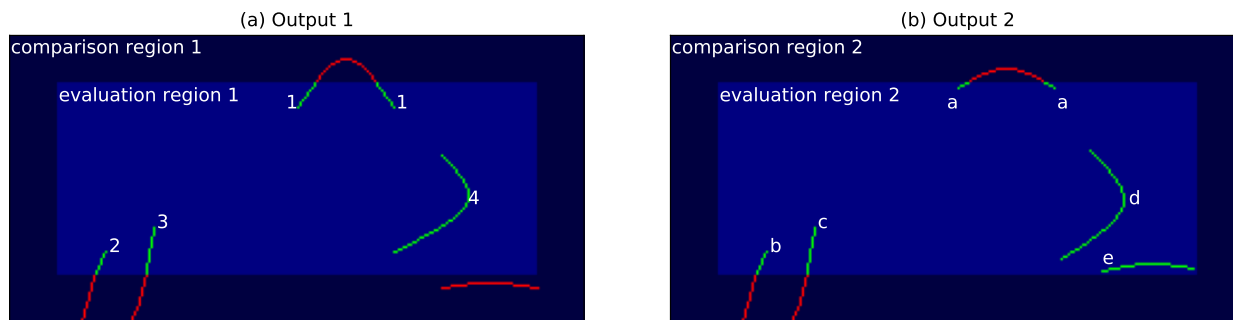*Supplement of*

Automated detection and classification of synoptic-scale fronts from atmospheric data grids

Stefan Niebler et al.

Correspondence to: Stefan Niebler (stnieble@uni-mainz.de)

Figure S1. Sketch of evaluation and comparison region used during CSI evaluation for two exemplary outputs (a) and (b). Green line segments are within the evaluation (and comparison) region, while red line segments can only be used during comparison but are not evaluated. Front segments connected within the comparison region are evaluated as a single front, even though they are not connected within the evaluation region alone.

Table S1. As Table 4 but matching can only happen against a single front.

Training region	NWS			DWD			Both		
	CSI	POD	SR	CSI	POD	SR	CSI	POD	SR
Binary	42.9 %	52.9 %	69.5 %	63.3 %	72.4 %	83.4 %	61.7 %	70.8 %	82.8 %
Warm	19.9 %	22.3 %	64.9 %	48.6 %	57.3 %	76.2 %	48.6 %	57.0 %	76.7 %
Cold	38.4 %	47.0 %	67.7 %	55.6 %	67.2 %	76.2 %	55.1 %	65.5 %	77.5 %
Occlusion	34.4 %	42.5 %	64.5 %	50.8 %	67.6 %	67.1 %	51.1 %	65.1 %	70.3 %

S1 CSI evaluation sketch

Fig. S1 shows a sketch of the evaluation and comparison regions used during evaluation of the CSI scores for two exemplary outputs. The comparison region fully contains the evaluation region plus the darker shaded blue region, while the evaluation region consists only of the brighter blue shaded region. Parts of a front are green if the segment is within the evaluation region, red otherwise.

During evaluation the front segments within the evaluation region of Output 1 (Panel a) are compared against the front segments within the comparison region of Output 2 (Panel b) and vice versa. That means only the green segments are evaluated while the combination of green and red segments serves as potential matches.
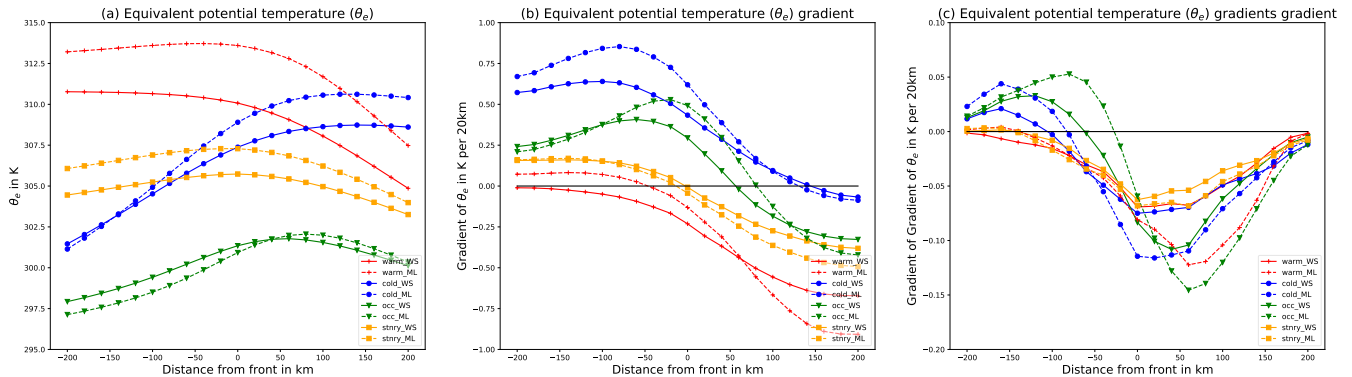
As front objects are determined within the comparison region, both green segments of front 1 (front a) are correctly identified as being parts of the same front. Not using the comparison regions, these segments would be counted as two separate fronts, potentially skewing the count of matched or unmatched fronts in the end. Additionally this method allows the algorithm to correctly match front e against the bottom right red front of Output 1. As this front is not located within the evaluation region, e would have otherwise been falsely counted as unmatched. This method is therefore useful to reduce errors in the evaluation introduced by the cropping of the outputs.

S2 Evaluation against individual fronts

We carried out an additional evaluation of the CSI, POD and SR scores, where each front is only allowed to match against one front of the corresponding type rather than the complete set. Table S1 displays the scores for the DWD, Table S2 for the NWS data-set.

Table S2. As Table 5 but matching can only happen against a single front.

Training region	NWS			DWD			Both		
	CSI	POD	SR	CSI	POD	SR	CSI	POD	SR
Binary	62.1 %	74.9 %	78.4 %	44.3 %	50.3 %	79.0 %	63.4 %	76.9 %	78.4 %
Warm	37.2 %	56.2 %	52.4 %	22.4 %	43.8 %	31.4 %	36.2 %	57.9 %	49.2 %
Cold	54.8 %	69.2 %	72.5 %	40.5 %	51.0 %	66.3 %	55.8 %	72.1 %	71.1 %
Occlusion	48.5 %	72.1 %	59.7 %	36.0 %	62.5 %	46.0 %	48.9 %	73.2 %	59.5 %
Stationary	42.3 %	56.7 %	62.5 %		–		40.7 %	53.1 %	63.5 %

Figure S2. Mean of (a) equivalent potential temperature (θ_e) and (b) θ_e gradient and (c) gradient θ_e gradient with front positions determined by NWS manual analysis (solid, WS) and by our network (dashed, ML). Physical quantities are evaluated at $850\mathrm{hPa}$.

The values of all three measures decrease in comparison to the evaluation of the fronts with matching to the complete set. For the SR and classification results this effect is less pronounced.
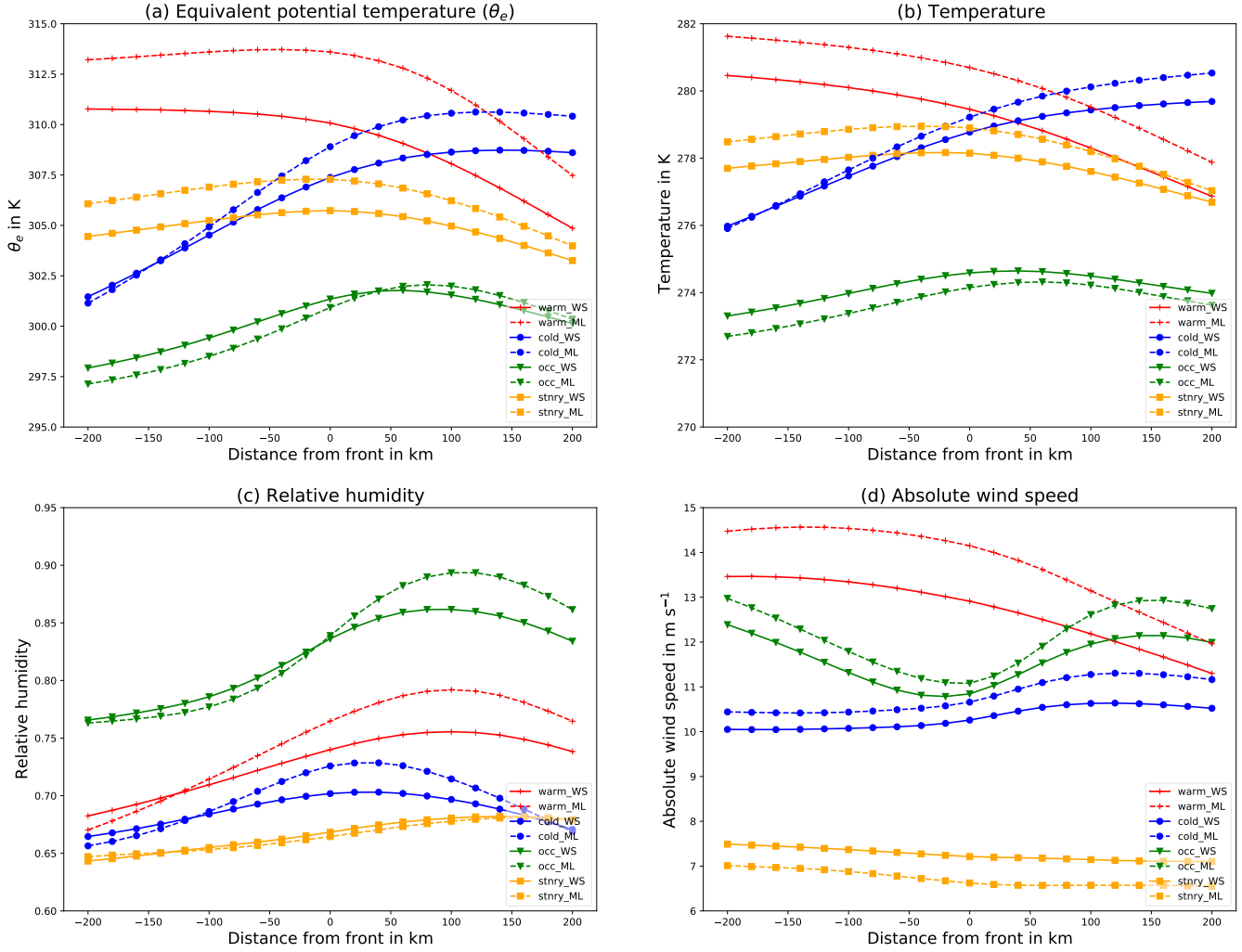
S3 Cross Sections on NWS Data

Here we present additional cross sections of physical properties (Fig. S2 and S3) relative to the front position provided by the NWS data-set. Qualitatively, the same behaviour as for the DWD fronts can be seen. Note that the NWS data contains stationary fronts, thus cross-sections for these are indicated here by a solid yellow line.

S4 Video Supplement

The provided video supplement shows the predicted and classified fronts for January 2016 at each hour. The background consists of the equivalent potential temperature at $850\mathrm{hPa}$. Color channels are chosen as follows:

- red: warm front
- blue: cold front
- pink: occlusion
- green: stationary front
- yellow: unclassified front

Figure S3. Mean of (a) equivalent potential temperature (θ_e), (b) temperature, (c) relative humidity and (d) absolute wind speed with front positions determined by NWS manual analysis (solid, WS) and by our network (dashed, ML). Physical quantities evaluated at $850hPA$.

Fronts are created as described in Section 2.4.2.

In some cases a classification may not be exclusive for a pixel, resulting in a potential overlap in the color channels. This effect may occur when one type of front transitions into another. E.g. a transition from a warm to a cold front may appear pink, due to mixing in the blue and red color channel. Weakly expressed fronts may appear fragmented, due to the filtering threshold.

S5 Connection between fronts and extreme precipitation

Additional information regarding the methodology used in Section 3.3 is provided here.

S5.1 Data and Definitions

For the determination of precipitation events, we use the surface precipitation as contained in the ERA5 data set (hourly 2D field). Extreme precipitation is defined as any precipitation event that exceeds the 99th percentile of precipitation at each grid point. Due to a limitation of available data we calculate this percentile using ERA5 data ranging from the years 2010 until 2018 (inclusive) using CDO.

We consider any grid point within an L2-distance of $2.5°$ (i.e. 10 grid points) to a front (extreme precipitation event) to be associated with a front (extreme precipitation event). We evaluate the connection between fronts and extreme precipitation using $N = 8784$ samples from the year 2016; we chose this year, as it was not used during the training of our network.

We need to define a few variables for the evaluation. For each grid point p we define the number $N_{evt}(p)$ of different events evt as the count of a evt occurring at p during 2016. For a grid point p, the counted events are as follows:

- epr: An extreme precipitation event occurred at p

- $a(epr)$: p is associated with an extreme precipitation event. (E.g. p is within a certain range (here $2.5°$) of a grid point where an extreme precipitation event occurred.)

- fr: A front occurred at p

- $a(fr)$: p is associated with a front. (E.g. p is within a certain range (here $2.5°$) of a grid point where a front occurred.)

- $x + y$: Events x and y occur at the same time at p. (e.g. $epr + a(fr)$ describes the event that an extreme precipitation event occurs at p while p is associated with a front.)

.

We further define the proportions $P_{evt}(p) = N_{evt}(p)/N(p)$ for events evt as defined above. Finally we also calculate the relations

- $R_1(p) = \frac{N_{a(fr)+epr}(p)}{N_{epr}(p)}$, describing the proportion of extreme precipitation events at grid point p that can be associated with a front

- $R_2(p) = \frac{N_{fr+a(epr)}(p)}{N_{fr}(p)}$, describing the proportion of fronts at grid point p that can be associated with an extreme precipitation event.

These definitions are slightly similar to the formulation of conditional probability.

We further define as high altitude regions any grid point within a 5 pixel distance from any grid point exceeding a height of 2000m. The height of a gridpoint is derived from the geopotential height variable of the ERA5 data-set

S5.2 Statistical Test

If we assume that both events, i.e. the occurrence of an extreme precipitation event and a front, are completely uncorrelated, we would expect R_1 to be similarly distributed as $P_{a(fr)}$, i.e. the frequency with which point p is associated with a front. For each grid point p poleward of $20°$ we calculate the frontal frequency $P_{a(fr)}(p)$. We then distribute all points p according to
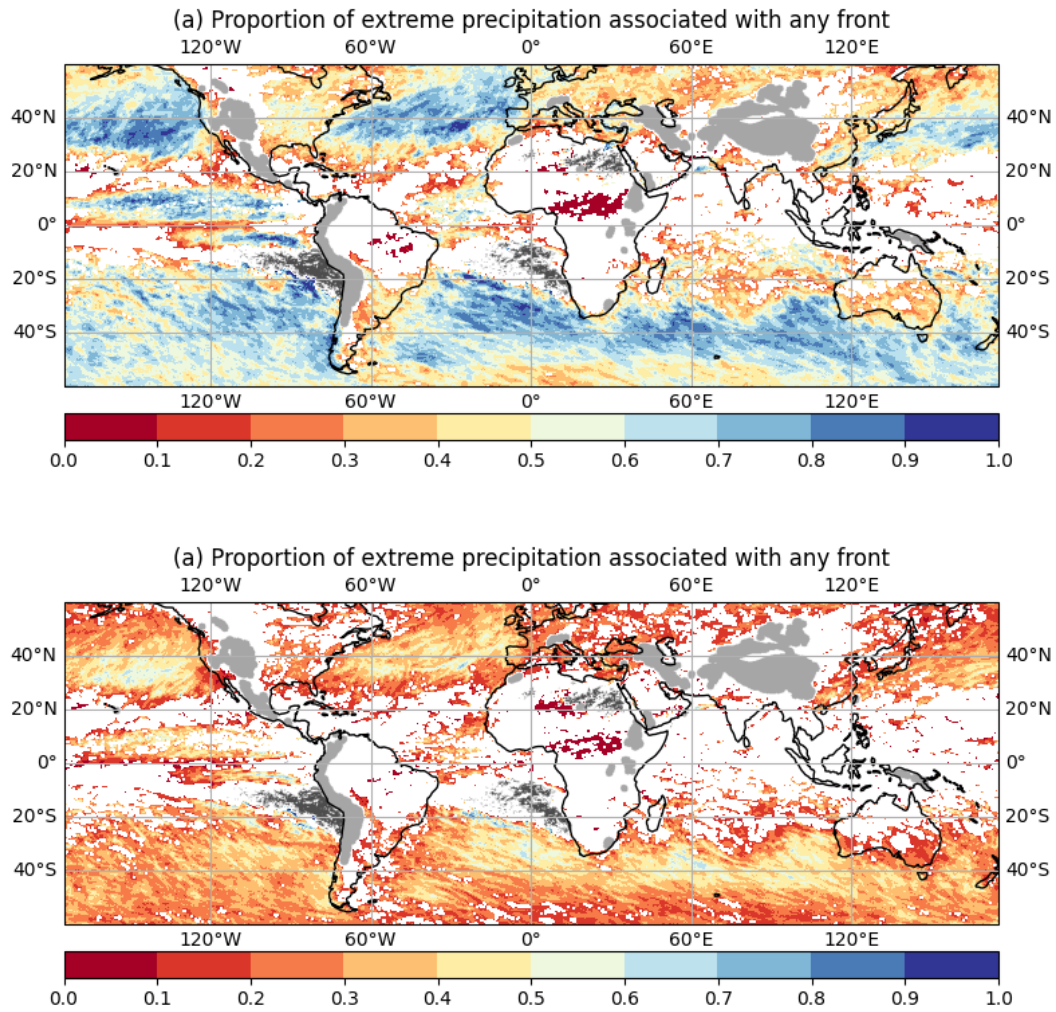
70   their respective frontal frequency, into bins of 1% width. For each bin with at least m entries we randomly select m grid points (base points) and create 1000 event lists, each containing k successive extreme precipitation events sampled at 6 points. Each of those points is located at the respective opposite hemisphere from the corresponding base point. k is chosen as 50 such that we obtain at least 300 samples of extreme precipitation events for each base point. As result for each frequency bin we obtain m sampled distributions of the proportion of extreme precipitation events occurring while the base point is associated with a
75   front. Taking the median of each of those samples we get a sample of m points per bin. We then apply a percentile regression on this data to obtain linear functions describing the $1st$ and $99th$ percentile of our data with respect to the frontal frequency. We then define that for each grid point p, where $R_1(p)$ is not within the limits described by these percentiles respecting the underlying frontal frequency, a significant connection between extreme precipitation and frontal frequency exists. We are then able to additionally mask all grid points where no significant connection is found.

80   For our test $m = 12$ was chosen as the maximum observed frontal frequency was around 53%. Ignoring bins with less than m entries a total of 576 base points were considered. This test will be used for the evaluation in different scenarios.

S5.3   Box Plots

The box plots in Fig. 11 display R_1 as a function of $P_{a(fr)}$. For this we divided all sample points into $k = 21$ bins. Each bin b_i with $0 \leq i < k, i \in \mathbb{N}$ contains all $R_1(p)$ for all grid points p within the midlatitudes, excluding high altitude regions, where
85    $(i - 1) \cdot 5\% < P_{a(fr)}(p) \leq (i) \cdot 5\%$.

S5.4   Refined attribution radius

In Fig. S4 we provide the same content as Fig. 10(a) for two smaller radii of $5px$ ($1.25°$) and $2px$ ($0.5°$), respectively. The qualitative features (i.e. the regions with high correlations) remain the same but the correlation magnitude is reduced due to the smaller radius of influence. Such investigations cannot be carried out with classical TFP methods, since they are restricted to
90   low resolution data sets.

Figure S4. As Fig. 10(a). Proportion of extreme precipitation events, which are also associated with a front, where the association radius is $5px$ ($1.25°$) (a) and $2px$ ($0.5°$) (b), respectively. Regions with high topography are shaded in light gray, while areas where no extreme precipitation events occurred in 2016 are shaded in dark gray. Regions where no significant correlation between extreme precipitation and fronts was found are blanked.