Short-range dynamics and prediction of mesoscale flow patterns in the MISTRAL field experiment

Rudolf O. Weber, Pirmin Kaufmann, Peter Talkner

Angaben zur Veröffentlichung / Publication details:

Weber, Rudolf O., Pirmin Kaufmann, and Peter Talkner. 1997. "Short-range dynamics and prediction of mesoscale flow patterns in the MISTRAL field experiment." *Tellus A: Dynamic Meteorology and Oceanography* 49 (4): 407–22. https://doi.org/10.3402/tellusa.v49i4.14680.

Short-range dynamics and prediction of mesoscale flow patterns in the MISTRAL field experiment

By RUDOLF O. WEBER*, PIRMIN KAUFMANN and PETER TALKNER, *Paul Scherrer Institute, CH–5232 Villigen PSI, Switzerland*

ABSTRACT

In a limited area of about 50 km by 50 km with complex topography, wind measurements on a dense network were performed during the MISTRAL field experiment in 1991/92. From these data, the characteristic wind fields were identified by an automated classification method. The dynamics of the resulting 12 typical regional flow patterns is studied in the present paper. It is discussed how transitions between the flow patterns take place and how well the transition probabilities can be described in the framework of a Markov model. Guided by this discussion, a variety of prediction models are described which allow a short-term forecast of the flow-pattern type. It is found that a prediction model which uses forecast information from the synoptic scale has the best forecast skill.

1. Introduction

For questions of emergency response planning and of air pollution control on a regional scale it is of great importance to know the local wind fields and their climatology in detail. In Weber and Kaufmann (1995) a method was developed to automatically classify local wind fields and to obtain classes of typical regional flow patterns. In Kaufmann and Weber (1996) this classification method was used to determine the characteristic flow patterns in an area around Basel, Switzerland. Only 12 classes of substantially different wind fields were found in that region with very complex topography. The dynamics of these twelve flow patterns is investigated following the ideas used in the analysis of synoptic weather types (Müller, 1961; Spekat et al., 1983; Van Dijk and Jonker, 1985; Mo and Ghil, 1987; De Swart and Grasman, 1987; Fraedrich, 1988; Nicolis, 1990; Vautard et al., 1990; Hannachi and Legras, 1995). Most of this discussion concentrates on the question whether the dynamics of synoptic weather types can adequately be described by a Markov model. An adequate description implies that the Markov model has good forecast skill.

As the weather types are characterized by a finite and discrete number, they can be described by a discrete state Markov model (a Markov chain). Such Markov chains were first used in meteorology to model the sequence of dry and wet days (see, for example, Gabriel and Neumann, 1962). For the modeling of precipitation, the original stationary 1st-order Markov models were extended to higher order (Bishnoi and Saxena, 1980; Gregory et al., 1992) and nonhomogeneous models (Rajagopalan, 1996). Similar higher order and nonstationary, continuous state, Markov models were developed to make short-term forecasts of wind speed (Brown et al., 1984; Huang and Chalabi, 1995).

For mesoscale wind fields, several authors have used automated classification schemes to obtain typical regional flow patterns (Hardy and Walton, 1978; Klink and Willmott, 1989; Green et al., 1992;

* Corresponding author.
e-mail: rudolf.weber@psi.ch

Kaufmann and Weber, 1996). So far, the dynamics of local flow patterns has not been investigated, although the possibility to make short-term forecasts of the flow patterns is of great importance for emergency response planning.

In Section 2, the twelve regional flow patterns of the area around Basel are presented. In Section 3, the transitions between the flow patterns are analyzed. Further it is discussed how well a Markov model can describe their dynamics. Several short-term forecast models are introduced and compared in Section 4.

2. Mesoscale flow patterns of the MISTRAL field experiment

In 1991/92, the MISTRAL field experiment ("Modell für Immissions-Schutz bei Transport und Ausbreitung von Luftfremdstoffen", model for impact prevention during transport and diffusion of air pollutants) took place in the region of Basel, Switzerland. This experiment is a part of REKLIP ("Regio-Klima-Projekt", Regio climate project) an international climatological project which takes place through the period 1989–1997 (Parlow, 1992). MISTRAL was aimed at measuring in detail the wind fields over a region of complex terrain. To this purpose 50 measurement sites with meteorological masts were operated for about 18 months (Kamber and Kaufmann, 1992). The MISTRAL area and the location of the masts are shown in Fig. 1. As the main objective of the meteorological measurements was the determination of regional flow patterns, the instruments were placed either on open space or on top of buildings at non-standard heights ranging from 5 to 69 m above ground. The anemometer at St. Chrischona ("C" in Fig. 1), operated by the Swiss Meteorological Institute, is located on a telecommunication tower at 262 m above ground. As the tower itself stands on a hill about 250 m above the Rhein valley, the wind observations from this station represent fairly well the larger-scale flow.

In Kaufmann and Weber (1996) a one year period from 1 September 1991 to 31 August 1992 was selected and 1-h means of the horizontal wind vectors were formed. The total of these 8784 observed wind fields was classified into 12 classes of typical regional flow patterns by means of an automated two-stage classification scheme (Weber and Kaufmann, 1995; Kaufmann and Weber 1996; Kaufmann, 1996).

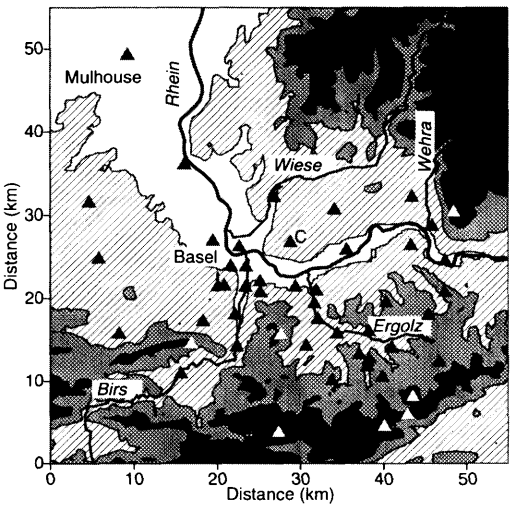

Fig. 1. The observation area of the MISTRAL project (55 km × 55 km) around the city of Basel with the 50 measurement sites (black and white triangles). Contour lines and shading give the height above sea level. White areas are below 300 m ASL, light shaded areas are 300 to 500 m ASL, medium shaded areas are 500 to 700 m ASL, dark shaded areas are 700 to 900 m ASL, and black areas are higher than 900 m ASL. Station locations higher than 700 m ASL are marked with a white triangle. The station labeled by a "C" is St. Chrischona with its anemometer placed on a tower 262 m above ground.

In a first step an agglomerative hierarchical cluster analysis (Anderberg, 1973) according to the complete linkage method was applied to find a reasonable number of clusters (12 in our case) and to obtain a first guess for the cluster membership of each wind field. The cluster analysis requires a distance matrix between all pairs of objects, with an object defined as the set of wind observations at all 50 stations at a specific time. The distance between the fields of horizontal wind observations at arbitrary times A and B is defined by

$$d_{AB} = \frac{1}{N} \sum_{j=1}^{N} [(\tilde{u}_{Aj} - \tilde{u}_{Bj})^2 + (\tilde{v}_{Aj} - \tilde{v}_{Bj})^2]^{1/2}, \qquad (1)$$

where j denotes the measurement sites (Weber and Kaufmann, 1995). The tilde in (1) indicates that the wind vectors are normalized in the following way. The mean absolute wind speed of a given

wind field at a specific time A is determined by

$$s_A = \frac{1}{N} \sum_{j=1}^{N} [u_{Aj}^2 + v_{Bj}^2]^{1/2}. \quad (2)$$

All individual wind vectors at the different locations j are divided by the mean speed of the wind field at a given time A:

$$\tilde{u}_{Aj} = u_{Aj}/s_A, \qquad \tilde{v}_{Aj} = v_{Aj}/s_A, \quad (3)$$

which yields wind fields with dimension-less wind speeds.

In a 2nd step, the classification in 12 classes was iteratively refined by a k-means method (Anderberg, 1973). To obtain the relevant features of the regional flow patterns more clearly, outliers (wind fields not well fitting in any of the groups) were excluded in this 2nd step (Kaufmann and Weber, 1996). A total of 1531 outliers was detected and excluded, and only the remaining 7253 wind fields were used to calculate class means of normalized wind vectors (3) and the distribution of wind direction within a class. To justify the choice of 12 classes, Kaufmann and Weber (1996) also discuss classifications with 7 and 16 classes, respectively. It was shown that in the case of 7 classes distinct flow patterns are merged into a single class, whereas in the case of 16 classes too many classes differ by very small details only.

For the investigation of the dynamics of the classes it is, however, more convenient to have a contiguous time series of classes. Therefore, the 1531 outlier wind fields are assigned to a class by the following allocation procedure (Kaufmann, 1996). For each wind field the distance (1) to all class means was calculated and the wind field was assigned to the class with the smallest distance. Table 1 summarizes the properties of the 12 classes after the allocation of all outliers. A comparison with Table 2 of Kaufmann and Weber (1996) shows that the wind speeds of the classes have become smaller after the allocation of the outliers. This reflects the fact that the outlier wind fields have in general small wind speed. Classes 1 and 6 still have the highest wind speeds. Fig. 2 shows the wind roses of the 12 classes after assigning of the outliers. A comparison with Fig. 9 of Kaufmann and Weber 1996 shows that the flow patterns of the classes remain very close to those obtained under exclusion of the outliers. The 12 flow patterns can be grouped according to four main wind directions of the St. Chrischona tower station (denoted by a "C" in the figures). Classes 1, 3, 6, and 7 have westerly wind on St. Chrischona tower. For classes 4, 8, 9, 10, and 11 the wind at St. Chrischona blows from the east, for classes 5 and 12 from the north, and for class 2 from south. The prevailing wind direction of each class is given in the 6th column of Table 1. A more detailed discussion of the 12 typical regional flow patterns can be found in Kaufmann and Weber (1996) and Kaufmann (1996). Some of the flow patterns show a pronounced diurnal cycle, reflecting the thermal forcing of local wind systems (see also Whiteman, 1990). The last column of Table 1 indicates whether the classes occur mainly during daytime (D) or nighttime (N). The nighttime flow patterns (2, 3, 4, 7 and 8) show strong cold air drainage flows in the valleys (Fig. 2).

3. Dynamics of the flow patterns

Based on the classification of the 8784 1-h wind fields of the MISTRAL experiment into 12 classes of typical regional flow patterns, we can analyze the dynamics of these patterns. Of main interest is the question whether there exist simple rules governing the transitions between the flow patterns. Such rules would allow to make short-term forecasts of the flow patterns.

To elucidate this question the transition probability matrix

$$\boldsymbol{P}(k) = (p_{ij}(k)), \quad (4)$$

is introduced. A matrix element $p_{ij}(k)$ gives the probability to get from the jth flow pattern to the ith flow patterns in k h. As one of the 12 flow patterns must occur at each time, the columns of the matrix (4) are normalized according to

$$\sum_{i=1}^{12} p_{ij}(k) = 1. \quad (5)$$

Table 2 gives the transition probabilities for a time lag of 1 h ($k = 1$). The diagonal elements of the matrix give the probability that the flow pattern remains the same one hour later, reflecting the persistence of the wind fields. This persistence ranges from 0.63 for class 7 to 0.80 for class 8. Class 7 has also a high probability of 0.20 to change into class 3, which has a similar flow pattern (Fig. 2) but more southwesterly than westerly winds. The transition from class 6 to 1 has another high probability, both classes representing

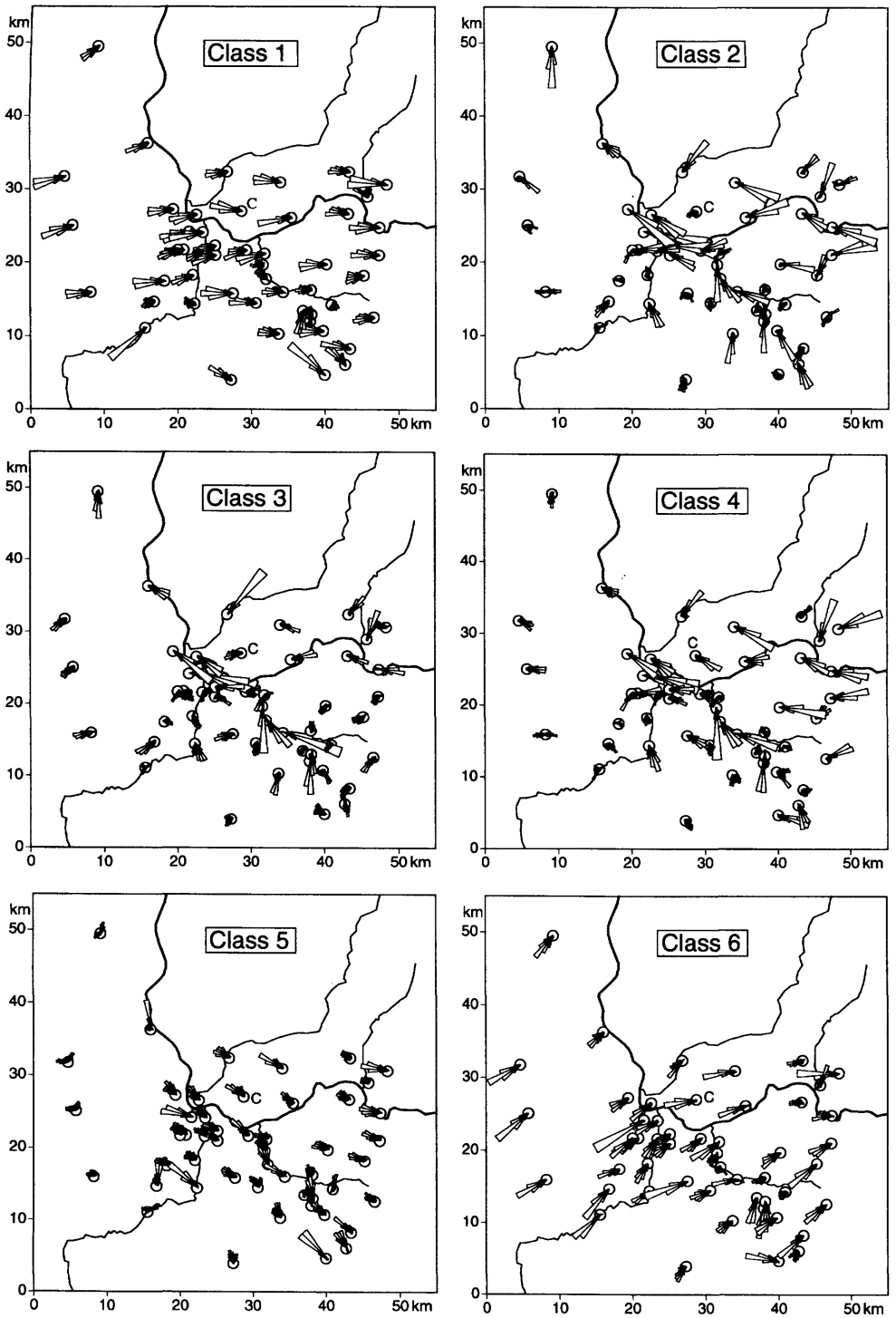

Fig. 2. Wind roses at all stations for the 12 classes of flow patterns after the allocation of all outliers to classes. The wind roses are divided in sectors of 10° whose length is proportional to the frequency. The circle around each measurement site represents the 5% frequency level. The "C" labels the station on the St. Chrischona tower.

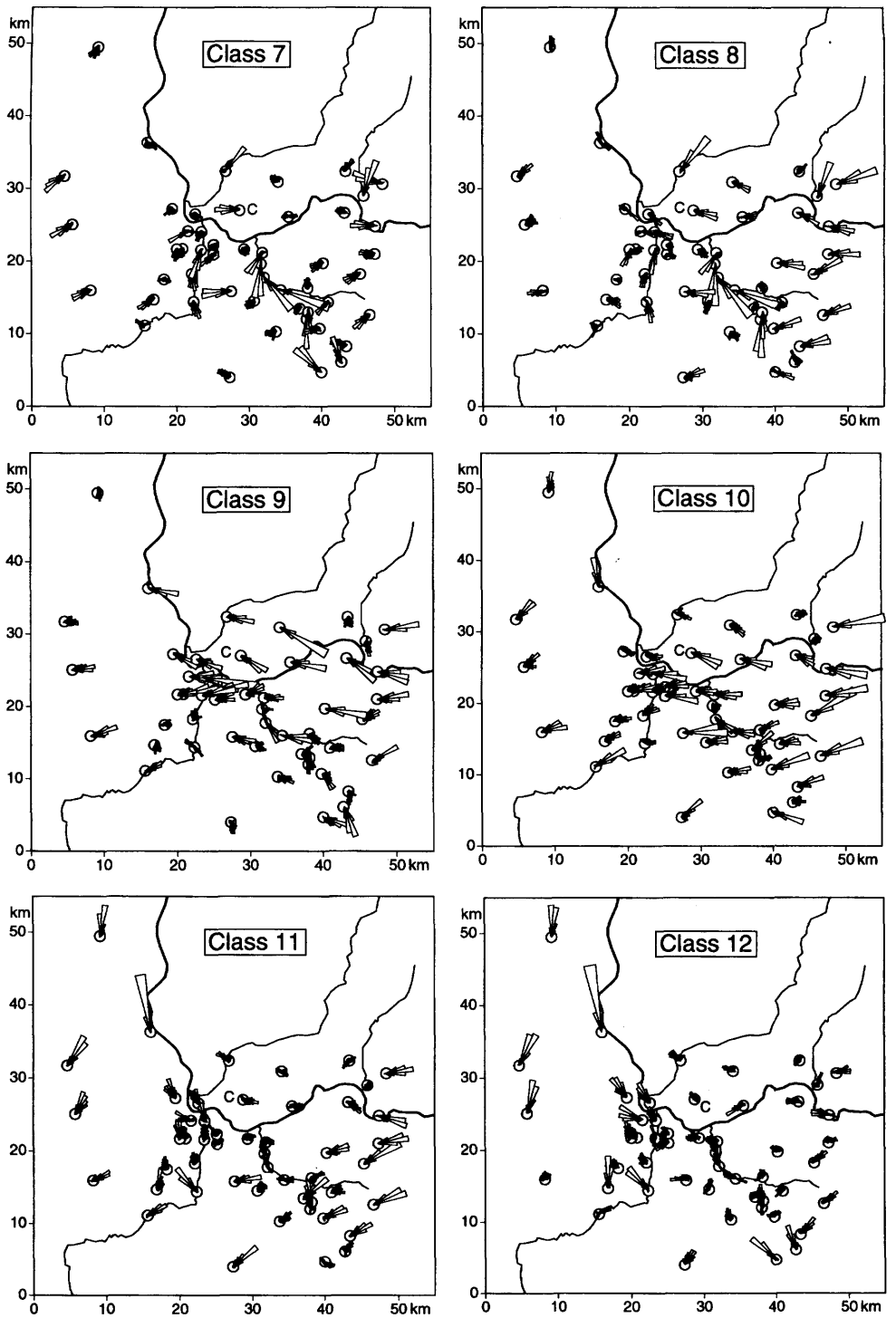
Class 7
Class 8
Class 9
Class 10
Class 11
Class 12
km
50
40
30
20
10
0
0
10
20
30
40
50 km
C

Fig. 2. (cont'd).

Table 1. *Summary of the 12 classes of all 1-h wind fields from 1 September 1991 to 31 August 1992 after the allocation of all outliers to classes*

Class	Counts	Relative frequency (%)	Mean wind speed ($m\,s^{-1}$)	Median wind speed ($m\,s^{-1}$)	Wind direction at station C	Day or night class
1	1230	14.0	4.0	3.7	W	—
2	908	10.3	2.2	2.1	S	N
3	901	10.3	1.8	1.7	W	N
4	686	7.8	2.2	2.2	E	N
5	799	9.1	2.4	2.2	N	D
6	590	6.7	4.4	4.0	W	—
7	694	7.9	1.7	1.7	W	N
8	713	8.1	1.7	1.7	E	N
9	596	6.8	2.6	2.6	E	D
10	539	6.1	3.2	3.0	E	D
11	583	6.7	2.2	2.0	E	D
12	545	6.2	1.9	1.8	N	—
total	8784	100	2.6	2.2		

Table 2. *Transition probabilities (%) for a time lag of 1 h*

To class	From class 1	2	3	4	5	6	7	8	9	10	11	12
1	**78**	*0*	*1*	*0*	12	15	*8*	*0*	*1*	*0*	*0*	*2*
2	*0*	**79**	**14**	*6*	*1*	*1*	*0*	*0*	*2*	*0*	*0*	*0*
3	*0*	8	**70**	*2*	*1*	*4*	**20**	*2*	*0*	*0*	*0*	*0*
4	*0*	*4*	*1*	**78**	*0*	*0*	*0*	7	7	*1*	*0*	*0*
5	*6*	*1*	*2*	*0*	**74**	*2*	*2*	*0*	*3*	*1*	*3*	8
6	7	*1*	*3*	*0*	*1*	**73**	*4*	*0*	*0*	*0*	*0*	*1*
7	8	*1*	8	*0*	*4*	6	**63**	*1*	*0*	*0*	*0*	*2*
8	*0*	*0*	*1*	*4*	*0*	*0*	*2*	**80**	*1*	7	*4*	*5*
9	*0*	*4*	*1*	8	*0*	*0*	*0*	*1*	**77**	5	*0*	*0*
10	*0*	*0*	*0*	*1*	*1*	*0*	*0*	*3*	4	**79**	**9**	*1*
11	*0*	*0*	*0*	*0*	*1*	*0*	*0*	5	5	7	**73**	8
12	*1*	*0*	*0*	*0*	6	*0*	*0*	*2*	*0*	*0*	**11**	**74**

Each matrix element p_{ij} gives the transition probability ($\times 100$) to get in one hour from class j to class i. Due to rounding errors the values of a column do not exactly sum up to 100. Transitions occurring more (less) likely than by mere chance at the 1% significance level are indicated by boldface (italics), according to the Monte Carlo test of Vautard et al. (1990).

flow patterns with strong westerly winds (Table 1, Fig. 2). A Monte Carlo method was used to determine confidence intervals of the transition probabilities. Starting from an arbitrary class, the transition matrix of Table 2 was used to randomly generate a time series of equal length (8784) as the original time series of class memberships. A total of 10,000 such random time series was generated. From these time series the transition probabilities were calculated and the 90% confidence intervals determined. The diagonal element of class 8 has a 90% confidence interval of (0.77, 0.82). The element p_{37} lies in the range 0.17 to 0.22, whereas a small matrix element like p_{12} ($=0.002$) lies in the interval (0.0, 0.005), hence has a much larger relative error.

To assess the statistical significance of the elements of the transition matrix, the Monte Carlo

method of Vautard et al. (1990) was used. The null hypothesis of this test is a random sequence of class memberships with fixed stationary probabilities of each class (3rd column of Table 1). Thus, the transition to any class i does not depend on the initial class j, and is given by the stationary probability p_i of class i. The following Monte Carlo method (Vautard et al., 1990) was used to test the null hypothesis. A randomly permuted time series of class numbers from 1 to 12 was generated under the restriction that each class occurs as often as in the observed data (2nd column of Table 1). This is repeated 10,000 times and the transition matrices of these random sequences are calculated. Then it is counted how many times a matrix element from the 10,000 random time series is larger than or equal to the observed matrix element p_{ij}. For a transition to be more likely than by mere chance at a significance level of 1%, this number must be less than 100. In the same way the number of times is counted when a matrix element is less than or equal to the observed matrix element, allowing to identify those transitions which are less likely than by mere chance. The transitions which are more (less) likely than by mere chance on the 1% significance level are indicated by boldface (italics) in Table 2.

To analyze the transitions from one flow pattern to the others, a *conditional* transition probability is defined as

$$\tilde{p}_{ij}(k) = \frac{p_{ij}(k)}{[1 - p_{jj}(k)]}, \tag{6}$$

which gives the probability to get in k h from class j to class i under the condition that a transition to another class takes place. The conditional transition probabilities for a time lag of one hour ($k = 1$) are listed in Table 3. Several pairs of classes like 2, 3 or 4, 9 exist, which have high conditional transition probabilities for transitions in both directions. The conditional transition probability to get from class 3 to class 2 is 0.45 and the reverse transition has a 0.39 chance. The same Monte Carlo method as above was used to determine 90% confidence intervals of the conditional transition probabilities. For p_{23} (=0.45) the confidence interval is (0.40, 0.50), for $p_{8,12}$ (=0.20) it is (0.15, 0.26). Again, large transition probabilities are more accurately determined than smaller ones. The significance of transitions that are more (less) likely than by mere chance on the 1% significance level is indicated by boldface (italics) in Table 3.

In Fig. 3, all conditional transition probabilities $\tilde{p} \geqslant 0.2$ are drawn as arrows between the classes. The flow patterns with westerly winds (1, 3, 6 and 7) form a group with mutual transitions between them. Similarly, the classes with easterly

Table 3. *Conditional transition probabilities (%) for a time lag of 1 h*

To class	From class 1	2	3	4	5	6	7	8	9	10	11	12
1	—	*1*	*4*	*1*	**47**	**55**	**22**	*2*	*2*	*0*	*1*	*8*
2	*0*	—	**45**	**27**	*2*	*3*	*1*	*0*	10	*0*	*0*	*0*
3	*2*	**39**	—	*7*	*3*	13	**54**	*8*	*2*	*2*	*0*	*0*
4	*0*	**19**	*5*	—	*0*	*1*	*0*	**34**	**29**	*5*	*1*	*1*
5	**27**	*7*	*7*	*0*	—	*6*	*6*	*1*	13	*3*	9	**31**
6	**31**	6	9	*2*	*3*	—	10	*0*	*2*	*1*	*1*	*3*
7	**36**	*4*	**26**	*2*	**14**	**21**	—	*5*	*0*	*0*	*0*	7
8	*0*	*1*	*2*	**19**	*1*	*1*	*5*	—	*2*	**32**	**14**	**20**
9	*0*	**21**	*2*	**36**	*1*	*0*	*0*	*3*	—	**24**	*1*	*0*
10	*0*	*1*	*1*	*4*	*2*	*0*	*0*	**14**	**18**	—	**33**	*2*
11	*0*	*1*	*0*	*2*	*3*	*0*	*1*	**25**	**21**	**33**	—	**29**
12	*4*	*0*	*0*	*1*	**22**	*1*	*1*	8	*2*	*1*	**41**	—

Each matrix element $\tilde{p}_{ij}$ gives the conditional transition probability ($\times 100$) to get in 1 h from class j to class i under the condition that a transition to another class takes place. Transitions occurring more (less) likely than by mere chance at the 1% significance level are indicated by boldface (italics), according to the Monte Carlo test of Vautard et al. (1990).

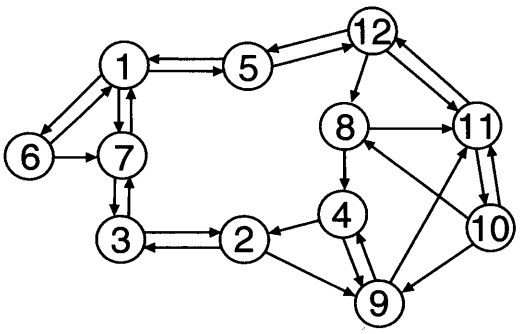

Fig. 3. Transition diagram for the 12 regional flow patterns. The arrows indicate one hour transitions with a conditional probability $\tilde{p} \geq 0.2$.

winds (4, 8, 9, 10 and 11) build a group. The transition between the westerly and the easterly flow types passes through the southerly (2) or the two northerly (5 and 12) flow patterns. Table 3 shows that other transitions between the westerly and easterly flow types are less likely than by mere chance. In Kaufmann (1996) all transitions between classes whose probability (from Table 2) exceed 0.05 are shown. The picture is nearly the same as in our Fig. 3.

In Table 4 the transition probabilities for a 12-h time lag are listed. Some of the flow patterns (1, 2, 3, 10 and 12) still show a persistence as their highest transition probability is on the diagonal. For classes 5 and 7 the diagonal element is close to the stationary probability (Table 1), which seems to indicate that these flow patterns occur independently after 12 h. The classes 6 and 7 mainly change to class 1 after 12 h, thus slightly turning to more northerly winds. Class 5 turns to class 3, changing from a daytime flow pattern to a night-time pattern (Table 1). An even more pronounced change between day- and night-time patterns is seen in the pairs 4, 9 and 8, 11, which have their highest probabilities for mutual transitions. For both classes 4 and 9 the second highest probability is to turn to class 2, which indicates a change from the easterly winds of classes 4 and 9 to the southerly winds of class 2. Classes 8 and 11 also form a pair of day- and night-time classes, and tend to turn to class 10. Although the pairs 4, 9 and 8, 11 have the same diurnal cycle and similar wind directions, they do not mix. If strong winds aloft and weak thermal winds in the valleys prevail, classes 8 and 11 alternate, for weaker winds aloft, but strong thermal winds, the classes 4 and 9 prevail (Kaufmann, 1996).

The behavior of the transition probabilities for larger time lags is discussed by showing a few representative examples. Fig. 4 shows the transition probability to go from class 1 to class 1 (solid curve) as a function of the time lag. Class 1

Table 4. *Transition probabilities (%) for a time lag of 12 h*

To class	From class 1	2	3	4	5	6	7	8	9	10	11	12
1	**34**	*11*	**17**	*4*	*10*	**34**	**23**	*4*	*2*	*2*	*2*	*5*
2	*6*	**28**	**13**	**14**	**13**	8	9	*4*	**14**	*4*	*2*	*2*
3	**14**	**14**	**17**	*6*	**19**	10	9	*2*	**13**	*3*	*3*	*2*
4	*2*	*5*	*2*	**13**	6	*1*	*4*	8	**30**	**13**	**11**	9
5	10	11	**6**	10	8	*3*	**20**	9	*2*	*1*	*2*	12
6	**13**	8	**10**	*2*	*3*	**32**	7	*0*	*0*	*1*	*0*	*1*
7	**13**	9	8	*4*	**17**	9	7	*2*	9	*2*	*5*	*5*
8	*2*	*1*	*2*	7	**10**	*1*	*2*	**16**	9	**19**	**26**	**18**
9	*1*	7	7	**20**	*1*	*2*	6	9	**16**	**12**	5	*2*
10	*1*	*2*	*3*	**10**	*1*	*0*	*2*	**16**	*1*	**31**	**14**	6
11	*2*	*2*	*3*	7	*3*	*0*	6	**19**	*2*	**10**	**21**	**14**
12	*4*	*1*	*4*	*4*	**9**	*0*	5	**13**	*2*	*2*	**11**	**26**

Each matrix element p_{ij} gives the transition probability ($\times 100$) to get in 2 h from class j to class i. The highest probability of each column is underlined. Transitions occurring more (less) likely than by mere chance at the 1% significance level are indicated by boldface (italics), according to the Monte Carlo test of Vautard et al. (1990).

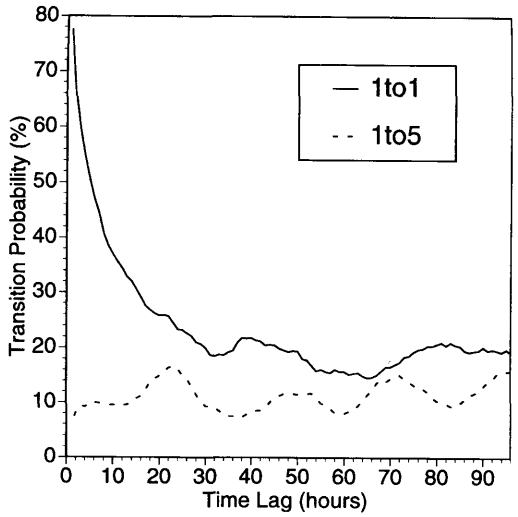

Fig. 4. Transition probability to get from class 1 to class 1 (solid line) and from class 1 to class 5 (dashed line) as a function of time lag.

with its strong westerly winds is a very stable class and decays only slowly. All transitions to the other classes have probabilities below p_{11} (solid line). An example is given by p_{51} (dashed line) whose oscillations reflect the diurnal cycle of the day-time flow pattern 5. Fig. 5 gives the transition probabilities to go from class 11 to class 8 (dashed line), to class 10 (dashed–dotted line) or to class 11 (solid line). The probability $p_{11,11}$ decays quite fast in the first 12 h and shows a period of 24 h as to be expected for a flow pattern occurring mainly during day. After 12 h, the transition to class 8 becomes most probable, reflecting the fact that classes 8 and 11 are alternating night- and daytime classes, respectively. The transition to class 10 shows maxima near multiples of 24 h, reflecting the fact that both are day-time flow patterns. Fig. 6 shows the transition probabilities to go from class 6 to class 1 (dashed line) or to class 6 (solid line), both classes with strong westerly winds. After 12 h, the class 6 has turned to class 1 with higher probability than to remain in 6.

Forecast models for the occurrence of the wind field classes would become much simpler if the classes could be adequately described by a Markov chain model (see, for example, Cox and Miller, 1965). The class number, with its discrete values from 1 to 12, is here considered as a random variable. Neglecting for the sake of simplicity diurnal and seasonal effects, the Markov property means that the transition probability matrix with time lag one ($k = 1$) determines all other transition matrices with larger time lag ($k > 1$) by

$$\boldsymbol{P}(k) = \boldsymbol{P}^k(1). \tag{7}$$

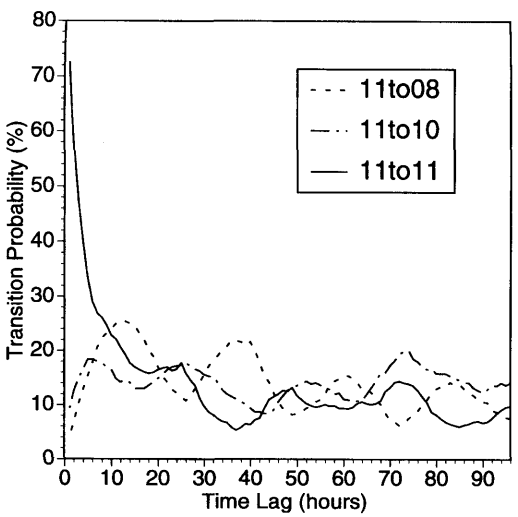

Fig. 5. Transition probability to get from class 11 to class 11 (solid line), to class 8 (dashed line), or to class 10 (dashed–dotted line) as a function of time lag.

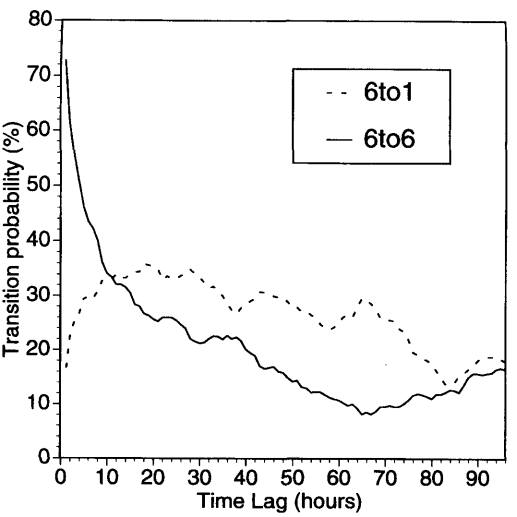

Fig. 6. Transition probability to get from class 6 to class 6 (solid line) and from class 6 to class 1 (dashed line) as a function of time lag.

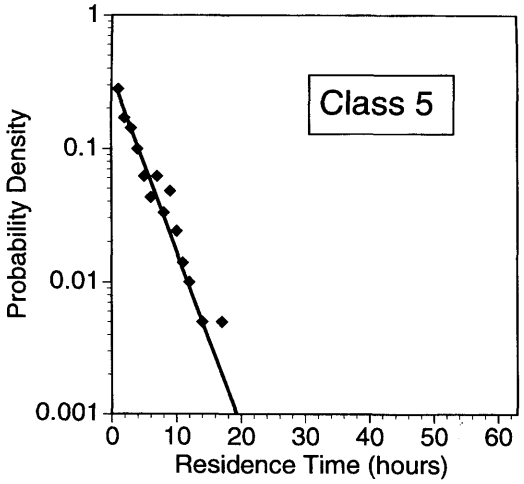

Fig. 7. Distribution of residence times for flow pattern class 5. The solid line represents the distribution of a Markov chain model (8).

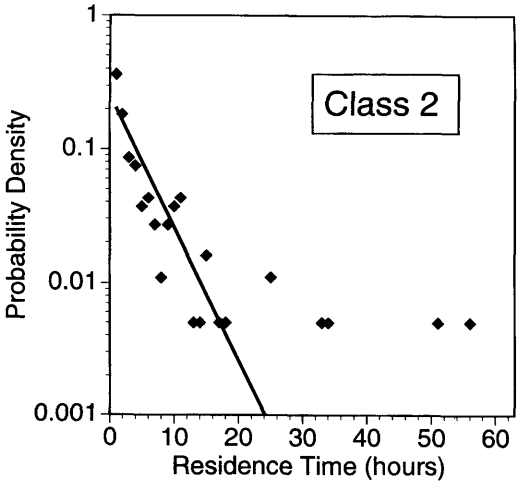

Fig. 8. Distribution of residence times for flow pattern class 2. The solid line represents the distribution of a Markov chain model (8).

Under the Markov assumption, it is possible to infer the distribution of the length of the flow pattern episodes (residence times in the classes) from the diagonal elements of $\boldsymbol{P}(1)$. The probability that class j uninterruptedly lasts m h becomes

$$\text{prob}_j(m) = p_{jj}^{m-1}(1 - p_{jj}). \tag{8}$$

This probability can directly be estimated from the sequence of class numbers by simply counting the episodes of all lengths. Fig. 7 shows the distribution of the residence times of class 5 as obtained from the data and from model (8) which makes only use of the estimated $p_{55}(1)$. The agreement between data and model is very good and we conclude that a Markov chain can well describe the time series of occurrences of class 5. Among all classes, class 5 is the one which is best approximated by a Markov chain model. The poorest coincidence between observed and modeled distribution of residence time is seen for class 2 (Fig. 8). For short times, less than 20 h, large scatter is already present between the observed residence times and the modeled ones. Very long episodes (up to 56 h) of this class occur, causing strong deviations from the Markov chain model and indicating an intermittent behavior of the occurrence of class 2. Similarly, long episodes (on time scales of days rather than hours) were observed by Müller (1961) and van Dijk and Jonker (1985) in their analyses of synoptic weather types.

The classes 1, 6, 10 and 11 also show long episodes of the same class. All episodes with residence times greater than 24 hours were selected to see whether the long periods show specific meteorological conditions. The mean wind speed of the long episodes was calculated and compared to the mean wind speed of the whole class in Table 5. The means are compared by Satterthwaite's test, which allows to compare means in case of unequal variances (Snedecor and Cochran, 1989) and uses a t-distributed test variable with adjusted degrees of freedom. Except for class 11 the long episodes have a significantly higher wind speed than the average of the classes. Class 11 has only 2 long episodes of 25 and 26 h length, just above the arbitrarily chosen limit of 24 h. Hence, these two episodes do not differ much

Table 5. *Mean wind speed for the classes and long episodes*

	Wind speed (m/s)		
Class	All cases	Long episodes	p-value
1	4.0	5.4	<0.00001
2	2.2	2.8	<0.00001
6	4.3	5.6	<0.00001
10	3.2	4.1	<0.00001
11	2.2	2.3	0.12

Table 6. *Mean distance (1) to the class mean for the classes and long episodes*

	Distance		
Class	All cases	Long episodes	*p*-value
1	0.41	0.29	<0.00001
2	0.54	0.48	<0.00001
6	0.47	0.34	<0.00001
10	0.49	0.39	<0.00001
11	0.64	0.51	<0.00001

from the other cases of class 11. In Table 6 the distance (1) to the class mean is shown for the classes with long episodes. For all 5 classes, the distance of the long episodes is significantly smaller than for the whole class. This shows that the wind fields of the long episodes represent the class mean well.

4. Prediction models

In this Section we develop several models to predict the flow pattern class some hours ahead from the knowledge of the present class or classes in the past. An improvement over these models can be gained if additional information about changes on the synoptic scale are included in the model.

First, the most simple prediction model is assuming that the wind field type remains the same in the near future, thus

$$C(t+k) = C(t), \tag{9}$$

where $C(t)$ is the class number of the flow pattern at time t. Landberg and Watson (1994) found this persistence model to be the best for forecasting station wind speeds over times up to 6 h.

In the 2nd model, the persistence of flow patterns is neglected and a prediction is made according to climatology. As class 1 has the highest frequency (Table 1), this class is used for each forecast in a climatology model

$$C(t+k) = 1. \tag{10}$$

In the 3rd model, we use a Markov chain model. As class 5 and other classes show an almost exponential distribution of residence times close to the ones a Markov chain model, a first order Markov chain model at time step k is built. To predict the class at time $t+k$ the transition matrix $\boldsymbol{P}(k)$ of time lag k is needed. This matrix can be obtained in 2 ways. In the first variant it is assumed that $\boldsymbol{P}(1)$ determines all other transition matrices and $\boldsymbol{P}(k)$ is obtained by (7). The class at time $t+k$ is predicted by determining the highest transition probability from class j at time t

$$C(t+k) = i \quad \text{such that} \quad p_{ij}(k) = \max. \tag{11}$$

As only 1-h transition probabilities are directly determined from the data, we call this model 1-h Markov model.

In the 4th model, we estimate the lag k transition matrix in a different way. Instead of using (7) to estimate $\boldsymbol{P}(k)$, the lag-k transition matrix can also be directly estimated from the data (denoted $\boldsymbol{Q}(k)$). The prediction is again based on the maximum probability

$$C(t+k) = i \quad \text{such that} \quad q_{ij}(k) = \max. \tag{12}$$

This model is termed k-h Markov model as the transition matrix at lag k is directly estimated from the data. Huang and Chalabi (1995) use similar 2 types of estimating the k-order autoregression coefficient in their models to predict single station wind speed.

In the 5th model, additional information besides the knowledge of the initial class is needed. As discussed above, some of the flow patterns show a distinct diurnal cycle and have thermally induced wind systems. In principle, this can be taken into account within the framework of Markov chains with time dependent transition probabilities (Huang and Chalabi, 1995; Rajagopalan et al., 1996) The diurnal effects on our flow patterns classes depend on the seasonal effects which cannot be estimated from 1 year of data. From Fig. 6 of Kaufmann and Weber (1996) the changing times between night- and day-time flow pattern were determined for each month of the year (Table 7). We distinguish 2 kinds of transitions: (1) from day to day or from night to night; (2) from day to night or from night to day. This gives a coarse description of the diurnal cycle of the transition.

As some of the local windsystems determining the flow patterns are thermally forced, it is important whether a day is sunny. For this reason the automatic sunshine duration measurements of the station at Basel–Binningen (operated by the Swiss

Table 7. *Time of change between night and day classes*

Month	Night/day	Day/night
January	12	18
February	11	18
March	10	19
April	9	20
May	8	20
June	8	20
July	8	20
August	9	19
September	9	18
October	9	18
November	10	17
December	11	17

The times are given in hours CET, where the hour given in the 2nd column still belongs to the night, and the hour given in the 3rd column still belongs to the day.

Meteorological Institute) are used. If more than 3 h of sunshine are registered for a day this day is declared as sunny.

A further factor determining the dynamics of the flow patterns, is a change of the synoptic scale wind. As an indication of such changes, the change of wind direction on the St. Chrischona tower is considered. If the wind direction remains constant during the forecast time within $\pm 10°$ the synoptic wind direction is declared as constant. Otherwise we distinguish whether a clockwise or a counterclockwise turn takes place, giving thus three cases of wind direction changes.

Combining the day/night distinction, sunshine and change of wind direction, a total of 12 cases is obtained. The matrices of transitions between the flow patterns given one of the 12 combinations of external parameters are calculated for different time lags. The forecast for a transition from class j is obtained as before by the maximum probability:

$$C(t+k) = i \quad \text{such that} \quad \hat{p}_{ij}(k) = \max, \qquad (13)$$

where $\hat{p}_{ij}(k)$ denotes the transition from state j to i in k h for given values of the day/night distinction, sunshine and change of wind direction. If the highest probability occurs simultaneously at several places in a column of the transition matrix, the diagonal element is preferred (if it is among the highest) or the stationary probability is used to determine the forecast. A similar division into cases was used by Bhan et al. (1994) who built classes according to station pressures and pressure differences in order to predict whether wind speed exceeds a threshold value.

To measure the accuracy of the forecast, a simple forecast skill is defined as the ratio of the number of correctly predicted classes to the number of all predictions made. Fig. 9 shows the forecast skill for the five models described above. The worst prediction is obtained by the climatology model (10), which always predicts class 1, independently of the initial class and the prediction time. The persistence model gives for short prediction times, up to 4 h, as accurate forecasts as the remaining three models do. The forecast skill of the persistence model (9) has a 24-h period, reflecting the diurnal cycle of most flow patterns. The forecast skill of the 1-h Markov model (11) follows closely the persistence model up to about 15 h prediction time. After 24 h, the product matrix (7) has become close to the stationary matrix and the forecast becomes equal to the climatological forecast. The k-hour Markov model (12) gives slightly better predictions in the 10–20 h range

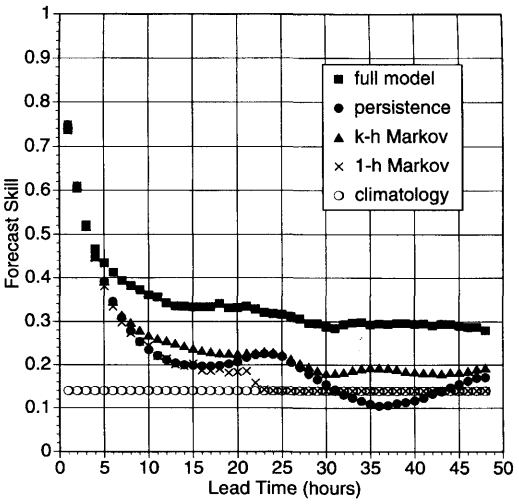

Fig. 9. Forecast skill of various prediction models for the flow pattern class. The full model (13) makes use of information about day/night changes, of the sunshine duration and of changes of synoptic scale wind direction, hence it needs forecasts on the synoptic scale. The other 4 models, the persistence model (9), the climatology model (10), the 1-h Markov model (11) and the k-h Markov model (12), make only use of the knowledge of the initial state.

than the 1-h Markov model and the persistence model. For prediction times longer than 5 h the full model (13) has significantly better forecast skills than any of the other 4 models.

Fig. 10 shows which of the synoptic scale properties is most important for the noticed improvement of forecast skill. Three models with either the day/night distinction or the sun/no sun distinction or the three cases of wind direction change are compared to the full model. If only changes of wind direction are taken into account in model (13), the best improvement of forecast skill compared to k-h Markov model (12) is obtained. However, the full model is still considerably better than any of the three models using only part of the synoptic information. If only the external parameters (day/night transition, sunshine, and wind direction changes) are used, but not the knowledge of the actual class, the forecast skill lies between 0.15 to 0.20 and is well below that of the full model. This low forecast skill is due to the fact that in this model only changes of the external parameters are used and not he absolute values of the external parameters, which would provide more information. This is done for comparison with the full model, where also changes of external parameters are considered. A model using the values of external parameters at prediction time

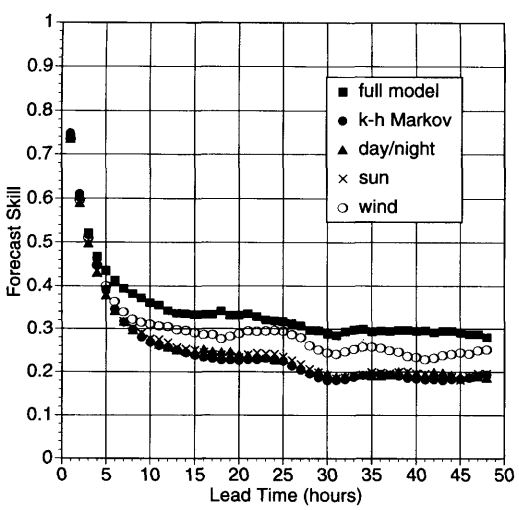

Fig. 10. Forecast skill for the full model (13) and 3 simplified models, which have only information about the day/night changes or the sunshine or the large scale wind direction changes included.

was also built. Wind direction at St. Chrischona (4 classes W, N, E, S were used), the distinction between sunny and cloudy days as defined above, and the information of day or night were used, giving 16 different cases. The forecast skill of this model is 0.34, independent of the time lag. This is, for large time lags, better than the full model. However, it must be kept in mind that the forecasts of the external parameters themselves are subject to a certain forecast error which will reduce the forecast skill in practical applications.

In conclusion it was demonstrated that the full model gives, for moderate forecast periods, the best forecast skill among a variety of models. However, the full model needs additional information. The day/night distinction depends only on the month and time of day of the initial state and the time of the forecast and can thus always be given according to Table 7. The distinction into sunny and not sunny days and days with or without changes of synoptic scale wind direction needs a forecast from a numerical weather prediction model, that operates on a larger scale.

The class number of the wind field at a given time indicates to which flow pattern this wind field belongs. However, as the classification analysis is based on normalized wind fields (3), additionally the field-mean wind speed (2) of the wind field must be known to give wind vectors in physical units. Therefore, several prediction model for the field-mean wind speed (2) are constructed and compared.

The most simple model is based on persistence

$$s(t+k) = s(t), \tag{14}$$

where $s(t)$ is the field-mean wind speed (2) over all 50 stations at time t.

A k-h Markov model for the prediction of the field-mean wind speed is defined by

$$s(t+k) = c(k)s(t), \tag{15}$$

where $c(k)$ is the autocorrelation coefficient of lag k h. Fig. 11 shows the autocorrelation function of the field–mean wind speed (2) for time lags up to 96 h. The correlation decays quite fast in the first 10 h, but remains for longer time lags at a non-zero level, in agreement with the behavior of station wind speeds (Brett and Tuller, 1991). A weak diurnal cycle is reflected in the shoulder at 24 h and the maxima at 48, 72 and 96 h.

As the distribution of the field–mean wind

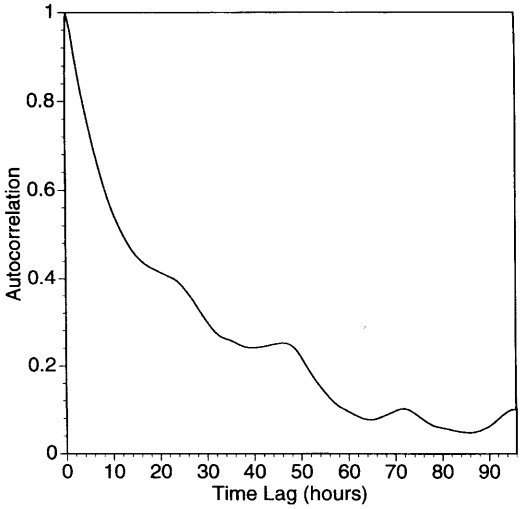

Fig. 11. Autocorrelation function of the field-mean wind speed (2) as a function of time lag.

speeds (2) is not Gaussian, the more refined model of Brown et al. (1984) was also used. In this model the wind speed is replaced by its square root, a transformation which ensures nearly Gaussian distributed variables. For each hour of the day, the mean is separately subtracted to remove the diurnal cycle of wind speed. Autoregressive models of order 1, 2, 3 and 4 where fitted to the transformed time series.

More refined models using non-stationary autoregressive models do exist (Huang and Chalabi, 1995), but were not considered here, as our field–mean wind speed does not show pronounced diurnal and seasonal cycles as Fig. 12 demonstrates, where the diurnal and seasonal distribution of the field–mean wind speeds are shown.

The models were tested by predicting the wind speed k h ahead taking all hours as starting points. The resulting predicted time series was compared with the observed time series by calculating the cross-correlation between them and by calculating the root mean square error of the predictions. It turned out that the Markov model (15) gives as good forecasts as the more complicated models do. For prediction times up to 6 h, even the simple persistence model (14) is as accurate as the other models in agreement with the findings of Landberg and Watson (1994) who developed several models to predict station wind speeds.

5. Conclusions

The 12 typical regional flow patterns observed in the MISTRAL area (Kaufmann and Weber, 1996) were used to allocate all 1-h mean wind fields of a whole year to a flow pattern class. The dynamics of the twelve classes, or regional flow patterns, was studied in the present paper. We showed that the transitions between the classes follow certain rules, depending on the time-scale considered. On the time-scale of 1 h, starting from a westerly class, a transition to a northerly and subsequently to an easterly and finally to a southerly class occurs with highest probability. Beginning with a southerly class transitions occur in reverse order with highest probability. Direct transitions from a southerly to a northerly or from an easterly to a westerly type have a very low probability to occur. On a time scale of 12 h pairs of flow patterns with similar winds aloft, but different valley wind systems were found. These classes represent thermally driven wind systems in the valleys with up-valley flow during daytime and down-valley flow during nighttime within the same synoptic-scale flow.

The length of the episodes with the same flow pattern class can well be described by a Markov chain model, except for very long episodes occurring for certain classes. According to the Markov chain model episodes of length 20 h or longer have a negligible probability to occur, whereas for some classes such very long episodes can be observed.

Guided by the discussion of the dynamics of flow patterns, several prediction models were developed that allow short-term forecasts of the flow pattern class. One group of models only uses the knowledge of the initial class and of the statistics of the observed time series of class numbers. Models allowing for a persistence of the classes, and for the climatology of the classes and a Markov chain model were discussed. Another model using additional information from synoptic-scale forecasts was introduced. The latter model showed the best forecast skill for prediction times longer than a few hours.

Similar models were used to predict the mean wind speed of the wind fields. A simple Markov model turned out to be sufficient for predictions up to lead times of 24 h. For shorter prediction times up to 6 h even a persistence model performed as well as the more complicated models.

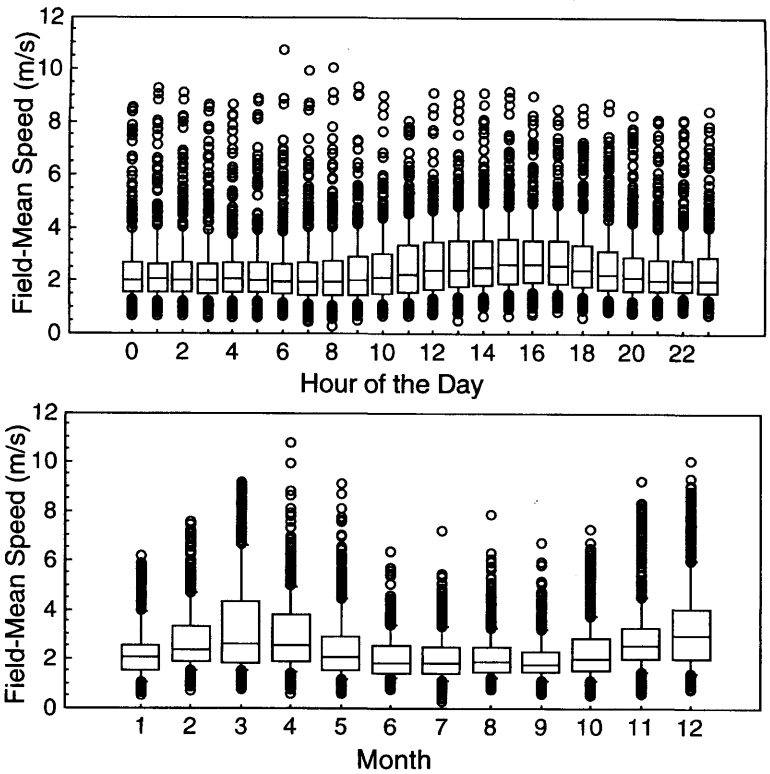

Fig. 12. Boxplots of the field-mean wind speed (2) as a function of hour of the day (upper chart) and month (lower chart). The 3 horizontal lines of the boxes give the 25th, 50th and 75th percentiles of the variable, the whiskers indicate the 10th and the 90th percentiles. All values below the 10th percentile or above the 90th percentile are plotted as circles.

6. Acknowledgements

The project REKLIP/MISTRAL is partly funded by the two cantons of Basel. Additional data were kindly provided by the Swiss Meteorological Institute, Zürich and the Geographical Institute of the University of Basel. The authors thank the referee for bringing the significance test for transition matrices of Vautard et al. (1990) to their attention.

REFERENCES

Anderberg, M. R. 1973. *Cluster analysis for applications.* Academic Press, 359 pp.

Bhan, S. C., Roy Bhowmik, S. K. and Sharma, R. V. 1994. An objective technique for forecasting wind speed over Bombay high area. *Mausam* **45**, 271–278.

Bishnoi, O. P. and Saxena, K. K. 1980. Precipitation simulation process with Markov chain modeling. *Statistical climatology. Developments in atmospheric science*, vol. 13, S. Ikeda, ed. Elsevier Science, 197–205.

Brett, A. C. and Tuller, S. E. 1991. The autocorrelation of hourly wind speed observations. *J. Appl. Meteor.* **30**, 823–833.

Brown, B. G., Katz, R. W. and Murphy, A. H. 1984. Time series models to simulate and forecast wind speed and wind power. *J. Climate Appl. Meteor.* **23**, 1184–1195.

Cox, D. R. and Miller, H. D. 1965. *The theory of stochastic processes.* Methuen & Co. Ltd., 398 pp.

De Swart, H. E. and Grasman, J. 1987. Effect of stochastic perturbations on a low-order spectral model of the atmospheric circulation. *Tellus* **39A**, 10–24.

Fraedrich, K. 1988. El Niño/Southern oscillation predictability. *Mon. Wea. Rev.* **116**, 1001–1012.

Gabriel, K. R. and Neumann, J. 1962. A Markov chain model for daily rainfall occurrence at Tel Aviv. *Quart. J. Roy. Meteor. Soc.* **88**, 90–95.

Gregory, J. M., Wigley, T. M. L. and Jones, P. D. 1992. Determining and interpreting the order of a two-state Markov chain: Application to models of daily precipitation. *Water Resour. Res.* **28**, 1443–1446.

Green, M. C., Myrup, L. O. and Flocchini, R. G. 1992. A method for classification of wind field patterns and its application to southern California. *Int. J. Climatol.* **12**, 111–135.

Hannachi, A. and Legras, B. 1995. Simulated annealing and weather regimes classification. *Tellus* **47A**, 955–973.

Hardy, D. M. and Walton, J. J. 1978. Principal components analysis of vector wind measurements. *J. Appl. Meteor.* **17**, 1153–1162.

Huang, Z. and Chalabi, Z. S. 1995. Use of time-series analysis to model and forecast wind speed. *J. Wind Eng. Ind. Aerodyn.* **56**, 311–322.

Kamber, K. and Kaufmann, P. 1992. Das Mistral-Messnetz, Konzeption, Aufbau und Betrieb. *Regio Basiliensis* **33**, 107–114 (in German).

Kaufmann, P., 1996. *Regionale Windfelder über komplexer Topographie.* PhD thesis no. 11565. Swiss Federal Institute of Technology (ETH), Zurich (in German).

Kaufmann, P. and Weber, R. O. 1996. Classification of mesoscale wind fields in the MISTRAL field experiment. *J. Appl. Meteor.* **35**, 1963–1979.

Klink, K. and Willmott, C. J. 1989. Principal components of the surface wind field in the United States. A comparison of analyses based upon wind velocity, direction, and speed. *Int. J. Climatol.* **9**, 293–308.

Landberg, L. and Watson, S. J. 1994. Short-term prediction of local wind conditions. *Boundary-Layer Meteor.* **70**, 171–195.

Mo, K. C. and Ghil, M. 1987. Statistics and dynamics of persistent anomalies. *J. Atmos. Sci.* **44**, 877–901.

Müller, W. 1961. Über die Häufigkeit der Üvbergänge bestimmter Wetterlagen in andere. *Meteor. Rundsch.* **14**, 88–91 (in German).

Nicolis, C. 1990. Chaotic dynamics, Markov processes and climate predictability. *Tellus* **42A**, 401–412.

Parlow, E., 1992. REKLIP — Klimaforschung statt Meinungsmache am Oberrhein. *Regio Basiliensis* **33**, 71–80 (in German).

Rajagopalan, B., Lall, U. and Tarboton, D. G. 1996. Nonhomogeneous Markov model for daily precipitation. *J. Hydrol. Eng.* **1**, 33–40.

Snedecor, G. W. and Cochran, W. G. 1989. *Statistical methods.* Iowa State University Press, 503 pp.

Spekat, A., Heller-Schulze, B. and Lutz, M. 1983. Über Großwetter und Markov-Ketten. *Meteor. Rundsch.* **36**, 243–248 (in German).

van Dijk, W. and Jonker, P J. 1985. Statistical remarks on European weather types. *Arch. Meteor. Geophys. Bioclimatol.* **35B**, 277–306.

Vautard, R., Mo. K. C. and Ghil, M. 1990. Statistical significance test for transition matrices of atmospheric Markov chains. *J. Atmos. Sci.* **47**, 1926–1931.

Weber, R. O. and Kaufmann, P. 1995. Automated classification scheme for wind fields. *J. Appl. Meteor.* **34**, 1133–1141.

Whiteman, C. D., 1990. Observations of thermally developed wind systems in mountainous terrain. *Atmospheric processes over complex terrain. Meteor. Monogr.* vol. 23, no. 45. W. Blumen, ed., Amer. Meteor. Soc., 5–42.