Advances in
Science & Research
Open Access Proceedings

Data-driven exploration of orographic enhancement of precipitation

L. Foresti[1], M. Kanevski[1], and A. Pozdnoukhov[2]

[1]Institute of Geomatics and Analysis of Risk, University of Lausanne, Switzerland
[2]National Centre for Geocomputation, National University of Ireland Maynooth, Ireland

Abstract. This study presents a methodology to analyse orographic enhancement of precipitation using sequences of radar images and a digital elevation model. Image processing techniques are applied to extract precipitation cells from radar imagery. DEM is used to derive the topographic indices potentially relevant to orographic precipitation enhancement at different spatial scales, e.g. terrain convexity and slope exposure to mesoscale flows. Two recently developed machine learning algorithms are then used to analyse the relationship between the repeatability of precipitation patterns and the underlying topography. Spectral clustering is first used to characterize stratification of the precipitation cells according to different mesoscale flows and exposure to the crest of the Alps. At a second step, support vector machine classifiers are applied to build a computational model which discriminates persistent precipitation cells from all the others (not showing a relationship to topography) in the space of topographic conditioning factors. Upwind slopes and hill tops were found to be the topographic features leading to precipitation repeatability and persistence. Maps of orographic enhancement susceptibility can be computed for a given flow, topography and forecasted smooth precipitation fields and used to improve nowcasting models or correct windward and leeward biases in numerical weather prediction models.

1 Introduction

The orographic precipitation enhancement is a complex atmospheric phenomenon which is the subject of many numerical (Rotunno and Houze, 2007) and observational studies (Gray and Seed, 2000; Panziera and Germann, 2010). High-resolution numerical weather prediction (NWP) models are computationally demanding to provide fast forecasts with appropriate data assimilation systems. Expert-based statistical approaches are developed to avoid these flaws. Such alternatives are successfully applied for thunderstorm nowcasting (Wilson and Gallant, 2000; Williams et al., 2008), and are also appearing in the context of orographic precipitation nowcasting (Panziera et al., 2010).

This study introduces an efficient computational alternative to analyse and to model orographic enhancement of precipitation from a sequence of radar images and a digital elevation model (DEM). The study considers how the terrain features such as terrain convexity and slope exposure to mesoscale flows help in explaining persistent patterns of orographic precipitation. Precipitation cells and the corresponding flow directions are extracted from radar images and attributed to the pre-computed underlying topographic variables. The orographic enhancement is defined as the ability of topography to enforce repeatability to particular precipitation patterns such as stationary cells, stable upslope ascent and localized thunderstorms. Evidences of high counts of cells repeatability reveal the topographic conditions and locations where the phenomenon is accentuated. This formulation allows characterizing precipitation enhancement using data-driven classification models. The system can be applied to simulate the localized enhancement under given flow and large scale precipitation patterns derived from nowcasting or NWP models.

Correspondence to: L. Foresti
(loris.foresti@unil.ch)

The paper is organized as follows. Section 2 explains the methodology. Section 3 describes the data preparation. Its exploration is shown in Sect. 4. The computational model of orographic enhancement is explained in Sect. 5.

2 Methodology

The methodology is illustrated in the work-flow diagram in Fig. 1. It can be summarized in four main steps:

1. Compute terrain indices such as convexity and gradients at different spatial scales from the DEM.

2. Estimate the motion vector field and extract the geographical location of precipitation cells from a representative sequence of radar images of orographic precipitation events. Compute indices for slope exposure to wind direction (*flow derivative*) using the motion vector field and terrain gradients.

3. Explore the dataset using methods of clustering to find natural partitions (classes) of mesoscale flows and exposure of cells with respect to the main Alpine ridge. Select the clusters presenting potential orographic conditions (windward clusters). Within these clusters, analyse the cells' repeatability to detect the places prone to precipitation persistence and those which are not.

4. Build a statistical classification algorithm separating orographic precipitation cells from non-orographic ones in the space of features. Based on new nowcasted precipitation fields, mesoscale flows and the underlying topography, compute the susceptibility of orographic enhancement.

More details on step 1, 2 and 3 can be found in Foresti and Pozdnoukhov (2010). Preliminary results of step 4 are presented in this paper.

3 Data preparation

Radar images used for testing the methodology concern the Swiss Alps in the period from 18 to 23 August 2005. This orographic precipitation event touched in particular the northern side of the Alps (Rotach et al., 2006). Precipitation amounts exceeded 200 mm in three days with return periods above 100 yr at several weather stations (Frei, 2006). The available radar imagery has a temporal resolution of 5 min and a spatial resolution of $1 \times 1\,km^2$ (Fig. 1, step 2a). It has been pre-processed to correct the vertical profiles in sheltered regions, to eliminate radar-rain gauge biases due to reduced visibility, to remove ground clutter and to account for the bright-band effect (Germann et al., 2006). The DEM used to derive the topographic information has a resolution of $250 \times 250\,m^2$. The topographic features are computed at the $1 \times 1\,km^2$ grid of the radar.

Data preparation passes through three main steps: the processing of the DEM, the estimation of motion vector field from subsequent radar images and the extraction of precipitation cells.

Feature extraction from DEM (Fig. 1, step 1a) was performed with Gaussian convolution filters to compute terrain convexity and terrain gradient (Fig. 1, step 1b). Features were derived at different spatial scales (degrees of smoothness) by applying convolution kernels with different bandwidths σ. More details about the extraction and the use of these features for meteorological applications can be found in Pozdnoukhov et al. (2009) and Foresti et al. (2011).

The motion vector field (Fig. 1, step 2b) is estimated from two consecutive radar images using the optical flow algorithm explained in Sun et al. (2008). Other studies consider variational techniques to the robust estimation of flow (Germann and Zawadzki, 2002). Common parameters of these algorithms allow controlling the trade-off between the precision and the spatial smoothness of the estimated field. In our approach we set the regularization parameters to have a smooth estimation of flow direction by minimizing the perturbations due to cell dissipation and growth particularly in convective situations. The flow derivative (FD) highlighting upwind slopes is computed from the terrain gradient and the motion vector field as follows:

$$FD(\boldsymbol{x},t) = \boldsymbol{\nabla} z(\boldsymbol{x}) \cdot \boldsymbol{u}(\boldsymbol{x},t) \tag{1}$$

where $\boldsymbol{\nabla} z(\boldsymbol{x})$ is the gradient vector of elevation evaluated at the (X, Y) spatial coordinates $\boldsymbol{x}$, $\boldsymbol{u}(\boldsymbol{x},t)$ is the flow vector with (u,v) components estimated at $\boldsymbol{x}$ at time t.

Several algorithms are available to detect precipitation cells from radar imagery (Lakshmanan et al., 2003; Wilson et al., 2004). In this study cells were identified by a simple method that finds the points of maxima of a smoothed precipitation field. It was done by subtracting two precipitation fields smoothed with different bandwidths σ. The resulting images describe precipitation anomalies and enable a robust selection of cells while filtering out most of clutter effects (Fig. 1, step 2b).

This processing is done on the dataset of radar images every 5 min within 6 days of precipitation (1728 images). The final dataset is composed of 28758 cells (observations) embedded in a space of 18 dimensions: [elevation | 7 convexities | 7 flow derivatives | precipitation | u,v flow components]. All input variables computed on the whole grid are stored in order to test the models under different weather situations. All data processing steps were implemented in Matlab.

4 Exploration of precipitation cells

The exploration of precipitation cells is done in two steps. First, the different flow situations (direction and strength) and the exposure of precipitation cells relative to the main Alpine ridge, described by a very large scale flow derivative, are
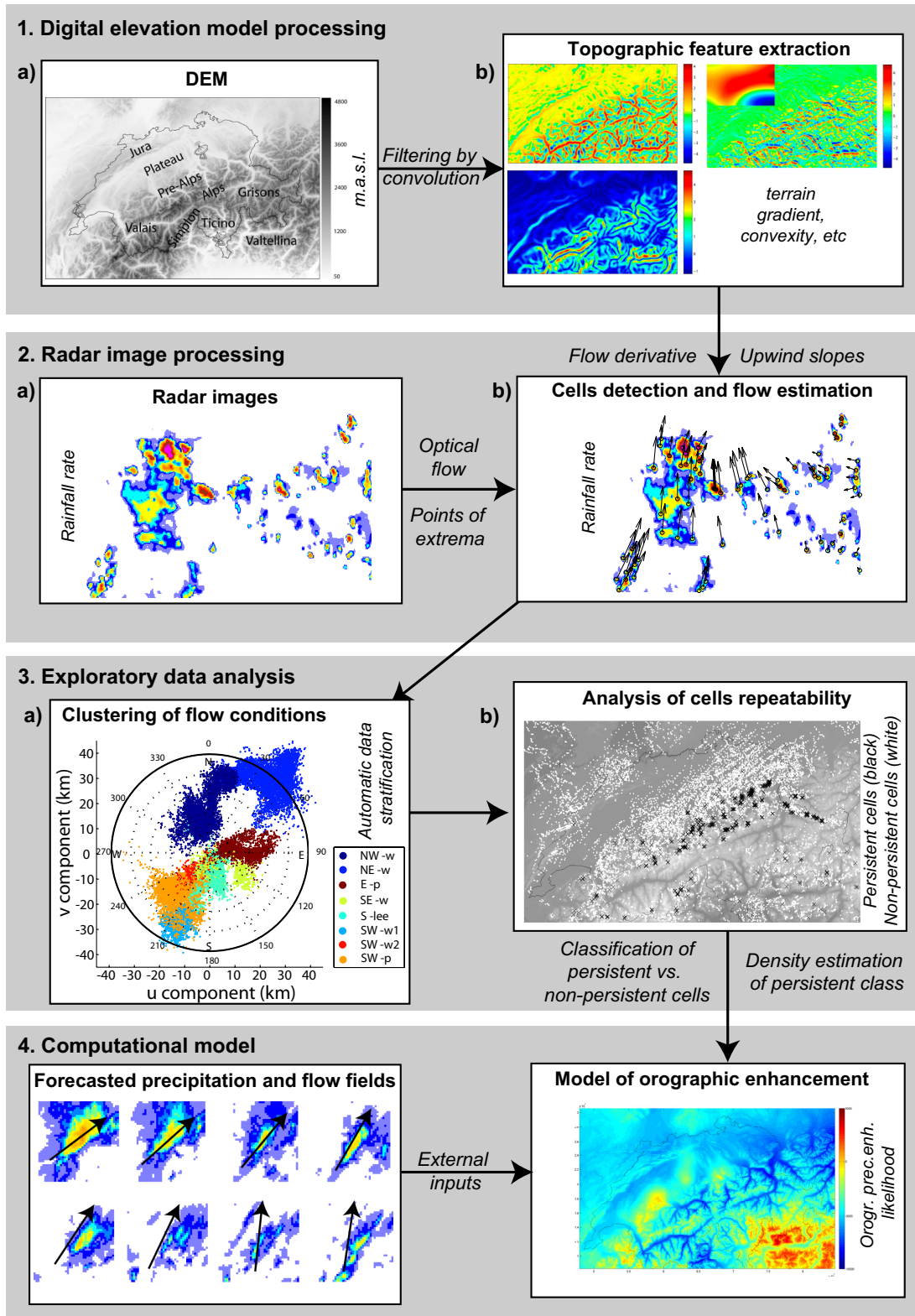
Figure 1. General scheme for data-driven modelling of orographic precipitation enhancement. External nowcasted precipitation and flow fields can be used as inputs for models of orographic precipitation enhancement, i.e. describing the likelihood of precipitation repeatability and persistence due to topography.

Table 1. Comparison of different models. AUROC and correspond-ing standard deviations are evaluated on 20 random splits.

Model	AUROC (st. dev.)
One-class SVM Linear	0.733 (0.008)
One-class SVM Gaussian	0.807 (0.011)
Two-class SVM Linear	0.694 (0.049)
Two-class SVM Gaussian	0.932 (0.007)

discriminated using a clustering algorithm such as k-means (Steinhaus, 1956; Hartigan and Wong, 1979) or spectral clus-tering (Ng et al., 2001). K-means can be used for delineat-ing convex-shaped clusters and is nowadays used as a bench-mark for weather types classification (Philipp et al., 2010). Spectral clustering (Ng et al., 2001) was used in this study because of the non spherical shape of clusters. This step is done to provide meteorological interpretability to the cells detected according to the direction of flow and their spatial location. Figure 1 step 3a plots the cells in polar coordinates according to flow direction. The different colours depict the cluster membership computed using spectral clustering in the 3-D space of (u,v) flow components an the largest scale flow derivative. Every cluster is homogeneous in terms of flow direction and relative location of cells (windward, leeward).

A detailed analysis is carried out within each cluster to recognize places which are prone to repeatability of precipi-tation cells. A number of counts measuring how many times a pixel is touched by a cell under similar flow conditions (same cluster) reveals a clear relationship with topography. A threshold on the counter of precipitation repeatability can be used to formulate a binary classification problem. The lo-cations exceeding the threshold are given to the *orographic class* and the other ones are given to the *non-orographic class*. Figure 1 (step 3b) plots the geographical distribution of the two classes corresponding to the threshold value of 4. This value was empirically selected to have a sufficient number of cells representing the orographic class while keep-ing low the number of potential non-orographic cells falling in the orographic class. Persistent precipitation cells (oro-graphic class) tend to concentrate in particular regions in geographical space (mainly Prealps, see Fig. 1, step 1a and step 3b) having specific topographic conditions, typically at the top of hills and on upwind slopes.

5 Computational model of orographic precipitation enhancement

The computational model of enhancement susceptibility is based on a classifier operating in the 16 dimensional space of the conditioning factors (u,v components were used only for clustering). Support vector machine (SVM, Vapnik, 1995) was selected as the classification method due to its robustness and explicit control over model's complexity. LibSVM tool-
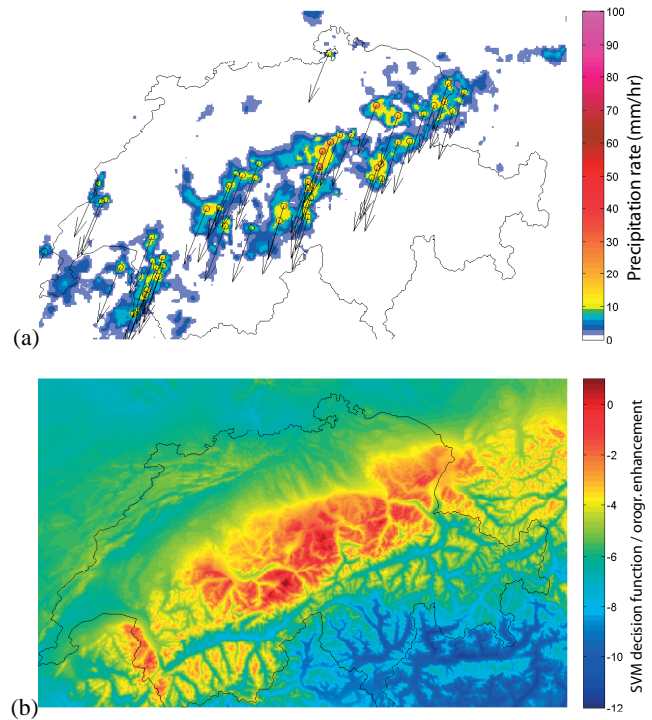

Figure 2. (a) Radar image with the detected cells and (b) the corre-sponding orographic enhancement characterized by the linear one-class SVM decision function.

box was used for the computations (Chang and Lin, 2001). It can be applied in a two-class and in a one-class settings (Schölkopf et al., 2001). The one-class setting considers the estimation of the support of the probability density function of the target class, the orographic cells, while discriminat-ing the other. Both linear and non-linear class separation can be achieved by changing the kernel function encoding data similarities (dot product for a linear or Gaussian radial basis function for a non-linear separation boundary). SVM's toler-ance to misclassification errors reduces the influence of the threshold value used to define the classes on the final results and allows to capture general tendencies of enhancement fac-tors from the data.

Data were randomly split into training (50% of the data), validation (25%) and testing (25%) datasets respectively for training, model selection and assessment purposes. Table 1 shows the areas under ROC curves (AUROC, Wilks, 2005) of the test dataset after parameters selection on the validation dataset. Maximum separability is obtained with an AUROC of 1, no separability between patterns with an AUROC of 0.5. The high AUROC values for all models considered point out that orographic and non-orographic classes are separable in the high-dimensional space of topographic features. Hence, the decision function of the classifiers can be interpreted as an indicator of orographic enhancement, i.e. the ability of producing repeatability effects and persistence on precipita-tion.
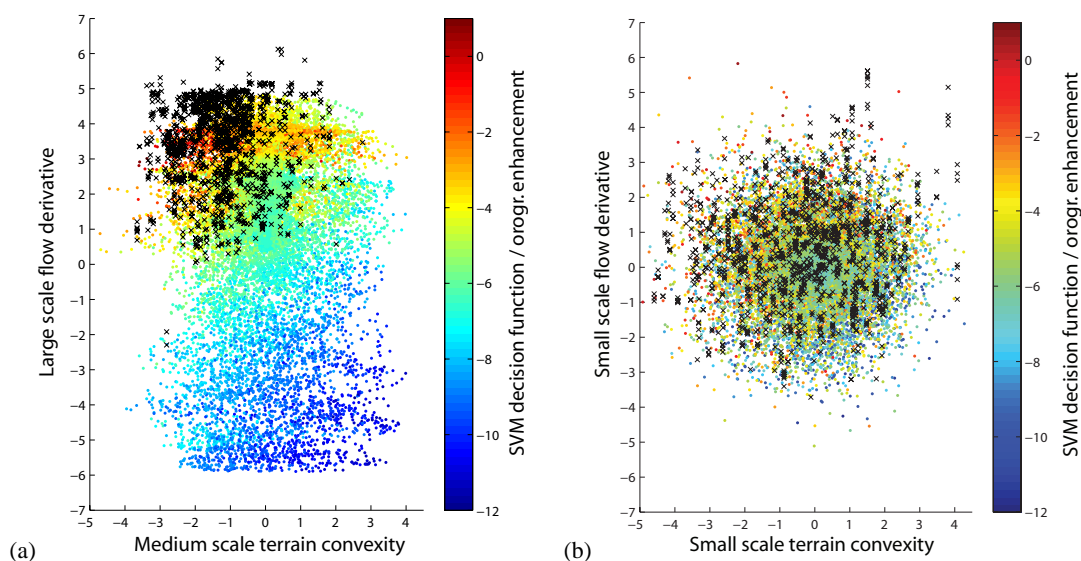
Figure 3. (a) Scatterplot of the orographic class (black crosses) in the space of features medium scale terrain convexity and large scale flow derivative. The SVM decision function is computed on the whole grid of Fig. 2b and is here displayed with the same colour scale. Orographic enhancement increases from the bottom right corner (valley bottom, leeward side of Alps) to the top left corner (hill top, windward side of Alps). **(b)** Same as (a) but with features smallest scale terrain convexity and smallest scale flow derivative. No patterns can be seen in this combination of features which is also neglected by the model.

Once the model is trained on a representative dataset it can be used for spatial predictions of precipitation enhancement under different flow and smooth precipitation patterns. Figure 2 shows an example of the system applied to characterize the orographic enhancement (Fig. 2b) with north-easterly flows and precipitation blocking in the north flank of the Alps (Fig. 2a). High enhancement values are found on the upwind northern side of Alps which is consistent with the blocking situation. Moreover, features due to the integration of elevation and terrain exposure can be noticed.

A key property of SVM is the ability of eliminating the irrelevant input information by weighting the different topographic and flow-related features. Thus, prediction maps are an optimal mixture of input features where the relevant ones dominate the spatial patterns and the irrelavant ones are simply filtered out. A close look to Fig. 2b indicates that patterns are likely to be produced only by a subset of the 16 features used. It suggests that terrain variables such as hills, ridges and upwind slopes need having a certain spatial scale (extension and size) to affect and be explanatory variables of precipitation persistence.

The study of features' relevance is better approached by plotting the orographic enhancement susceptibility indicator in the space of features. Figure 3 shows the same predictions of Fig. 2b but visualised respectively in a space composed of explanatory features (Fig. 3a) and in a space of irrelevant features (Fig. 3b). The SVM decision function in Fig. 3a depicts very well the membership to the orographic class constructed from the available persistent cells as a function of terrain convexity and flow exposure. On the other hand, no clear patterns can be seen in Fig. 3b.

6 Conclusions

This study introduced a generic data-driven methodology to study the orographic enhancement of precipitation. It aimed at discovering the persistent topography-related patterns of precipitation repeatability from high resolution radar images without using computationally demanding numerical models.

The extraction of precipitation cells, the estimation of mesoscale flows from radar images and the understanding of their connection to the underlying topography was the key point of the work. It allowed to reveal relevant variables for explaining patterns of orographic precipitation at different spatial scales. The exploratory analysis of the dataset with a clustering algorithm highlighted similar *weather types* in terms of mesoscale flows and exposure to the main Alpine crest (windward or leeward). Additional analyses whithin these clusters were performed to detect geographical locations prone to precipitation persistence, i.e. the places which were repeatedly touched by precipitation cells. Such places were found to be located at the top of hills and on upwind slopes. The patterns of precipitation repeatability and persistence were observed in the range of spatial scales represented by terrain features, i.e. between the micro- and the meso-gamma scales. However, only a subset of the considered scales were found to be relevant to orographic precipitation.

The evidence of separability of precipitation cells patterns motivated the construction of data-driven classification models in the high-dimensional space of conditioning variables such as topographic and flow features. The classification

of cells into orographic and non-orographic, defined using a threshold on precipitation repeatability, was approached using support vector machines and provided remarkable empirical performances. The SVM decision function, which can be interpreted as a susceptibility indicator of orographic enhancement, represents how likely topographic, flow and large scale precipitation conditions produce repeatability effects on small scale precipitation patterns.

The data-driven modelling of small scale precipitation enhancement patterns in complex topography provides observational support to operational NWP including the convection-permitting models (Migliorini et al., 2011). Radar-based susceptibility maps of orographic precipitation could be used to correct the windward and leeward quantitative precipitation estimation biases present in many NWP models (Bauer et al., 2011). An important issue for the future work is to analyse larger datasets of precipitation persistence and to construct more robust predictive data-driven models which are representative of a broader set of flow and atmospheric conditions.

Edited by: M. Dolinar
Reviewed by: M. Stoll and another anonymous referee

References

Bauer, H.-S., Weusthoff, T., Dorninger, M., Wulfmeyer, V., Schwitalla, T., Gorgas, T., Arpagaus, M., and Warrach-Sagi, K.: Predictive skill of a subset of models participating in D-PHASE in the COPS region, Q. J. Roy. Meteor. Soc., 137, 287–305, 2011.

Chang, C.-C. and Lin, C.-J.: LIBSVM: a library for support vector machines, Tech. rep., Software available at: http://www.csie.ntu.edu.tw/~cjlin/libsvm/, 2001.

Foresti, L. and Pozdnoukhov, A.: Exploration of alpine orographic precipitation patterns with radar image processing and clustering techniques, Meteorol. Appl., submitted, 2010.

Foresti, L., Tuia, D., Kanevski, M., and Pozdnoukhov, A.: Learning wind fields with multiple kernels, Stoch. Env. Res. Risk A., 25(1), 51–66, 2011.

Frei, C.: Eine Länder übergreifende Niederschlagsanalyse zum August Hochwasser 2005. Ergänzung zu Arbeitsbericht 211., Tech. rep., Arbeitsbericht MeteoSchweiz Nr. 213, 2006.

Germann, U. and Zawadzki, I.: Scale-dependence of the Predictability of Precipitation From Continental Radar Images. Part I: Methodology, Mon. Weather Rev., 130, 2859–2873, 2002.

Germann, U., Galli, G., Boscacci, M., and Bolliger, M.: Radar precipitation measurement in a mountainous region, Q. J. Roy. Meteor. Soc., 132(618), 1669–1692, 2006.

Gray, W. R. and Seed, A. W.: The characterisation of orographic rainfall, Meteorol. Appl., 7(2), 105–119, 2000.

Hartigan, J. A. and Wong, M. A.: A k-means clustering algorithm, Appl. Stat., 28, 100–108, 1979.

Lakshmanan, V., Rabin, R., and DeBrunner, V.: Multiscale storm identification and forecast, Atmos. Res., 67-68, 367–380, 2003.

Migliorini, S., Dixon, M., Bannister, R., and Ballard, S.: Ensemble prediction for nowcasting with a convection-permitting model – I: description of the system and the impact of radar-derived surface precipitation rates, Tellus A, 63(3), 468–496, doi:10.1111/j.1600-0870.2010.00503.x, 2011.

Ng, A. Y., Jordan, M. I., and Weiss, Y.: On spectral clustering: analysis and an algorithm, in: Advances on Neural Information Processing Systems, Vol. 14, 2001.

Panziera, L. and Germann, U.: The relation between airflow and orographic precipitation on the southern side of the Alps as revealed by weather radar, Q. J. Roy. Meteor. Soc., 136, 222–238, 2010.

Panziera, L., Germann, U., Gabella, M., and Mandapaka, P. V.: NORA – Nowcasting of orographic rainfall by means of analogs, Q. J. Roy. Meteor. Soc., submitted, 2010.

Philipp, A., Bartholy, J., Beck, C., Erpicum, M., Esteban, P., Fettweis, X., Huth, R., James, P., Jourdain, S., Kreienkamp, F., Krennert, T., Lykoudis, S., Michalides, S. C., Pianko-Kluczynska, K., Post, P., Álvarez, D. R., Schiemann, R., Spekat, A., and Tymvios, F. S.: Cost733cat – A database of weather and circulation type classifications, Phys. Chem. Earth, Parts A/B/C, Special Issue: Classifications of Atmospheric Circulation Patterns – Theory and Applications, 35(9–12), 360–373, 2010.

Pozdnoukhov, A., Foresti, L., and Kanevski, M.: Data-driven topoclimatic mapping with machine learning methods, Nat. Hazards, 50, 497–518, 2009.

Rotach, M., Appenzeller, C., and Albisser, P. E.: Meteoschweiz: 2006, Starkniederschlagsereignis August 2005, Tech. Rep. 211, Arbeitsberichte der MeteoSchweiz, 211, 63 pp., 22 pp., 2006.

Rotunno, R. and Houze, R. A.: Lessons on orographic precipitation from the Mesoscale Alpine Programme, Q. J. Roy. Meteor. Soc., 133, 811–830, 2007.

Schölkopf, B., Platt, J., Shawe-Taylor, J., Smola, A. J., and Williamson, R. C.: Estimating the support of a high-dimensional distribution, Neural Comput., 13, 1443–1471, 2001.

Steinhaus, H.: Sur la division des corps en parties, Bulletin de l'academie polonaise de sciences, C1 II vIV, 801–804, 1956.

Sun, D., Roth, S., and Black, M.: Learning optical flow, in: European Conference on Computer Vision, Part 3, 83–97, 2008.

Vapnik, V.: The Nature of Statistical Learning Theory, Springer-Verlag Berlin, 1995.

Wilks, D. S.: Statistical Methods in the Atmospheric Sciences, 2nd Edn., Academic Press, 2005.

Williams, J. K., Ahijevych, D. A., Kessinger, C. K., Saxen, T. R., Steiner, M., and Dettling, S.: A machine learning approach to finding weather regimes and skillful predictor combinations for short-term storm forecasting, 6th Conference on Artificial Intel-

ligence Applications to Environmental Science, 13th Conference on Aviation, Range and Aerospace Meteorology, J1.4, 2008.

Wilson, J. P. and Gallant, J. C. (Eds.): Terrain analysis: Principles and applications, Wiley, 2000.

Wilson, J. W., Ebert, E. E., Saxen, T. R., Roberts, R. D., Mueller, C. K., Sleigh, M., Pierce, C. E., and Seed, A.: Sydney 2000 forecast demonstration project: convective storm nowcasting, Weather and Forecast., 19, 131–150, 2004.