Supplementary Information for "Predicting Avalanche Danger in Northern Norway Using Statistical Models"

Kai-Uwe Eiselt[1] and Rune Grand Graversen[1,2]

[1]Department of Physics and Technology, University of Tromsø
[2]Norwegian Meteorological Institute

S1 Artificial neural network

The artificial neural network (ANN; e.g., LeCun et al., 2015) is essentially a generalisation of the logistic regression (e.g., Dreiseitl and Ohno-Machado, 2002). ANNs consist of an input layer, an output layer, and an arbitrary number of hidden layers in between. The layers are made up of so-called perceptrons whose activation is regulated by a specific activation function (e.g., the sigmoid function as in Eq. 2). Note that in the case of no hidden layer and if the sigmoid function is used as the activation function, the neural network is equivalent to a logistic regression (e.g., Dreiseitl and Ohno-Machado, 2002). The weights of the individual perceptrons and layers are determined during training via the minimisation of the cross-entropy error. The activation function of the hidden layer used here is the rectified linear unit (ReLU):

$$f(z) = \max(z, 0), \tag{1}$$

which has become the most popular choice as it accelerates the learning process compared to other methods (e.g., LeCun et al., 2015). The activation function for the output layer depends on the number of classes in the target variable. In the case of a binary classification, the sigmoid function is used. It predicts the probability *p(x)* of the binary outcome for a set of predictors x_i (i.e., the probability that the outcome is 1 or true)

$$p(x_i) = \frac{1}{1 + e^{-z}}, \tag{2}$$

where z is a linear combination of the predictor values x_i and the model's predicted coefficients β_i (for $i > 0$) and intercept β_0:

$$z = \beta_0 + \beta_1 x_1 + \beta_2 x_2 + ... + \beta_r x_r. \tag{3}$$

The β_i are obtained by maximum likelihood estimation, maximising the log-likelihood function. Solving Eq. 2 for z, one may interpret z as the logarithm of the *odds* of outcome 1 (i.e., the probability of success over the probability of failure):

$$z = \log\left(\frac{p(x)}{1 - p(x)}\right). \tag{4}$$

Eq. 2 gives a probability between 0 and 1, and values < 0.5 are categorised as 0 and values > 0.5 as 1 for our binary case.

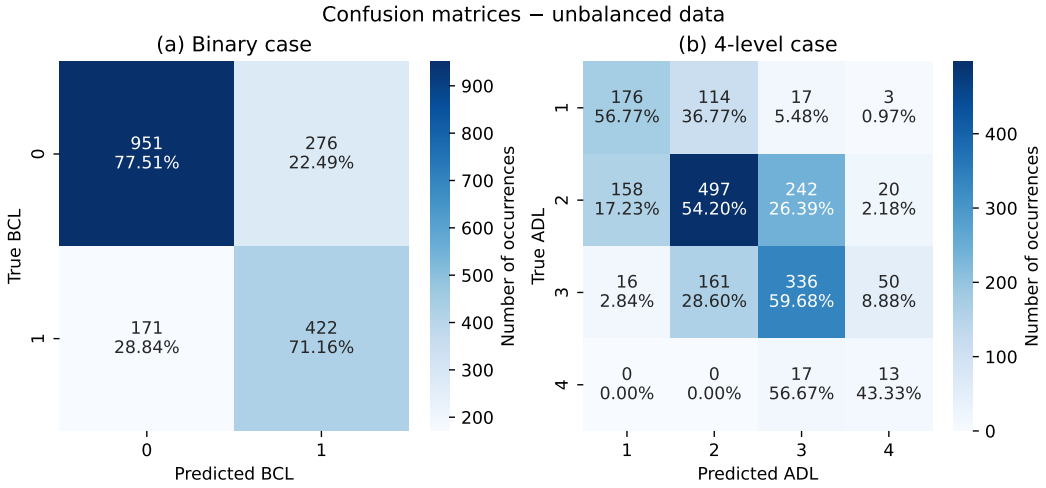

Figure S1. As Fig. 5, but for the artificial neural network (ANN).

If the target variable consists of multiple classes the number of nodes in the output layer must equal the number of classes. In this case we follow Sharma et al. (2023) in using the SoftMax function as the activation function for the output layer:

$$\text{SoftMax}\,(c_j) = \frac{e^{c_i}}{\sum_j e^{c_j}}, \tag{5}$$

where $\text{SoftMax}\,(c_j)$ corresponds to the probability of the class c_j of the target variable. Hence, in our 4-level case, we need an output layer with four nodes and j ranges from 1 to 4. Since our focus here is on the full integer danger levels, the ADL with the highest predicted probability is considered the predicted ADL.

Here we apply the ANN structure suggested by Sharma et al. (2023) (see their Table 4). Note that in contrast to the RF, the data were zero-transformed by subtracting the mean and dividing by the standard deviation for their usage with the ANN.

A weakness of the ANN is that the weights are difficult to interpret, making it a "black-box" model (e.g., Dreiseitl and Ohno-Machado, 2002), similar to the RF. An advantage of the ANN is that it may find and represent structures in the data that simpler linear models are missing. Several studies have applied ANNs to predict ADL (Schirmer et al., 2009; Dekanová et al., 2018; Fromm and Schönberger, 2022; Sharma et al., 2023; Blagovechshenskiy et al., 2023), however with varying success, as discussed in section 1.

S2 Artificial neural network — Evaluation

Both in the binary and the 4-level case the ANN exhibits a performance very similar to the random forest (RF) model, with overall accuracies of 75 % and 56 %, respectively. Due to the inherent randomness of an ANN, these accuracies vary by ±2 % when repeatedly training the model on the same data. The confusion matrices and classification reports for the ANN and RF are similar as well (compare Fig. 5 with Fig. S1 and Tables 5 & 6 with Tables S1 & S2, respectively). The ANN appears more

Table S1. As Table 5 but for the artificial neural network (ANN).

level	precision	recall	f1-score	support
0	0.85	0.78	0.81	1227
1	0.60	0.71	0.65	593
accuracy			0.75	1820
macro avg	0.73	0.74	0.73	1820
weighted avg	0.77	0.75	0.76	1820

Table S2. As Table 6 but for the artificial neural network (ANN).

level	precision	recall	f1-score	support
1	0.50	0.57	0.53	310
2	0.64	0.54	0.59	917
3	0.55	0.60	0.57	563
4	0.15	0.43	0.22	30
accuracy			0.56	1820
macro avg	0.46	0.58	0.48	1820
weighted avg	0.58	0.56	0.57	1820

capable of classifying ADLs 1 and 4 correctly, however at the expense of greater misclassification of ADLs 2 and 3, resulting in the overall similar performance. A consequence of this is a stronger underestimation of the frequency of ADL 2 (see Fig. S2). Also, the ANN appears to have slightly greater tendency to a classification difference larger 1 (compare Table 7 with Table S3).

In summary, the RF and the ANN of the structure suggested by Sharma et al. (2023) perform similarly when comes to predicting both ADL and binary-case level (BCL).

S3 Artificial neural network — Hindcast

Like the RF (section 6), we employ the binary-case ANN to perform a hindcast of the BCL for the NORA3 period (1970-2023). We define the binary-case avalanche activity (BCA) as the number of days with BCL 1 per season. The general results are consistent with the results from the RF: There is no trend in the full-season BCA with partly significant weak negative and positive trends in winter (December through February) and spring (March through May), respectively (not shown). The BCA peak in the 1990s (especially in winter) is also evident in the results from the ANN (Fig. S3). However, there are some

Table S3. As Table 7, but for the artificial neural network (ANN).

difference	unbalanced	balanced
-3	0.0	0.0
-2	0.88	0.57
-1	18.46	25.82
0	56.15	54.85
1	22.31	17.12
2	2.03	1.47
3	0.16	0.16

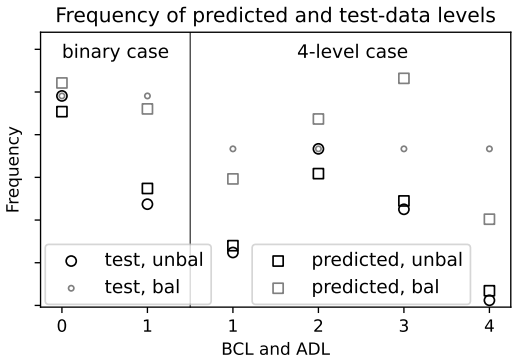

Figure S2. As Fig. 6, but for the artificial neural network (ANN).

differences. For example, the ANN generally hindcasts a higher BCA (compare Figs. S3 and S10), which is mainly due to higher winter BCA. Notably, the correlation of the Arctic Oscillation (AO) with the ANN-derived BCA is even stronger than with the RF-derived BCA on an annual basis (Pearson R = 0.45-0.5) and similar on for 7-year rolling means (compare Figs. 8b, d and S4b, d). Finally, it appears that the ANN-based BCA hindcast exhibits more of an increasing trend in the decade 2010-2020 than the RF-based hindcast, being more consistent with the AO trend (compare Figs. 8c and S4c). This may be connected to the ANN-based BCA following more closely the accumulated snow than the RF-based BCA (compare Figs. 9 and S5).

In summary, the ANN-based BCA hindcast is similar to the RF-based hindcast and appears to confirm the impact of the AO on northern Norwegian avalanche activity. However, given the differences especially in the decade 2010-2020, the robustness of this relationship remains uncertain. Decomposing the ADLs and BCLs according to the different avalanche problems may provide more clarity regarding the connection between northern Norwegian BCA and the AO, since some types of avalanches are likely more strongly determined by conditions related to, e.g., a high AO index than other types.

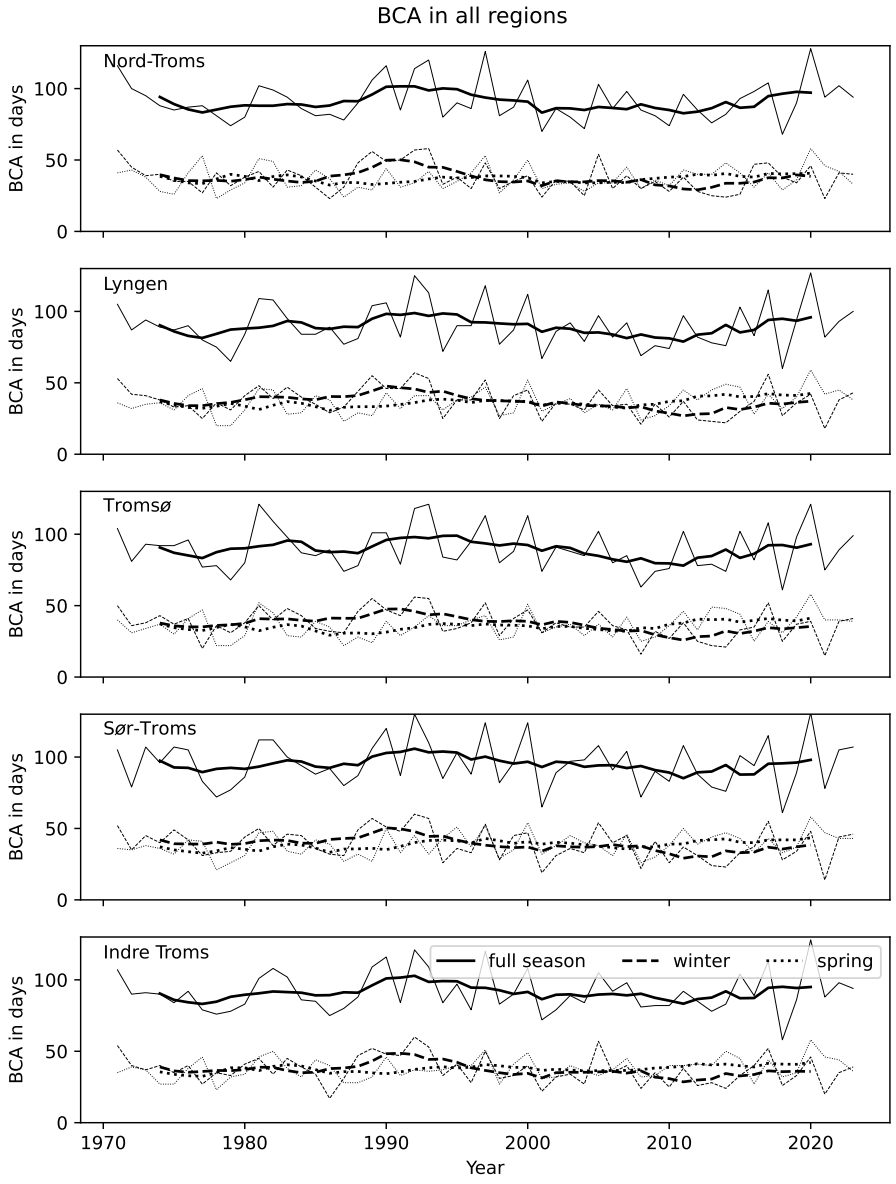

Figure S3. As Fig. 7, but for all regions and based on the artificial neural network (ANN) instead of the random forest (RF).

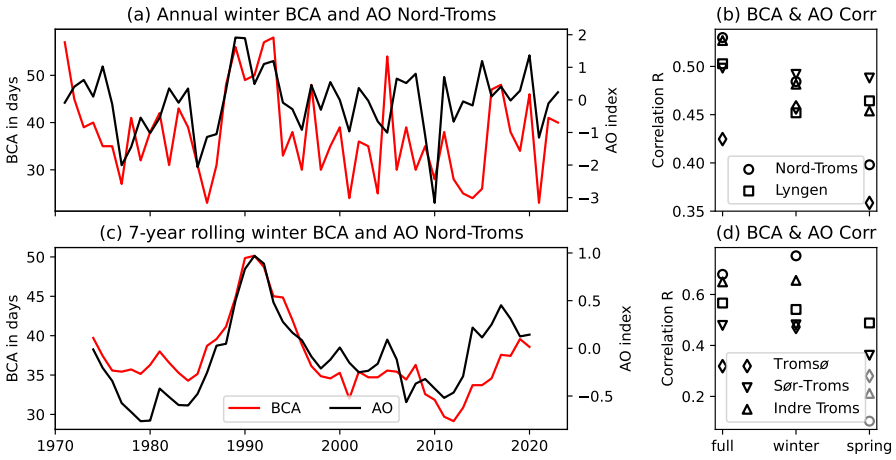

Figure S4. As Fig. 8, but for the artificial neural network (ANN) instead of the random forest (RF).

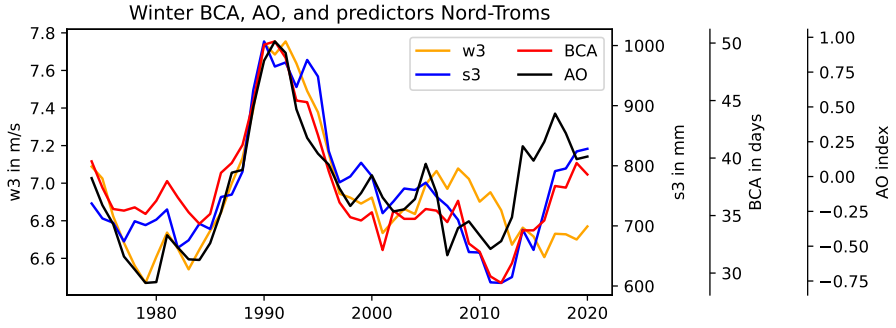

Figure S5. As Fig. 9, but for the artificial neural network (ANN) instead of the random forest (RF).

Table S4. As Table 5 but with the balanced data.

level	precision	recall	f1-score	support
0	0.77	0.79	0.78	1227
1	0.78	0.76	0.77	1227
accuracy			0.77	2454
macro avg	0.77	0.77	0.77	2454
weighted avg	0.77	0.77	0.77	2454

Table S5. As Table 6 but with the balanced data.

level	precision	recall	f1-score	support
1	0.74	0.53	0.62	917
2	0.46	0.60	0.52	917
3	0.48	0.65	0.55	917
4	0.89	0.55	0.68	917
accuracy			0.58	3668
macro avg	0.64	0.58	0.59	3668
weighted avg	0.64	0.58	0.59	3668

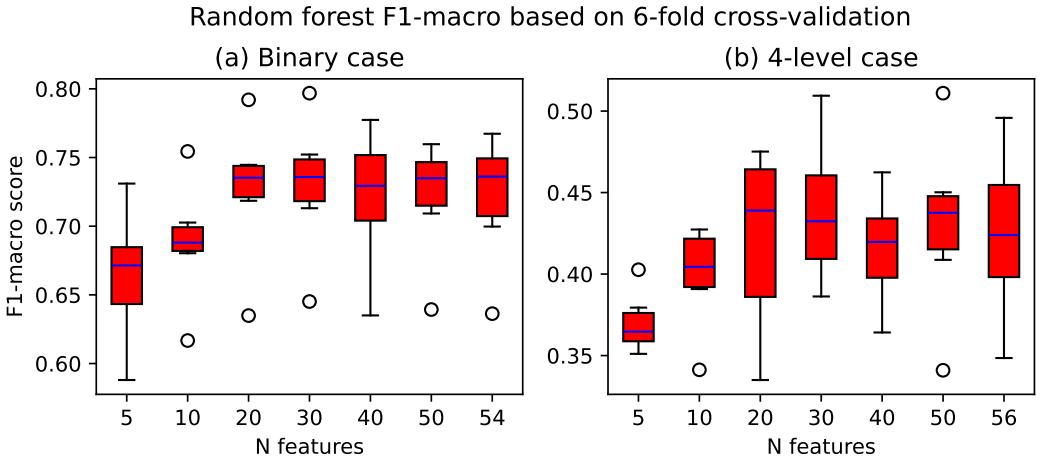

Figure S6. As Fig. 4, but with 6-fold cross-validation.

Table S6. Slope, Pearson correlation R, and p value (Wald test with a t distribution) of linear regressions from of avalanche activity hindcast with the random forest model from 1970 to 2023 for all regions.

	full	winter	spring
Nord-Troms			
slope	-0.11 $\pm$ 0.11	-0.14 $\pm$ 0.08	0.03 $\pm$ 0.06
R	-0.14	-0.24	0.07
p	0.33	0.09	0.63
Lyngen			
slope	-0.04 $\pm$ 0.10	-0.15 $\pm$ 0.08	0.14 $\pm$ 0.06
R	-0.06	-0.26	0.29
p	0.69	0.06	0.03
Tromsø			
slope	-0.08 $\pm$ 0.09	-0.22 $\pm$ 0.08	0.16 $\pm$ 0.05
R	-0.13	-0.37	0.40
p	0.36	0.01	0.00
Sør-Troms			
slope	-0.04 $\pm$ 0.10	-0.15 $\pm$ 0.08	0.14 $\pm$ 0.06
R	-0.05	-0.24	0.32
p	0.72	0.09	0.02
Indre Troms			
slope	0.00 $\pm$ 0.11	-0.05 $\pm$ 0.08	0.07 $\pm$ 0.06
R	0.00	-0.09	0.16
p	0.98	0.51	0.24

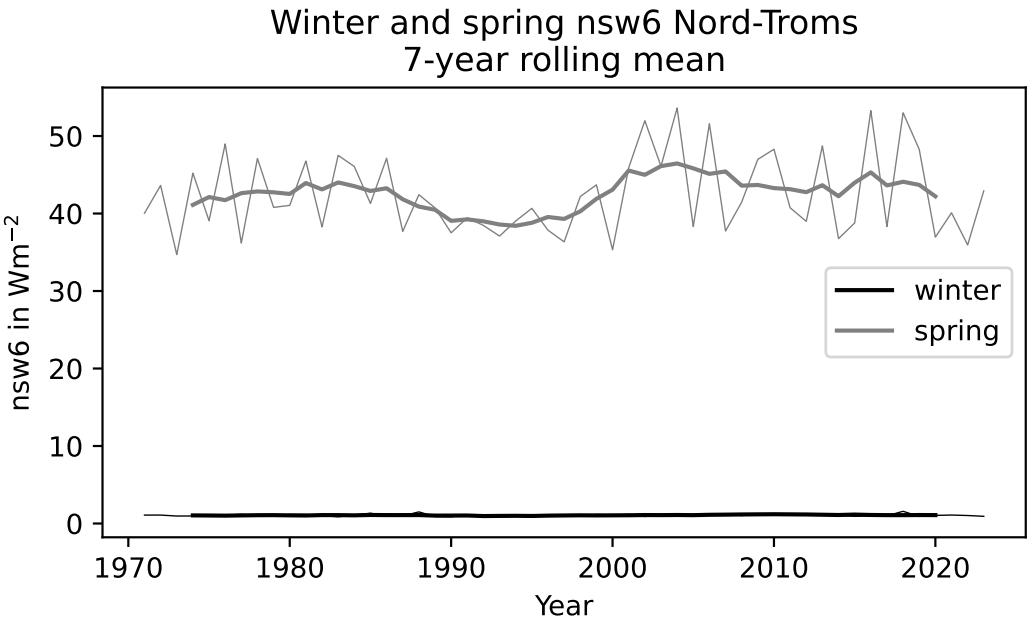

Figure S7. Winter and spring nsw6 in Nord-Troms on (a) an annual basis and (b) for 7-year rolling means.

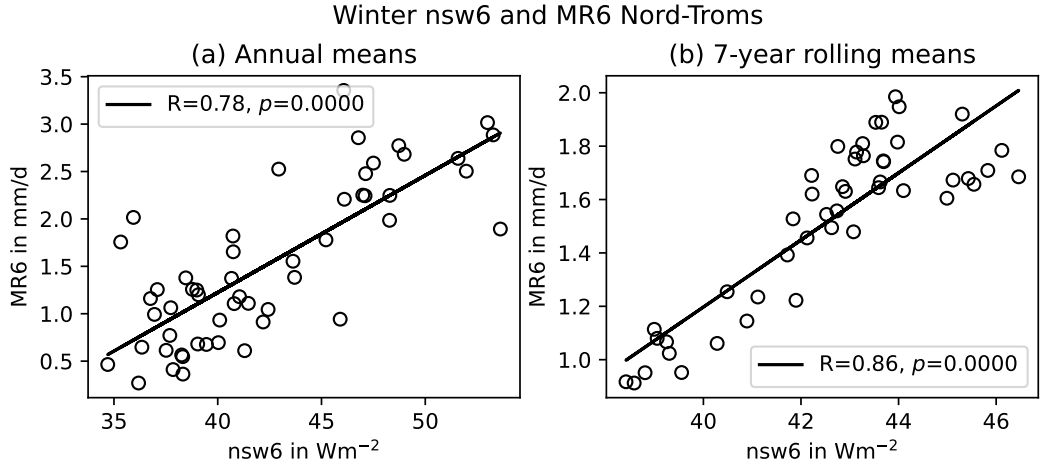

Figure S8. Correlation of spring nsw6 and MR6 in Nord-Troms.

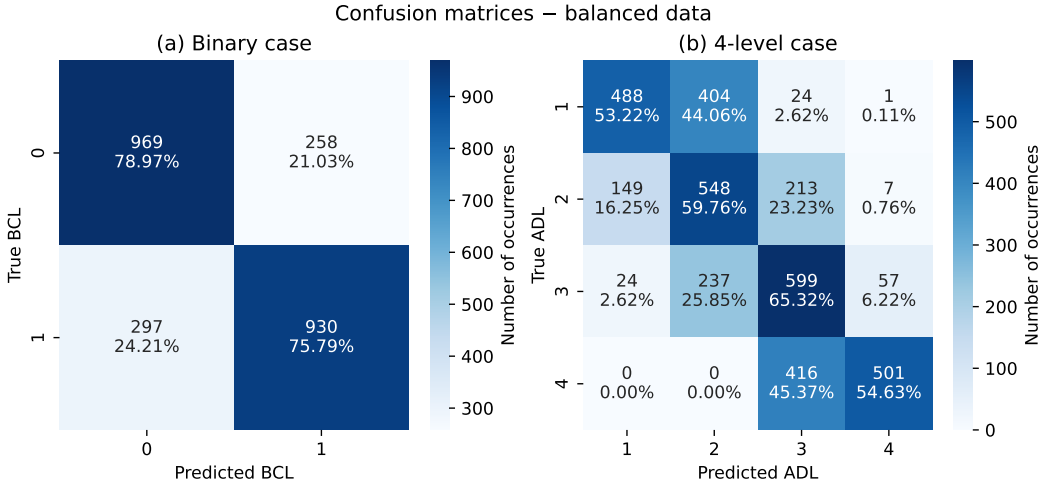

Figure S9. As Fig. 5, but for the balanced data.

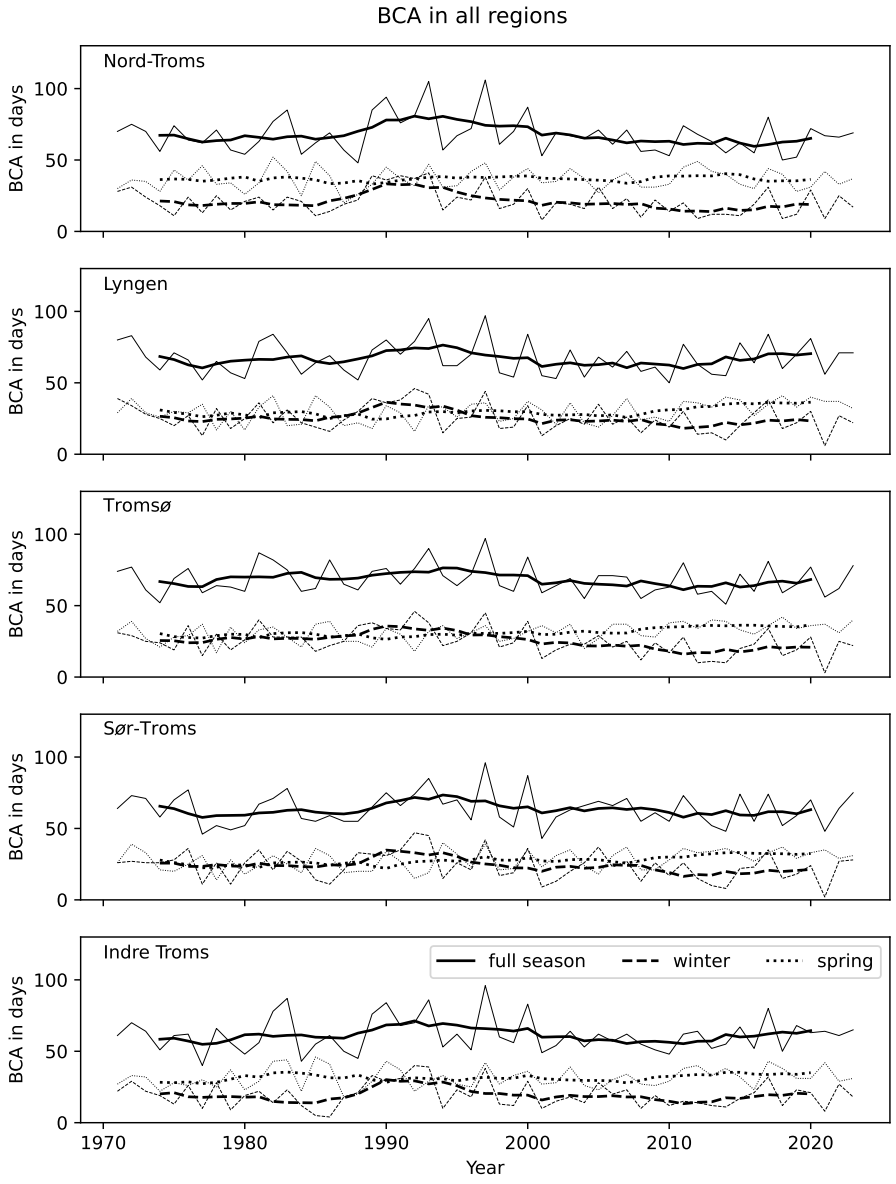

Figure S10. As Fig. S3, but for the random forest (RF).

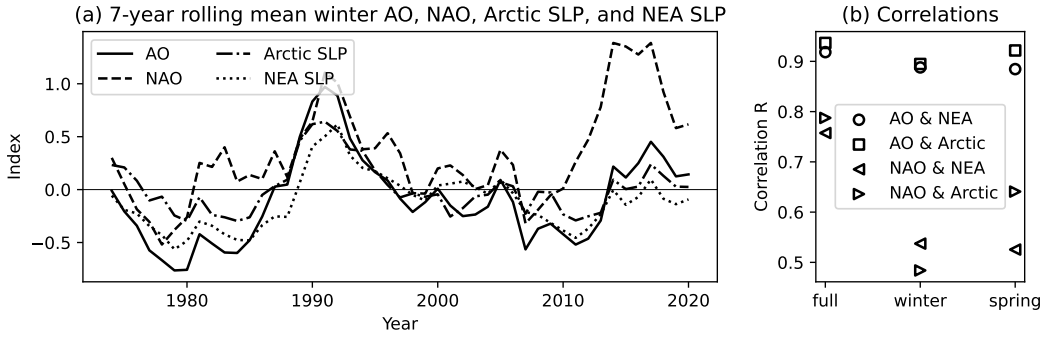

Figure S11. (a) Time series and (b) correlations of Arctic Oscillation (AO) and North Atlantic Oscillation (NAO) with Arctic ($> 75\ ^{\circ}$N) and north-east Atlantic (NEA; 70-80 $^{\circ}$N, 10-20 $^{\circ}$E) sea-level pressure (SLP). The SLP data is taken from the ERA5 reanalysis. Note that all correlations exhibit p values < 0.01 based on a Wald test with a t distribution.

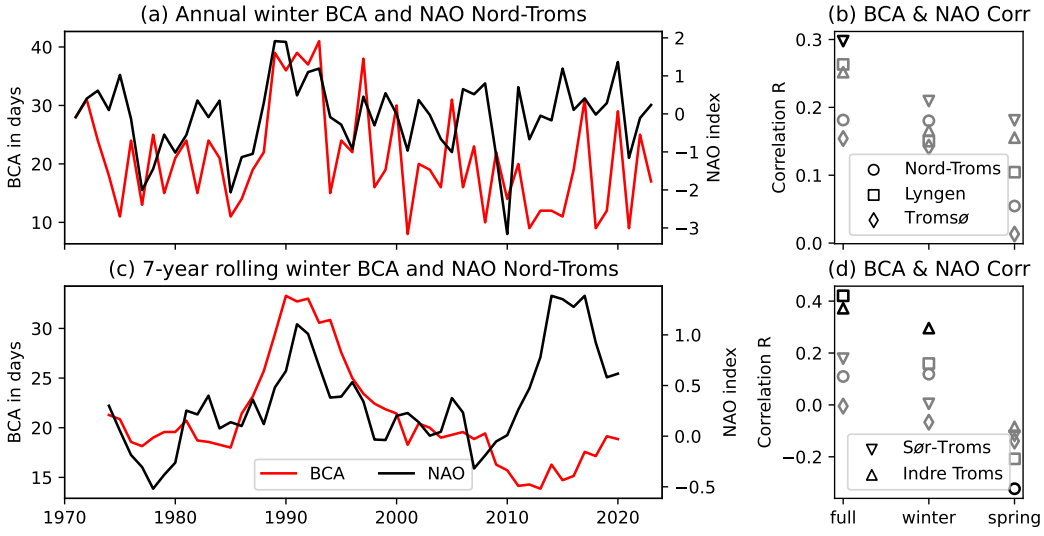

Figure S12. As Fig. 8, but for the North Atlantic Oscillation (NAO) instead of the Arctic Oscillation (AO).

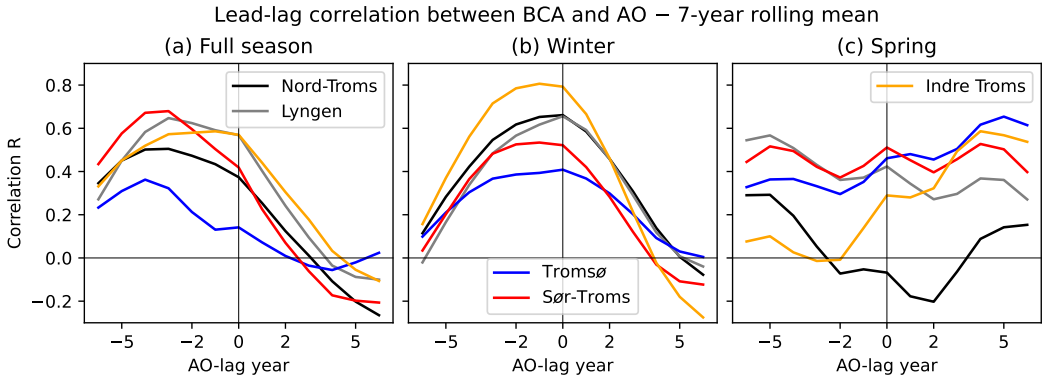

Figure S13. Lead-lag correlations between Arctic Oscillation (AO) and binary-case avalanche activity (BCA) for (a) full season, (b) winter, and (c) spring on an annual basis.

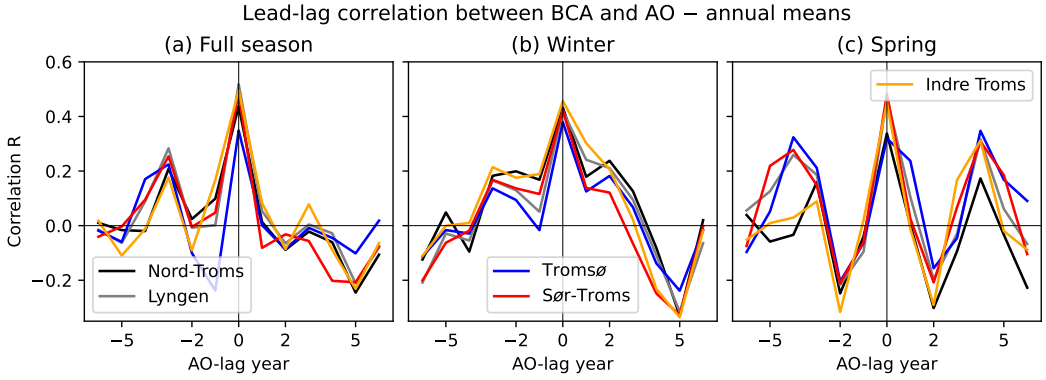

Figure S14. As Fig. S13, but for annual means.

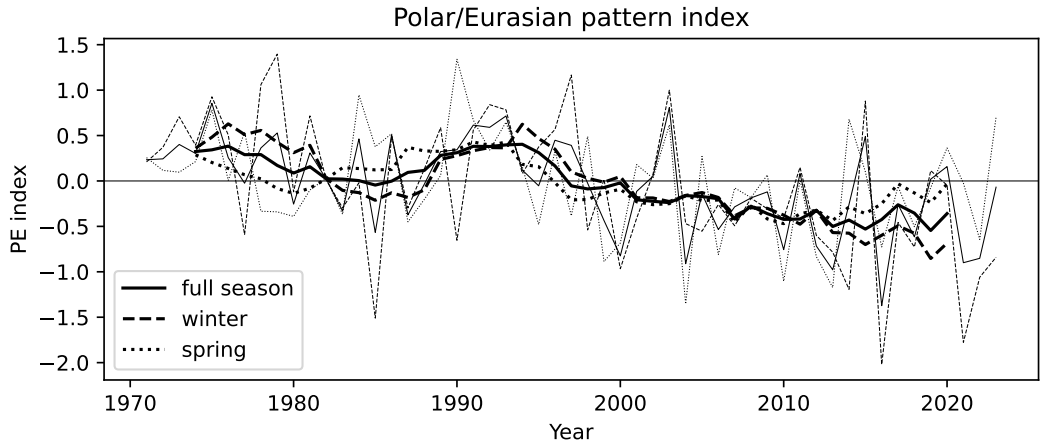

Figure S15. Polar/Eurasian pattern index for full season (continuous), winter (dashed), and spring (dotted). Shown are annual means (thin) and 7-year rolling means (thick). The index was downloaded from https://www.cpc.ncep.noaa.gov/data/teledoc/poleur.shtml.

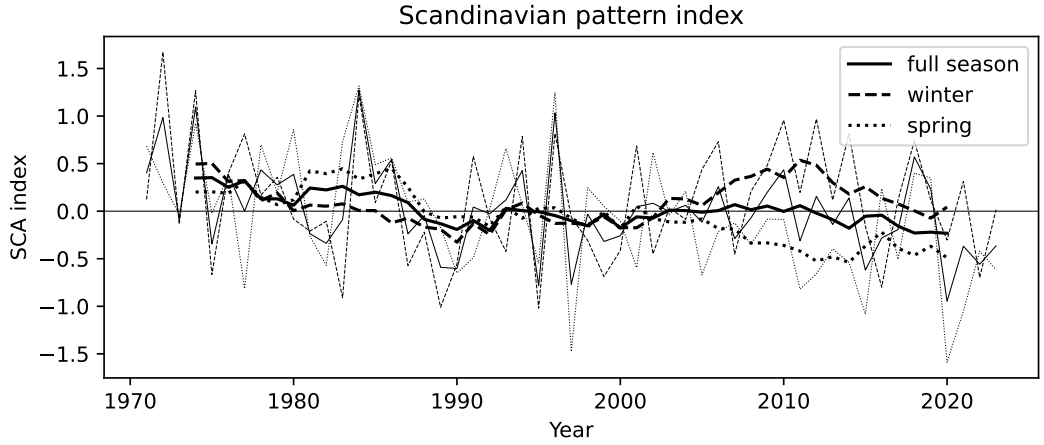

Figure S16. Scandinavian pattern index for full season (continuous), winter (dashed), and spring (dotted). Shown are annual means (thin) and 7-year rolling means (thick). The index was downloaded from https://psl.noaa.gov/data/timeseries/month/SCAND/.

References

Blagovechshenskiy, V., Medeu, A., Gulyayeva, T., Zhdanov, V., Ranova, S., Kamalbekova, A., and Aldabergen, U.: Application of Artificial Intelligence in the Assessment and Forecast of Avalanche Danger in the Ile Alatau Ridge, Water, 15, 1438, https://doi.org/doi.org/10.3390/w15071438, 2023.

Dekanová, M., Duchoň, F., Dekan, M., Kyzek, F., and Biskupič, M.: Avalanche forecasting using neural network, In Proceedings of the IEEE ELEKTRO, Mikulov, Czech Republic, https://doi.org/10.1109/ELEKTRO.2018.8398359, 2018.

Dreiseitl, S. and Ohno-Machado, L.: Logistic regression and artificial neural network classification models: a methodology review, Journal of Biomedical Informatics, 35, 352–359, https://doi.org/10.1016/S1532-0464(03)00034-0, 2002.

Fromm, R. and Schönberger, C.: Estimating the danger of snow avalanches with a machine learning approach using a comprehensive snow cover model, Mach. Learn. Appl., 10, 100405, https://doi.org/10.1016/j.mlwa.2022.100405, 2022.

LeCun, Y., Bengio, Y., and Hinton, G.: Deep learing, Nature, 521, 436–444, https://doi.org/10.1038/nature14539, 2015.

Schirmer, M., Lehning, M., and Schweizer, J.: Statistical forecasting of regional avalanche danger using simulated snow-cover data, J. Glaciol., 55, 761–768, https://doi.org/10.3189/002214309790152429, 2009.

Sharma, V., Kumar, S., and Sushil, R.: A neural network model for automated prediction of avalanche danger level., Nat. Hazards Earth Syst. Sci., 23, 2523–2550, https://doi.org/10.5194/nhess-23-2523-2023, 2023.