Final response

Our responses are marked in *italic and blue* and were directly inserted below each comment of the referees. Changes in the manuscript are described below each comment and are highlighted in the attached version of the paper.

We thank both referees for their helpful comments and the time they spent with the review.

Comments from Referee #1

1. Introduction: The authors list studies done for many different high-altitude sites. However, one observatory which is not mentioned at all is Pic du Midi in the French Pyrenees. There was a recent study by Hulin et al. (2019) on atmospheric composition and the detection of thermally driven circulations with different methods, which should be cited as well.

Thanks for pointing out that study. In the new manuscript version, the study is cited in the introduction.

2. p. 4, l. 22: "available at 1 min intervals". Are the values averaged over 1 min intervals or instantaneous?

The values are averaged over 1 min intervals. In the new manuscript version, we use the phrase "available as 1 min averages".

3. p. 5, l. 2: What is the temporal resolution of the meteorological data?

The standard meteorological data was available as 10 min averages except at the Schneefernerhaus where 1 min averages were available. This is now mentioned in the manuscript.

4. p. 5, l. 3-5: How are the aerosol layer heights detected? With the manufacturer software or with a algorithm developed by the authors?

The manufacturer software was used, which is now mentioned in the manuscript.

5. p. 5, l. 4: ">1": What is the unit? dB, B,...?

This statement was taken from (Heese et al., 2010) who expressed the signal-to-noise ratio in a dimensionless number.

6. p. 5, l. 7: Is daylight saving time taken into account?

We checked whether the time stamp referred to daylight saving time during the summer half year but this was not the case. However, some measurements were recorded with UTC time, which was converted to local standard time.

7. p. 5, l. 8: Why did the authors use a threshold of 66 % and not something else for the availability?

We agree that the minimum data availability of 66 % per interval is somewhat arbitrary. However, it guarantees that the averaged values are representative for most of the time interval and not only for a small part of the interval in the case of data gaps.

8. p. 6, l. 1ff: Why is there a time offset between the sites? Where the sites not synced to a time server? If the sites are not synced was there a shift of the time offset with time? E.g. what is the difference between Sept. 2013 and March 2014? How high is the correlation coefficient?

After contacting the manufacturer of the weather stations, we realized that the time stamps are synced to a time server and should not differ much between the sites. The initially supposed time offsets based on crosscorrelation may be due to an inexact horizontal alignment of the pyranometers. Therefore, we repeated the analyses without shifting the time and updated the text and the figures. This had a negligible effect on $\Delta\theta_v$ and the statistical air mass classification because initially, a time offset had only be supposed for the sites ZPLT, Schachen, Kreuzalm, and Felsenkanzel and not for GAP, Brandwiese, and ZSG. Omitting the time shift, had a small effect on the mechanistic classification and slightly improved the agreement between the mechanistic and the statistical classifications.

9. p. 6, l. 19: "exp" should not be italic.

Thanks, this is now corrected.

10. p. 7, l. 16: Why only 89 % at the beginning? What happened to the remaining 11 %.

During 11 % of the time, the ceilometer could not determine any aerosol layer height, most likely due to fog or precipitation.

11. p. 8, l. 2: Why four classes? In the following lines only three air mass classes are defined?

"Four" was a mistake, it should be three. We have corrected this now.

12. p. 8, l. 24: What are "most suitable variables"? How are they defined? What are the criteria?

To be more precise, we changed the phrase to "variables with an expected unambiguous link to vertical transport processes and a high data availability". These variables include CO, $\mathrm{CH_4}$, $\mathrm{CO_2}$, $\mathrm{O_3}$, *specific humidity, air pressure at Garmisch-Partenkirchen, and* $\Delta\theta_v$ *and are now mentioned directly afterwards in the manuscript.*

13. p. 9, l. 17: The figures should be referred to in the correct order, i.e. 8 not before 4.

Figure 8b is now Fig. 3b so that the figures are referred to in the correct order.

14. p. 9, l. 19: What are the remaining measurements?

The remaining measurements refer to the gases $\mathrm{NO_y}$, $\mathrm{NO_x}$, $^{222}\mathrm{Rn}$, $^{7}\mathrm{Be}$, HCHO, $^{222}\mathrm{Rn}$, *the aerosol quantities* N_{90}, *eBC,* $\mathrm{PM_{10}}$, *and the standard meteorological variables precipitation, relative humidity, temperature, global radiation. Because these measurements are already mentioned in Sect. 2.2, the new manuscript version summarizes them as "remaining chemical (e.g.* $\mathrm{NO_y}$, $\mathrm{NO_x}$, $^{222}\mathrm{Rn}$*) and standard meteorological measurements (e.g. precipitation, relative humidity)".*

15. p. 9, l. 25: I was surprised to read about a marine boundary layer considering the location of Zugspitze. Where does the marine air mass come from? The authors speculate about that later in the manuscript, but I think a hint about its origin should already be given here.

We adopt this suggestion and now mention the Atlantic Ocean as a potential origin of air masses in that part of the manuscript.

16. p. 10, l. 8ff: What is the typical stability distribution? This could e.g. be checked with a histrogram.

We address this question by adding a short section in the supplement, including histograms of $\Delta\theta_v$ *for the periods June–July and December–January (Fig. S1a). Note that we have additionally changed the sign convention for* $\Delta\theta_v$ *(see comment 13 of Referee #2).* $\Delta\theta_v$ *was almost always positive, indicating stable conditions. In June and July,* $\Delta\theta_v$ *ranged between* $-5\ \mathrm{K}$ *and* $+21\ \mathrm{K}$ *whereas in December and January, it was generally more positive with values between* $4\ \mathrm{K}$ *and* $31\ \mathrm{K}$*. Additionally, we included the histogram, from which the threshold of* $\Delta\theta_v = 8\ \mathrm{K}$ *between anabatic and katabatic winds was determined (Fig. S1b).*

17. p. 10, l. 12: GAP is a valley floor station. Why is this station used to detect strong synoptic forcing? Wouldn't it make more sense to use high-elevation sites?

We restricted the wind velocities at ZPLT and GAP to $< 3\ \mathrm{m\,s^{-1}}$ *when detecting katabatic winds. This criterion was based on the idea that katabatic winds are favored by a weak synoptic forcing and are expected to have low wind velocities. After applying the wind direction and stability criteria for katabatic winds, the wind velocity still ranged up to approximately* $10\ \mathrm{m\,s^{-1}}$ *at ZPLT while it was almost always* $< 3\ \mathrm{m\,s^{-1}}$ *at GAP (Fig. R1). Cases with high wind velocities (*$\geq 3\ \mathrm{m\,s^{-1}}$*) were suspected to be synoptically driven and were discarded. For almost all of these cases, the wind velocity threshold was exceeded at the high-elevation site ZPLT, not at GAP. For the reason of consistency, the wind velocity treshold was also applied to GAP, even if only very few cases were affected.*

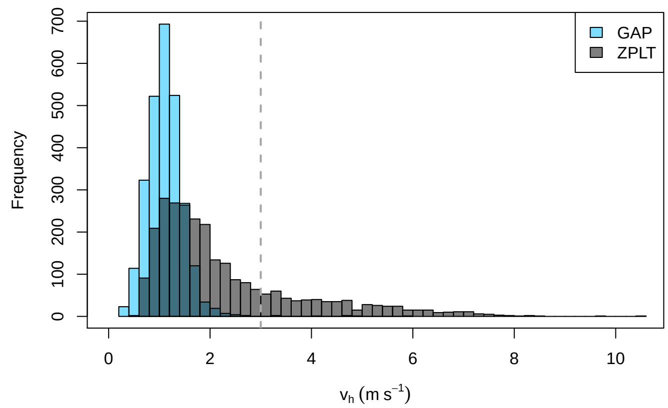

Figure R1. Histograms of the horizontal wind velocities (v_h) at GAP and ZPLT for potential katabatic winds, i.e. for downvalley wind direction sectors at both stations and strongly stable conditions ($\Delta\theta_v \geq 8$ K). The dashed line shows the threshold of $3\,\mathrm{m\,s^{-1}}$ for katabatic winds.

18. p. 10, l. 16: Why does condition a) requires anabatic wind OR UFS below $\mathrm{MLH_{GAP}}$? Why not AND?

The mixing layer height can be heterogeneous, especially in mountainous terrain. $\mathrm{MLH_{GAP}}$ is measured in the center of the valley atmosphere. Due to thermally driven upslope winds, boundary layer air can reach higher altitudes at the mountain slopes compared to the valley center (Henne et al., 2004; Gohm et al., 2009). Thus, the UFS could be influenced by the boundary layer while $\mathrm{MLH_{GAP}}$ is below the UFS level. Additionally, a residual layer could result in boundary layer air masses at the UFS during night while anabatic winds are absent.

19. p.10, l. 30-31: The authors state that PC1 was always a meaningful indicator while PC2 did not always allow for an unambiguous interpretation. On what is that assumption based?

This statement is explained in the following paragraph in which the PC loadings are discussed. To make that clear, we inserted the phrase ", which will be explained in the following".

20. Sect. 3.1: I found this part of the result section hard to follow and it might be difficult for readers not familiar with PCA to understand the interpretation of the results. It might be helpful to give a detailed example at the beginning on how to read and understand the loading diagrams in Fig. 6. What does a low score mean? Low absolute values or large negative values? In the text (p. 11, l. 1-6), the authors talk about scores while in Fig. 6 loadings are shown. How does this relate?

With low score, we mean large negative linear combinations according to Eq. 5. In the new manuscript, we explain that variables with a high absolute loading largely determine the PC scores and give an example on how loadings affect PC scores. If an original variable is much higher than its 2-month mean value and its loading on PC1 is strongly negative then the variable will strongly contribute to a large negative score of PC1.

21. p. 11, l. 32: "... air masses (Fig. 7a-g): compared ..." To make clear that the explanation why it was consistent is following.

We adopted the suggestion of inserting a colon.

22. p. 14, l. 15: This additional criteria of no clouds below 4 km, should be moved to Sect 2.5.

As suggested, we moved this criterion to Sect. 2.5, in which the mechanistic approach is described.

23. p. 14, l. 21: " .. this winds were not thermally induced BUT the MLH_{GAP} suggested non-ML air masses." BUT does not make sense. It should be AND.

We agree and replace BUT by AND. The same applies to two similar sentences in the same section and in Sect. 2.5.

24. p. 14, l. 28ff: What about LRMD and MBL/UFT air mass classes?

LRMD and MBL/UFT are subclasses of HYBRID. To make that clear, we write "HYBRID including LRMD and MBL/UFT" in the new manuscript version.

25. p. 15, l. 27: What differed between the six 2-month periods?

The first two principal components, their interpretation, and thus, the mapping of air mass regimes to air mass classes differed between the 2-month periods. We changed the sentence into the following one: "but the principal components and their interpretation differed between the six 2-month periods".

26. Fig. 7: Why not stick to the numbers I-IX for the regimes instead of introducing the long names for the regimes. This would make it more comparable with Figs. 3 and 8. The colours for the air mass classes should be brighter like in Fig. 8 and 9. Adjust the scales of the subplots to make them clearer to read (e.g. Figs. 7o, 7p, 7q, 7r).

To make Fig. 7 more comparable with Fig. 3 and 8, we replaced the long regime names by the numbers I to IX. The same was done for the figures in the supplement. Also, the colours and scales in Fig. 7 were adjusted as suggested.

27. Fig. 8: I probably understand it wrong, but how can regime VI belong to ML? In Fig. 3 there is no connector between ML and regime IV. Also, how can I and II belong to HYBRID? It would be good to enhance the boxes of the label to make the hatched areas better visible.

In Fig. 3, the mapping between of air mass regimes to air mass classes is only shown for February and March. In other 2-month periods, the mapping was different, which is now explained in the figure caption. In Fig. 8, the legend is now bigger and the hatched areas are better visible.

28. Fig. 9: Maybe add "UFS below MLH..." or "UFS above..." to the label and maybe refer to the text (p. 10, l. 14ff) where the 3 conditions are explained.

We adopted these suggestions.

Comments from Referee #2

(1) Pre-processing of the data before use in the statistical classification method is limited to standardization (that is, adjustment of the sample mean to 0 and of the sample variance to 1). I am wondering whether any slightly more sophisticated pre-processing could be beneficial.
For instance:
(a) Some of the variables in the data matrix have well-defined seasonal and diurnal cycles. Would it be possible to determine average annual and daily cycles, and to remove them from the data set? Performing the analysis on deviations from the average cycles might improve classification results.

In our analysis, the seasonal course was roughly removed by subtracting the 2-month mean in each of the 2-month periods when standardizing the data. We added a sentence in Sect. 2.4.2 to make that clear. In principle, it would be beneficial to remove the seasonal course more accurately so that the leading principal components would only depend on the shorter-term variability of the time series. In the beginning of the data analysis, we considered the determination and removal of the seasonal course with a spectral approach such as a wavelet filter. However, this idea was abandoned because it requires knowledge of the time series within a window centered at the time of interest and is thus not applicable to real-time operational mode. It would be possible to remove an average seasonal cycle from each variable based on a larger, multi-year data set. Due to interannual variability, however, this approach would only partly remove the seasonal course. For example, Scheel et al. (1999)

showed that the monthly mean O_3 mixing ratios at Mt. Zugspitze can strongly differ between individual years and the 10-year ensemble. Therefore and because a large fraction of the seasonal variability is already removed by subtracting the 2-month mean values, we doubt that the removal of an average seasonal cycle would substantially improve the classification results.

The mean diurnal cycles of the analyzed variables reflect shifts between air masses that we aim to identify. Afternoon maxima of atmospheric constituents such as water vapor, CO, ^{222}Rn, and NO_y at high-alpine sites have been explained by thermally induced uplift processes including anabatic winds (Forrer et al., 2000; Zellweger et al., 2003; Griffiths et al., 2014). These conditions are typically accompanied by afternoon minima of atmospheric stability and air pressure. The latter is associated with plain to mountain winds in the Northern Alps (Lugauer and Winkler, 2005). Removing the average diurnal cycle would complicate the classification of air masses in the case of thermally induced vertical transport.

(b) PCA does not require the data to follow multivariate normal distributions, but its results can often be interpreted more easily if they do. It strikes me that most variables are concentrations, therefore their PDF will certainly be markedly non-Gaussian. Would a cleverly designed variable transformation allow bringing more variance into the leading principal components?

Among the PCA input variables, especially CO, CH_4, and in the winter half year also CO_2 tend to have right-skewed PDFs. A logarithmic transformation, for example, would reduce the influence of the "long tails" of the PDFs on the principal component loadings. However, a variable transformation would not completely remove the "long tails" because the PDFs rarely follow an idealized distribution such as a log-normal distribution and they vary with the 2-month periods. Additionally, a variable transformation has the following side effect. If a logarithmic transformation is used for example and the original variable increases by a certain amount of units, then the transformed variable will not increase by a certain amount of units but by a certain factor. Although a variable transformation might somewhat increase the fraction of variance explained by the leading principal components in some 2-month periods, we think that it would not significantly improve the classification of the entire data set.

(2) The matching between air-mass regimes (I-IX) and air-mass classes (ML, UFT/SIN, HYBRID) is different in each two-month period (see Figure 8). The manuscript text contains little or no information about the overarching logic. Why was this necessary? What criteria were used to attribute regimes to classes, how did these criteria change with the season?

In my opinion, the ad-hoc tuning of the method is a serious shortcoming. It is clearly a subjective component of the classification, and as such it cannot be exported to other sites. The authors do not explain this point in a satisfactory manner, and they probably should. Why wasn't it possible to design a fully objective classification rule? Formal methods to identify classification rules exist and could be used (see for instance chapter 14 in Wilks, 2011, Statistical Methods in the Atmospheric Sciences. DOI: 10.1016/B978-0-12-385022-5.00014-2).

We realize that the mapping of air mass regimes to air mass classes was only shortly described in the manuscript. Now, more details are included in Sect. 2.4.3. The loadings of the leading principal components changed with the 2-month periods (Fig. 6), which can be explained by seasonal changes in chemical processes (e.g. CO_2 emissions and uptake, photochemical O_3 production) and atmospheric dynamics (e.g. thermally induced uplift). Therefore, the characteristics of the air mass regimes (I-IX) changed with the 2-month periods and required a separate interpretation in each 2-month period. The air mass regimes were assigned to the air mass classes by visually comparing boxplots of the original input variables between the air mass regimes (Fig. 7). In the winter half year, the class ML (now renamed as BL, see comment 4 of Referee #2) was assigned if CO, CH_4, CO_2, and q were relatively high and $\Delta\theta_v$, p_{GAP} and O_3 were relatively low compared to the other regimes; note that the sign of $\Delta\theta_v$ was changed so that low values indicate a low static stability (see comment 13 of Referee #2). The class UFT/SIN was assigned in the opposite case and the remaining regimes were assigned to the class HYBRID. Apart from CO_2 and O_3, the same criteria were used in the summer half year; CO_2 was required to be relatively low for the class BL and O_3 was not considered.

We agree that this subjective mapping of regimes to classes is a shortcoming. Nevertheless, it is based on typical qualitative differences between lifted and subsided air masses. The present study has the character of a pilot study to develop a novel multivariate approach for air mass classification. In future studies, a more objective and robust mapping of regimes to classes could be achieved by using a metric such as the difference between the median of a regime and the overall median to define a threshold for "relatively high/low" characteristics and by using data from multiple years or more observatories. The supervised *classification methods, which are described in Chapter 14 in Wilks (2011), could only be applied to our problem if a reliable test data set with known air mass classes was available and included a variety of meteorological conditions.*

(3) Note: the line numbering in pages 2-end is wrong, i.e., the 6th line from the top is labelled as "5" and so on, as if the the first line were 0. In what follows, I'm using this unusual "zero-based" system.

This issue has been corrected.

(4) Nomenclature. The first air-mass class is labelled ML, for "mixing layer". I'm wondering if this is appropriate. A mixed layer, by definition, has nearly adiabatic lapse rate. The boundary layer (BL) is not always well-mixed, especially at night. On page 1, line 12, it is stated that "the terms ML and BL can be defined synonymously ... ". In my mind, the two concepts are quite distinct. I'd rather say that the BL might sometimes include a ML. I don't really have a strong opinion on this matter, but anyway I suggest renaming the first air-mass class to BL, for "boundary layer". A similar comment applies to "mixing layer height" (MLH). This should probably become "boundary layer height" (BLH), in particular because, according to the description of the wavelet detection algorithm, MLH/BLH potentially includes multiple aerosol layers. These typically develop in connection with inversion layers, i.e., non-mixed parts of the atmosphere.

We agree that the definitions of the boundary and mixing layers are based on different concepts. The boundary layer (BL) is affected by surface forcings such as the turbulent transfer of momentum, heat, and matter while the mixing layer specifically refers to the dispersion of surface-emitted atmospheric constituents (Stull, 1988; Seibert et al., 2000). In the new manuscript version, we cite definitions of the boundary and mixing layers. We point out that we consider a residual layer and elevated aerosol layers, which were influenced by the surface within a time scale of one diurnal cycle, as parts of the boundary and mixing layers, as suggested by Reuten et al. (2007). Because the term mixing layer is usually restricted to the well-mixed layer adjacent to the surface, we renamed the air mass class ML as BL. However, we keep the term mixing layer height (MLH) because the ceilometer measures the aerosol backscatter profile, which reflects the dispersion of surface-emitted particles.

(5) Page 1, title. "Discrimination" or "classification"? The two terms have slightly different meanings. See again chapter 14 in Wilks 2011.

We now use the word "classification" instead of "discrimination" because discrimination would be based on training data, for which the groups (air mass classes) are already known, while classification refers to the attribution of test data to groups (Wilks, 2011).

(6) Page 1, lines 8 and 12. Use of the word "classifiable". I believe these statements should be formulated more clearly. As they are now, they seem to allude to the intrinsic "ability" of the methods to separate the events, and seem to suggest that the statistical method permits to obtain a meaningful classification much more often than the mechanistic one (78 % of the time as opposed to 25 %). Instead, these two percentage only represent the availability of the input data for the two methods. Please use something like: "Due to data gaps, only x % of the investigated year could be classified".

To avoid misunderstandings, we avoid the expression "classifiable cases" and now use the phrases "... input data was available in 78 % ... " and "Due to data gaps, only 25 % of the cases could be classified ... ". In the rest of the paper, we use "classified cases" instead of "classifiable cases".

(7) Page 2, lines 10-11. Foehn flows are listed among processes that cause air mass lifting. This is inexact and quite confusing. Foehn is a fall wind: its dynamics are inextricably tied to air-mass descent (not lifting!) on the lee side of a mountain range. That said, intense foehn certainly causes mechanical mixing of the lower atmosphere, which may result in transport of chemical species from the PBL to the free troposphere. Zellweger et al (2003, cited in the manuscript) list foehn among the meteorological conditions in which free-troposphere air masses are mixed with PBL air masses (page 781, top of second column). Correctly, Zellweger et al (2003) do not mention "lifting" in relation to foehn. Please revise. The comment also applies to page 2, line 23 and to page 12, line 15.

In the case of south foehn, Mt. Zugspitze is located on the lee side of the Alps where the air flow descends. Nevertheless, foehn events can be associated with lifting on the windward side of the mountain range and thus transport air masses from the BL to *high-alpine sites. If the southern Alps experience north foehn, Mt. Zugspitze is located on the windward side of the Alps where the air flow typically ascends and lifts BL air masses. In the new manuscript version, we list foehn among transport processes that cause air mass lifting or mixing. We also mention that foehn winds descend on the lee side of a mountain range and can be associated with air mass lifting on the windward side. On page 12, we now speak of an "influence of BL air masses" with respect to foehn instead of an "uplift of BL air masses".*

(8) Page 2, line 30. "... because the MLH is a meteorological quantity". Wording could be more careful here. MLH/BLH is not a directly measured quantity, but rather an estimate obtained from measurements of other quantities. The determination of MLH/BLH from a vertical profile can be quite tricky, too. I'd rather say something like: "... because determination of the MLH from vertical profiles of measured quantities requires a-priori meteorological knowledge".

We adopted this suggestion.

(9) Page 5, line 17. "... set zero" → "... set to zero".

Done.

(10) Page 5, line 21. Please delete the blank space before the full stop.

Done.

(11) Page 6, line 17. "... using the Clausius Clapeyron equation". Or rather a numerical approximation? There are many such formulas: Goff-Gratch, Magnus-Tetens, Bolton ... which one?

We used the following form of the Clausius Clapeyron equation (e.g. Wallace and Hobbs, 2006),

$$\frac{1}{e_s}\frac{de_s}{dT} = \frac{L_v}{R_v T^2}, \tag{R1}$$

where e_s (hPa) *is saturation vapor pressure,* T (K) *is temperature,* L_v $(\mathrm{J\,kg^{-1}})$ *is latent heat of evaporation, and* R_v $(\mathrm{J\,kg^{-1}\,K^{-1}})$ *is specific gas constant for water vapor. Integration of Eq. R1 yields*

$$e_s(T) = e_s(T_0)\exp\left\{\frac{L_v}{R_v}\left(\frac{1}{T_0} - \frac{1}{T}\right)\right\}, \tag{R2}$$

where T_0 (K) *is a reference temperature with known* e_s*.*

(12) Page 6, line 19. I think the standard notation should be either e^x or $\exp\{x\}$. Also, please replace T_v by $\bar{T}_v$, to indicate the vertical averaging.

Done.

(13) Page 7, line 2. I personally find the sign convention counterintuitive. Although static stability conventionally corresponds to $d\theta/dz > 0$, here $\Delta\theta$ is greater than zero when θ decreases with height. Why not computing $d\theta/dz$?

Noting that the sign convention for $\Delta\theta_v$ *was counterintuitive, we changed this sign convention. Now, positive values indicate stable conditions. This change corresponds to multiplying* $\Delta\theta_v$ *by* -1*. Consequently, the sign of the PC loadings of* $\Delta\theta_v$ *changed as well (Fig. 6). Dividing* $\Delta\theta_v$ *by* Δz *would be an alternative but would not significantly change the results because* Δz *does not vary much.*

(14) Page 7, lines 20-21. "The MLH attribution was based on the idea that the MLH varies only gradually". I am not sure this is always appropriate over mountains. Horizontal advection of aerosol layers due to mountain venting can cause spatial and temporal discontinuities in MLH.

We notice this shortcoming and now mention it in the manuscript. For most of the time, however, we expect a gradual variation of the MLH.

(15) Page 10, lines 9-10. Why using -8 K to discriminate "weak" and "strong" static stability?

The threshold of -8 K (now $+8$ K because of the changed sign convention, see comment 13 of Referee #2) was determined by comparing the histograms of $\Delta\theta_v$ for winds coming from the upvalley or downvalley wind direction sectors at GAP and ZPLT. The histograms are now included in the supplement as Fig. S1b. Using the intersect of the histograms corresponds to minimizing the number of data points that are excluded from potential anabatic and katabatic winds.

(16) Page 11, line 29. Are CH_4 and CO_2 pollutants?

Originally, we considered CH_4 and CO_2 as pollutants. However, the paragraph under consideration was removed because the long names of the air mass regimes were replaced by roman numbers. See comment 26 of Referee #1.

Please find below the marked-up manuscript versions of the main article and its supplement.

References

Forrer, J., Rüttimann, R., Schneiter, D., Fischer, A., Buchmann, B., and Hofer, P.: Variability of trace gases at the high-Alpine site Jungfraujoch caused by meteorological transport processes, Journal of Geophysical Research: Atmospheres, 105, 12 241–12 251, https://doi.org/10.1029/1999JD901178, 2000.

Gohm, A., Harnisch, F., Vergeiner, J., Obleitner, F., Schnitzhofer, R., Hansel, A., Fix, A., Neininger, B., Emeis, S., and Schäfer, K.: Air Pollution Transport in an Alpine Valley: Results From Airborne and Ground-Based Observations, Boundary-Layer Meteorology, 131, 441–463, https://doi.org/10.1007/s10546-009-9371-9, 2009.

Griffiths, A. D., Conen, F., Weingartner, E., Zimmermann, L., Chambers, S. D., Williams, A. G., and Steinbacher, M.: Surface-to-mountaintop transport characterised by radon observations at the Jungfraujoch, Atmospheric Chemistry and Physics, 14, 12 763–12 779, https://doi.org/10.5194/acp-14-12763-2014, 2014.

Heese, B., Flentje, H., Althausen, D., Ansmann, A., and Frey, S.: Ceilometer lidar comparison: backscatter coefficient retrieval and signal-to-noise ratio determination, Atmospheric Measurement Techniques, 3, 1763–1770, https://doi.org/10.5194/amt-3-1763-2010, 2010.

Henne, S., Furger, M., Nyeki, S., Steinbacher, M., Neininger, B., de Wekker, S. F. J., Dommen, J., Spichtinger, N., Stohl, A., and Prévôt, A. S. H.: Quantification of topographic venting of boundary layer air to the free troposphere, Atmospheric Chemistry and Physics, 4, 497–509, https://doi.org/10.5194/acp-4-497-2004, 2004.

Lugauer, M. and Winkler, P.: Thermal circulation in South Bavaria climatology and synoptic aspects, Meteorologische Zeitschrift, 14, 15–30, https://doi.org/doi:10.1127/0941-2948/2005/0014-0015, 2005.

Reuten, C., Steyn, D. G., and Allen, S. E.: Water tank studies of atmospheric boundary layer structure and air pollution transport in upslope flow systems, Journal of Geophysical Research: Atmospheres, 112, D11 114, https://doi.org/10.1029/2006JD008045, 2007.

Scheel, H., Sladkovic, R., and Kanter, H.: Ozone variations at the Zugspitze (2962 m asl) during 1996-1997, WIT Transactions on Ecology and the Environment, 35, 264 – 268, 1999.

Seibert, P., Beyrich, F., Gryning, S.-E., Joffre, S., Rasmussen, A., and Tercier, P.: Review and intercomparison of operational methods for the determination of the mixing height, Atmospheric Environment, 34, 1001–1027, https://doi.org/10.1016/S1352-2310(99)00349-0, 2000.

Stull, R. B.: An Introduction to Boundary Layer Meteorology, Kluwer Acad. Publ., Dordrecht, Boston, London, 1988.

Wallace, J. M. and Hobbs, P. V.: Atmospheric Thermodynamics, in: Atmospheric Science, edited by Wallace, J. M. and Hobbs, P. V., pp. 63–111, Academic Press, San Diego, 2 edn., https://doi.org/https://doi.org/10.1016/B978-0-12-732951-2.50008-9, 2006.

Wilks, D.: Chapter 14 - Discrimination and Classification, in: Statistical Methods in the Atmospheric Sciences, edited by Wilks, D. S., vol. 100 of *International Geophysics*, pp. 583 – 602, Academic Press, https://doi.org/https://doi.org/10.1016/B978-0-12-385022-5.00014-2, 2011.

Zellweger, C., Forrer, J., Hofer, P., Nyeki, S., Schwarzenbach, B., Weingartner, E., Ammann, M., and Baltensperger, U.: Partitioning of reactive nitrogen (NO_y) and dependence on meteorological conditions in the lower free troposphere, Atmospheric Chemistry and Physics, 3, 779–796, https://doi.org/10.5194/acp-3-779-2003, 2003.

Multivariate statistical air mass ~~discrimination~~ classification for the high-alpine observatory at the Zugspitze mountain, Germany

Armin Sigmund[1,*], Korbinian Freier[2,3], Till Rehm[4], Ludwig Ries[5], Christian Schunk[6,+], Anette Menzel[6,7], and Christoph K. Thomas[1]

[1]Micrometeorology Group, University of Bayreuth, Bayreuth, Germany
[2]Bavarian Environment Agency, Augsburg, Germany
[3]Research Unit Sustainability and Global Change, Center for Earth System Research and Sustainability (CEN), University of Hamburg, Hamburg, Germany
[4]Environmental Research Station Schneefernerhaus, Zugspitze, Germany
[5]German Environment Agency, GAW Global Observatory, Zugspitze-Hohenpeissenberg, Germany
[6]Ecoclimatology, Technical University of Munich, Freising, Germany
[7]Institute for Advanced Study, Technical University of Munich, Garching, Germany
[*]now at: School of Architecture, Civil and Environmental Engineering, Swiss Federal Institute of Technology, Lausanne, Switzerland
[+]now at: Safety and Radiation Protection, Technical University of Munich, Garching, Germany

Correspondence: A. Sigmund (armin.sigmund@epfl.ch)

Abstract. To assist atmospheric monitoring at high-alpine sites, a statistical approach for distinguishing between the dominant air masses was developed. This approach was based on a principal component analysis using five gas-phase and two meteorological variables. The analysis focused on the site Schneefernerhaus at Mt. Zugspitze, Germany. The investigated year was divided into 2-month periods, for which the analysis was repeated. Using the 33.3 % and 66.6 % percentiles of the first two principal components, nine air mass regimes were defined. These regimes were interpreted with respect to vertical transport and assigned to the air mass classes ~~ML~~ BL (recent contact with the ~~mixing~~ boundary layer), UFT/SIN (undisturbed free troposphere or stratospheric intrusion), and HYBRID (influences of both the ~~mixing~~ boundary layer and the free troposphere or ambiguous). The input data were available in 78 % of the investigated year~~were classifiable. ML~~. BL accounted for 31 % of the cases with similar frequencies in all seasons. UFT/SIN comprised 14 % of the cases but ~~were~~ was not found from April to July. HYBRID (55 %) mostly exhibited intermediate characteristics, whereby 17 % of HYBRID suggested an influence of the marine boundary layer or the lower free troposphere. The statistical approach was compared to a mechanistic approach using the ceilometer-based mixing layer height from a nearby valley site and a detection scheme for thermally induced mountain winds. ~~Only~~ Due to data gaps, only 25 % of the cases ~~were classifiable~~ could be classified with the mechanistic approach. Both approaches agreed well, except in the rare cases of thermally induced uplift. The statistical approach is a promising step towards a real-time ~~discrimination~~ classification of air masses. Future work is necessary to assess the uncertainty arising from the standardization of real-time data.

1 Introduction

High-alpine observatories such as the Environmental Research Station Schneefernerhaus (UFS) at Mt. Zugspitze, Germany, play an important role in studying changing concentrations of atmospheric constituents such as greenhouse gases, aerosols, and persistent organic pollutants (POPs), which have critical impacts on the climate, environmental integrity, or human health (McClure et al., 2016; Kirchner et al., 2016). In particular, high-alpine observatories frequently offer the opportunity to sample well-mixed air masses of the free troposphere. These air masses are representative for large spatial areas and thus suitable for the determination of large-scale and global trends (Yuan et al., 2019). Therefore, the observational network of the Global Atmosphere Watch Program of the World Meteorological Organization includes many high-alpine observatories. At times, however, high-alpine sites can be affected by local anthropogenic emissions on the mountain or regional emissions if air masses of the ~~mixing layer (ML~~boundary (BL) are lifted or mixed with the free troposphere by processes such as synoptic lifting (e.g. at fronts), thermally induced anabatic winds, and foehn flows (Zellweger et al., 2003). Foehn winds descend on the lee side of a mountain range and can be associated with air mass lifting on the windward side. In relatively rare cases, the air masses at high-alpine sites originate from the stratosphere, e.g. on approximately 6 % of the days at Mt. Zugspitze (Stohl et al., 2000).

The ~~terms ML and boundary layer~~ BL can be defined ~~synonymously as the sum of all atmospheric layers that exchange air with the surface during~~ "as the part of the troposphere that is directly influenced by the presence of the earth's surface, and responds to surface forcings with a time scale of about an hour or less" (Stull, 1988). Reuten et al. (2007) proposed to use a time scale of one diurnal cycle ~~(Seibert et al., 2000; Reuten et al., 2007). This definition was proposed~~ for this definition to explicitly include the residual layer above a stably stratified nocturnal boundary layer and elevated aerosol layers that can result from thermally driven upslope flows ~~(Reuten et al., 2007; Gohm et al., 2009)~~(Gohm et al., 2009). For instance, on typical fair weather days in summer, Henne et al. (2004) observed a two-layer structure of the ~~ML~~ BL in deep Alpine valleys where upslope flows lifted air from a polluted lower layer to a moderately polluted injection layer that reached well above the crest height.

In air pollution studies, the terms mixing layer or mixed layer are more common than BL and were defined as "the layer adjacent to the ground over which pollutants or any constituents emitted within this layer or entrained into it become vertically dispersed by convection or mechanical turbulence within a time scale of about an hour" (Seibert et al., 2000). Over flat terrain, the mixing layer typically coincides with the BL, at least during daytime. Following Reuten et al. (2007), we use a time scale of one diurnal cycle and consider a residual layer and elevated aerosol layers as parts of the BL and the mixing layer.

Basically, there are two different approaches for air mass ~~discrimination~~ classification at high-alpine sites. Mechanistic approaches investigate directly atmospheric transport processes using meteorological measurements or trajectory models, whereas statistical approaches infer the influences of the air layers from the air mass composition at the site and, possibly, meteorological auxiliary data.

Zellweger et al. (2003) used a mixed mechanistic and statistical approach at the high-alpine site Jungfraujoch, Switzerland. They identified three ~~uplift~~ transport processes, namely foehn events, synoptical lifting, and thermally induced uplift, and attributed the other cases to the undisturbed free troposphere. Foehn events and synoptic lifting were detected with mechanistic approaches using standard meteorological measurements and the height of back-trajectories, respectively. Thermally induced uplift was determined with a statistical approach based on the diurnal variation of the sum of oxidized nitrogen species (NO_y) or, in the case of data gaps, the aerosol surface area concentration or specific humidity.

In a recent study, Hulin et al. (2019) explored different methods for selecting days with thermally driven circulations influencing the Pic du Midi high-alpine observatory in the French Pyrenees. A statistical method based on the diurnal cycle of specific humidity seemed to be most reliable. Wind profiler measurements above the nearby plain indicated that a plain-mountain circulation with an upper-level return-flow played an important role although this circulation was sometimes not visible in the near-surface in-situ wind measurements in the plain.

Other studies used a ground-based lidar or ceilometer that was installed near a high-alpine site at a lower altitude to determine the mixing layer height (MLH) from the vertical aerosol backscatter profile (Gallagher et al., 2012; Ketterer et al., 2014). This method is considered as a mechanistic approach because the ~~MLH is a meteorological quantity~~determination of the MLH from vertical profiles of measured quantities requires a-priori meteorological knowledge. Recently, Poltera et al. (2017) used a tilted configuration of a ceilometer and demonstrated that the Jungfraujoch was rarely embedded in the local convective boundary layer but much more frequently in an above lying injection layer with slightly higher aerosol concentrations compared to the free troposphere.

Some statistical approaches defined a threshold for the concentration of a surface-emitted atmospheric constituent or a ratio of constituents. At the Mt. Bachelor Observatory, USA, a seasonal or monthly threshold for the water vapor mixing ratio was used to distinguish between free-tropospheric and ~~ML~~ BL influenced air masses (Ambrose et al., 2011; Zhang and Jaffe, 2017). The threshold was chosen such that the water vapor mixing ratios below this threshold had the same seasonal or monthly mean as the data from National Weather Service soundings that were launched at a lower elevated site.

For the Jungfraujoch, Herrmann et al. (2015) compared a mechanistic approach, which was based on back-trajectories and an inventory of carbon monoxide (CO) emissions, with two simple statistical approaches that used a constant threshold of the radon-222 concentration and the ratio of CO to NO_y. The CO/NO_y threshold appeared to achieve the best distinction between free-tropospheric and ~~ML~~ BL influenced air masses but the authors noted that a single threshold cannot account for varying degrees of ~~ML~~ BL influence (Herrmann et al., 2015).

Other statistical approaches aimed at selecting baseline (also called background) concentrations of a trace gas at different kinds of remote sites and were based on outlier removal techniques (e.g. Ruckstuhl et al., 2012) or the fact that well-mixed air masses result in a small temporal variability of the trace gas mixing ratio (e.g. Yuan et al., 2018). Baseline concentrations refer to a given species in a well-mixed air mass with a minimal influence of anthropogenic impurities of relatively short lifetime (Calvert, 1990; Yuan et al., 2018), which is associated with the free troposphere at high-alpine sites.

In contrast to mechanistic approaches, statistical air mass classifications are not able to distinguish between different uplift processes but require only local data. So far, however, statistical approaches were only based on a single constituent or a ratio of constituents, although several atmospheric constituents are typically monitored at high-alpine observatories.

This study proposes a novel statistical approach based on a principal component analysis (PCA) using seven chemical and meteorological variables. This approach is intended for a later use in real-time operational mode to enable an automated sampling of ambient air with respect to different air masses using a multi-channel sampling system for monitoring of POPs (Kirchner et al., 2016). The objectives were to (i) develop a statistical classification scheme for the site UFS at Mt. Zugspitze and to (ii) validate this approach, as far as possible, with a mechanistic approach based on ceilometer and standard meteorological measurements.

2 Methods

2.1 Measurement sites

The UFS (47°25'00" N, 10°58'47" E, 2650 m a.s.l.) is located on a steep south-facing slope, approximately 300 m below the summit of Mt. Zugspitze (2962 m a.s.l.), which is the highest mountain in the German Alps and represents the first real barrier for northwesterly advection from the Alpine foreland. At the UFS, westerly and easterly wind directions dominate due to the local topography (Risius et al., 2015). Because of trace gas and aerosol measurements, the UFS reduced its emissions of these substances to a minimum. Nevertheless, the measurements can be influenced by local emissions from the direct surrounding, which is a highly frequented tourist area all year round. Local emission sources include nearby cable car stations at the Zugspitze summit (ZSG) and the Zugspitzplatt (ZPLT) which is a gently sloping plateau below the UFS (Fig. 1). Additionally, a skiing area is situated at the ZPLT. The large metropolitan area of Munich is approximately 90 km north of the study site.

Beside the UFS, seven weather stations at different altitudes at a maximum horizontal distance of 11 km from the UFS were available and included in this study (Fig. 1). One of these stations is located at ZSG. Another weather station is located on the ZPLT which is surrounded by mountain ridges except towards the east where the plateau leads to the narrow and deep valley Reintal. The weather stations Schachen (1830 m a.s.l.), Kreuzalm (1600 m a.s.l.), Felsenkanzel (1250 m a.s.l.), and Brandwiese (900 m a.s.l.) are part of the project KLIMAGRAD (Schuster et al., 2014). The site Schachen is situated on a plateau above the middle Reintal in close vicinity of some trees. Kreuzalm is situated on a meadow on a mountain saddle. Brandwiese is located on a meadow surrounded by forest, where a tributary valley reaches the Reintal from the west. At Felsenkanzel, the measurements are made on a steep south-facing slope northwest of the town of Garmisch-Partenkirchen (GAP). The site GAP (720 m a.s.l) is located in the western periphery of the town where a broad west-east oriented valley turns to northeast towards the alpine foreland (Fig. 1).

2.2 Instrumentation and data set

The data set that was analyzed in this study spans a period of 1 year from 22 August 2013 to 21 August 2014, which was selected for high data availability. Table 1 gives an overview of the chemical measurements, associated instruments, measurement principles, and research institutions that provided the data. Most of the atmospheric constituents are measured at the UFS, including carbon mono- (CO) and dioxide ($\mathrm{CO_2}$), methane ($\mathrm{CH_4}$), ozone ($\mathrm{O_3}$), the sum of oxidized nitrogen species ($\mathrm{NO_y}$), nitrogen oxides ($\mathrm{NO_x = NO + NO_2}$ with NO and $\mathrm{NO_2}$ being nitrogen mono- and dioxide, respectively), formaldehyde (HCHO), the ambient particle number size distribution ($dN\ (dlog\ d_p)^{-1}$) for particle diameters (d_p) between 10 nm

and 600 nm, the mass concentration of particulate matter with $d_p < 10$ µm (PM_{10}), and equivalent black carbon (eBC). The radioisotopes beryllium-7 (^{7}Be) and radon-222 (^{222}Rn) are sampled at ZSG and the mountain ridge (2825 m a.s.l.) directly above the UFS, respectively. At the remaining sites, only meteorological parameters are recorded.

While most of the chemical data were available ~~at~~ as 1 min ~~intervals~~averages, HCHO and the particle size distribution were provided at 10 min intervals and ^{222}Rn and ^{7}Be were available ~~at~~ as 2 h and 12 h ~~intervals~~averages, respectively (Table 1). The HCHO data contained a large data gap between 15 December 2013 and 15 July 2014 and the ^{222}Rn data was only available since 1 January 2014. The ^{222}Rn data were downloaded from the World Data Centre for Greenhouse Gases (WDCGG, 2019). Because ^{7}Be attaches to aerosol particles, it is measured by collecting the carrier aerosol with a glass fiber filter. While NO can be measured directly, NO_y and NO_2 have to be converted to NO before the detection. This conversion of NO_y and NO_2 is performed with photolysis and gold/CO converters, respectively.

The standard meteorological measurements included air temperature (T), relative humidity (rH), global radiation (R_g), and horizontal wind velocity and direction at all sites. Air pressure (p) was available at the sites UFS, ZSG, and GAP. Additionally, year-round precipitation measurements with an electronic weighing system and a windbreak ring (Sommer Messtechnik, Koblach, Austria) at the site ZPLT were used. At the ZPLT and the four KLIMAGRAD sites, the wind data is measured by propeller anemometers (Wind Monitor 05103, Young, Traverse City, USA) while at GAP and ZSG, cup anemometers and wind vanes (SK-565 and SK-566, respectively, Thies Clima, Göttingen, Germany) are used. At the UFS, an ultrasonic anemometer (model 2D, Thies Clima, Göttingen, Germany) is used. The standard meteorological data were available as 10 min averages except at the UFS (1 min averages).

At GAP, a ceilometer (CHM 15k, Lufft, Fellbach, Germany) provides 15 s averages of up to three aerosol layer heights (ALHs) with a vertical resolution of 15 m ~~at 15 intervals~~ using a wavelet algorithm of the manufacturer, which detects strong gradients in the range corrected attenuated backscatter profile. In the absence of clouds, the signal-to-noise ratio of the ceilometer is typically > 1 up to a height of 4 to 5 km a.g.l. in the daytime and up to greater heights at night (Heese et al., 2010).

2.3 Data post-processing and quality control

All time stamps were converted to local standard time. The data were aggregated on a 30 min basis while requiring a data availability of ≥ 66 % in each interval unless stated differently. The ^{222}Rn and ^{7}Be concentrations were interpolated using nearest neighbor interpolation to match them with the 30 min intervals used for the analysis.

2.3.1 Atmospheric constituents

The chemical measurements were quality controlled by the research institutions that provided the data. Invalid data, which resulted from calibration, instrument repair, or power failure, were discarded.

Due to measurement uncertainties, NO_y occasionally exhibited a lower mixing ratio than NO_x. If NO_y was lower than 75 % of NO_x and NO_x was at least 0.03 ppb, both quantities were treated as artifacts and were discarded. If NO_y remained still lower than NO_x after averaging on a 30 min basis, both quantities were assumed to be equal and were replaced by the mean of NO_x and NO_y. Negative eBC concentrations were treated in a similar way. Values below $-0.05\ \mu g\,m^{-3}$ were discarded. If eBC remained still negative after averaging on a 30 min basis, the concentration was set to zero.

The processing of the particle number size distributions is described in Birmili et al. (2016) and includes, among others, a multiple charge inversion and corrections for particle losses. Following Herrmann et al. (2015), the number concentration of accumulation mode particles (N_{90}) was used as an indicator for ~~ML~~ BL air masses and approximated by including particle diameters between 90 nm and 600 nm.

Data of most trace gases and PM_{10} had been flagged manually with respect to locally polluted air masses. These air masses are a special kind of ~~ML~~ BL air masses because they do not result from uplift processes but from human activities on the mountain. Local emissions were evident from short but pronounced peaks in trace gas and aerosol concentrations and were most frequently observed for NO_x and NO_y. The flag for local NO_x emissions was approximately reproduced by selecting data with a 30 min NO_x standard deviation (σ_{NO_x}) of > 0.4 ppb (not shown). According to this criterion, local pollution events affected 9 % of the NO_x data but half of these events lasted no longer than one time interval, i.e. 30 min. Consequently, local pollution events are generally so short that they cannot be predicted on a 30 min basis in real-time applications. Therefore and because the classification scheme was developed for a later use in real-time operational mode, the air mass classification does not account for local pollution events.

2.3.2 Standard meteorological data

~~Possible time offsets between the sites were determined and corrected by maximizing the cross-correlation coefficient between global radiation at the UFS and the other sites on 23 September 2013 and 20 March 2014, two days with clear-sky conditions. The cross-correlation was evaluated in a time window where morning and evening hours were excluded because of shadowing effects depending on the site. Using data at 10 min intervals, the time offsets ranged between −20 min and +10 min.~~

The wind data at ZSG and GAP were discarded for wind velocities below the starting threshold of the wind vane of $0.3\ \mathrm{m\,s^{-1}}$ for a displacement of 90° (Löffler, 2012). For the wind vane of the propeller anemometers at the ZPLT and the KLIMAGRAD sites, the starting threshold was $1.1\ \mathrm{m\,s^{-1}}$ (Young, 2018). To achieve a trade-off between a high data quality and a high data availability, the wind data of these anemometers were discarded for wind velocities of $< 0.7\ \mathrm{m\,s^{-1}}$.

The standard meteorological data were aggregated on a 30 min basis using the vector mean for wind direction, the sum for precipitation, and the arithmetic mean for the other variables. Specific humidity q ($\mathrm{g\,kg^{-1}}$) and virtual potential temperature θ_v (K) were calculated for each site as follows (Foken, 2008)~~:~~,

$$q = 0.622\,\frac{e}{p - 0.378\,e} \tag{1}$$

$$\theta_v = (1 + 0.608\,q)\,T\left(\frac{p_0}{p}\right)^{R_L/c_p}, \tag{2}$$

where e (hPa) is vapor pressure, p (hPa) is air pressure, p_0 is 1000 hPa, T (K) is air temperature, R_L ($\mathrm{J\,kg^{-1}\,K^{-1}}$) is the gas constant of dry air, and c_p ($\mathrm{J\,kg^{-1}\,K^{-1}}$) is specific heat capacity of dry air. The vapor pressure was calculated from relative humidity (rH) and T using the Clausius Clapeyron equation. Because p was only measured at the sites ZSG, UFS, and GAP, it was calculated for the other sites using the hypsometric equation~~:~~,

$$p = p_{ref}\, \cancel{exp^{-\frac{g_0\,(Z-Z_{ref})}{R_L\,T_v}}}\, \exp\left\{ -\frac{g_0\,(Z-Z_{ref})}{R_L\,\overline{T}_v} \right\}, \tag{3}$$

where p (hPa) and p_{ref} (hPa) are the air pressures at the site of interest and a reference site, respectively, Z (m) and Z_{ref} (m) are the geopotential heights that were approximated with the altitudes of the respective sites, g_0 is $9.8\ \mathrm{m\,s^{-2}}$, R_L ($\mathrm{J\,kg^{-1}\,K^{-1}}$) is the gas constant of dry air, and $\cancel{T_v}\overline{T}_v$ (K) is the mean virtual temperature for the air layer between the two sites under consideration. p was computed in two steps. First, p was approximated using the dry-bulb temperature instead of $\cancel{T_v}\overline{T}_v$. Second, $\cancel{T_v}\overline{T}_v$ was calculated with the approximated p and used to recalculate p. The UFS was used as the reference site to calculate p at the next lower site (ZPLT), which was afterwards used as the reference site to calculate p at the next lower site (Schachen). This procedure was continued until Kreuzalm. Similarly, GAP was used as the reference site to determine p at the next higher site and this procedure was continued until Schachen. At Kreuzalm and Schachen, the two p estimates based on the next lower and next higher sites were averaged.

In order characterize the static stability of the valley atmosphere with a single quantity, the range of the pseudo-vertical profile of virtual potential temperature ($\Delta\theta_v$ (K)) was defined as follows, considering all stations except Schachen~~:~~,

$$\Delta\theta_v = \begin{cases} \theta_v^{\max} - \theta_v^{\min}, & \text{if } z_{\theta_v^{\min}} \leq z_{\theta_v^{\max}} \\ \theta_v^{\min} - \theta_v^{\max}, & \text{if } z_{\theta_v^{\min}} > z_{\theta_v^{\max}} \end{cases}, \tag{4}$$

where $\theta_v^{\min}$ (K) and $\theta_v^{\max}$ (K) are minimum and maximum virtual potential temperature of the pseudo-vertical profile, respectively, and $z_{\theta_v^{\min}}$ (m) and $z_{\theta_v^{\max}}$ (m) are the associated altitudes. Schachen was excluded from this calculation because θ_v was particularly high at Schachen if the global radiation was high and the wind velocity was low, which suggested that the naturally aspirated thermometer was affected by radiative errors (not shown). $\theta_v^{\min}$ was mostly found at the low-elevation sites GAP and Brandwiese, which applied to ~~57~~58 % and 37 % of the cases, respectively. $\theta_v^{\max}$ occurred mostly at the high-elevation sites ZSG, UFS, or ZPLT, which applied to 68 %, 18 %, and 13 % of the cases, respectively. The values of $\Delta\theta_v$ ranged between -5 K and $+31$ K, where positive values indicate stable conditions (see supplement, Fig. S1a).

2.3.3 Ceilometer data

To determine the MLH at GAP from the ALH retrievals of the ceilometer, several post-processing steps were necessary (Fig. 2). ALHs below 600 m a.g.l., corresponding to 1320 m a.s.l., were regarded as artifacts and were discarded because of high measurement uncertainties in this part of the backscatter profile resulting from an incomplete overlap of the laser beam and the field of view of the ceilometer (Flentje et al., 2010). The backscatter signal only reflects aerosol concentrations below clouds. Thus, ALHs that were greater or equal to the cloud base height, which was measured by the same instrument, were discarded. For this reason, it was not possible to determine the correct MLH if clouds were present at the top of or within the ~~ML~~mixing layer. These first two post-processing steps strongly reduced the data availability from 89 % to 45 % of the time.

If two or three ALHs were found at the same time, the following procedure was applied. For each ALH, the number of ALHs in its vicinity, i.e. within centered intervals of 300 m and 30 min in the vertical and temporal dimensions, respectively, was determined. The ALH with the maximum number of ALHs in its vicinity was attributed to the MLH while the other ALHs were discarded. If each of the ALHs had less than three ALHs in its vicinity, a missing value was registered. The MLH attribution, as described above, was based on the idea that the MLH varies only gradually so that similar MLHs are expected shortly before and after the time under consideration. In particular situations, in which elevated aerosol layers enter or leave the field of view of the ceilometer, this assumption may not be valid.

From the resulting time series at 15 s intervals, outliers were removed iteratively as follows. From each pair of subsequent data points with an absolute rate of change exceeding $240\ \mathrm{m\ min^{-1}}$, one data point was removed so that the standard deviation of the MLH in the preceeding 15 min and the following 15 min was most strongly reduced. This procedure was repeated until the absolute rate of change was maximum $240\ \mathrm{m\ min^{-1}}$ corresponding to four vertical averaging intervals $(4 \cdot 15\ \mathrm{m})$ per time interval (15 s).

The MLHs were aggregated using the 30 min median while requiring a data availability of ≥ 50 % and a standard deviation of $\leq 170\ \mathrm{m}$ corresponding to the 90 $\%$ percentile of all 30 min standard deviations. From the aggregated time series, outliers were removed in the same way as described above but with a maximum allowed absolute rate of change of $400\ \mathrm{m\ h^{-1}}$. Finally, the MLH was available for 34 % of the time.

2.4 Statistical classification approach

2.4.1 Definition of the air mass classes

The statistical analysis aimed at distinguishing the following ~~four~~ three classes of air masses: (a) ~~ML~~BL: Air masses with a recent contact with the ~~ML~~BL, characterized by increased concentrations of surface-emitted atmospheric constituents suggesting a recent uplift with a dominating influence of regional or local sources; (b) UFT/SIN: Air masses of the undisturbed free troposphere or stratospheric intrusions, characterized by very low concentrations of surface-emitted constituents suggesting recent subsidence or horizontal advection; and (c) HYBRID: Air masses that are influenced by both the ~~ML~~ BL and the free troposphere or that exhibit ambiguous characteristics. The class HYBRID was anticipated, for example, if the air mass had been exported from the ~~ML~~ BL to the free troposphere at a significant horizontal distance from the UFS and had been mixed with the free troposphere on the further trajectory, resulting in intermediate concentrations of surface-emitted constituents.

Stratospheric intrusions are characterized by high $^{7}\mathrm{Be}$ and $\mathrm{O_3}$ concentrations and a low humidity and have been observed at Mt. Zugpitze on approximately 5 $\%$ of the days (Stohl et al., 2000). Because of this low frequency and the coarse temporal resolution of $^{7}\mathrm{Be}$ of 12 h, stratospheric intrusions were not classified individually but attributed to the same class as air masses of the undisturbed free troposphere, which exhibit similar characteristics.

2.4.2 Principal component analysis

A PCA allows for reducing the number of dimensions of a data set while maintaining as much variance as possible. Principal components (PCs) are uncorrelated standardized linear combinations of the original variables (Mardia et al., 1979). If the variance of the original variables is mainly caused by shifts between the air mass classes, the first few PCs will be suitable indicators for air mass ~~discrimination~~classification. The PCA was performed separately for 2-month periods to (i) account for seasonally changing relationships between the variables and to (ii) largely eliminate the influence of the seasonal variability of atmospheric constituents because the seasonality reflects not only the frequency of ~~ML~~ BL influenced air masses but also other factors such as the source and sink strength (e.g. residential heating in winter, photosynthetic $\mathrm{CO_2}$ removal in summer), chemical reactivity, and deposition. The original variables were standardized by subtracting the arithmetic mean and dividing by the standard deviation of the respective 2-month period in order to avoid biased results due to different units and ranges of the data. The subtraction of the 2-month mean roughly removes the seasonal cycle.

Only ~~the most suitable variables~~ variables with an expected unambiguous link to vertical transport processes and a high data availability were used as input variables of the PCA~~whereas the other variableswere~~ . These variables include the gases CO, $\mathrm{CH_4}$, $\mathrm{CO_2}$, $\mathrm{O_3}$, and water vapor. Among the meteorological variables, the air pressure at GAP and $\Delta\theta_v$ were considered most suitable for the PCA because they are physically linked to uplift (low pressure, low static stability) and subsidence processes (high pressure, high static stability). The remaining variables were only used for validation purposes. Aerosol measurements were excluded from the PCA because low aerosol concentrations do not necessarily indicate UFT/SIN air masses but can also result from wet deposition during the uplift of ~~ML~~ BL air masses. $^{222}\mathrm{Rn}$ and HCHO were only used for validation purposes due to low data availability. $\mathrm{NO_y}$ and $\mathrm{NO_x}$ were excluded from the PCA because their variability was particularly strongly influenced by local emissions, which are not indicative of uplift processes. $\mathrm{O_3}$ was only used as PCA input variable in the winter half year because high $\mathrm{O_3}$ mixing ratios are not only caused by a subsidence of UFT/SIN air masses but can also result from photochemical $\mathrm{O_3}$ production in ~~ML~~ BL air masses, which typically contain high precursor concentrations. In the winter half year, photochemical $\mathrm{O_3}$ production was assumed to play a minor role at high-alpine sites because of the generally low solar irradiance and the weak thermally induced uplift. This argument is supported by trajectory residence time statistics of Kaiser et al. (2007), which demonstrated that in winter, the $\mathrm{O_3}$ mixing ratio at European high-alpine sites was generally lower if the air mass originated from lower altitudes, whereas in summer, this was often not the case. ~~Among the meteorological variables, the air pressure at GAP and $\Delta\theta_v$ were considered most suitable because they are physically linked to uplift (low pressure, low static stability) and subsidence processes (high pressure, high static stability).~~

The 2-month periods were defined as December to January, February to March, ..., so that the winter and summer half years included the months with the lowest and highest solar forcing, respectively. The PCs were computed from the following five gas-phase variables and two meteorological variables, where $\mathrm{O_3}$ was only included in the 2-month periods of the winter half year (October to March)~~.~~,

$$\mathrm{PC}_i = a_{1i}\,\widehat{[\mathrm{CO}]} + a_{2i}\,\widehat{[\mathrm{CH_4}]} + a_{3i}\,\widehat{[\mathrm{CO_2}]} + a_{4i}\,\widehat{[\mathrm{O_3}]} + a_{5i}\,\widehat{q} + a_{6i}\,\widehat{\Delta\theta_v} + a_{7i}\,\widehat{p_{\mathrm{GAP}}}. \tag{5}$$

Here, PC_i represents the scores of the i-th PC; $\widehat{[\mathrm{CO}]}$, $\widehat{[\mathrm{CH_4}]}$, $\widehat{[\mathrm{CO_2}]}$, and $\widehat{[\mathrm{O_3}]}$ are the standardized mixing ratios of the respective trace gases at the UFS; $\widehat{q}$ is the standardized specific humidity at the UFS; $\widehat{\Delta\theta_v}$ is the standardized range of the pseudo-vertical profile of virtual potential temperature; $\widehat{p_{\mathrm{GAP}}}$ is the standardized air pressure at GAP; and $a_{1i},\ldots,a_{7i}$ are the loadings of the i-th PC. All quantities in Eq. 5 are dimensionless due to standardization. The loadings represent standardized eigenvectors of the correlation matrix of the original variables while the associated eigenvalues correspond to the variances of the PC scores (Mardia et al., 1979).

2.4.3 Isolating regimes and classes of air masses

In each 2-month period, the air masses were divided into nine regimes using the 33.3 % and 66.6 % percentiles of the first two PCs as thresholds (Fig. 3~~, 8b). These regimes~~). The division of each of the two PCs into three equally large subsets follows the same logic as the definition of the three air mass regimes. ML and UFT/SIN air masses are expected on opposite edges of the distributions of the PC1 and PC2 scores while a mixture of these air masses may be found in the middle part of the distributions. The nine regimes were interpreted with respect to vertical transport and attributed to the three air mass classes ~~ML~~BL, UFT/SIN, and HYBRID by visually comparing summary statistics of the PCA input variables between the regimes. ~~The attribution was~~ In the winter half year, the regimes were assigned to the class BL if CO, $\mathrm{CH_4}$, $\mathrm{CO_2}$, and q were relatively high and $\Delta\theta_v$, p_{GAP} and $\mathrm{O_3}$ were relatively low compared to the other regimes. The class UFT/SIN was assigned in the opposite case and the remaining regimes were assigned to the class HYBRID. Apart from $\mathrm{CO_2}$ and $\mathrm{O_3}$, the same criteria were used in the summer half year. Due to photosynthesis, $\mathrm{CO_2}$ was required to be relatively low for BL air masses compared to other air masses during the summer half year. $\mathrm{O_3}$ was not used as an input variable in the summer half year.

This subjective mapping of regimes to classes of air masses was validated using summary statistics of the remaining ~~measurements.~~ chemical (e.g. $\mathrm{NO_y}$, $\mathrm{NO_x}$, $^{222}\mathrm{Rn}$) and standard meteorological measurements (e.g. precipitation, relative humidity). Long-range transport of mineral dust (LRMD) was regarded as a subclass of HYBRID and was identified with the following three criteria: (i) a high $\mathrm{PM_{10}}$ concentration of $\geq 13\ \mathrm{\mu g\,m^{-3}}$, (ii) a relatively low 30 min standard deviation of the $\mathrm{PM_{10}}$ of $\leq 0.2\ \mathrm{PM_{10}}$ to avoid the attribution of local pollution events, and (iii) an air mass regime that does not suggest a current uplift of ~~ML~~ BL air masses. The $\mathrm{PM_{10}}$ threshold of 13 $\mathrm{\mu g\,m^{-3}}$ was motivated by a local minimum in the $\mathrm{PM_{10}}$ histogram for ~~non-ML~~ non-BL air masses in June–July (not shown) and the plausibility was checked by visual inspection of the whole $\mathrm{PM_{10}}$ time series. Ambiguous air mass regimes, which carried the fingerprint of the marine boundary layer or the lower free troposphere (among others, very low CO and $\mathrm{CH_4}$ mixing ratios but high q and low $\mathrm{O_3}$ mixing ratios) were attributed to MBL/UFT, another subclass of HYBRID, unless the criteria for LRMD were fulfilled (Fig. 3). Characteristics of the marine boundary layer can be observed in the Alps if the air mass was lifted above the Atlantic Ocean and transported to the Alps within several days (Balzani Lööv et al., 2008).

2.5 Mechanistic classification approach

In search of criteria for thermally induced anabatic and katabatic winds that influence the UFS, the wind directions of all sites were investigated. The wind direction at the UFS was not useful because upslope and downslope flows were not clearly visible but appeared to be superimposed by synoptically driven winds due to the relatively exposed location of the UFS (not shown). Only at GAP, Felsenkanzel, and ZPLT, the wind directions exhibited pronounced diurnal patterns that indicated thermally induced mountain winds, especially in summer (Fig. 4). However, the wind data from Felsenkanzel were not used because of problems with shifts of the sensor orientation. Upvalley and downvalley winds were characterized by northeasterly and southwesterly wind direction sectors at GAP, respectively, and by easterly and westsouthwesterly wind direction sectors at the ZPLT, respectively (Fig. 4). The wind patterns at the ZPLT were consistent with the study of Gantner et al. (2003) who observed a thermal circulation above the same plateau with an easterly inflow of a few hundred meters depth in the daytime and a westerly outflow at night under fair weather conditions in summer. Modeling suggested that the inflow ascended the narrow and deep valley Reintal before reaching the ZPLT.

The wind direction sectors at the ZPLT and GAP were used as a criterion for anabatic and katabatic winds (Table 2). Additionally, anabatic and katabatic winds were restricted to cases with a weak static stability (~~$\Delta\theta_v > -8$~~ $\Delta\theta_v < 8$ K) and a strong static stability (~~$\Delta\theta_v \le -8$~~ $\Delta\theta_v \ge 8$ K), respectively, which was justified by the intersect between the frequency distributions of $\Delta\theta_v$ for the wind direction sectors described above (~~not shown~~see supplement, Fig. S1b). For katabatic winds, a low wind velocity of $< 3\ \mathrm{m\,s^{-1}}$ was required at the ZPLT and GAP to exclude cases with a strong synoptic forcing. Additionally, both anabatic and katabatic winds were required to persist for at least $1\ \mathrm{h}$, i.e. two time intervals (Table 2).

The detection of these winds and the ceilometer-based MLH at GAP were combined by a mechanistic approach that primarily accounted for thermally induced vertical transport~~and~~. For this purpose, the $\mathrm{MLH_{GAP}}$ was discarded if clouds were detected at a height of $< 4\ \mathrm{km\ a.s.l.}$ because low-level clouds increase the uncertainty of the MLH retrieval. The mechanistic approach distinguished between the following three conditions: (a) Anabatic winds occur or the UFS is below $\mathrm{MLH_{GAP}}$, which suggests ~~ML~~ BL air masses; (b) Katabatic winds occur and the UFS is above $\mathrm{MLH_{GAP}}$, which suggests UFT/SIN or HYBRID air masses; (c) The UFS is above $\mathrm{MLH_{GAP}}$ ~~but~~ and the winds are not thermally induced, which also suggests UFT/SIN or HYBRID air masses under the assumption of identical MLHs at the UFS and GAP. This approach was limited by the low availability of the MLH ~~of 34~~(22 % after discarding times with low-level clouds) but still allowed for a partial verification of the statistical approach.

3 Results and discussion

3.1 Principal components and their seasonal dependence

The PCA converts several input variables into the same number of PCs while the variance and thus the importance of the PCs is highest for the first few PCs. In the winter half year, the percentage of explained variance of the PCs decreased strongly from PC1 to PC2 and more slightly towards higher-order PCs. In the summer half year, the explained variance decreased approximately linearly with increasing PC number (Fig. 5). The first two PCs explained a total of 60 % to 72 % and 53 % to 58 % of the variance in the winter half year and summer half year, respectively. Hence, most of the variance was maintained when reducing the number of considered PCs from six or seven to two, especially in the winter half year.

The loadings of the first two PCs were similar among the 2-month periods of the winter half year but ~~more variable within~~ different and more variable in the summer half year (Fig. 6). Although the loadings differed between the 2-month periods, PC1 was always a meaningful indicator for air mass ~~discrimination~~classification, which will be explained in the following. PC2, as well as higher-order PCs, did not always allow for an unambiguous interpretation.

According to Eq. 5, variables with a high absolute loading determine the PC scores, i.e. the linear combinations, to a large extent. If, for example, an original variable is much higher than its 2-month mean value and its loading on PC1 is strongly negative then the variable will strongly contribute to a large negative score of PC1. In the winter half year, ~~low~~ large negative scores of PC1 primarily reflected high CO, CO_2, and CH_4 mixing ratios ~~, a rather high~~ and rather low values of $\Delta\theta_v$ (i.e. rather low static stability) ~~, and a rather low~~ and the air pressure at GAP (p_{GAP}), which suggested an uplift of ~~ML~~ BL air masses, while ~~high~~ large positive scores of PC1 reflected the opposite characteristics suggesting a subsidence of UFT/SIN air masses (Fig. 6). PC2 primarily separated air masses with low q, high O_3, and rather high CH_4 mixing ratios from air masses with opposite characteristics in the winter half year. On its own, PC2 was a less reliable indicator for vertical transport because low q and high O_3 mixing ratios suggested a subsidence of UFT/SIN air masses while high CH_4 mixing ratios suggested an uplift of ~~ML~~ BL air masses. Possibly, PC2 reflects not only vertical but also horizontal concentration gradients.

In the summer half year, the loading of CO_2 on PC1 had the opposite sign compared to the loadings of CH_4 and CO on PC1. This observation suggests that low CO_2 mixing ratios were generally indicative for ~~ML~~ BL air masses during the vegetation period due to the CO_2 removal by photosynthesis. Additionally, p_{GAP} contributed more strongly to the first two PCs than in the winter half year and was anti-correlated with q. $\Delta\theta_v$ exhibited very small absolute loadings on the first two PCs in the periods April–May and June–July, indicating a poor correlation with the other variables.

In April and May, an uplift of ~~ML~~ BL air masses was indicated by ~~low~~ large negative scores of PC1, which primarily reflected high q, low CO_2 mixing ratios, and low p_{GAP}, and by ~~low~~ large negative scores of PC2, which primarily reflected high CO and CH_4 mixing ratios.

In June and July, PC1 separated cases with high CH_4, low CO_2, and low p_{GAP}, which indicated an uplift of ~~ML~~ BL air masses, from cases with opposite characteristics, which indicated a subsidence of UFT/SIN air masses. In the same period, PC2 could not be interpreted unambiguously with respect to vertical transport because high CO mixing ratios, which suggested ~~ML~~ BL air masses, but also low q and high p_{GAP}, which suggested UFT/SIN air masses, contributed to ~~low~~ large negative scores of PC2.

In August and September, PC1 separated cases with high CO and CH_4 ~~,~~ and low CO_2 ~~, and high~~ and $\Delta\theta_v$ (i.e. low static stability), which was typical for ~~ML~~ BL air masses, from cases with opposite characteristics, which was typical for UFT/SIN air masses. Again, PC2 was a less reliable air mass indicator because low p_{GAP} and high q, which suggested an uplift of ~~ML~~ BL air masses, but also high CO_2 mixing ratios, which were typical for UFT/SIN air masses, increased the scores of PC2 in August and September (Fig. 6).

3.2 Air mass regimes in February and March

The interpretation of the air mass regimes, which were confined by the 33.3 % and 66.6 % percentiles of the first two PCs, differed among the 2-month periods except for the periods December–January and February–March. As an example, the period February–March is discussed in detail using summary statistics. A case study illustrating the classification results in the measured time series is shown in the supplement Sect. ~~S2. The air mass regimes were named as "x,y", where x and y are prominent features of PC1 and PC2, respectively. In the period February–March, x was called "polluted", "moderately polluted" (mod. poll.), or "unpolluted" with respect to the pollutants , , and while y was called "high q", "intermediate q" (interm. q), or "low/interm. q". The~~ S3. The joint data availability of the PCA input variables was 87 % of the time in February–March.

For the regimes ~~"polluted, high q", "polluted, interm. q", and "polluted, low/interm. q",~~ I, II, and III, the summary statistics of all seven PCA input variables were consistent with a recent uplift of ~~ML~~ BL air masses (Fig. 7a–g)~~.~~: Compared to the other regimes, CO, CH_4, and CO_2 generally exhibited high mixing ratios and q was also relatively high while O_3 and p_{GAP} were low or intermediate. $\Delta\theta_v$ indicated a weakly stable stratification with a median of approximately ~~8 K for the "polluted" regimes~~ 8 K for the regimes I to III.

The other chemical and meteorological measurements supported the interpretation of the ~~"polluted" regimes~~ regimes I to III (Fig. 7h–r). Relatively high values for rH (medians between 80 % and 90 %) and total precipitation were consistent with a current uplift of air masses. For NO_y and NO_x, local pollution events ($\sigma_{NO_x} > 0.4$ ppb) were excluded while for ^{7}Be, N_{90}, eBC, and PM_{10}, cases with precipitation at the ZPLT were excluded from the summary statistics because local emissions or wet deposition obscure the chemical signature resulting from vertical transport processes. Nevertheless, the data availability of these variables was 75 % to 78 % of the time. The ~~"polluted" regimes~~ regimes I to III exhibited high NO_y, NO_x, and ^{222}Rn and low ^{7}Be concentrations compared to the other regimes, which underlines the strong influence of the ~~ML~~BL. The high NO_x mixing ratios suggest a strong influence of combustion processes. High ^{222}Rn concentrations are typical for the continental ~~ML~~ BL because of natural emissions from the ice-free land surface (Griffiths et al., 2014). ^{7}Be concentrations are generally low in the ~~ML~~ BL due to the formation by cosmic rays in the stratosphere and upper troposphere (Stohl et al., 2000). The aerosol concentrations N_{90}, eBC, and PM_{10} were somewhat elevated for the ~~"polluted" regimes~~ regimes I to III but not as strongly as the gas-phase measurements. This finding could be due to wet deposition during the uplift of the air masses. At ZSG, southeasterly to southerly winds with varying velocities were frequently observed for the three regimes under consideration (see supplement, Fig. ~~S1a~~S2a,d,g). The observations suggest that ~~an uplift of ML~~ the influence of BL air masses was mostly caused by ~~a low pressure system or a south foehn event~~ low pressure systems or south foehn events in February and March.

The regimes ~~"mod. poll., high q", "mod. poll., interm. q", and "mod. poll., low/interm. q"~~ IV, V, and VI exhibited intermediate CO, CH_4, and CO_2 mixing ratios and an intermediate p_{GAP} (Fig. 7a–g). For these air masses, vertical transport processes were not inferable but the intermediate mixing ratios suggest an influence from both the ~~ML~~ BL and the free troposphere. Thus, the three regimes under consideration were attributed to the air mass class HYBRID, which was in line with intermediate NO_y, NO_x, and ^{222}Rn concentrations (Fig. 7l–n).

~~The regime "unpolluted, high q"~~ Regime VII was generally characterized by the lowest CO, CH_4, and CO_2 mixing ratios and a strongly stable stratification (median $\Delta\theta_v$ of ~~-15 K~~15 K). However, low O_3 mixing ratios, high q, and an intermediate and strongly variable p_{GAP} indicated that the air masses did not originate from the upper troposphere or stratosphere but from the lower troposphere or the marine boundary layer (Fig. 7a–g). This interpretation was in line with intermediate $^{7}\mathrm{Be}$ and $^{222}\mathrm{Rn}$ concentrations and low $\mathrm{NO_y}$, $\mathrm{NO_x}$, N_{90}, eBC, and $\mathrm{PM_{10}}$ concentrations (Fig. 7l–r). Intermediate rH with a median of 54 $\%$ corroborated some recent influence of the free troposphere. The air temperature was high compared to the other regimes (Fig. 7i), the wind direction at ZSG almost always had a south component, and the wind velocity at ZSG showed the highest mean of 10.22 $\mathrm{m\,s^{-1}}$ among the regimes (see supplement, Fig. ~~S1c~~S2c). The summary statistics did not allow for drawing a final conclusion on the influence of the marine boundary layer and thus, ~~the regime "unpolluted, high q"~~ regime VII was attributed to the air mass class HYBRID and the subclass MBL/UFT unless the criteria for LRMD were fulfilled.

Balzani Lööv et al. (2008) clustered back-trajectories for the Jungfraujoch, another mountain site in the Alps, and demonstrated that some air masses originated from the marine boundary layer above the tropical Atlantic Ocean and were transported to the Jungfraujoch inside the free troposphere within 5 to 15 days, resulting in rather high rH (average of 67 %), high temperatures, and low O_3, CO, $\mathrm{NO_y}$, and $\mathrm{NO_x}$ mixing ratios. Air masses with similar trajectories may also reach Mt. Zugspitze and represent, at least partly, the regime ~~"unpolluted, high q"~~VII.

For the regimes ~~"unpolluted, interm. q" and "unpolluted, low/interm. q",~~ VIII and IX, the CO, CH_4, and CO_2 mixing ratios were generally low but not as low as for ~~the regime "unpolluted, high q"~~regime VII. O_3 was intermediate for ~~the regime "unpolluted, interm. q"~~ regime VIII and high for ~~the regime "unpolluted, low/interm. q".~~ regime IX. Together with a strongly ~~negative~~ positive $\Delta\theta_v$ (median of approximately ~~-15 K~~15 K), i.e. a strongly stable stratification, and a generally high p_{GAP}, these observations indicated a subsidence of UFT/SIN air masses for the two regimes under consideration unless the criteria for LRMD were fulfilled (Fig. 7a–g). This interpretation was supported by a generally low rH (medians of 45 $\%$ and 25 $\%$), a total precipitation of almost zero, low $\mathrm{NO_y}$, $\mathrm{NO_x}$, N_{90}, eBC, and $\mathrm{PM_{10}}$ concentrations, and high $^{7}\mathrm{Be}$ concentrations compared to the other regimes (Fig. 7h–r).

For all regimes except ~~"polluted, low/interm. q",~~ III, the boxplots of global radiation (R_g) were similar with medians close to zero (Fig. 7k), indicating that the air mass characteristics were generally independent from day- and nighttime in February and March.

3.3 Seasonal frequencies of the air mass classes

In the entire year, the data availability limited the percentage of ~~classifiable~~ classified cases to 78$\%$ of the time. Data gaps occurred predominantly in the periods June–July and October–November (Fig. 8). On average, ~~ML~~BL, UFT/SIN, and HYBRID air masses accounted for 31 $\%$, 14 $\%$, and 55 $\%$ of the ~~classifiable~~ classified cases, respectively. The percentage of UFT/SIN air masses is in reasonable agreement with the study of Yuan et al. (2019), in which 13.6 $\%$ of the long-term CO_2 data from the UFS were selected as baseline concentrations using an univariate statistical approach called "adaptive diurnal minimum variation selection". For earlier CO_2 data from ZSG and a pedestrian tunnel, approximately 70 m above the UFS, the percentage of baseline concentrations was 19.5 $\%$ and 9.9 $\%$, respectively, according to the same study.

The air mass class ~~ML~~ BL was attributable to three of the nine air mass regimes in all 2-month periods except June–July and October–November when the class ~~ML~~ BL comprised two regimes (Fig. 8). This finding implies that air masses are lifted up to the UFS with similar frequencies in all seasons although the underlying processes may vary. However, the frequency of the air mass classes was constrained to some extent by using the 33.3 % and 66.6 % percentiles of the first two PCs as thresholds between the air mass regimes. These percentiles were somewhat arbitrary but allowed for a distinction between more unambiguous air masses at both edges of the distribution and more ambiguous air masses in the middle part of the distribution.

UFT/SIN air masses were attributable to two regimes in the winter half year but only to one regime in the period August–September and to none of the regimes in the periods April–May and June–July. UFT/SIN air masses were only evident if the summary statistics of an air mass regime indicated a dominating influence of subsidence. From April to July, subsidence appeared to be very rare at the UFS so that UFT/SIN air masses could not be isolated by splitting the data at the 33.3 % and 66.6 % percentiles of the first two PCs. This result can be explained by the fact that subsidence in high-pressure systems is counteracted by thermally induced uplift in the daytime, especially in the summer half year. However, the data availability in June–July was only 56 %, which questions the long-term representativeness for this period. Other statistical classifications at the high-alpine sites Jungfraujoch and Mt. Bachelor Observatory indicated that air masses of the undisturbed free troposphere occur year-round but less frequently in summer than in winter (Herrmann et al., 2015; Ambrose et al., 2011). In principle, these air masses do not have to descend but can also be advected horizontally to high-alpine sites. Such cases may be included in the class HYBRID and most likely in the subclass MBL/UFT.

HYBRID air masses including MBL/UFT comprised four to five regimes from August to March and even six and seven regimes in the periods April–May and June–July, respectively. In most of the 2-month periods, MBL/UFT air masses accounted for one regime (Fig. 8). Seven LRMD events were identified in the whole year, predominantly from April to July, which is in line with the study of Flentje et al. (2015) that found 5 to 15 Sahara dust events per year at the mountain site Hohenpeissenberg, approximately 40 km north of Mt. Zugspitze.

3.4 Comparing the statistical and mechanistic classifications

The ceilometer-based MLH at GAP exhibited a pronounced diurnal cycle with a maximum in the late afternoon or early evening on days with a high total global radiation (not shown). Little diurnal variations of the MLH were observed for days with a low total global radiation. Thermally induced anabatic winds were most frequently observed during summer and daytime whereas thermally induced katabatic winds occurred most frequently at night and in the first few hours after sunrise (not shown). The patterns described above were expected and confirmed the plausibility of the mechanistic approach.

~~When using the mechanistic approach for validation purposes, the MLH_{GAP} was discarded if clouds were detected at a height of $<$ 4 km a.s.l. because low level clouds increase the uncertainty of the MLH retrieval. Finally,~~ 25 % of the statistically ~~classifiable~~ classified cases were evaluated with the mechanistic approach. In 3 % of the statistically ~~classifiable~~ classified cases, the mechanistic approach suggested an influence of the local ~~ML~~ BL on the UFS, mostly due to thermally induced anabatic winds. It should be noted that the mechanistic approach is not able to detect synoptic uplift processes because low-level clouds and precipitation result in data gaps of the $\mathrm{MLH}_{\mathrm{GAP}}$. In 5 % of the statistically ~~classifiable~~ classified cases, katabatic winds and a low $\mathrm{MLH}_{\mathrm{GAP}}$ compared to the UFS level suggested ~~non-ML~~ non-BL air masses. In another 17 % of the statistically ~~classifiable~~ classified cases, the winds were not thermally induced ~~but~~ and the $\mathrm{MLH}_{\mathrm{GAP}}$ suggested ~~non-ML~~ non-BL air masses.

When the mechanistic approach suggested ~~ML~~ BL air masses, the statistical approach agreed in ~~54~~56 % of the cases and yielded HYBRID air masses in almost all other cases (Fig. 9). This finding indicates a poor performance of the statistical approach in the case of thermally induced uplift. On the other hand, thermally induced uplift only accounts for a small fraction of the time according to the mechanistic approach. Other uplift processes including non-local uplift were not identifiable with the applied mechanistic approach. LRMD was found in 3 % of the cases, which were mechanistically classified as influenced by the ~~ML~~BL, likely due to an overestimation of the ceilometer-based $\mathrm{MLH}_{\mathrm{GAP}}$ if the advected dust layer merged with the ~~ML~~mixing layer.

In the presence of katabatic winds and a low $\mathrm{MLH}_{\mathrm{GAP}}$ compared to the UFS level, the statistical classes UFT/SIN and HYBRID including LRMD and MBL/UFT were both considered to agree with the mechanistic approach (illustrated by the thick black line in Fig. 9). Thus, the level of agreement was 91 % for katabatic winds and a relatively low $\mathrm{MLH}_{\mathrm{GAP}}$, whereby UFT/SIN and HYBRID air masses accounted for approximately half of these cases, respectively. When the $\mathrm{MLH}_{\mathrm{GAP}}$ was lower than the UFS level ~~but~~ and the winds were not thermally induced, the level of agreement between the classifications was a bit lower (~~83~~86 %) with UFT/SIN and HYBRID air masses accounting for ~~31~~34 % and 52 %, respectively.

Uncertainties of the mechanistic approach arise mainly from the assumption of identical MLHs at the UFS and GAP and the lack of information on non-local uplift followed by horizontal advection of air masses. Uncertainties of the statistical approach result primarily from the lack of objective thresholds between the air mass classes and a varying significance of sources and sinks of the gases used. In view of these uncertainties, a reasonable agreement between the approaches was found when the mechanistic approach indicated ~~non-ML~~ non-BL air masses.

3.5 Implications for real-time operational mode

To classify air masses in real-time, the scores of the first two PCs can be approximated by using the loadings that were calculated in this study. Thus, the same thresholds as described here can be applied to determine the regime and the class of the current air mass. To standardize the original variables, the arithmetic mean and standard deviation of the current 2-month period are required but these statistics are only known at the end of the 2-month period. Due to the long-term trends of CO_2, CH_4, and CO and the interannual variability of all variables, it is advisable not to use the 2-month means and standard deviations from the year investigated in the present study but to estimate these statistics from recent multi-year records. To make use of the known part of the current 2-month period, the mean (μ) and standard deviation (σ) could be updated regularly using weighted averages,

$$\mu = (1-f)\mu_l + f\mu_c \tag{6}$$

$$\sigma = (1-f)\sigma_l + f\sigma_c \, , \tag{7}$$

where μ_l and σ_l are the 2-month mean and standard deviation, respectively, estimated from the long-term record, μ_c and σ_c are the mean and standard deviation of the known part of the current 2-month period, and f is the time fraction that is known in the current 2-month period. In real-time operational mode, additional uncertainties can arise from uncorrected measurement artifacts.

4 Conclusions

In this study, a novel statistical approach was developed to ~~discriminate~~ distinguish between the air mass classes ~~ML~~ BL (recent contact with the ~~mixing~~ boundary layer), UFT/SIN (undisturbed free troposphere or stratospheric intrusion), and HYBRID (influences of both the ~~mixing~~ boundary layer and the free troposphere or ambiguous characteristics) at the Schneefernerhaus at Mt. Zugspitze. A main purpose of the classification scheme is a later use in real-time operational mode. The scheme was based on the first two principal components, which were calculated from five gas-phase (CO, CH_4, CO_2, O_3, q) and two meteorological ($\Delta\theta_v$, p_{GAP}) variables but the principal components and their interpretation differed between the six 2-month periods. Additionally, the PM_{10} concentration and its 30 min standard deviation were needed to identify long-range transport of mineral dust. In retrospect, local pollution events on the mountain were evident from a high 30 min standard deviation of NO_x in approximately 9 % of the time. But these events are too short to be predictable on a 30 min basis in real-time applications and are thus neglected by the classification scheme.

While the first principal component was a suitable air mass indicator throughout the year, the second principal component left room for ambiguities. ~~ML~~ BL air masses were detected in all seasons with similar frequencies (average of 31 % of the ~~classifiable~~ classified cases). UFT/SIN air masses were predominantly found in the winter half year (23 % of the ~~classifiable~~ classified cases from October to March) but subsidence was so rare from April to July that the class UFT/SIN was not determinable in these months. HYBRID air masses (average of 55 % of the ~~classifiable~~ classified cases) mostly exhibited intermediate characteristics leaving room for ambiguities. For 17 % of HYBRID, it remained unclear whether the air mass originated from the lower undisturbed free troposphere or the marine boundary layer. To achieve a distinction between these two areas of origin, future work with trajectory models is necessary. 5 % of HYBRID was explained by long-range transport of mineral dust.

Independent chemical and standard meteorological measurements such as NO_y, ^{222}Rn, and precipitation were in line with the statistical approach. A mechanistic classification based on ceilometer and standard meteorological measurements was feasible in 25 % of the statistically ~~classifiable~~ classified cases and predominantly suggested ~~non-ML~~ non-BL air masses in these cases, which was in good agreement with the statistical approach. In the rare cases of thermally induced uplift, however, the statistical approach often misclassified ~~ML~~ BL air masses as HYBRID.

In principle, the statistical classification scheme can be used in real-time operational mode if the unknown arithmetic means and standard deviations of the observational variables in the current 2-month period can be estimated with sufficient accuracy. Future work should test the real-time applicability and quantify the uncertainty arising from the use of approximated means and standard deviations for standardization.

The framework of the presented statistical classification might also be useful at other high-alpine sites because it is based on common measurements.

Data availability. The $^{222}\mathrm{Rn}$ data are available from the World Data Centre for Greenhouse Gases at https://gaw.kishou.go.jp/search/file/0019-6031-6002-01-01-9999 after free registration. The other data can be requested from the authors.

Author contributions. AS, KF, CT, and TR developed the research concept and design. LR, CS, and AM collected measurement data and controlled the data quality. AS performed the data analysis. AS prepared the manuscript with contributions from all co-authors.

Competing interests. The authors declare that they have no conflict of interest.

Acknowledgements. This study was supported by the Bavarian State Ministry for the Environment and Consumer Protection in the framework of the project PureAlps (VH-ID: 70667 / TNT01 NaT-70667) and was performed in close coordination with the Bavarian Environment Agency. In this work, data from climate stations, which were set up during the project KLIMAGRAD (funded by the Bavarian State Ministry for the Environment and Consumer Protection, ZKL01Abt7_18458) were used. We thank Germany's National Meteorological Service for providing meteorological data from the sites ZSG, UFS, and GAP, ceilometer measurements, Rn^{222}, and $^{7}\mathrm{Be}$ data. Meteorological data from the UFS were provided through the Alpine Environmental Data Analysis Centre (AlpEnDAC.eu). We thank the AlpEnDAC staff for their support. Henry Schmölz (Bavarian Avalanche Warning Service) is acknowledged for providing meteorological data for the site ZPLT. We thank Markus Hermann and his team (Leibniz Institute for Tropospheric Research) for providing the SMPS data. Ralf Sohmer and Cedric Couret (German Environment Agency) are acknowledged for their assistance in the measurements and quality control of the trace gases. We thank Karl Lapo for his helpful comments on the manuscript. Manfred Kirchner and Matthias Mauder are acknowledged for valuable discussions. We additionally thank two anonymous reviewers for their comments, which helped to improve this manuscript.

References

Ambrose, J., Reidmiller, D., and Jaffe, D.: Causes of high O_3 in the lower free troposphere over the Pacific Northwest as observed at the Mt. Bachelor Observatory, Atmospheric Environment, 45, 5302 – 5315, https://doi.org/https://doi.org/10.1016/j.atmosenv.2011.06.056, 2011.

Balzani Lööv, J. M., Henne, S., Legreid, G., Staehelin, J., Reimann, S., Prévôt, A. S. H., Steinbacher, M., and Vollmer, M. K.: Estimation of background concentrations of trace gases at the Swiss Alpine site Jungfraujoch (3580 m asl), Journal of Geophysical Research: Atmospheres, 113, D22 305, https://doi.org/10.1029/2007JD009751, 2008.

Birmili, W., Weinhold, K., Rasch, F., Sonntag, A., Sun, J., Merkel, M., Wiedensohler, A., Bastian, S., Schladitz, A., Löschau, G., Cyrys, J., Pitz, M., Gu, J., Kusch, T., Flentje, H., Quass, U., Kaminski, H., Kuhlbusch, T. A. J., Meinhardt, F., Schwerin, A., Bath, O., Ries, L., Gerwig, H., Wirtz, K., and Fiebig, M.: Long-term observations of tropospheric particle number size distributions and equivalent black carbon mass concentrations in the German Ultrafine Aerosol Network (GUAN), Earth System Science Data, 8, 355–382, https://doi.org/10.5194/essd-8-355-2016, 2016.

Calvert, J. G.: Glossary of atmospheric chemistry terms (Recommendations 1990), Pure and Applied Chemistry, 62, 2167–2219, 1990.

Flentje, H., Heese, B., Reichardt, J., and Thomas, W.: Aerosol profiling using the ceilometer network of the German Meteorological Service, Atmospheric Measurement Techniques Discussions, 3, 3643–3673, https://doi.org/10.5194/amtd-3-3643-2010, 2010.

Flentje, H., Briel, B., Beck, C., Coen, M. C., Fricke, M., Cyrys, J., Gu, J., Pitz, M., and Thomas, W.: Identification and monitoring of Saharan dust: An inventory representative for south Germany since 1997, Atmospheric Environment, 109, 87 – 96, https://doi.org/https://doi.org/10.1016/j.atmosenv.2015.02.023, 2015.

Foken, T.: Micrometeorology, Springer, Berlin, Heidelberg, 2008.

Gallagher, J. P., McKendry, I. G., Cottle, P. W., Macdonald, A. M., Leaitch, W. R., and Strawbridge, K.: Application of Lidar Data to Assist Airmass Discrimination at the Whistler Mountaintop Observatory, Journal of Applied Meteorology and Climatology, 51, 1733–1739, https://doi.org/10.1175/JAMC-D-12-067.1, 2012.

Gantner, L., Hornsteiner, M., Egger, J., and Hartjenstein, G.: The diurnal circulation of Zugspitzplatt: observations and modeling, Meteorologische Zeitschrift, 12, 95–102, https://doi.org/doi:10.1127/0941-2948/2003/0012-0095, 2003.

Gohm, A., Harnisch, F., Vergeiner, J., Obleitner, F., Schnitzhofer, R., Hansel, A., Fix, A., Neininger, B., Emeis, S., and Schäfer, K.: Air Pollution Transport in an Alpine Valley: Results From Airborne and Ground-Based Observations, Boundary-Layer Meteorology, 131, 441–463, https://doi.org/10.1007/s10546-009-9371-9, 2009.

Griffiths, A. D., Conen, F., Weingartner, E., Zimmermann, L., Chambers, S. D., Williams, A. G., and Steinbacher, M.: Surface-to-mountaintop transport characterised by radon observations at the Jungfraujoch, Atmospheric Chemistry and Physics, 14, 12 763–12 779, https://doi.org/10.5194/acp-14-12763-2014, 2014.

Heese, B., Flentje, H., Althausen, D., Ansmann, A., and Frey, S.: Ceilometer lidar comparison: backscatter coefficient retrieval and signal-to-noise ratio determination, Atmospheric Measurement Techniques, 3, 1763–1770, https://doi.org/10.5194/amt-3-1763-2010, 2010.

Henne, S., Furger, M., Nyeki, S., Steinbacher, M., Neininger, B., de Wekker, S. F. J., Dommen, J., Spichtinger, N., Stohl, A., and Prévôt, A. S. H.: Quantification of topographic venting of boundary layer air to the free troposphere, Atmospheric Chemistry and Physics, 4, 497–509, https://doi.org/10.5194/acp-4-497-2004, 2004.

Herrmann, E., Weingartner, E., Henne, S., Vuilleumier, L., Bukowiecki, N., Steinbacher, M., Conen, F., Collaud Coen, M., Hammer, E., Jurányi, Z., Baltensperger, U., and Gysel, M.: Analysis of long-term aerosol size distribution data from Jungfraujoch with emphasis on free tropospheric conditions, cloud influence, and air mass transport, Journal of Geophysical Research: Atmospheres, 120, 9459–9480, https://doi.org/10.1002/2015JD023660, 2015JD023660, 2015.

Hulin, M., Gheusi, F., Lothon, M., Pont, V., Lohou, F., Ramonet, M., Delmotte, M., Derrien, S., Athier, G., Meyerfeld, Y., Bezombes, Y., Augustin, P., and Ravetta, F.: Observations of Thermally Driven Circulations in the Pyrenees: Comparison of Detection Methods and Impact on Atmospheric Composition Measured at a Mountaintop, Journal of Applied Meteorology and Climatology, 58, 717–740, https://doi.org/10.1175/JAMC-D-17-0268.1, 2019.

Kaiser, A., Scheifinger, H., Spangl, W., Weiss, A., Gilge, S., Fricke, W., Ries, L., Cemas, D., and Jesenovec, B.: Transport of nitrogen oxides, carbon monoxide and ozone to the Alpine Global Atmosphere Watch stations Jungfraujoch (Switzerland), Zugspitze and Hohenpeissenberg (Germany), Sonnblick (Austria) and Mt. Krvavec (Slovenia), Atmospheric Environment, 41, 9273 – 9287, https://doi.org/https://doi.org/10.1016/j.atmosenv.2007.09.027, 2007.

Ketterer, C., Zieger, P., Bukowiecki, N., Collaud Coen, M., Maier, O., Ruffieux, D., and Weingartner, E.: Investigation of the Planetary Boundary Layer in the Swiss Alps Using Remote Sensing and In Situ Measurements, Boundary-Layer Meteorology, 151, 317–334, https://doi.org/10.1007/s10546-013-9897-8, 2014.

Kirchner, M., Jakobi, G., Körner, W., Levy, W., Moche, W., Niedermoser, B., Schaub, M., Ries, L., Weiss, P., Antritter, F., Fischer, N., Henkelmann, B., and Schramm, K.-W.: Ambient Air Levels of Organochlorine Pesticides at Three High Alpine Monitoring Stations: Trends and Dependencies on Geographical Origin, Aerosol and Air Quality Research, 16, 738–751, https://doi.org/10.4209/aaqr.2015.04.0213, 2016.

Leuchner, M., Ghasemifard, H., Lüpke, M., Ries, L., Schunk, C., and Menzel, A.: Seasonal and Diurnal Variation of Formaldehyde and its Meteorological Drivers at the GAW Site Zugspitze, Aerosol and Air Quality Research, 16, 801–815, https://doi.org/10.4209/aaqr.2015.05.0334, 2016.

Löffler, H.: Meteorologische Bodenmesstechnik (vormals: Instrumentenkunde), vol. 6 of *Leitfaden für die Ausbildung im Deutschen Wetterdienst*, Selbstverlag des Deutschen Wetterdienstes, Offenbach am Main, 3 edn., https://www.dwd.de/DE/leistungen/pbfb_verlag_leitfaeden/l_einzelbaende/leitfaden6.html, 2012.

Mardia, K., Kent, J., and Bibby, J.: Multivariate analysis, Probability and mathematical statistics, Academic Press, 1979.

McClure, C. D., Jaffe, D. A., and Gao, H.: Carbon Dioxide in the Free Troposphere and Boundary Layer at the Mt. Bachelor Observatory, Aerosol and Air Quality Research, 16, 717–728, https://doi.org/10.4209/aaqr.2015.05.0323, 2016.

Poltera, Y., Martucci, G., Collaud Coen, M., Hervo, M., Emmenegger, L., Henne, S., Brunner, D., and Haefele, A.: PathfinderTURB: an automatic boundary layer algorithm. Development, validation and application to study the impact on in situ measurements at the Jungfraujoch, Atmospheric Chemistry and Physics, 17, 10 051–10 070, https://doi.org/10.5194/acp-17-10051-2017, 2017.

Reuten, C., Steyn, D. G., and Allen, S. E.: Water tank studies of atmospheric boundary layer structure and air pollution transport in upslope flow systems, Journal of Geophysical Research: Atmospheres, 112, D11 114, https://doi.org/10.1029/2006JD008045, 2007.

Risius, S., Xu, H., Di Lorenzo, F., Xi, H., Siebert, H., Shaw, R. A., and Bodenschatz, E.: Schneefernerhaus as a mountain research station for clouds and turbulence, Atmospheric Measurement Techniques, 8, 3209–3218, https://doi.org/10.5194/amt-8-3209-2015, 2015.

Ruckstuhl, A. F., Henne, S., Reimann, S., Steinbacher, M., Vollmer, M. K., O'Doherty, S., Buchmann, B., and Hueglin, C.: Robust extraction of baseline signal of atmospheric trace species using local regression, Atmospheric Measurement Techniques, 5, 2613–2624, https://doi.org/10.5194/amt-5-2613-2012, 2012.

Schuster, C., Kirchner, M., Jakobi, G., and Menzel, A.: Frequency of inversions affects senescence phenology of Acer pseudoplatanus and Fagus sylvatica, International Journal of Biometeorology, 58, 485–498, https://doi.org/10.1007/s00484-013-0709-0, 2014.

Seibert, P., Beyrich, F., Gryning, S.-E., Joffre, S., Rasmussen, A., and Tercier, P.: Review and intercomparison of operational methods for the determination of the mixing height, Atmospheric Environment, 34, 1001–1027, https://doi.org/10.1016/S1352-2310(99)00349-0, 2000.

Steinkopff, T., Frank, G., and Salvamoser, J.: Measurement of radionuclides within the Global Atmosphere Watch Programme of WMO, Kerntechnik, 77, 193–196, 2012.

Stohl, A., Spichtinger-Rakowsky, N., Bonasoni, P., Feldmann, H., Memmesheimer, M., Scheel, H., Trickl, T., Hübener, S., Ringer, W., and Mandl, M.: The influence of stratospheric intrusions on alpine ozone concentrations, Atmospheric Environment, 34, 1323 – 1354, https://doi.org/https://doi.org/10.1016/S1352-2310(99)00320-9, 2000.

Stull, R. B.: An Introduction to Boundary Layer Meteorology, Kluwer Acad. Publ., Dordrecht, Boston, London, 1988.

Sun, J., Birmili, W., Hermann, M., Tuch, T., Weinhold, K., Spindler, G., Schladitz, A., Bastian, S., Löschau, G., Cyrys, J., Gu, J., Flentje, H., Briel, B., Asbach, C., Kaminski, H., Ries, L., Sohmer, R., Gerwig, H., Wirtz, K., Meinhardt, F., Schwerin, A., Bath, O., Ma, N., and Wiedensohler, A.: Variability of Black Carbon Mass Concentrations, Sub-micrometer Particle Number Concentrations and Size Distributions: Results of the German Ultrafine Aerosol Network Ranging from City Street to High Alpine Locations, Atmospheric Environment, https://doi.org/https://doi.org/10.1016/j.atmosenv.2018.12.029, 2018.

WDCGG: World Data Centre for Greenhouse Gases, https://gaw.kishou.go.jp, last access: 2 February, 2019.

Young: Wind Monitor - Model 05103, http://www.youngusa.com/products/7/5.html, last access: 23 December, 2018.

Yuan, Y., Ries, L., Petermeier, H., Steinbacher, M., Gómez-Peláez, A. J., Leuenberger, M. C., Schumacher, M., Trickl, T., Couret, C., Meinhardt, F., and Menzel, A.: Adaptive selection of diurnal minimum variation: a statistical strategy to obtain representative atmospheric CO_2 data and its application to European elevated mountain stations, Atmospheric Measurement Techniques, 11, 1501–1514, https://doi.org/10.5194/amt-11-1501-2018, 2018.

Yuan, Y., Ries, L., Petermeier, H., Trickl, T., Leuchner, M., Couret, C., Sohmer, R., Meinhardt, F., and Menzel, A.: On the diurnal, weekly, and seasonal cycles and annual trends in atmospheric CO_2 at Mount Zugspitze, Germany, during 1981–2016, Atmospheric Chemistry and Physics, 19, 999–1012, https://doi.org/10.5194/acp-19-999-2019, 2019.

Zellweger, C., Forrer, J., Hofer, P., Nyeki, S., Schwarzenbach, B., Weingartner, E., Ammann, M., and Baltensperger, U.: Partitioning of reactive nitrogen (NO_y) and dependence on meteorological conditions in the lower free troposphere, Atmospheric Chemistry and Physics, 3, 779–796, https://doi.org/10.5194/acp-3-779-2003, 2003.

Zhang, L. and Jaffe, D. A.: Trends and sources of ozone and sub-micron aerosols at the Mt. Bachelor Observatory (MBO) during 2004-2015, Atmospheric Environment, 165, 143 – 154, https://doi.org/https://doi.org/10.1016/j.atmosenv.2017.06.042, 2017.

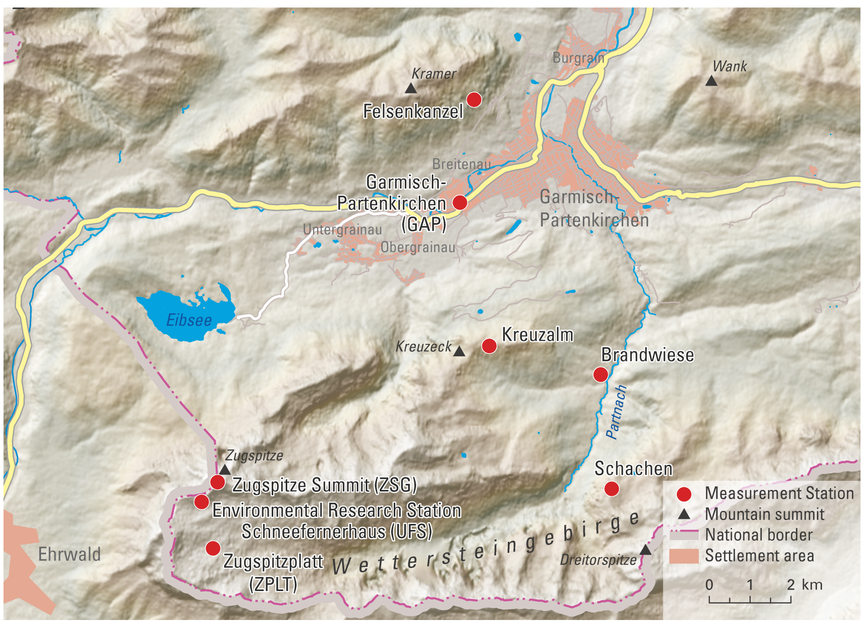

Figure 1. Location of the measurement sites in the surrounding of Mt. Zugspitze. North is at the top.

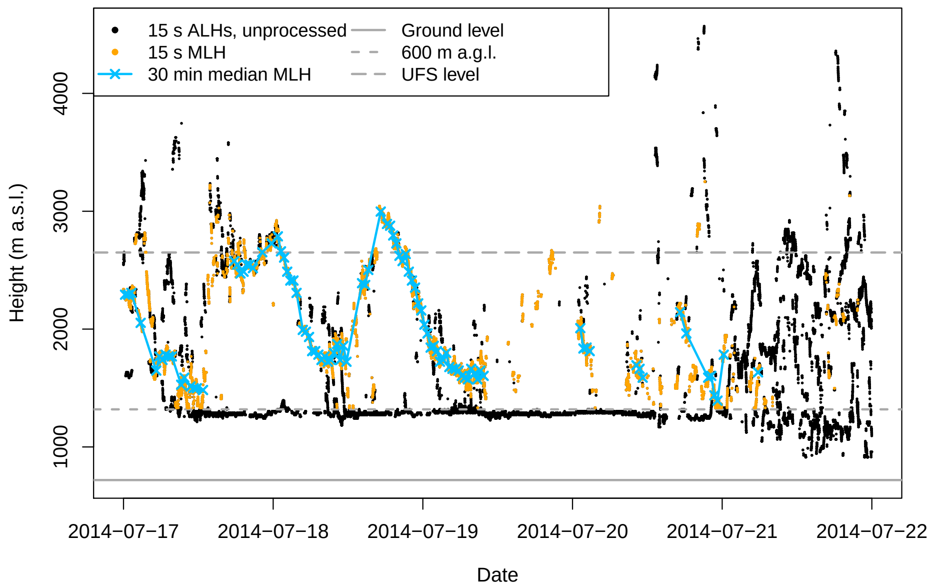

Figure 2. Post-processing of the ceilometer-based aerosol layer heights (ALHs), which were used to determine the mixing layer height (MLH) at GAP, illustrated for the period from 17 to 21 July 2014. From the MLH, outliers have been removed. Points of the 30 min median MLH were connected by a line if separated by maximum 3 h.

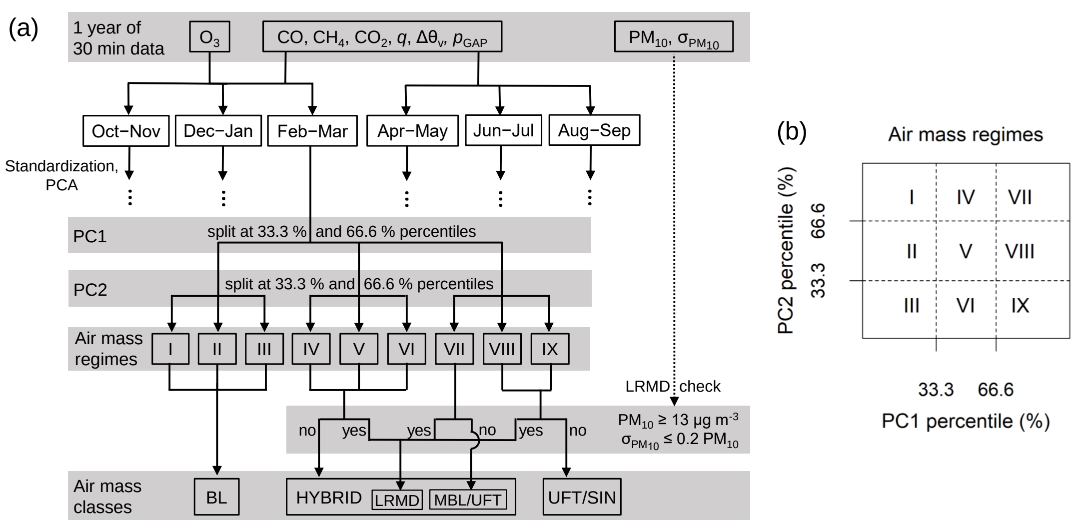

Figure 3. (a) Scheme for statistical air mass ~~discrimination~~ classification at the UFS using the first (PC1) and second principal components (PC2) of a 6- or 7-dimensional data set~~in each 2-month period, as illustrated~~. The mapping of air mass regimes to air mass classes is only shown for February and March and was different for other 2-month periods. ~~ML~~ BL denotes a recent contact with the ~~mixing~~ boundary layer, UFT/SIN includes the undisturbed free troposphere and stratospheric intrusions, and HYBRID reflects influences from both ~~ML~~ BL and UFT/SIN and includes the subclasses LRMD (long-range transport of mineral dust) and MBL/UFT (influence of the marine boundary layer or UFT). (b) Numbering of the air mass regimes.

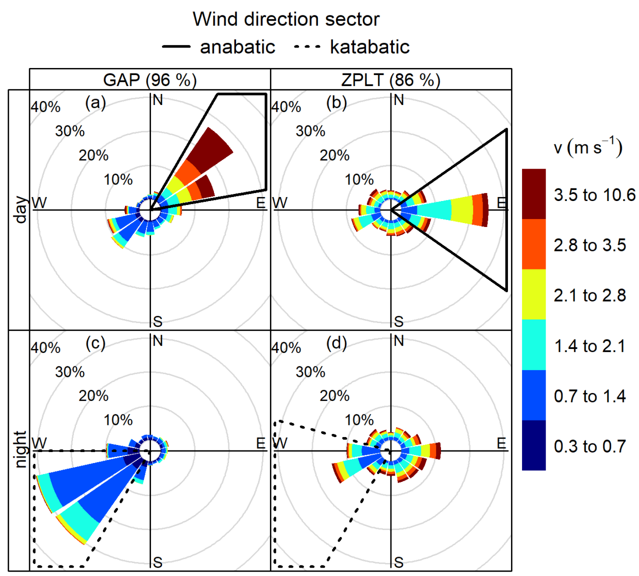

Figure 4. Wind rose plots showing the frequency of counts by wind direction and velocity (v) at the sites GAP (a,c) and ZPLT (b,d) during day- (a,b) and nighttime (c,d) in summer (June to August). Daytime was defined by a global radiation of $> 5\ \mathrm{W\,m^{-2}}$ at minimum one of the sites used in this study; all other cases were treated as nighttime. The plots were used to define wind direction sectors for thermally induced anabatic and katabatic winds. The percentage in brackets specifies the data availability.

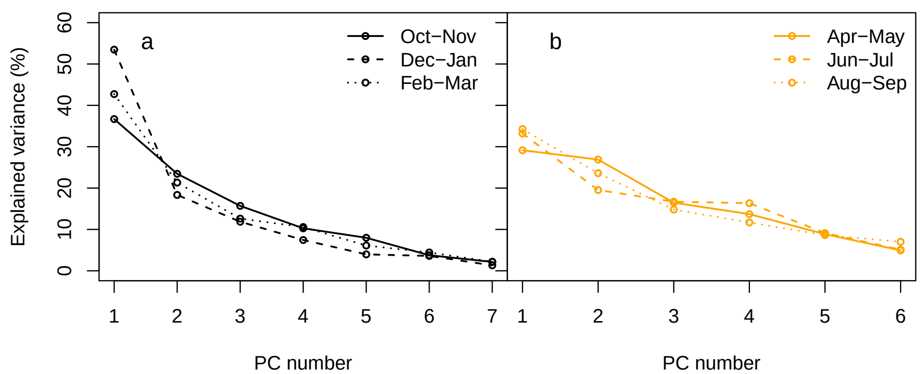

Figure 5. Percentage of explained variance as a function of the principal component (PC) number for each 2-month period: (a) winter half year and (b) summer half year.

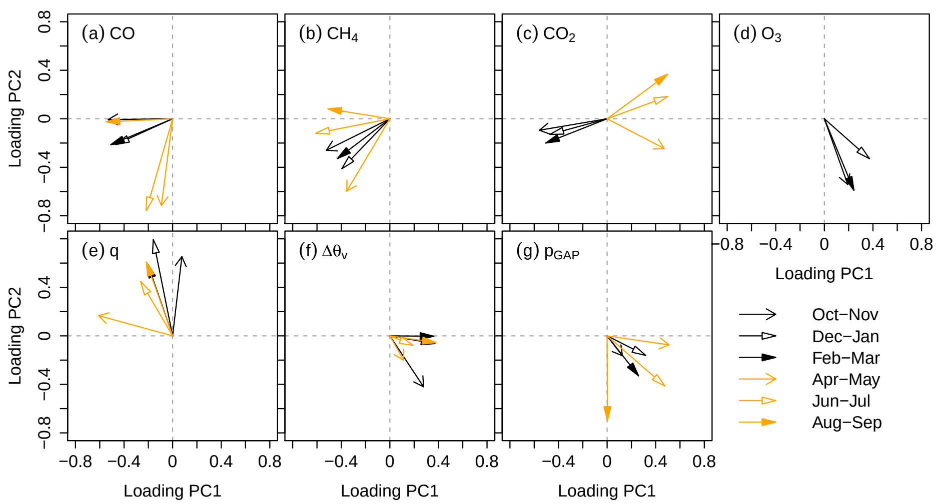

Figure 6. Loadings of the input variables on the first (PC1) and second principal component (PC2) for each 2-month period. O_3 was only used as an input variable in the winter half year (black arrows).

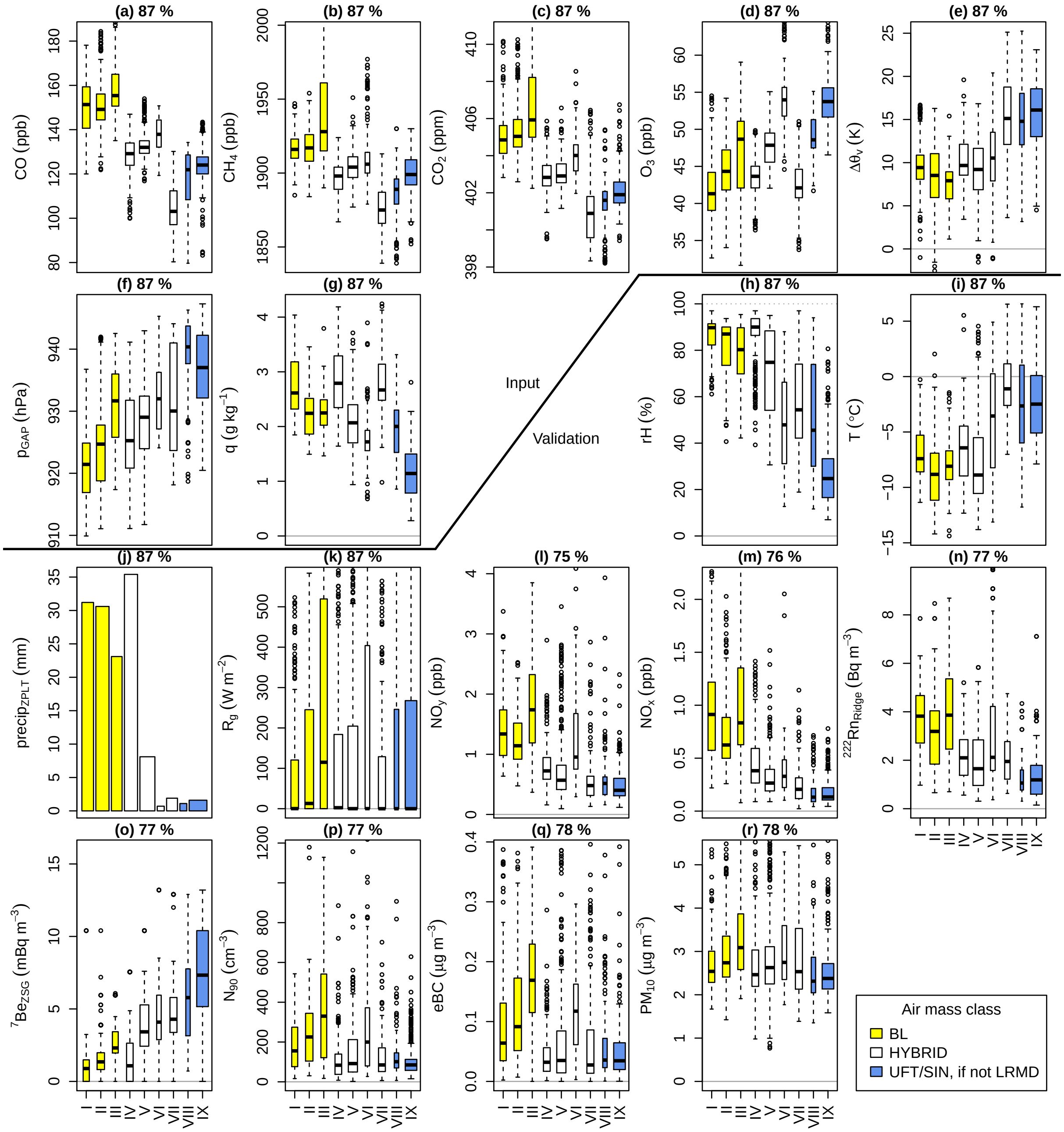

Figure 7. Summary statistics of PCA input variables (a–g) and validation variables (h–r) for the nine statistical air mass regimes in February and March 2014. The boxes show the quartiles and the boxwidth is proportional to the number of data points; each whisker is limited to a length of 1.5 times the interquartile range. Total precipitation ($\mathrm{precip_{ZPLT}}$) is depicted in a bar plot. Unless labeled differently, the data was measured at the UFS except for $\Delta\theta_v$ that represents the maximum difference among the sites. Symbols are explained in Sect. 2.2 and 2.3. Above the plots, the data availability is specified as fraction of time. For $\mathrm{NO_x}$ and $\mathrm{NO_y}$, local pollution events were excluded. For $^7\mathrm{Be_{ZSG}}$, N_{90}, eBC, and $\mathrm{PM_{10}}$, cases with precipitation ($\mathrm{precip_{ZPLT}} > 0$) were excluded. The axis of ordinate was limited to ~~$\mu \pm 3\sigma$~~ $m \pm 3\mathrm{MAD}$, where ~~μ~~ m and ~~σ~~ MAD are ~~mean~~ median and ~~standard~~ median absolute deviation of individual regimes, respectively.

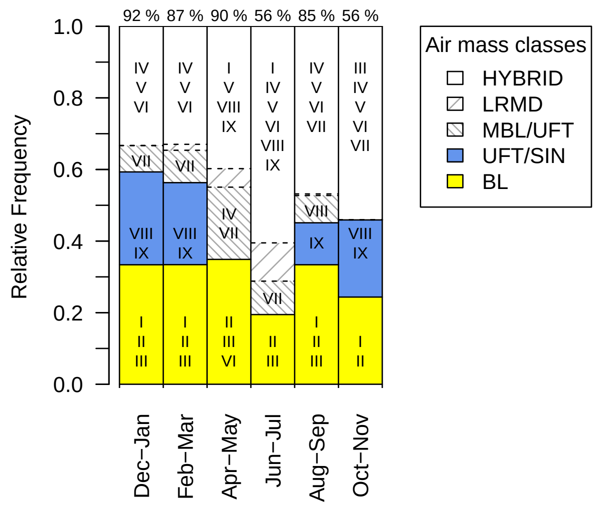

Figure 8. ~~(a)~~ Relative frequency of the air mass classes as fraction of the ~~classifiable~~ classified cases in each 2-month period. Above the plot, the ~~percentages of classifiable cases are~~ joint data availability is given as fraction of time. ~~(b) Numbering of the nine underlying air mass regimes.~~ LRMD (long-range transport of mineral dust) represents certain cases within the regimes that were not attributed to ~~ML~~ BL (recent contact with the ~~mixing~~ boundary layer). LRMD and MBL/UFT (influence of the marine boundary layer or UFT) are part of HYBRID (influences of both the ~~mixing~~ boundary layer and the free troposphere or ambiguous). UFT/SIN denotes undisturbed free troposphere or stratospheric intrusion.

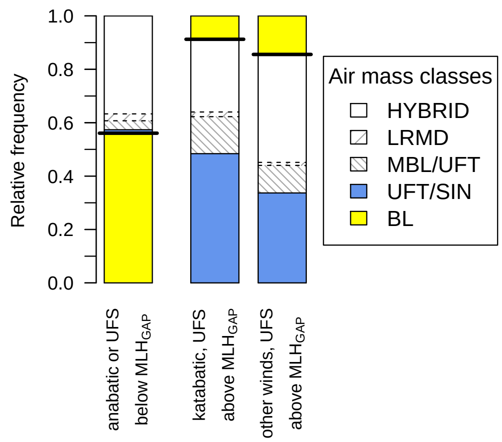

Figure 9. Comparison between the statistical and mechanistic classifications of air masses. The mechanistic classes were based on the ~~mixing~~ boundary layer height at GAP ($\mathrm{MLH_{GAP}}$) and thermally induced anabatic and katabatic winds (Sect. 2.5) and accounted for 3 %, 5 %, and 17 % of the statistically ~~classifiable~~ classified cases (from left to right). The thick black lines show the percentages considered as an agreement between the approaches. $\mathrm{MLH_{GAP}}$ was discarded if clouds were detected below 4 km a.s.l.. LRMD and MBL/UFT are subclasses of HYBRID.

Table 1. Chemical measurements used in this study. The given interval is the temporal interval at which the data were provided by the respective institution. UV denotes ultraviolet light. The symbols for the measured quantities are explained in the text.

Symbol	Unit	Interval	Instrument	Measurement principle	Institution
CO	ppb	1 min	AL5002, Aero-Laser, Garmisch-Partenkirchen, Germany	UV fluorescence	UBA[e]
CH_4	ppb	1 min[a]	EnviroSense 3000i, PICARRO, Sunnyvale, USA	cavity ringdown spectroscopy	UBA[e]
CO_2	ppm	1 min[a]	EnviroSense 3000i, PICARRO, Sunnyvale, USA	cavity ringdown spectroscopy	UBA[e]
O_3	ppb	1 min	Model 49*i*, Thermo Scientific, Waltham, USA	UV absorption	UBA[e]
NO_y,NO_x	ppb	1 min	CraNOx II, ECO PHYSICS, Dürnten, Switzerland	chemiluminescence	UBA[e]
HCHO	ppb	10 min	AL4021, Aero-Laser, Garmisch-Partenkirchen, Germany	fluorometric Hantzsch reaction[c]	TUM[f]
^{222}Rn	$Bq\,m^{-3}$	2 h	Radon sphere 270, IGU, Wörthsee, Germany	electrostatic deposition followed by alpha particle spectrometry[d]	DWD[g]
^{7}Be	$mBq\,m^{-3}$	12 h	Glass fiber filter and pump, Tracerlab, Köln, Germany	gamma spectrometry	DWD[g]
$dN\ (dlog\ d_p)^{-1}$	cm^{-3}	10 min	SMPS[b] model 3936, TSI Inc., Shoreview, USA	separation of charged particles in an electric field	TROPOS[h]
PM_{10}	$\mu g\,m^{-3}$	1 min	FH 62 C14, Thermo Scientific, Waltham, USA	attenuation of beta radiation	UBA[e]
eBC	$\mu g\,m^{-3}$	1 min	MAAP[b] model 5012, Thermo Scientific, Waltham, USA	light attenuation and reflection by particle-laden quartz fiber filters	UBA[e]

[a] 30 min in the year 2014; [b] see Birmili et al. (2016) and Sun et al. (2018) for details on the Scanning Mobility Particle Sizer (SMPS) and the Multi Angle Absorption Photometer (MAAP); [c] see Leuchner et al. (2016) for details; [d] see Steinkopff et al. (2012) for details; [e] German Environment Agency; [f] Ecoclimatology, Technical University of Munich; [g] German Meteorological Service; [h] Experimental Aerosol and Cloud Microphysics, Leibniz Institute for Tropospheric Research

Table 2. Criteria for thermally induced anabatic and katabatic winds occurring simultaneously at GAP and ZPLT. φ is wind direction, v is wind velocity, and $\Delta\theta_v$ is the range of the pseudo-vertical profile of virtual potential temperature.

Wind class	φ_{GAP}	φ_{ZPLT}	$\Delta\theta_v$	v_{ZPLT} and v_{GAP}	Duration
Anabatic	30° to 80°	55° to 125°	$\cancel{> 8}\ \leq 8\,\mathrm{K}$		$\geq 1\ \mathrm{h}$
Katabatic	210° to 270°	210° to 285°	$\cancel{\leq 8}\ \geq 8\,\mathrm{K}$	$< 3\,\mathrm{m\,s^{-1}}$	$\geq 1\ \mathrm{h}$

Supplement

S1 Atmospheric stability

The range of the pseudo-vertical profile of virtual potential temperature ($\Delta\theta_v$) was almost always positive, indicating stable conditions. In June and July, $\Delta\theta_v$ ranged between -5 K and $+21$ K whereas in December and January, it was generally more positive with values between 4 K and 31 K (Fig. S1a). Upvalley winds at both GAP and ZPLT were generally associated with a less negative $\Delta\theta_v$ compared to downvalley winds, which gave rise to the threshold of 8 K for the distinction between anabatic and katabatic winds (Fig. S1b).

~~S1~~ S2 Wind patterns at Zugspitze summit in February and March

Figure S2 shows the wind patterns at ZSG for each of the nine air mass regimes in the 2-month period February–March. The ~~three "polluted" regimes indicated that ML~~ regimes I , II, and III indicated that BL air masses were often associated with southeasterly to southerly winds with varying velocities (Fig. S2~~a,d,g~~a-c). Some of these air masses may reflect south foehn events, especially in the case of strong southerly winds.

For all other regimes, the wind direction was more variable. Strong southerly winds were also included in the regimes ~~"mod. poll., high q", "unpolluted, high q", and "mod. poll., interm. q"~~ IV, V, and VII (Fig. S2~~b,c,e~~d,e,g), which suggests that foehn flows are not always associated with a strong uplift on the windward side of the Alps (Seibert, 1990) and can result in varying air mass characteristics. ~~The regime "unpolluted, high q"~~ Regime VII exhibited the highest mean wind velocity (10.22 $\mathrm{m\,s^{-1}}$) among the regimes, which would be in line with a fast transport of the air masses from the marine boundary layer to the UFS (Fig. S2~~e~~g).

The regimes ~~"unpolluted, low/interm. q" and "unpolluted, interm. q",~~ VIII and IX, which were dominated by UFT/SIN air masses, exhibited similar mean wind velocities compared to the regimes of ~~ML~~ BL air masses (Fig. S2).

~~S2~~ S3 Statistical classification: Case study in March

In a case study, the regimes and classes of air masses were highlighted in the measured time series to gain insight in the transport processes involved and to check the plausibility of the classification. The period from 1 to 13 March 2014 mainly included two phases with contrasting air mass characteristics (Fig. S3). From 1 to 6 March, the three air mass regimes, which were attributed to the class ~~ML~~BL, dominated (Fig. S3a) – mainly due to high CO (Fig. S3b), CH_4 (Fig. S3c), and CO_2 mixing ratios (Fig. S3d) that peaked on 5 and 6 March. From 7 to 12 March, the two air mass regimes, which were attributed to the class UFT/SIN in the absence of LRMD, dominated – due to low CO, CH_4, and CO_2 mixing ratios, predominantly low q (Fig. S3e), and high O_3 mixing ratios (Fig. S3f). On 13 March, ~~the regime "unpolluted, high q"~~ regime VII indicated ambiguous air masses that originated either from the lower free troposphere or the marine boundary layer (Fig. S3a).

$\Delta\theta_v$ showed weak and strong diurnal variations in the phases from 1 to 6 March and from 7 to 13 March, respectively (Fig. S3g), indicating a shift from cloudy to clear-sky conditions, as confirmed by R_g measurements (Fig. S3s). p_{GAP} reached a minimum on 3 March, increased continuously and strongly until 6 March, and remained high until 13 March (Fig. S3h). These observations suggest that the ~~ML~~ BL air masses were lifted by a low pressure system and associated fronts whereas the UFT/SIN air masses descended in a high pressure system. This interpretation was supported by high rH (Fig. S3p), precipitation (Fig. S3r), and a low R_g (Fig. S3s) during the strong pressure increase and low rH, absent precipitation and high R_g during the high pressure phase.

The remaining chemical measurements were in line with the classification. $\mathrm{NO_y}$, $\mathrm{NO_x}$ (Fig. S3m), and $^{222}\mathrm{Rn}$ concentrations (Fig. S3n) were substantially higher and $^{7}\mathrm{Be}$ concentrations (Fig. S3o) were much lower for the ~~ML~~ BL than for the UFT/SIN air masses in the case study. The eBC (Fig. S3j) and N_{90} concentrations (Fig. S3k) tended to be higher for the ~~ML~~ BL than for the UFT/SIN air masses but temporary wet deposition resulted in strong variations (Fig. S3j,k). The $\mathrm{PM_{10}}$ concentration was only slightly higher for the ~~ML~~ BL air masses (median of 3.0 $\mu\mathrm{g\,m^{-3}}$) than for the UFT/SIN air masses (median of 2.3 $\mu\mathrm{g\,m^{-3}}$) (Fig. S3i). On 11 March 2014, a stratospheric intrusion was evident from exceptionally high Be-7 (Fig. S3o) and O_3 concentrations (Fig. S3f) of 26 $\mathrm{mBq\,m^{-3}}$ and 79 ppb, respectively, and a low relative humidity of 12 % (Fig. S3p).

The ceilometer-based MLH at GAP was mostly missing in the period from 1 to 6 March because the uplift of ~~ML~~ BL air masses was associated with low-level clouds and precipitation (Fig. S3v). From 7 to 12 March, the MLH at GAP mostly showed diurnal variations with afternoon maxima but remained at least 300 m lower than the level of the UFS, which is in line with the statistical classification. Wind direction (Fig. S3t) and velocity (Fig. S3u) at ZSG did not differ significantly between ~~ML~~ BL and UFT/SIN air masses in the case study.

References

Seibert, P.: South Foehn Studies Since the ALPEX Experiment, Meteorology and Atmospheric Physics, 43, 91–103, 1990.

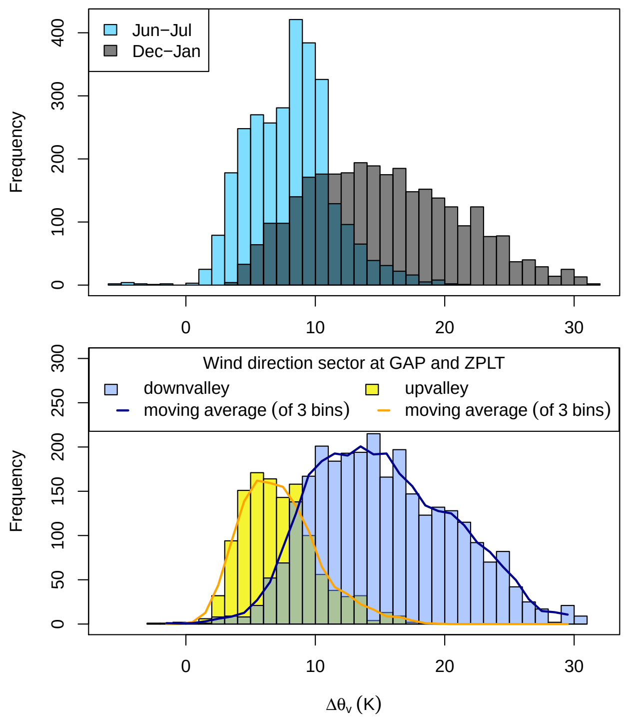

Figure S1. (a) Histograms of the range of the pseudo-vertical profile of virtual potential temperature ($\Delta\theta_v$) for (a) the periods from June to July and from December to January and for (b) upvalley and downvalley wind direction sectors at the stations GAP and ZPLT. Positive values represent stable conditions.

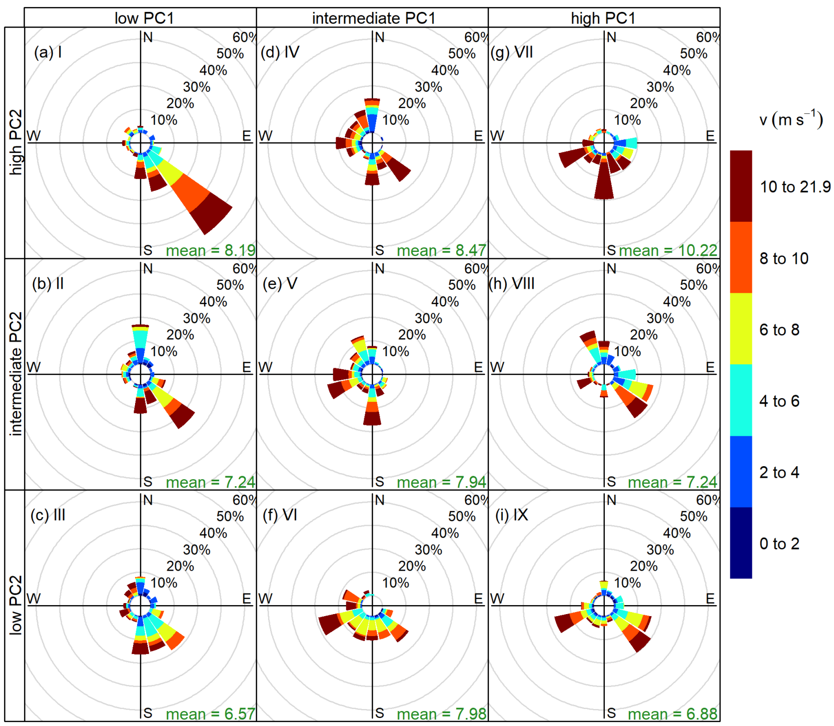

Figure S2. Windrose plots showing the frequency of counts by wind direction and velocity (v) at Zugspitze summit for the nine air mass regimes (I to IX) in February and March 2014. PC1 and PC2 denote the first and second principal components, respectively.

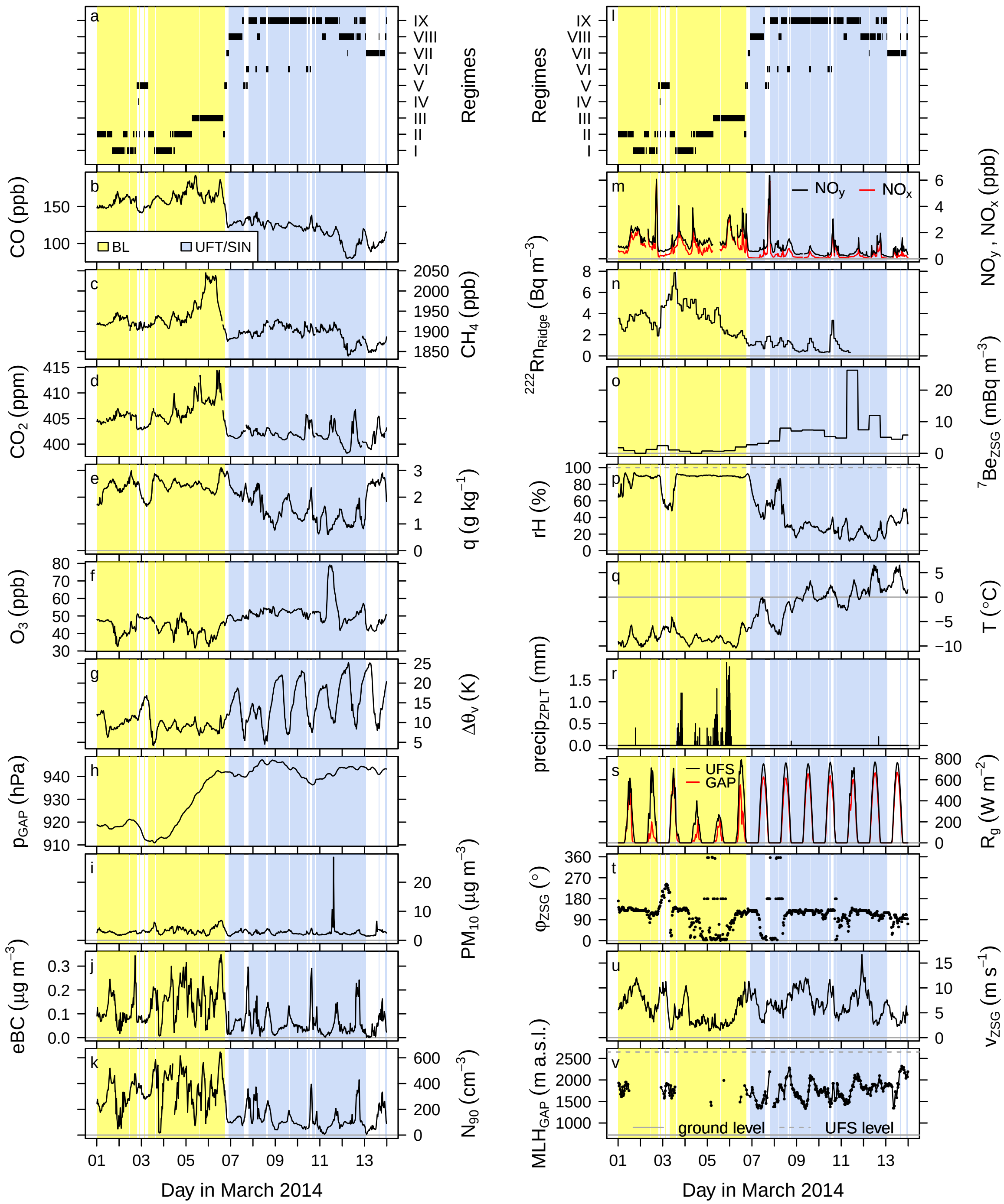

Figure S3. Time series of air mass regimes (a,l), input variables for the classification (b–i), and validation variables (j,k,m–v) from 1 to 13 March 2014. Unless labeled differently, the data was measured at the UFS except for $\Delta\theta_v$ that represents the maximum difference among the sites. The color shading highlights the most important air mass classes. Symbols are explained in Sect. 2.2 and 2.3 of the main article.