AUTOMATIC INTERPRETATION OF FORECAST CHARTS

BY

O. LÖNNQVIST

THE SWEDISH METEOROLOGICAL AND HYDROLOGICAL INSTITUTE

NORRKÖPING

SWEDEN

A. General aspects on weather interpretation of numerical forecast charts

1. Introduction

We have all our hobbies. One hobby of mine is crossword puzzles - either to do them or to make them. Recently I have made a crossword for the staff journal of the Swedish Meteorological and Hydrological Institute, and since I want to try something new, that crossword puzzle happens to be one in three dimensions - the words going across, down and inwards (Fig. A1).

Next step is a crossword in 4 dimensions. It will be addressed to Swedish meteorologists, but you can have a first glance (Fig. A2). Although it only contains three-letter-words it was fairly difficult to compile. You find 27 words across, 27 words down, 27 words inwards and 27 words sideways. Each letter, that means each box or sub-cube in this 4 dimensional cube, takes part in 4 words which makes the puzzle much easier to solve than to construct.

Now to meteorology and forecast interpretation. This 4-dimensional cube is very useful for demonstrating the method of interpretation by type climatology which I have used as background for the forecast chart interpretation now used since about 12 years in the Swedish weather service. Let us use, as in next figure (Fig. A3), the horizontal direction in each one of the squares for the west-to-east component of the geostrophic wind, divided in 3 classes, and the vertical direction for the south-to- north component. Further, let the inward direction be the axis for atmospheric pressure, also divided in 3 classes. Finally let the thickness be represented by going sideways from page to page from low, via medium to high values of the 1000-500 mb thickness.

The numbers you find inserted in the boxes show the minimum temperature in December at Göteborg on the

Swedish West Coast.

I call this a type-classification in 3x3x3x3=81 types. Since the historical data only comprised 16 years the number of cases is 496, which means that no more than around 6 cases fall in each box; therefore this is to go a little too far when using the method. Since the crossword contains 108 words, this figure demonstrates 108 functional relationships.

Let us use the crossword terminology and look at "13 across" Here we read 4 0 -2. This means, that provided the pressure is normal and the thickness is low and we have a zonal situation, then westerly winds give the maximum temperature +4^{O} and easterly winds -2^{O}C.

Next, looking at "7 sideways" we read the word

4 6 8

which menas: With southwesterly winds and low pressure, the maximum temperature varies like that with increasing thickness.

And so on.

May I ask you, for instance, to compare "29 inwards" with "35 inwards". - Yes, in one case temperature increases with pressure, in the other it decreases with pressure.

2. The historical material

My intention is to give you various aspects on forecast chart interpretation. I shall first discuss the data needs, then the choice of predictors and predictands. I shall also give my personal views and experiences on the statistical methods to be used to find out the relationships which can best be used for interpretation.

Interpretation is based on experience. The experience can be grasped from climatological data, including past day-to-day observations and saved data from previous computer-made analysis and forecast charts. I prefer the refer to all these data as "the historical material".

When the historical data have been analysed for interpretation purposes and have been stored as an "experience bank" we have obtained what can be called "the interpretation climatology".

The necessary number of cases in the historical material varies from one problem to the other: For simple and fairly well-known relationships one year may suffice; for details in precipations, for instance, I know that even 16 years seems to be too short a period. It is important that all sorts of weather situations are represented in the material, including a sufficient number of extreme cases. Pilot studies can be carried out by limited data, using a sample of selected data, e.g. where extreme cases are over-represented. Usually you cannot afford to spare part of the historical material for verification purposes, but some sort of verification must be incorporated in the process.

Before starting, you have to decide whether to use grid-point data, or data from real stations. You must also know whether you prefer single station data, areal mean values or perhaps areal extreme values.

3. Predictors

The term "predictand" is used for the parameter to be predicted by interpretation; a "predictor" on the other hand is any parameter used for the interpretation. When the interpretation is carried out, most predictors are extracted from the forecast charts and they take different values on different time steps. But there are other predictors as well; first of all "persistency", that is the last observed value of the predictand itself.

Others could be elevation of the sun, length of the night, humidity of the soil, etc.

When using the historical material, the predictor values simultaneous to the values of the predictand, can either be found in records of observed values or be extracted from analysis charts. This is the so-called "perfect prog method". Or they can be found on forecast charts - the MOS method, "model output statistics". I am sure we will talk a lot during this seminar on the pros and cons of these two methods. Let me start the discussion.

If all forecasts charts were correct, the perfect prog method is the obvious method to use. Therefore, one merit of this method is that the interpretation is more accurate the better the forecast chart is. That means that if the forecaster improves the forecast chart manually and sends i back to the computer for re-interpretation he is guaranteed a better result. Another advantage is that the interpretation statistics can be worked out once and for ever and really just works better and better each time the numerical models are improved. The main disadvantage is that one has to be careful in the choice of predictors not to use those which do not come out very well by the numerical technique.

The advantages of the perfect prog method are of course disadvantages of the MOS-method, and vice versa. To make it more clear, the main advantage of MOS is that the bias and the inaccuracy of the forecast method is automatically taken into account in an optimum way. But as a result the interpretation is in fact slightly better in those areas of the forecast charts where you have normal errors, and slightly worse in those areas where the forecast happens to be almost correct. Furthermore a new model must be run for some time until you get enough data for an interpretation climatology, and strictly speaking this climatology must be reviewed every time there are changes made in the operational model.

Personally I favour the perfect prog method at least for short forecasts. For extended forecasts up to 10 days, MOS might be preferable, using maybe one set of climatology for 3-5 days, another set for 6-10 days. The best answere might be, if you can afford it, to develop and use both methods.

Coming back to the choice of predictors, priority should be given to those which by physical reasoning can be expected to effect the predictand. It might be observed or computed quantities. That computed values of vertical velocity should be included, if possible, in predicting precipitation and cloudiness, and by that probably also in predicting temperature is obvious from the following table, showing the relation between vertical velocity (computed on the analysis chart) and the average probability of precipitation at a number of Swedish stations.

Vertical wind (1000-500 mb) class limits (mm/s)	Probability				
	no precip.	<0.5 mm	1-2 mm	3-4 mm	>5mm/12h
	12 %	13	25	13	37
+9.0					
	36	20	24	10	10
+2.5					
	62	21	13	1	3
-2.5					
	69	21	8	1	1
-9.5					
	74	17	8	1	0

But you can also try, as I have done with some success, other and more odd predictors, such as the longitudinal position of the main ridge or the main trough over NW Europe, the past weather at the starting point of trajectories ending at the point of prediction, convectivity functions and whatever you think makes sense.

Since the pressure gradient, or say the components of the gestrophic wind, are important predictors for most purposes, one important question is over which distances these components should be measured. The following figure (Fig. A4) gives some examples

Hovmöller, who once introduced a manual interpretation method in Sweden in 1959, worked with 500 mb charts and determined the gradient over 20 latitude degrees, that is over 2,200 km. When I introduced automatic interpretation based on surface forecast charts, in 1966, my choice was to compute the pressure gradient by a station-cross over a distance of about 500 km. In a recent study which I shall present in my next lexture, I tried to compute the geostrophic wind over a triangle with a size of abour 80 km.

It is no doubt convenient to chose the normal grid size for this purpose but I think it is still an open question what is really the optimum distance for various purposes.

4. Predictands

Both for predictors and predictands you can either use values observed or computed at a certain time or, say, the past 6 hour changes in the same parameters. For temperature interpretation I have used both in such a way that the final forecast is a weighted combination of a directly forecast temperature and the temperature obtained by adding the forecast change to the latest observed or forecast value.

$$T = (1-\alpha)(T_{normal} + A_{forecast}) + \alpha(T_{observed} + dT_{forecast})$$

Here A is the typical temperature anomaly and dT the typical change in temperature. For short forecasts α approches 1. For forecasting dT then, either, as an example, the pressure p or its change dp can be used as predictors. One has to find out whether in summer, it is low pressure or decreasing pressure that are best correlated with a drop in temperature.

Speaking now particularly on predictands, I would like to argue for using combined parameters for precipitation and cloudiness. I have previously suggested, and I don't think it is originally my own idea, a function which I have called PCF (which stands for Precipitation/Cloudiness Function).

In case of precipitation, PCF = 10 log R_{12} ,

where R_{12} is the relative amount of precipitation during 12 hours (compared with monthly normal).

In case of no precipitation, PCF = -10 - 2 (8-N) ,

where N is the amount of clouds in octas.

It is easier to apply regression techniques to this function than to precipitation (or cloudiness) alone.

It is also worth while to treat continuous precipitation separate from convective precipitation, since the best predictors are certainly not the same for these two phenomenon.

5. Statistical analysis

Predictors and predictands being chosen, next step is to decide which method shall be used for establishing the relationship between each predictand and its predictors. Here, the traditional choice is regression methods. In most cases I have found it more fruitful to use type-classification, that is to group together cases with similar predictor values. It is in a way a systematic development of the crude method of searching for analogue cases.

Another possible technique is to use empirical orthogonal functions.

Before deciding on the technique it is necessary to be quite clear on the aim of the interpretation. The normal thing is to try to minimize the squares of the errors. But the interest of the customer might justify the use of a very special verification table, giving credit to success in forecasting what is particularly important and punishing for dangerous failures. In those cases the errors should be minimized according to the verification table. One example is forecasting aviation weather, where there are cases when a too simple verification table erroneously gives maximum credit to a forecaster who always gives the forecast: "good weather".

Also, you have to decide whether the interpretation shall indicate one single best value, or probabilities for, say, precipitation above 1 mm, temperature below zero, winds above 14 m/s, and cloud base and visibility below the closing limits at an airport.

Maybe, you should not always let the whole statistical job be done in one step. If not, I see two possible alternative ways.

One alternative is to sort first the historical material into groups; maybe one group of data for snow on the surface, the other for cases with bare ground. Or one group when the station is situated in a warm-sector, one when the station is affected by a frontal passage, and a third group for cases when the station was situated far from and on the polar side of the main frontal system, and so on. Such a procedure was once suggested by Bergeron (1930). It can also be worth while to single out all nocturnal cases with clear sky in order to study separately the minimum temperature in those cases and the conditions for formation of radiation fog. This will lead to a procedure where interpretation is supplemented by a special treatment when ever clear nights are forecast. It could be a statistical method or a physical method like the one suggested by Lönnqvist (1945).(Fig.A5). At that time the answer was a nomogram. Now it must be computerized.

The other alternative is to correct the data for certain more or less well-known effects like the normal diurnal variation of the wind or temperature, before starting the general statistical analysis. I shall give examples of this in my following lectures.

I will conclude this section on statistical analysis methods by returning to the 4-dimensional crossword puzzle. This time the 3x3x3x3 sub-square scheme is used for 24 hour precipitation in December at the station Borås near the climatological rain maximum of southwestern Swedish inland. (Fig. A6) This time the predictors are the following:

- geostrophic west wind component (<u>across</u>)
- geostrophic south wind component (<u>down</u>)
- atmospheric pressure (<u>inwards</u>)
- vorticity or "cyclonality" (<u>sideways</u>).

It is clearly seen that precipitation mainly occurs with a southwesterly or westerly current, and that there falls a higher amount in low pressure, cyclonic situations than in high pressure, anticyclonic ones. This clear-cut features show up very well although the historical material is strictly speaking too sparse for such a detailed analysis. No more than 496 cases but distributed over 81 boxes means just around 6 cases in each box. Figure A7 shows how the number of cases were really distributed. Due to the high correlation between pressure and cyclonality, the number of cases are higher along one of the diagonals than in the opposite corners. In some boxes there are only 1 or 2 cases, in others there is a complete lack of information. To overcome this difficulty I have used a technique where information is borrowed from the close-by boxes in a systematic way and to an extent depending on the number of cases in the box itself. It was possible to select a higher or lower degree of such smothing. For clarification I should mention that for box number 6, where no case is falling, information is thus taken from the 5 surrounding boxes, namely numbers 3, 5 and 9, 15 and 33. This means an important improvement of the type classification method.

As I see it, one clear advantage of the type classification method is that it gives directly the frequency distribution of the predictand in each box, and by that an immediate answer to the question of probabilities.
Fig. A 8 shows the probabilities of getting more than

2 mm/24 hours. The probability extends from zero in 21 boxes to 100 percent in 9 boxes.

When using this type of diagram for operational interpretation it is possible to introduce a sort of smothing in order to avoid the high steps from one box to the other. Such interpolation in all four directions is easier to program than to demonstrate with the aid of this 4-dimensional cube.

A last question concerns the choice between the type classification, as demonstrated by this crossword scheme, and multiple regression analysis. Let us go back to Fig. A6. If we use a linear function of the four predictors we obtain instead the results shown in Fig. A9.

Compare, for instance, the three squares to the left (anticyclonic cases). The agreement is not striking. For some boxes in the bottom square the fit is really bad. Using also squares of the predictors for the regression the fit becomes only slightly better. The use of cross-products as well, and maybe also some third degree terms, will certainly give a much closer agreement.

6. Conclusions

Before starting an interpretation project the following questions must be answered:

- Size of the historical sample

  How may years? Should a specially composed sample be used ? Should part of the data be spared for verification?

- Type of data

  Grid point or station data? Single place or areal data?

- Predictands

  Actual time data or time derivatives? Observed or computed values? Combined (composed) parameters?

- Predictors

  Actual time data or time derivatives? Perfect Prog Method or Model Output Statistics? Use of persistency? Use of "odd" predictors?.

- Size of grid

  Over which distance should wind components be determined?

- Method for statistical analysis

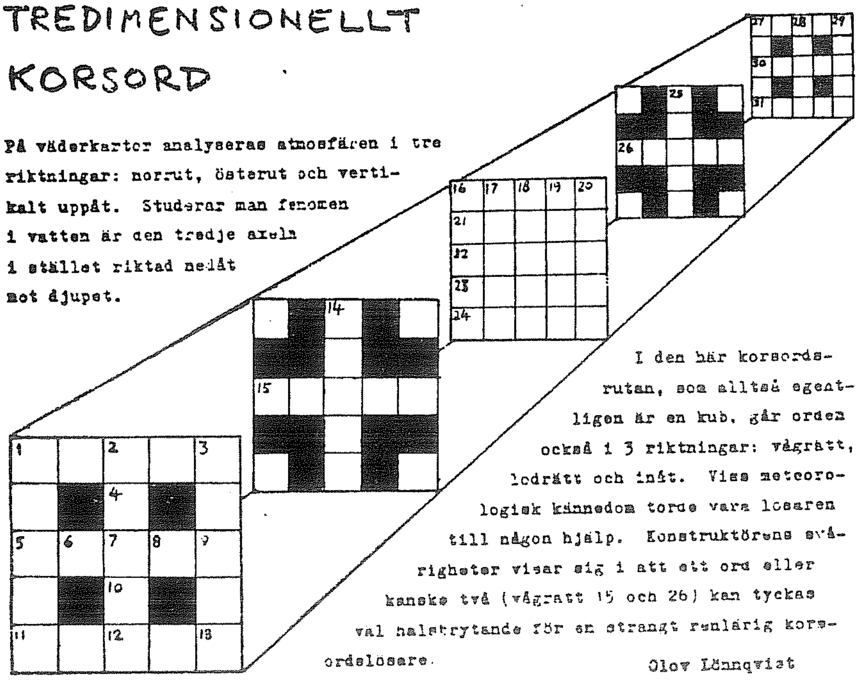
  Regression, weather types, E.O.F. or others? Should parts of historical sample be dealt with in different ways? Data first corrected for known diurnal variation?

- Adaption to customer wishes

  Should special verification tables be used? Result in probabilities?

TREDIMENSIONELLT KORSORD

På väderkartor analyseras atmosfären i tre riktningar: norrut, österut och vertikalt uppåt. Studerar man fenomen i vatten är den tredje axeln i stället riktad nedåt mot djupet.



I den här korsordsrutan, som alltså egentligen är en kub, går orden också i 3 riktningar: vågrätt, lodrätt och inåt. Viss meteorologisk kännedom torde vara läsaren till någon hjälp. Konstruktörens svårigheter visar sig i att ett ord eller kanske två (vågrätt 15 och 26) kan tyckas väl halsbrytande för en strängt renlärig korsordslösare.

Olov Lönnqvist

VÅGRÄTT

1. Ämnar.
5. Vid dess stränder anlades staden Cesar-Augusta.
11. Slände.
15. Påbörjad dödskalle, skulle man möjligen kunna säga.
16. Gör en som observerar.
21. Alnus incana.
22. Håller igång med pace.
23. Bakom visarnas gång.
24. Ur sump.
26. Vara utan med fel eller vara utan fel.
27. Måste linet.
30. Blir hund av hund.
31. Glittrande droppar.

LODRÄTT

1. Manege.
2. Borde finnas på terrarium.
3. Berömd.
14. Siare.
16. För ökenstoft.
17. Varg var väl en sådan.
18. Bar.
19. Öva tal.
20. Stoppljud.
25. Låta best.
27. En ren våtvara.
28. Påståenden.
29. Betecknas med S eller (c).

INÅT

1. Menar och gås.
2. Vått om fötterna.
3. Ristas.
4. Träsk.
5. Bakom 'Spiralen'.
6. Plattform som inte motiverar utropet.
7. Gör aska och stubb.
8. Gamla skatter.
9. Smyckad.
10. Kos föregångare.
11. Har många mödrar.
12. Gör kanske observatören både i tjänsten och på semestern.
13. Var herrar i Skåne förr.

Fig. A1 : A three-dimensional crossword puzzle published in Sweden.

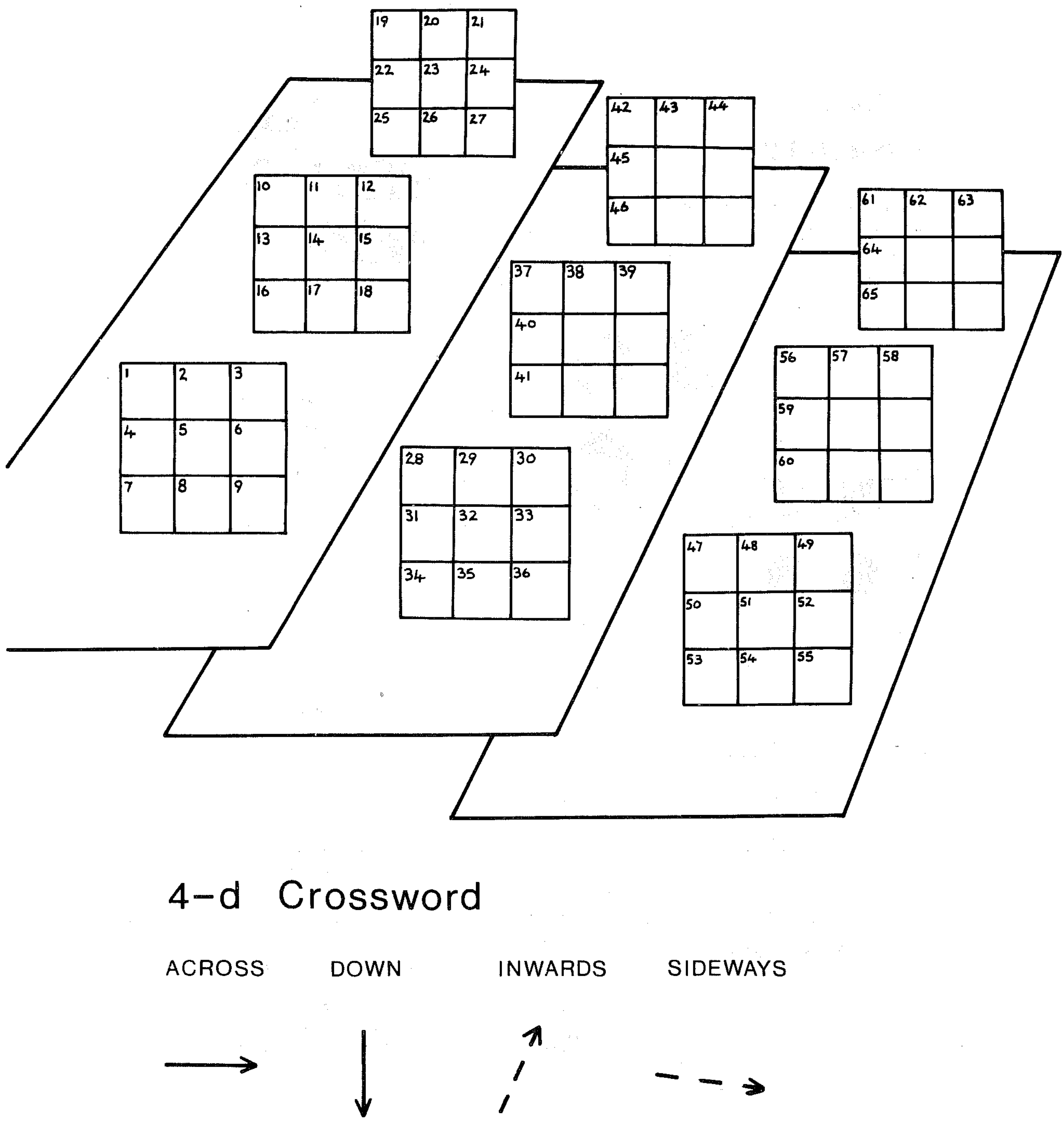


Fig. A 2 : A four-dimensional crossword puzzle containing 108 three-letter words.

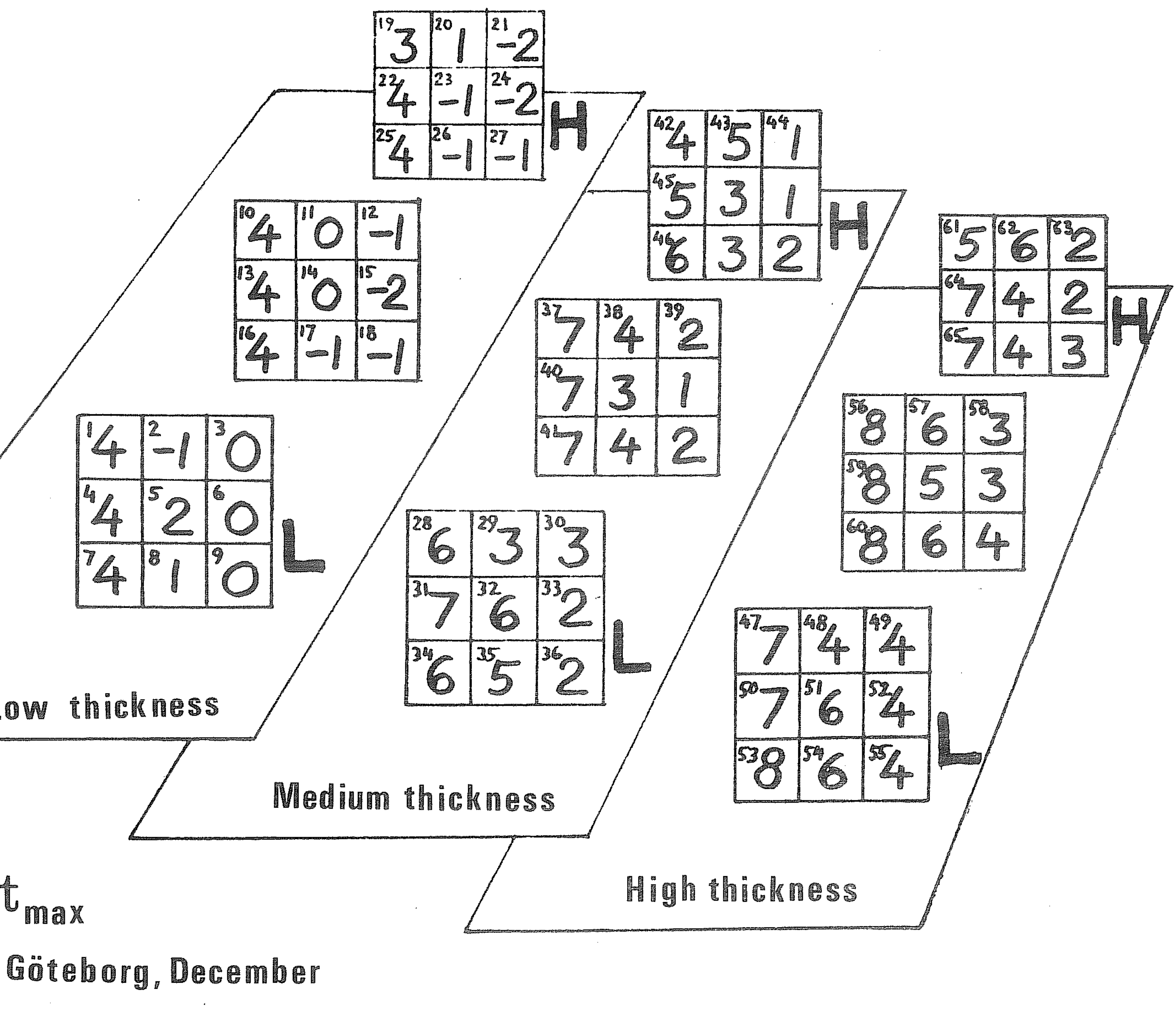


Fig. A3: The 4-d crossword used to demonstrate how the maximum temperature at Göteborg in December varies with geostrophic wind (across and down), pressure (inwards) and thickness (sideways).

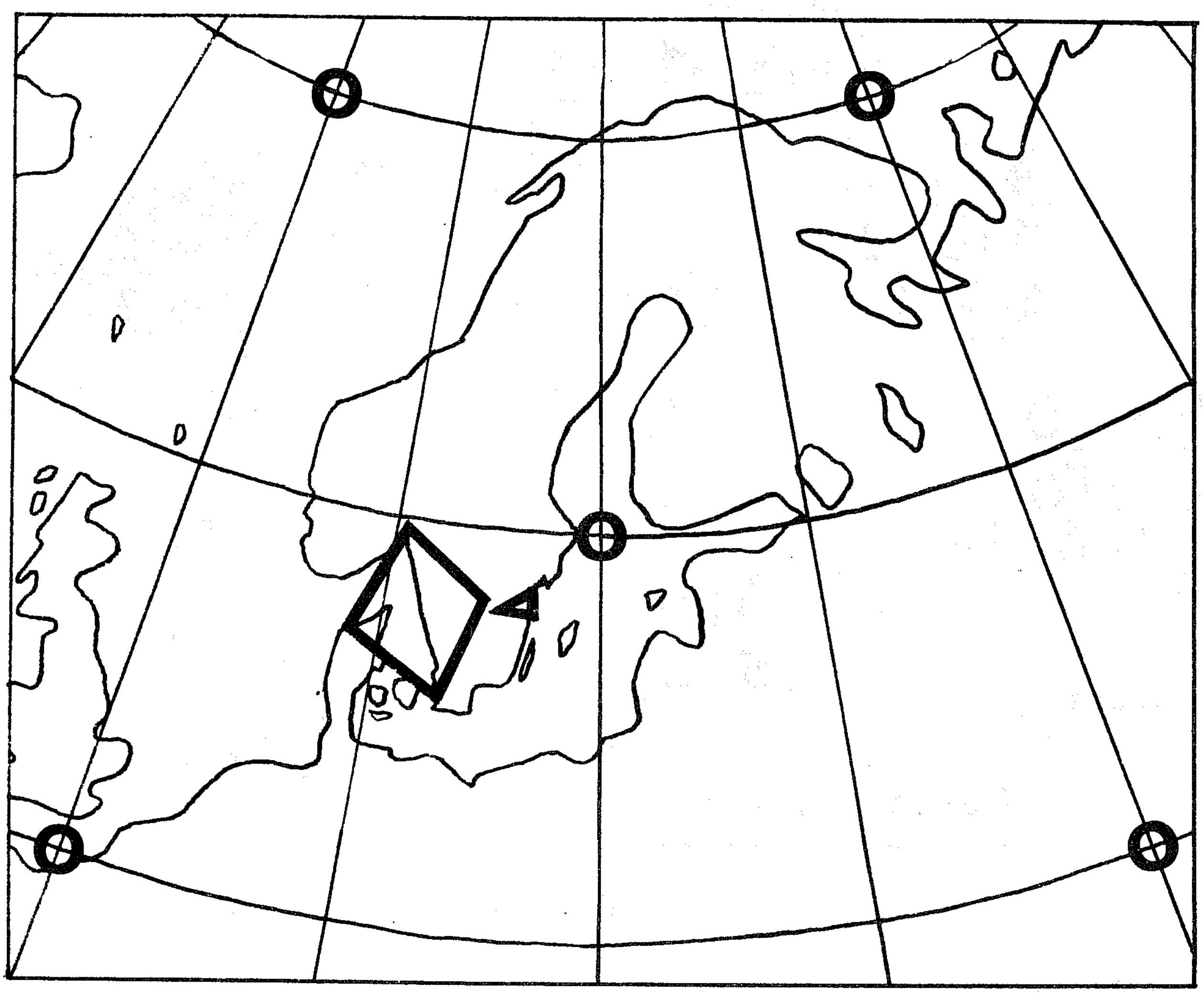


Fig. A 4 : Three possible grid sizes used for determining the zonal and meridional components of the current, or on the smaller scale, of the geostrophic wind :

2200 km (Hovmöller), 500 km (Lönnqvist) or 80 km (Lönnqvist, pilot study).

6 ARKIV FÖR MATEMATIK, ASTRONOMI O. FYSIK. BD 32 A. N:O 13.

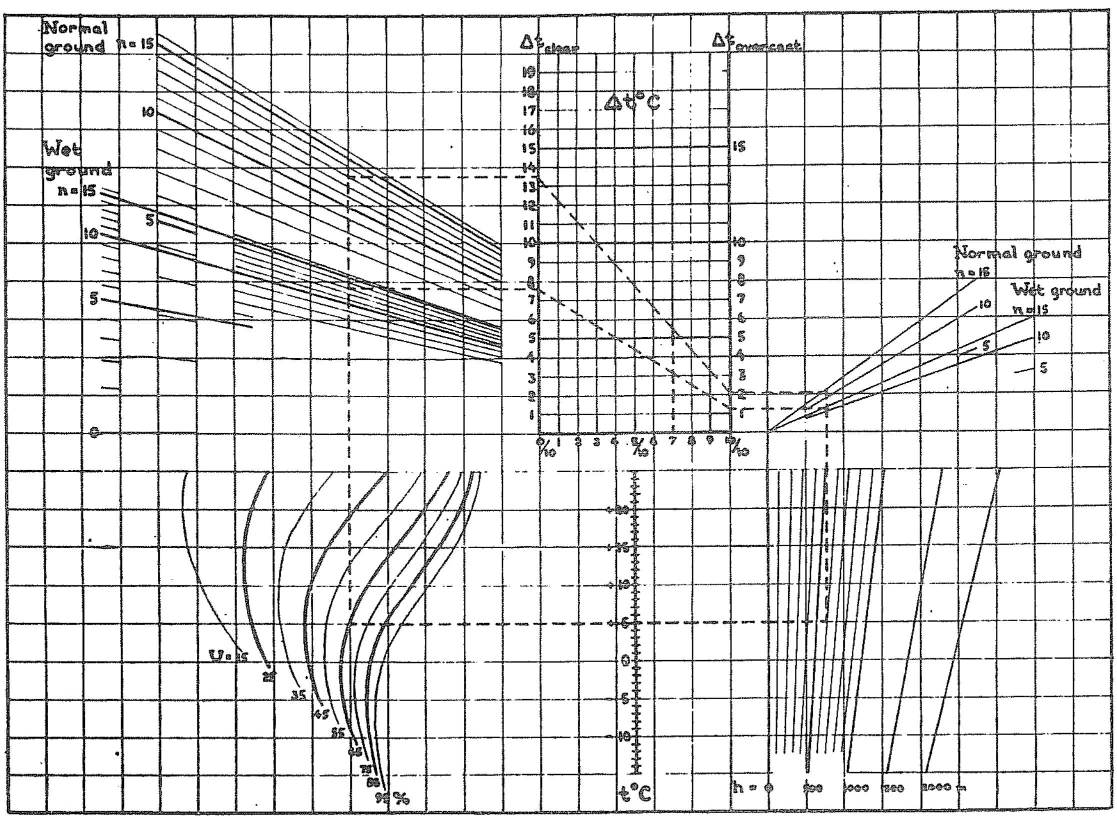


Fig. A5: A frost-prediction diagram
(Lönnqvist,1945).

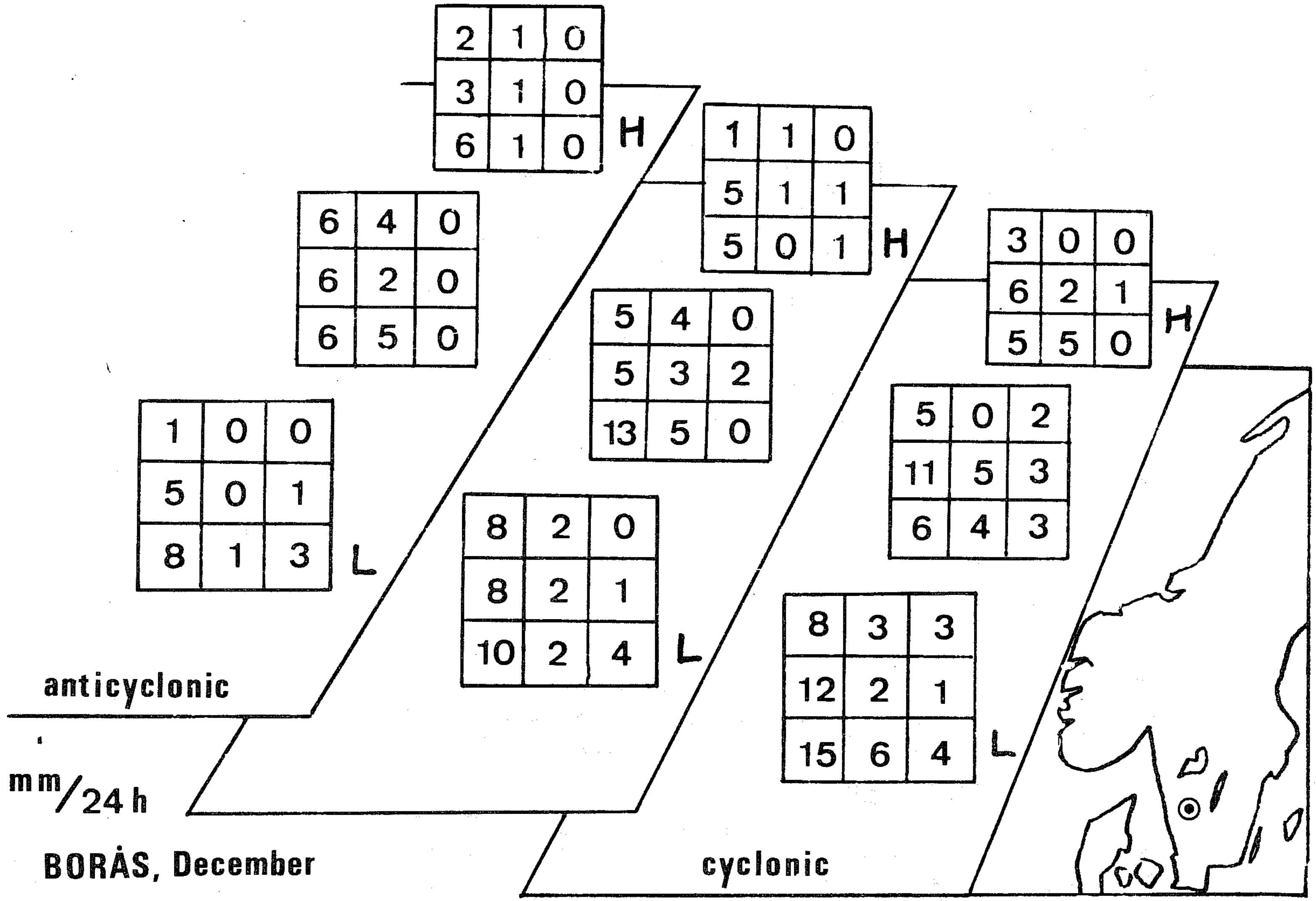


Fig. A6: The 4-d crossword used to demonstrate how the 24 hour precipitation at Borås in December varies with geostrophic wind (across and down), pressure (inwards) and cyclonality (sideways).

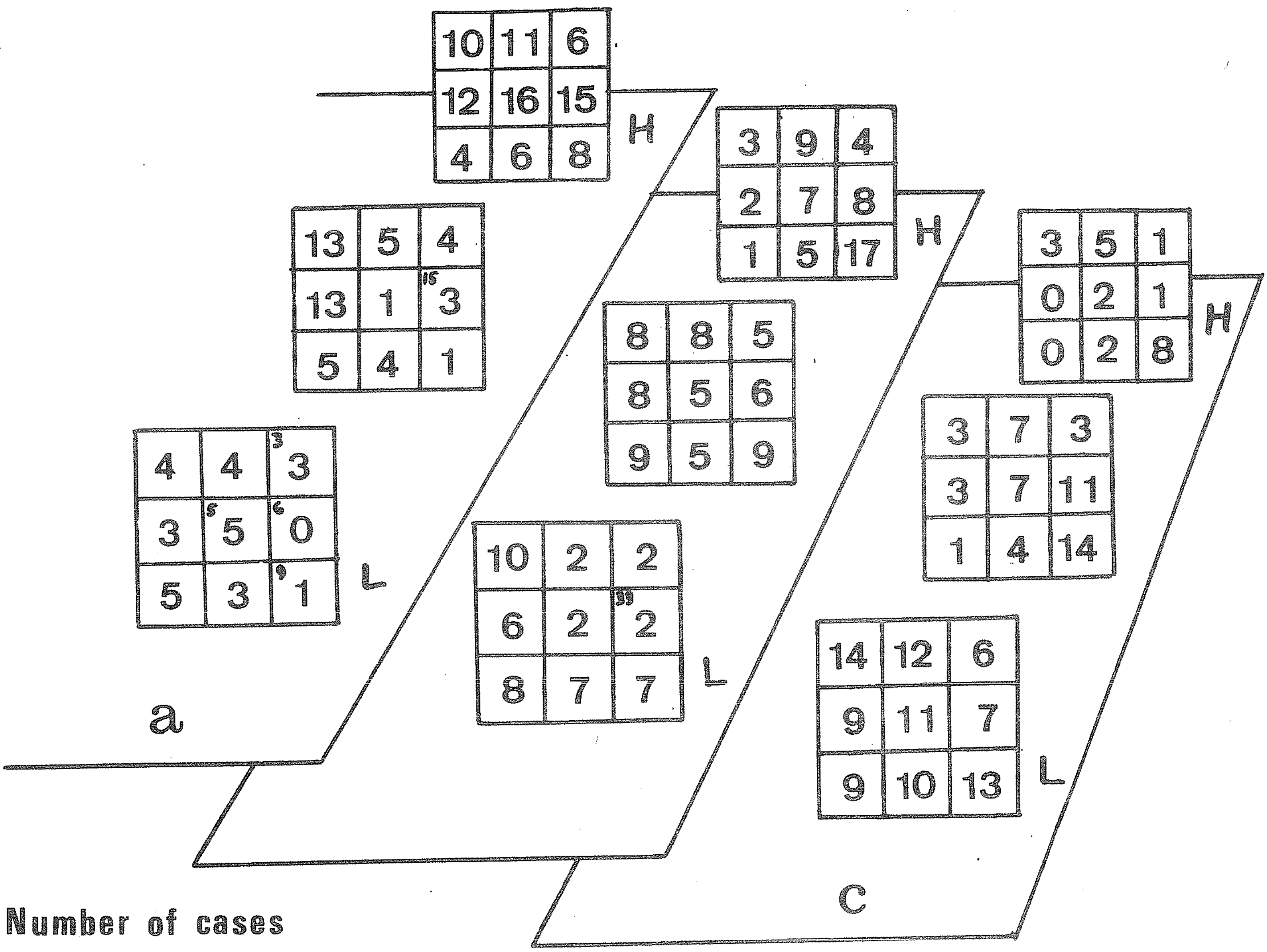


Fig. A7 : The number of cases in each box used for the study presented in Fig. A6.

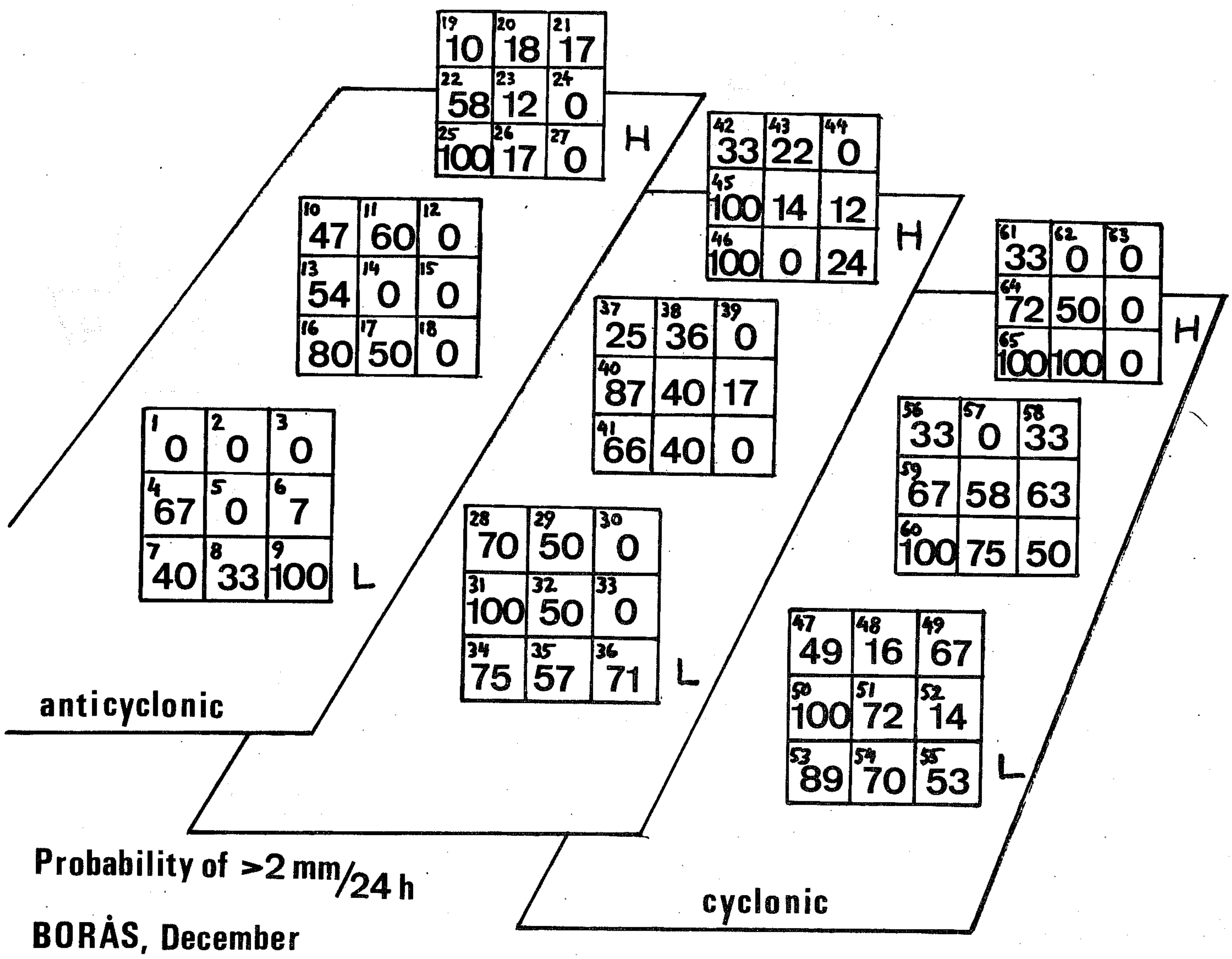


Fig. A8 : Probability figures (> 2mm/24h) obtained in the study presented in Fig. A6.

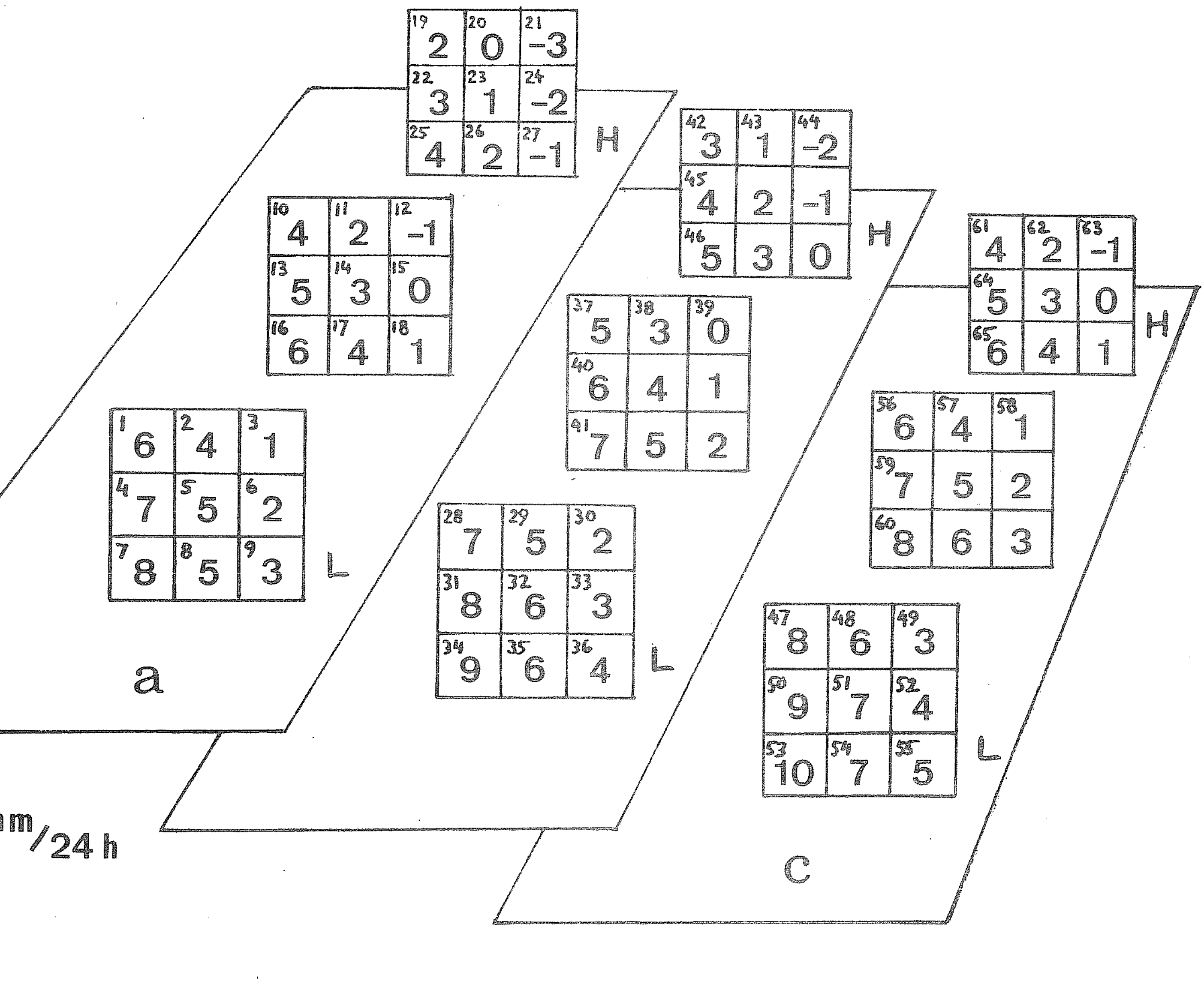


Fig. A9 : Figures for the 24 hour precipitation in Borås obtained by linear regression. These figures should be compared with the figures given in Fig. A6, obtained by the author's type-classification method. Note in particular the differences between Figs. A6 and A9 as to the squares to the left (anticyclonic cases) and the bottom square to the right (cyclonic, low pressure cases).