1)})$ was substituted for $E(\pi_{i,j}^{(n)})$ in the numerical experiments. The consideration of the difference focusses on the local effect through the substitution of a $\pi_{i,j}^{(n)}$. When using $E(p)$ in the form of eq. (3), the differences between the competitive values for a single $p_{i,j}$ would be obscured through the overwhelming contribution of the many grid points with unaltered $p - \overline{p_N}$ being irrelevant at the moment of investigating the grid point (i, j).

4) *A control temperature T and an annealing regime.* The parameter T is decreased stepwise during the SA process. A large value is chosen at the beginning, causing $(-E/T)$ to be close to 0 in any instance and resulting in almost equal probabilities for the competitive pressure values. The fact that for T large enough a pressure value causing a larger E has a considerable probability to be selected is responsible for the ability of the SA approach to by and large avoid entrapment in local minima. As T is slowly lowered, smoother configurations gradually exhibit higher probabilities P. This process continues until the system is „frozen“, i.e. no alternative configuration is accepted by the SA procedure anymore. The analogy to the freezing of a metal (where the name „annealing“ stems from) is obvious, as is the reason for calling T a control temperature. *Dittmer* and *Szymanski* (1995) recommended to start with $\Delta E/T \approx 1$, where ΔE is the difference in E due to a change in one parameter of the configuration (in our case: the value of $p(\Theta_e)$ at a single grid point). In order to estimate the values of ΔE which are to be expected, the effect of a few random changes in the parameters can be inspected before the actual annealing procedure is started. *Dittmer* and *Szymanski*'s rule was given for the Metropolis algorithm, but $\Delta E/T \approx 1$ turned out to be an appropriate first guess also for the present application of

the heat-bath algorithm. For too low temperatures, E would steadily decrease and the system would tumble down to a local minimum rather soon. Working with extremely small T is equivalent to visiting the grid points sequentially in random order and always selecting the value $\pi_{i,j}^{(n)}$ which produces the lowest local contribution to E (this method of rapidly achieving a decreased E will be referred to as „straight local minimization“, abbreviated SLM, in the following).

Stopping rules have to be defined as criteria when T should be further decreased resp. when the whole procedure is to be terminated. Let N_A be the number of grid points where an ambiguity in p is observed. N_S be the number of events where, for the current T, the random selection process of the heat-bath scheme alters the hypothesis about the shape of the Θ_e-surface. The used rules can then be described as follows:

1) Decrease the temperature further when either N_S reaches N_A or each grid point was inspected ten times. The actually employed cooling regime was $T_{i+1}=0.9\times T_i$.
2) Terminate the procedure when $N_S=0$ after visiting 10 times N_A grid points, i.e. for the configuration just under inspection, there is at each grid point no random number in 10 trials which leads to the preference of an alternative pressure value.

3. Identification of the smoothest Θ_e-surface: Experience with actual data

The outlined SA algorithm was applied to LAM data, specifically to 170x100-element fields of the model ALADIN operated by Météo-France in Toulouse. The data were provided on a latitude-longitude grid with 0.175° longitudinal distance and 0.15° latitudinal distance. In the SA experiments, however, differences in latitudinal and longitudinal metric distances were ignored. Since the aim was only to check the applicability of the concept, it has been considered adequate to simply view the data as if they were given on a regular orthogonal grid. This was already anticipated in eqs. (2) and (3) where the irrelevant grid point distance was dropped. The data comprised model output on eight isobaric levels from 925 hPa reaching upwards to 50 hPa.

For one particular example with ambiguities at 666 of the 17000 grid points, an E of 24226.4 hPa resulted when applying the GPH rule. Starting with this configuration and carrying out SLM, the system in repeated runs always „froze“ to a configuration with E=21517.8 hPa. Finally, with the SA approach E=16352.9 hPa was reached, i.e. the algorithm did not fall into the local minimum of the abruptly minimizing procedure but found a (considerably) better solution. Fig. 2 highlights a special feature of this particular experiment. The GPH rule produced two outstanding peaks of larger p-values. SLM had some success in bridging the gap visible in the back of the GPH plot but the picture in the region around the two spikes changed only slightly. The large contribution to E due to those spikes vanished in the outcome of the SA procedure since the spikes disappeared in a smooth surface achieved through selecting larger p-values at several surrounding grid points. SA did moreover completely fill the valley in the back, producing a surface whose superior smoothness is obvious.

Fig. 2. Wire-mesh representations of Θ_e-surfaces derived from ALADIN data (excerpt of the investigated domain). Left: The GPH rule, which in many other cases produces satisfactory results, yields a surface which exhibits spikes at two grid points with unique $p(\Theta_e)$ (to the right of the center). Middle: Result of SLM applied to the starting configuration provided by the GPH rule. Right: The result of the SA procedure.

The usual case, however, are smaller differences in E which suggests that the SA algorithm does only in special situations significantly outperform the simple GPH rule with subsequent SLM. It has to be mentioned that calculations for a single Θ_e-surface lasted several hours on the used SUN Sparc 10 workstation, and this slow convergence of the SA procedure prohibited a quite extensive examination. Therefore, the percentage of cases with considerable benefit of SA is perhaps larger than the present investigation has indicated, but there are also reasons to believe that such cases are indeed rare. One factor favouring the selection of smaller p-values in our experiments emerges from the fact that a lower surface is prone to crossing the bottom level. At the points where this occurs, the situation shifts from an ambiguity comprising a lower and a higher value to a unique $p(\Theta_e)$ positioned at upper levels, as on the right side of Fig. 1. The algorithm is therefore strongly discouraged to select the larger value of p for the grid point with the ambiguity since then one enormous jump in p is produced. The outcome is well acceptable if the disappearance of a coordinate surface into the ground is judged as adverse, as it is usually done.

The use of pressure coordinates was also a factor which supported a good performance of the GPH rule. Note, for example, that the same difference in geometric height is numerically diminished in pressure coordinates when occurring at higher altitudes. Undoubtedly, cases can be found where minimization of $E(p)$ in pressure coordinates yields a different (higher) location of the Θ_e-surface than the consideration of an analogous $E(z)$ in metric coordinates. As there was already no reasonable physical justification provided for the smoothness criterion as such, there is no point in discussing which coordinate system is the correct one to minimize the spatial variability

in. Being aware that the choice of the coordinate system influences the outcome, it can obviously not be claimed that a definitely best solution exists and can be found through the SA algorithm. The latter just helps to imitate the subjective analysis a meteorologist would most likely carry out in a prespecified coordinate system if he/she could handle the wealth of data properly.

The problem of height ambiguity may also occur when constructing potential temperature or potential vorticity (*PV*) surfaces. For the potential temperature, however, many of the possible heights are already eliminated through the reasonable rule to not draw coordinate surfaces through superadiabatic (i.e. unstable) layers. The remaining ambiguities are probably only small in number and SA is therefore unlikely to be more economic than a complete combinatorial examination. A similar statement may be true also for the potential vorticity. Usually only one particular value such as *PV*=2 PVU (potential vorticity unit = 10^{-6} $m^2s^{-1}K$ kg^{-1}) is considered which functions as a criterion to locate the tropopause. Some levels with *PV*=2 PVU due to occasional low layers of high stability (hence high potential vorticity) can be rejected by means of the simple rule that *PV* has to increase with height at the tropopause. If it were postulated that there should be no low-*PV* layers above the tropopause, it would immediately follow that the greatest possible height for *PV*=2 has to be selected. It can be expected that this would on the whole also result from an SA smoother because there are certainly points where the *PV*=2 - height is unambiguously defined, and many lower-situated alternatives at grid points with ambiguities are disfavoured through an effect analogous to that when a Θ_e-surface crosses the ground. At the ZAMG, a GPH rule is in fact in operational use and fields of satisfactory appearance are obtained. The *PV*=2 - height fields of *Røsting et al.* (1995) showed some peculiarities (areas encircled by densely packed isolines) which, according to *Bjørn Røsting* (1996, personal communication), are to be attributed to wrong choices from two (or more) candidates for $p(PV=2)$. It is interesting to note that those optically salient features were not at all interpreted during the meteorological discussion but, in fact, subjectively filtered. This underlines that a human interpreter tends to disbelieve in too uneven fields, which was the starting hypothesis here.

4. Infra-red cloud motion vector field construction

For finding the smoothest infra-red cloud motion vector configuration, the cost function *E* to be minimized was defined as

$$E(u,v) = \sum_{i=1}^{k}\sum_{j=1}^{l}\left(\left|u_{i,j} - \overline{u_N}\right| + \left|v_{i,j} - \overline{v_N}\right|\right), \tag{5}$$

where *u* and *v* denote zonal and meridional cloud motion vector component, respectively. In any other aspect, the SA procedure was identical to that introduced in

section 2. Using the standard cross-correlation tracking techniques, the vectors were calculated on a grid being regular on the original satellite image. As before, no correction was applied to account for differences in metric grid point distances.

For Fig. 3, four different IR cloud motion vector fields were produced by variation of the target size resp. the considered time interval (two of these fields are shown in Figs. 3a and 3b). The cost function E was larger than 10000 ms^{-1} in any instance. Application of the SA algorithm for selecting vectors from the four initial fields in order to obtain a smooth composite yielded a field with E=5092.45 ms^{-1} (the result of this run is shown in Fig. 3e). Using each of the original vector fields once for initialization, cost functions of 5414.05, 5537.68, 5816.02 and 6084.4 ms^{-1} were achieved for four SLM runs. Through the quick stagnation in a nearby local minimum, the four final fields resembled their respective starting fields but differed among each other pronouncedly (see the two examples of Figs. 3c and 3d). The more positive impression of SLM gained from Θ_e-surface computations is solely due to the fact that there we are in a position to produce an already near-optimum starting field. An appropriate SA scheme is not sensitive to the chosen starting field since during the first steps with high T, the initial configuration is effectively destroyed.

The SA algorithm could be exploited to produce a larger number of reasonable IR vectors than can be achieved with one particular target size alone. The original motivation behind the experiments, however, was somewhat different. Discrimination between cloudy and cloud free areas is a rather delicate task in northern latitudes in winter and late autumn because of cold land surface temperatures and high incidence angles of the geostationary satellite's measurements (*Olofsson et al*., 1996). Knowing the characteristic appearance of infra-red cloud motion vectors in clear-sky areas, the idea emerged that one can perhaps make use of this to identify the cloud free regions. Little displacement in infra-red features should be observed in the absence of clouds or, alternatively, obviously false vectors are obtained when an infra-red land pattern disappears below clouds in the subsequent image, for example. First experiments to find the optimal target size were rather discouraging. For small target sizes, it frequently happens that temporal changes in some cloudy areas result in the production of incorrect vectors also in the presence of clouds. After considerably increasing the target size, the comparatively stable medium-scale patterns lead to acceptable vectors, but this is likewise true for many such targets being centered over land and containing several bright cloud pixels at the border. The region of reasonable vectors reflecting a displacement of the cloudy portion of the respective targets is thereby extended far into the cloudless areas. Thus, the cloud/no-cloud discrimination must fail if there is no upper restriction to the target size. It may be hypothesized that cloudy areas are characterized by the possibility to *somehow* achieve a reasonable small-target IR cloud motion vector. From several experiments with varying target area size and different sets of images (i.e., forward or backward calculation as in Figs. 3a, 3b), SA could distil the

Fig. 3. Upper row: Raw Meteosat IR cloud motion vector fields for 7 November 1996, noon, obtained by tracking 20x20-pixel targets between 1130 UTC and 1200 UTC (left image) resp. between 1200 UTC and 1230 UTC (right image). Two additional vector fields (not shown) were obtained from analogous calculations with 16x16-pixel targets and used for the experiments illustrated in panels c) - f). The target sizes as well as the grid point distance of about 5 pixels (≈25 km) were chosen somewhat arbitrarily and should not be understood as recommended values for any application mentioned in the text.
Middle row: Results of SLM starting with the configurations depicted in panels a) (for the left picture) resp. b) (right) and using the other three computed vector fields to find alternatives. Some obvious differences suggest that in at least one case the final result is far from the global optimum.
Bottom row: Results from two independent SA runs. Although there are a few differences in the two examples, they are far less significant than for the SLM runs (with the only notable exception around 49/5).

most reasonable cloud motion vectors. The outcome of the procedure would then be evaluated by criteria yet to establish in order to distinguish „cloud vectors“ from „no-cloud vectors“. Clearly, numerous experiments with different sets of target sizes have to follow in order to clarify whether the concept really works in practice (stationary fog and barrage cloud - the former being moreover structureless - are expected to cause problems). Though the CMV quality is said to degrade outside a 55 degree great circle arc around the sub-satellite point because the viewing geometry distorts the imagery there (*Schmetz et al.*, 1993), areas like Scandinavia or the Norwegian Sea can probably also be covered with the outlined prospective cloud/no-cloud discrimination. Ultimately, it is only the spatial continuity of the field which is checked whereby a somewhat lower accuracy of cloud motion vectors should not disturb too much.

5. Discussion

The SA procedure to extract the smoothest possible field from a finite, but vast, number of possible configurations has operated successfully, but there still have been slight differences in the final E for repeated SA runs (cf. Figs. 3e and 3f). So, SA has obviously not always found the global minimum (however, for the example of Fig. 3 even the worst SA run yielded an appreciably lower E than the best SLM run), and one might try to refine the annealing schedule to hit the global minimum more reliably (e.g. through slower cooling). It became nevertheless even with that small imperfection evident that the presented methodology of constructing a field with a specified optimum property enables some particular investigations being otherwise hardly feasible. For example, despite the arguments why the greatest possible altitude should lead to an almost optimum $E(p)$ for the Θ_e-problem, one cannot be sure about this point without verifying it on real atmospheric data. It is not clear how this should be accomplished were there not the powerful method of stochastic optimization.

The long time the SA scheme required to return an answer was found to be less satisfactory. In meteorology, perhaps more often than in any other field of natural science, the notion of certain relations is unsharp and as remedy parametrized empirically. As an example, one could envisage the introduction of weighting factors taking into account that higher variability can be allowed between IR cloud motion vectors from adjacent sectors with highly different IR greyshades ($\Rightarrow$ distinct atmospheric levels with possibly distinct displacement vectors). Carrying out a trial-and-error procedure in order to achieve an appropriate parametrization becomes exhausting when one verification run on an average workstation requires the time of one working day to finish. Likewise, weather forecasters cannot accept to wait for any SA analyses that long. Several potential modifications to the SA procedure, which was employed here in an *ad hoc* fashion, can be found in literature. It is to be hoped that future developments (and/or use of faster computers) will improve the practicability of

the approach which in principle offers interesting research opportunities, as was demonstrated here.

An obvious extension of the Θ_e-problem is to analyze a stack of Θ_e-surfaces. A surface with a higher Θ_e should everywhere lie in greater height since the crossing of surfaces with different temperatures must be avoided. This can be accomplished within the SA framework through minimizing E for all considered surfaces simultaneously, withdrawing temporarily all options of p which are smaller than the currently assumed p-value of a „warmer“ surface. The complexity of the algorithm is thereby only little increased. However, the GPH rule ascertains the correct stacking of Θ_e-surfaces without any additional provisions, and as results were until now in most cases not much worse in terms of $E(p)$, it does not seem warrantable to recommend the computationally far more expensive SA approach for operational moist-isentropic analyses. Occasionally, SA even gave poorer results than the GPH rule with subsequent SLM. As the SA scheme did in principle fulfil its task, this has to be considered as an indication that often GPH almost provides the global optimum solution, with sub-optimum choices at a few points being curable through SLM. SA with the inherent stochastic component gives a very good solution (virtually always smoother than GPH alone) but may just before „freezing“ in a certain region suffer a slight random change for the worse. SA then fails to reach the global minimum, which GPH+SLM may have attained in those instances where this approach was superior. *Note:* The focus was on pressure coordinates in the present study because this framework is the preferred one in meteorological practice. As already indicated, the made observations are not necessarily transferable to other coordinate systems.

Finally, a few comments are due on the introduction of the somewhat unusual form of the smoothness criterion (1). Smoothing terms usually contain squared derivatives in their definition (e.g. *Wahba* and *Wendelberger*, 1980, or *Birkenheuer*, 1996), so one might be tempted to use

$$E(p) = \iint\limits_{\text{Area}} \left(\nabla^2 p\right)^2 \mathrm{d}x\,\mathrm{d}y\,, \tag{6}$$

like the correspondent expression of eq. (1) treating negative and positive values of the Laplacian equivalently and moreover attributing more weight to larger values. The last-mentioned property, however, has turned out to be not desirable since it favours the occurrence of metastable configurations during the SA procedure. To illustrate this point, let us consider the simple example of Fig. 4, where the value of p=1 was allowed at each grid point while the alternative of p=2 was admissible only at some locations. Fig. 4a shows the outcome of one run of the simulated annealing algorithm using a cost function of type (6). SA has in this case failed to yield the obvious true solution to the minimization problem with everywhere p=1. Considering the balances at the lettered points of Fig. 4b for p(B)=1 resp. p(B)=2 (Table 1), it becomes apparent that it is

difficult to break up the metastable rectangular portion of p=2 due to the stabilizing overemphasis expression (6) lays on grid points A and E with increasing local contributions to E. With cost function (3), p(B)=1 and p(B)=2 are considered as equivalent options. Consequently, it is easier to corrode the region with p=2, and there did never occur problems in finding the correct solution. Theory teaches that the SA procedure should - for an appropriate cooling regime and a sufficiently large number of iterations - always be able to converge towards a global minimum (e.g. *Rothman*, 1986). A reliable minimization of (6) via SA is therefore possible. However, the measures to be taken result in an unnecessarily lengthy process for finding a smooth surface, compared with the minimization of (1).

Fig. 4. An example of an SA failure caused by an adverse formulation of the cost function. a) The area where two options (p=1 and p=2) are available is grey-shaded. Outside this area only p=1 is possible. Numbers at the grid points indicate the values selected by one particular SA run which should minimize expression (6) (with the exception of the altered cost function, the SA procedure was identical to that described in section 2). b) Zoomed region from a), for which numerical evaluations are provided in Table 1.

Table 1. Local contributions to the cost functions of types (1) resp. (6) at the points labelled in Fig. 4b, with the Laplacian approximated as indicated by eq. (2). The table lists the balances for the two options p(B)=1 and p(B)=2, all other values being left unchanged by the SA algorithm.

Point	**eq. (1) (p(B)=2)**	**eq. (1) (p(B)=1)**	**eq. (6) (p(B)=2)**	**eq. (6) (p(B)=1)**
A	0.25	0.5	0.0625	0.25
B	0.5	0.5	0.25	0.25
C	0.25	0	0.0625	0
D	0.25	0	0.0625	0
E	0.25	0.5	0.0625	0.25
Total	1.5	1.5	0.5	0.75

6. References

Bader, M.J., G.S. Forbes, J.R. Grant, R.B.E. Lilley and A.J. Waters (Eds.), 1995. *Images in weather forecasting*; Cambridge University Press: Cambridge.

Billings, S.D., 1994. Simulated annealing for earthquake location. *Geophys. J. Int.*, **118**, 680-692.

Birkenheuer, D., 1996. Applying satellite gradient moisture information to local-scale water vapor analysis using variational methods. *J. Appl. Meteor.*, **35,** 24-35.

Carlson, T.N., 1980. Airflow through midlatitude cyclones and the comma cloud pattern. *Mon. Wea. Rev.*, **108,** 1498-1509.

Dittmer, J.K. and J.E. Szymanski, 1995. The stochastic inversion of magnetics and resistivity data using the simulated annealing algorithm. *Geophysical Prospecting*, **43**, 397-416.

Kirkpatrick, S., C.D. Gelatt, Jr. and M.P. Vecchi, 1983. Optimization by simulated annealing. *Science*, **220**, 671-680.

Mahringer, G. and V. Zwatz-Meise, 1993. A semi-operational synoptic diagnosis method. *Beitr. Phys. Atmosph.*, **66,** 89-106.

Metropolis, N., A.W. Rosenbluth, M.N. Rosenbluth, A.H. Teller and E. Teller, 1953. Equation of state calculations by fast computing machines. *J. Chem. Phys.*, **21**, 1087-1092.

Olofsson, P.-O., L. Häggmark and D.B. Michelson, 1996. Analysis of clouds at high latitudes using Meteosat IR imagery with observed and forecasted surface temperatures. Proceedings of the 1996 Meteorological Satellite Data Users' Conference, September 16-20, Vienna, Austria, 137-141.

Press, W.H., S.A. Teukolsky, W.T. Vetterling and B.P. Flannery, 1992. *Numerical recipes in FORTRAN. The art of scientific computing* (2nd ed.); Cambridge University Press: Cambridge.

Røsting, B., J. Sunde and K.H. Midtbø, 1995. Application of satellite data in monitoring NWP models. DNMI Research Report No.13, Det norske meteorologiske institutt, Oslo.

Rothman, D.H., 1986. Automatic estimation of large residual statics correction. *Geophysics*, **51**, 332-346.

Sambridge, M. and G. Drijkoningen, 1992. Genetic algorithms in seismic waveform inversion. *Geophys. J. Int.*, **109**, 323-342.

Schmetz, J., K. Holmlund, J. Hoffman, B. Strauss, B. Mason, V. Gärtner, A. Koch and L. van de Berg, 1993. Operational cloud-motion winds from Meteosat infrared images. *J. Appl. Meteor.*, **32**, 1206-1225.

Tarvainen, M. and T. Tiira, 1995. Study of the locating capability of genetic algorithms. *Geophysica*, **31**(2), 1-17.

Wahba, G. and J. Wendelberger, 1980. Some new mathematical models for variational objective analysis using splines and cross validation. *Mon. Wea. Rev.*, **108**, 1122-1143.

Wilson, W.G. and K. Vasudevan, 1991. Application of the genetic algorithm to residual statics estimation. *Geophys. Res. Lett.*, **18**, 2181-2184.