Improved objective identification of meteorological fronts: a case study with ERA-Interim

Philip G. Sansom[1,2] and Jennifer L. Catto[1]

[1]Faculty of Environment, Science and Economy, University of Exeter, North Park Road, Exeter, EX4 4QE, UK

[2]Met Office, FitzRoy Road, Exeter, EX1 3PB, UK

Correspondence: Jennifer L. Catto (j.catto@exeter.ac.uk)

Abstract. Meteorological fronts are important for their associated surface impacts, including extreme precipitation and extreme winds. Objective identification of fronts is therefore of interest in both operational and research settings. We have implemented a number of changes in a widely used objective front identification algorithm, and present the improvements associated with these changes. First, we show that a change to the order of operations from applying a mask then joining frontal points to contouring the thermal field then applying the mask, yields smoother fronts with fewer breaks. Next we address the selection of the identification parameters, including the thresholds and number of smoothing passes. This allows a comparison between datasets of differing resolutions. Finally, we have made a number of numerical improvements in the implementation of the algorithm, such as more accurate finite differencing, direct calculation of the wet-bulb potential temperature, and better handling of short fronts, which yield further benefits in smoothness and number of breaks. This updated version of the algorithm has been made fully portable and scalable to different datasets in order to enable future climatological studies of fronts and their impacts.

1 Introduction

Atmospheric fronts are of great importance for the day-to-day variability of weather in the mid-latitudes. They are associated with a large proportion of total and extreme precipitation, as demonstrated with numerous case studies (Browning, 2004), modelling, and more recently, long-term climatologies (Berry et al., 2011b; Parfitt et al., 2017b; Schemm et al., 2017). They are also strongly linked to extreme wind events (Dowdy and Catto, 2017; Catto et al., 2019; Raveh-Rubin and Catto, 2019; Catto and Dowdy, 2021), and are key for air-sea interaction (Parfitt et al., 2017b). With a wealth of global gridded observationally-constrained and model-produced data, there is a desire to be able to objectively identify these frontal features in the gridded data. This avoids the huge time requirements of a manual analysis, and allows the features to be linked to high impact weather, such as extreme precipitation or winds (Catto et al., 2012; Catto and Pfahl, 2013; Dowdy and Catto, 2017). The application of

the methods to model data of historical and future climate also allows the models to be evaluated for their ability to capture the dynamical features and their connection to precipitation events (Leung et al., 2022), and to investigate the future of such features and how they may impact on water resources and natural hazards (Catto et al., 2014).

A number of methods have been developed to perform such objective identification in recent years. Hewson (1998) compiled
25 a summary of methods prior to then, used to identify frontal features in gridded data, and further developed the methods based on a thermal front parameter. Thomas and Schultz (2019) highlighted the three main factors required in identifying fronts with such a thermal front parameter. First, the thermal variable and vertical level to be considered, such as temperature, potential temperature, or equivalent (or wet-bulb) potential temperature at, say, $850\,\mathrm{hPa}$. The second, a function of the variable, such as the simple gradient or some second or third derivative. And third, some thresholds. They found that different thermal variables
30 each had pros and cons, and could be selected depending on the purpose of the study. The study by Jenkner et al. (2010) used a thermal variable (equivalent potential temperature), and the second derivative of that variable to place the frontal lines. This results in the fronts lying in the centre of a frontal zone, rather than at the leading edge as a synoptic meteorologist would typically put them. A very similar method was developed by Berry et al. (2011b), which directly applied the methods of Hewson (1998) to gridded data at $2.5° \times 2.5°$ resolution, and placing fronts on the warm side of the strong temperature
35 gradient. This also included the addition of a numerical line-joining algorithm, which is a final piece of the front identification puzzle.

Other methods have used dynamical information to identify fronts. Simmonds et al. (2012) use information solely on wind shifts. This method was found to work better in the Southern Hemisphere than the Northern Hemisphere by Schemm et al. (2015). A combination of this and the thermal method was used by Bitsa et al. (2021) in the Mediterranean, to identify cold
40 fronts with the method benefiting the smaller spatial scale of fronts in this region. Parfitt et al. (2017b) use a combination of vorticity and temperature, requiring both a thermal gradient and a wind shift. While each method has its advantages and disadvantages, many of the methods typically identify many of the same features (Hope et al., 2014).

A major difficulty in applying objective front identification is the many datasets and differing resolutions. This is particularly an issue when using gradients of thermal properties, since the resolution of the data will have a large impact on these gradients.
45 The thresholds used to define fronts need to be varied depending on the resolution. For example, the threshold used in Berry et al. (2011b) was $-8 \times 10^{-12}\,\mathrm{K\,m^{-2}}$, while in Dowdy and Catto (2017) on the higher resolution data was $-5 \times 10^{-11}\,\mathrm{K\,m^{-2}}$. Recently, Soster and Parfitt (2022) investigated the sensitivity of results to the use of different datasets and found a large difference in front frequency between the datasets. Higher resolution datasets consistently show higher frequency of frontal points, with the difference reduced when re-gridding to a common grid. This can impact the attribution of precipitation to
50 fronts.

Despite the many methods of identifying fronts, and issues and uncertainties associated with each of them, the thermal front parameter method of Hewson (1998) has been successful in identifying the key climatological features of front frequency and the link to other variables in a number of studies (Berry et al., 2011b, a; Catto and Pfahl, 2013; Dowdy and Catto, 2017). The aim of this study is to create a portable implementation of the front identification method of (Hewson, 1998) that is able to scale
55 to contemporary high resolution (re-)analyses. We demonstrate a method that can be used to avoid the subjective tuning of the

thresholds. First the data used are described in Section 2. Section 3 gives a description of the thermal front parameter method, and the new techniques and improvements implemented in the current version of the algorithm. In Section 4 we compare the front climatology using the new method with previous methods and different datasets. We finish in Section 5 with a discussion of the benefits and challenges associated with such objective identification methods.

2 Data

The updated front identification procedure is applied to European Centre for Medium-Range Weather Forecasts' (ECMWF) reanalysis ERA-Interim (Dee et al., 2011). The data used here have a resolution of $0.75° \times 0.75°$ on a regular longitude-latitude grid. The 6-hourly instantaneous air temperature and specific humidity fields at the $850\,\mathrm{hPa}$ level were used to compute the wet-bulb potential temperature θ_W in order to identify fronts. The wet-bulb potential temperature θ_W is computed using the direct method of Davies-Jones (2008, Equation 3.8). The 6-hourly eastward and northward wind components at $850\,\mathrm{hPa}$ were used to compute the front speed for classification into cold, warm or quasi-stationary fronts. ERA-Interim was chosen over the more recent ERA5 reanalysis for the primary analysis since the updated procedure is of greatest benefit in middle and low resolution models, and the resolution of ERA-Interim is equal to that of the highest resolution standard CMIP6 GCMs. Our baseline for comparison is the global climatology of fronts in ERA-Interim at $0.75° \times 0.75°$ produced by Dowdy and Catto (2017) using the method of Berry et al. (2011b). We also present a high resolution climatology based on applying the updated front identification procedure to the ERA5 reanalysis (Hersbach et al., 2020) using the same 6-hourly fields as ERA-Interim but with a grid spacing of $0.25° \times 0.25°$.

3 Methodology

Following Hewson (1998) and Berry et al. (2011b), fronts are identified in the wet-bulb potential temperature field θ_W at $850\,\mathrm{hPa}$. As described in Hewson (1998, their Equation 5) and used in Berry et al. (2011b), fronts are located as the zero contour of:

$$\nabla \cdot \nabla |\nabla \theta_W| = 0 \quad \text{or} \quad \nabla^2 |\nabla \theta_W| = 0. \tag{1}$$

In one dimension, this is simply the third derivative of the wet-bulb potential temperature θ_W (see Figure 3 of Hewson (1998) for an intuitive explanation). We will refer to Equation 1 as the Thermal Front Locator (TFL).

Hewson (1998) defined two additional criteria that must be met in order for a zero contour of the Equation 1 to be considered a front. First, the rate of change of θ_W across the front in the direction of cold air must exceed some threshold value K_1. This criterion was formalised in Equation 9 of Hewson (1998) as:

$$\nabla |\nabla \theta_W| \cdot \frac{\nabla \theta_W}{|\nabla \theta_W|} < K_1 \quad \text{where} \quad K_1 \leq 0\,\mathrm{K\,m^{-2}}. \tag{2}$$

This is the Thermal Front Parameter (TFP) defined by Renard and Clarke (1965). In one dimension, this criterion simply states that the second derivative of θ_W must be negative, placing the front on the warm side of the gradient. Second, the gradient of θ_W in the adjacent baroclinic zone (ABZ) must be greater than some threshold value K_2. This criterion was formalised in Equation 11 of Hewson (1998) as:

$$|\nabla\theta_W|_{ABZ} > K_2 \quad \text{where} \quad K_2 \geq 0\,\mathrm{K\,m^{-1}} \tag{3}$$

with

$$|\nabla\theta_W|_{ABZ} = |\nabla\theta_W| + m\chi|\nabla|\nabla\theta_W||$$

where $m = 1/\sqrt{2}$ and χ is the grid length. In one dimension, this criterion simply states that the magnitude of the first derivative of θ_W must be greater than K_2 a fraction of a grid length in the direction of increasing gradient, i.e. inside the adjacent baroclinic zone.

Fronts are identified as warm, cold or quasi-stationary using the front speed defined by Equation 13 of Hewson (1998), which is given here as:

$$\frac{\mathbf{V}\cdot\nabla|\nabla\theta_W|}{|\nabla|\nabla\theta_W||} \tag{4}$$

where $\mathbf{V} = (u,v)^T$ is the vector wind field at $850\,\mathrm{hPa}$. Following Berry et al. (2011b), we adopt a threshold of $K_3 = 1.5\,\mathrm{m\,s^{-1}}$ such that front points are defined as belonging to warm fronts if they have speed exceeding $1.5\,\mathrm{m\,s^{-1}}$, and as belonging to cold fronts if they have speed less than $-1.5\,\mathrm{m\,s^{-1}}$. All other front points are defined as belonging to quasi-stationary fronts.

The automatic front identification method described by Equations 1, 2, 3 and 4 has been reimplemented in the R statistical computing language (R Core Team, 2021). The new implementation includes one key methodological improvement, as well as number of numerical updates compared to that of Berry et al. (2011b).

3.1 Methodological changes

The key methodological change we present is in the order of operations when locating the linear front objects. Berry et al. (2011b) take what we will call a "mask-then-join" approach. First they locate all those grid boxes that satisfy Equation 2 to form a mask (the criterion in Equation 3 is not used). Zero points of Equation 1 are located by an exhaustive search using linear interpolation between only those grid boxes included in the mask defined Equation 2. Finally, a line-joining algorithm is used to join the zero points of Equation 1 into lines representing fronts. Points are joined to their nearest neighbour if the two points are separated by less than a threshold amount specified in degrees.

Instead, we propose a "contour-then-mask" approach. We simultaneously identify zero points in the complete field defined by Equation 1 and join them into lines using a contouring algorithm, specifically the `contourLines()` function in R. Zero points are again located by linear interpolation, but only joins between adjacent grid boxes are considered. We then interpolate
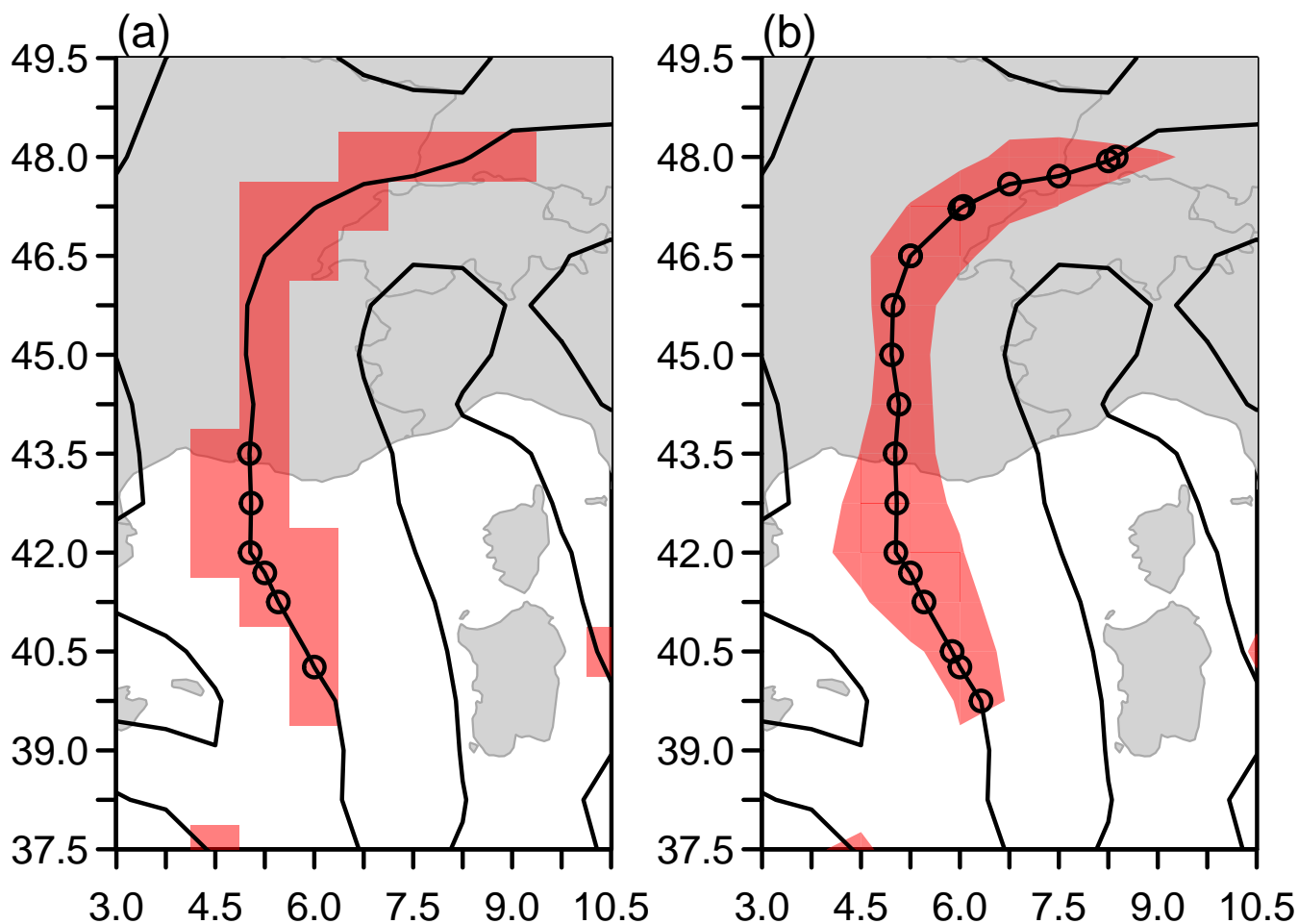
Figure 1. Front identification in ERA-Interim at 00:00 on 2001-01-01. (a) Using the mask-then-join approach, and (b) using the contour-then-mask approach. Thick black lines are contours of the $\nabla \cdot \nabla |\nabla \theta_W| = 0$. Red shading indicates regions where TFP $\leq -5 \times 10^{-11}\,\mathrm{K\,m^{-2}}$ by masking in (a) and interpolation in (b). Points indicate front points located by each algorithm.

the values of the fields defined by Equations 2 and 3 onto the points located by the contouring algorithm. Only points that meet the criteria defined by Equations 2 and 3 are retained, leaving a set of pre-joined line segments representing fronts.

115   The two approaches are compared in Figure 1. Zero points in Equation 1 usually occur between grid points. That means that adjacent grid boxes meeting the criteria in Equation 2 are required in order to find zero points using Equation 1 by the mask-then-join approach. In even moderately high resolution analyses such as ERA-Interim, the region that satisfies Equation 2 is often narrow, frequently only one grid box wide. Therefore, the mask-then-join approach frequently fails to locate front points when the zero contour strays to the edge of a grid box, or grid boxes are adjacent diagonally. This may result in some features not being identified at all or, more frequently, gaps in what should be continuous features. The line-joining algorithm used by 120 Berry et al. (2011b) attempts to mitigate this by using a search radius larger than one grid length, but this is only partially

effective. The weak front extending from the Mediterranean to the east and north of the Alps (Figure 1) is a typical example. The mask-then-join approach in Figure 1(a) flags the northern half as a possible front but is unable to identify front points due to the tight masking. The contour-then-mask approach approach in Figure 1(b) is able to successfully identify the front as a

125 single object. Overall, the contour-then-mask approach results in more fronts and front points identified, and fewer breaks.

3.2   Choosing the thresholds

Although automated methods offer the promise of objective feature identification, it is still usually necessary to set some key parameters. For front identification there are three parameters we can change: the amount of smoothing applied to the θ_W field, the TFP threshold K_1, and the gradient threshold K_2. Some studies have compared outputs with manual analyses

130 by meteorologists to calibrate the parameters. This is difficult, time consuming and calibrates the algorithm to the subjective judgement of a single meteorologist. Also, all three parameters depend on the resolution of the data. Therefore, the calibration must be repeated for each new dataset, or datasets brought to a common resolution for comparison. Instead, we offer some suggestions for objective calibration criteria.

  We first address the smoothing problem, since the amount of smoothing applied to θ_W affects the choice of K_1 and K_2. The

135 purpose of smoothing is to remove local minima and maxima that might break up otherwise smooth features. We particularly wish to avoid local extrema in the TFL field defined by Equation 1. Local extrema will appear as short closed contours of $\mathrm{TFL} = 0\,\mathrm{K\,m^{-3}}$. Therefore, it makes sense to examine the effect of smoothing on the average length of the contours of $\mathrm{TFL} = 0\,\mathrm{K\,m^{-3}}$. Previous studies applying the method of Berry et al. (2011b) to ERA-Interim used $n = 2$ passes of a simple five-point average to smooth the θ_W field. In testing on ERA-Interim, it was found that the average length of the contours

140 of $\mathrm{TFL} = 0\,\mathrm{K\,m^{-3}}$ initially increases rapidly with the number of passes of the five-point smoother, but after 6–10 passes, the effect of further smoothing diminishes. Therefore, we settled on $n = 8$ passes of a five-point smoother.

  It is common to define weather phenomena as events exceeding some percentile of the climatological distribution. Therefore, we propose a quantile based approach to setting the thresholds K_1 and K_2. When considering the choice of threshold, it may be necessary to consider both location and season. In practice, there is little monthly variation in the climatological averages of

145 TFP or the magnitude of the gradient (Equation 3). However, there is considerable spatial variation, as illustrated in Figure 2. The climatological zonal averages of both the TFP and magnitude of gradient are relatively constant in the extratropical regions, but very different in the tropics and polar regions. Since most fronts occur in the extratropical regions, we will focus our attention there. We set the first threshold K_1 to the 25th percentile of the climatological distribution of the TFP. In the Northern Hemisphere extra-tropics this is around $-1.6 \times 10^{-11}\,\mathrm{K\,m^{-2}}$. We take the 25th percentile of the TFP since we require that

150 $K_1 < 0\,\mathrm{K\,m^{-2}}$ in order that the location represents a minimum, and the 50th percentile of TFP is approximately $0\,\mathrm{K\,m^{-2}}$. We set the second threshold K_2 equal to the 50th percentile of the climatological distribution of the magnitude of the gradient of θ_W. In the Northern Hemisphere extra-tropics (23.4°N–66.6°N) this is around $7.5 \times 10^{-6}\,\mathrm{K\,m^{-1}}$. In the absence of strong physical reasoning, these quantiles have a simple symmetry and were found to identify continuous fronts in good agreement with published charts for the North Atlantic and Europe. For comparison, previous studies applying the method of Berry et al.

155 (2011b) to ERA-Interim used thresholds of $K_1 = -5 \times 10^{-11}\,\mathrm{K\,m^{-2}}$ and $K_2 = 0\,\mathrm{K\,m^{-1}}$ after $n = 2$ smoothing passes. Our
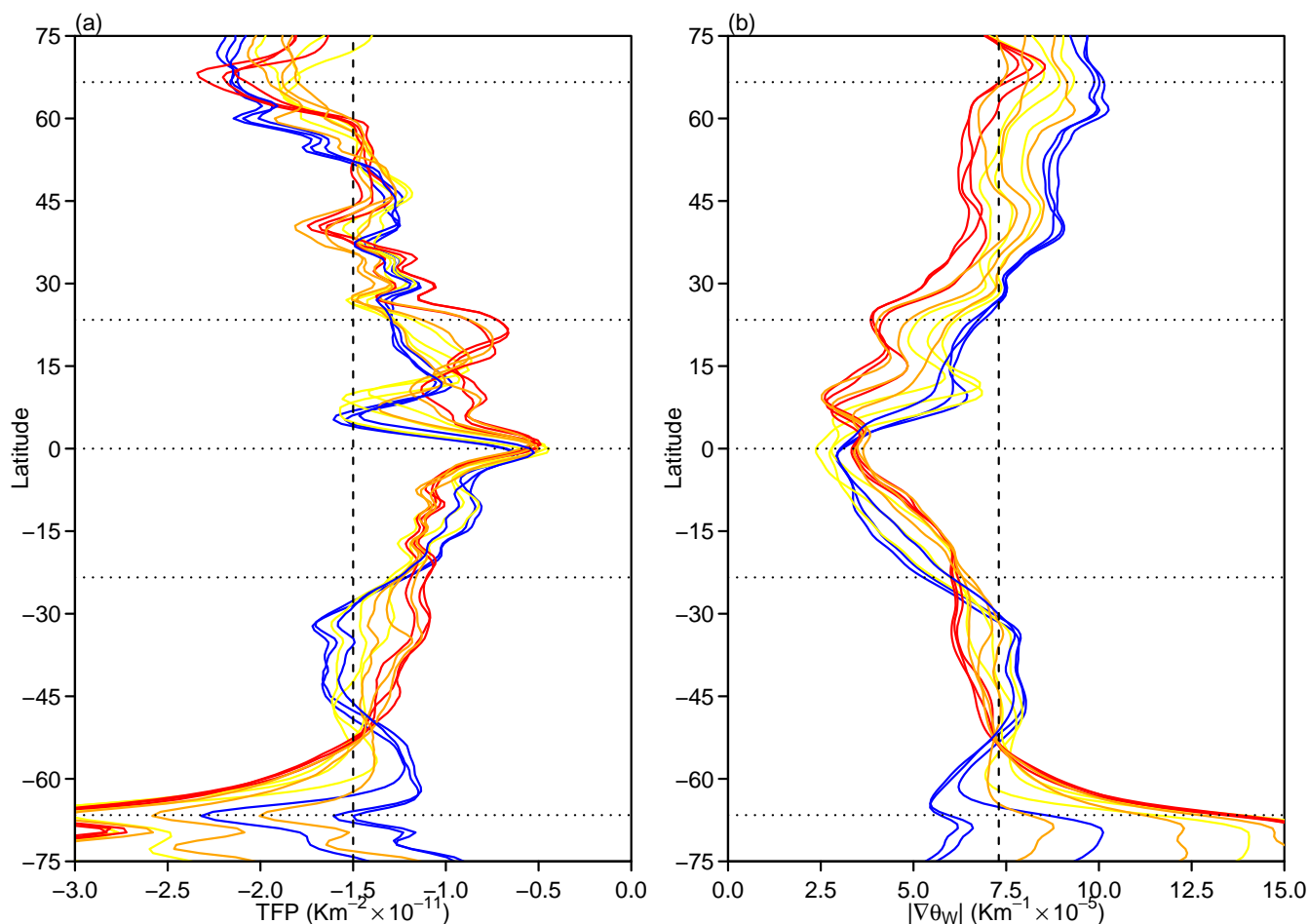
Figure 2. Choosing the thresholds K_1 and K_2. (a) The 25th percentile of the zonal TFP, and (b) the 50th percentile of $|\nabla \theta_W|$. Each coloured line represents a different month: blue for DJF, yellow for MAM, red for JJA, and orange for SON. Horizontal dotted lines represent the major circles of latitude. Vertical dashed lines indicate the thresholds chosen in the text.

threshold K_1 is lower in part due to the additional smoothing, and in part due to the addition of the second threshold K_2 which was previously set to 0 and thus had no effect.

3.3 Comparing fronts from different datasets

When comparing analyses from different weather and climate datasets, the most common approach is to interpolate all the
160 datasets to a common resolution, usually the lowest resolution among the datasets of interest. For some features such as fronts that are more easily identified in higher resolution data, this can be limiting. The objective calibration criteria described in Section 3.2 provide one route by which fronts could be identified at the native resolution of each dataset and then compared. The quantile based criteria will identify the same fraction of grid boxes potentially containing front points at any reasonable resolution or number of smoothing passes. However, computing the required climatologies is time consuming. An alternative is

165 to keep the thresholds K_1 and K_2 constant, and adjust the number of smoothing passes so that the climatological distributions of the TFP and magnitude of gradient fields are similar between datasets. Specifically, the quantiles used to set the thresholds should be similar. In testing, it was found that matching the threshold percentile of the TFP field provided a more consistent comparison than the gradient field. The percentiles can be compared for only a small subset of the data, provided the same subset is used for each dataset, avoiding the need to compute a long climatology. In testing, one month of data in the Northern
170 Hemisphere extra-tropics was sufficient to consistently determine an appropriate number of smoothing passes.

3.4 Numerical updates

The biggest single numerical update relates to the computation of the quantities in Equations 1, 2 and 3. Berry et al. (2011b) used repeated applications of a simple central finite difference approximation to the first derivative. The simple approximation uses one grid box either side of the object to approximate the derivative to second order accuracy. However, repeated
175 applications of the simple approximation degrades its accuracy for higher derivatives. In contrast, we use explicit central finite difference approximations for the second derivatives in order to maintain second order accuracy throughout. The increased accuracy comes at minimal computation cost, but is valuable due to the small scale of the quantities of interest, e.g., $K_1 = -5 \times 10^{-11}\,\mathrm{K\,m^{-2}}$ in Dowdy and Catto (2017).

Other numerical differences include using the mixed-phase parametrization of relative humidity from the ECMWF Integrated
180 Forecasting System (ECMWF, Section 7.4.2) rather than the table based approach from the NCAR Command Language (NCL). In the new implementation, the wet-bulb potential temperature (θ_W) is computed using the direct method of Davies-Jones (2008, Equation 3.8), rather than an iterative method. The final numerical difference between the two implementations is how short fronts are handled. In the original application, Berry et al. (2011b) reject any fronts less than three points long. In later applications this was updated to a distance based criteria where fronts whose end points are less than $250\,\mathrm{km}$ apart are rejected.
185 In our implementation, we sum the distance between all adjacent points in each front and reject fronts whose total length is less than $250\,\mathrm{km}$.

4 Results

4.1 Comparison with previous implementations

Figure 3 illustrates the difference between the mask-then-join and contour-then-mask methods, and the effect of the updated
190 parameter choices (i.e. smoothing passes, K_1, and K_2) in ERA-Interim at 00:00 on 2001-01-01. The mask-then-join approach using the original parameters in Figure 3(a) is clearly identifying fronts, but they are fractured with frequent gaps. The contour-then-mask in Figure 3(b) results in much smoother front features with fewer gaps, and more fronts identified. Figure 3(c) shows the results of the updated parameters with more smoothing cycles and stronger thresholds. Compared to the original parameters, the front features are smoother, with fewer breaks and many spurious local fronts have been removed. One feature
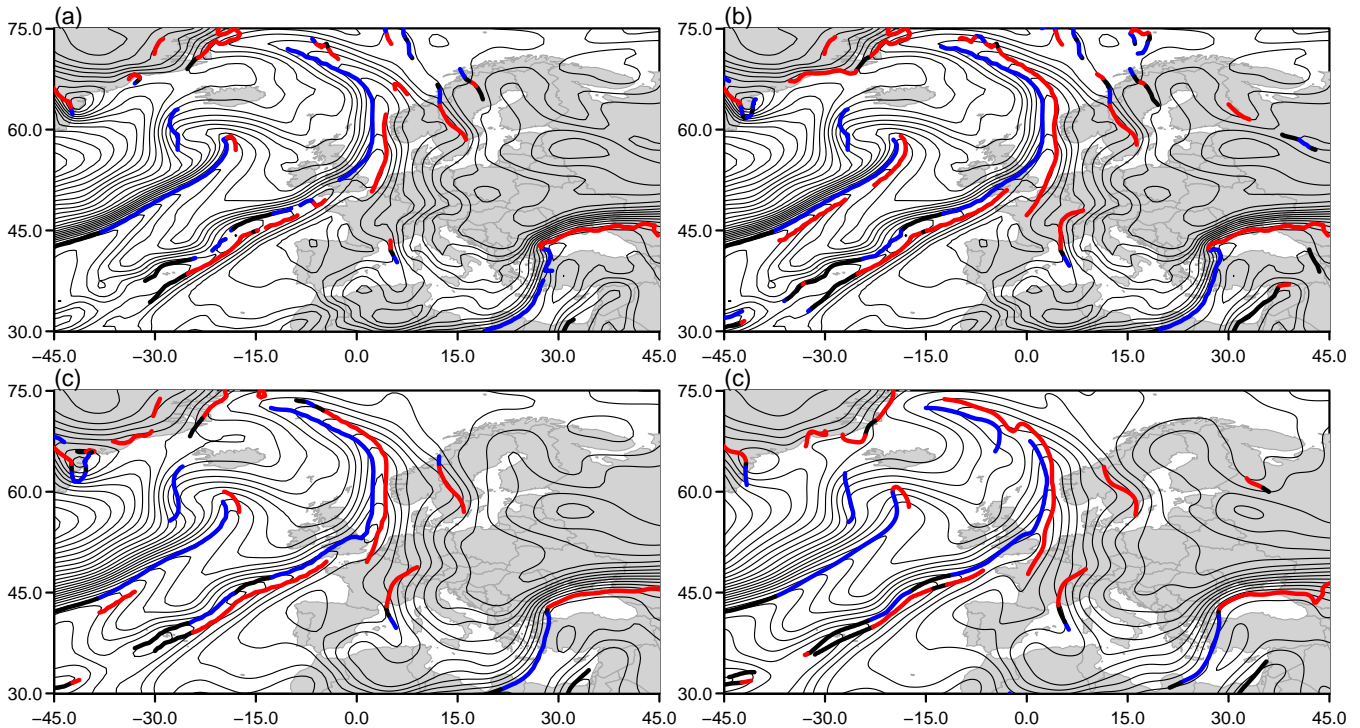
Figure 3. Comparison of methods. (a) Mask-then-join with $n = 2$, $K_1 = -5 \times 10^{-11}\,\mathrm{K\,m^{-2}}$ and $K_2 = 0\,\mathrm{K\,m^{-1}}$, (b) contour-then-mask with $n = 2$, $K_1 = -5 \times 10^{-11}\,\mathrm{K\,m^{-2}}$ and $K_2 = 0\,\mathrm{K\,m^{-1}}$, (c) contour-then-mask with $n = 8$, $K_1 = -1.6 \times 10^{-11}\,\mathrm{K\,m^{-2}}$ and $K_2 = 7.5 \times 10^{-6}\,\mathrm{K\,m^{-1}}$, (d) ERA5 with $n = 96$, $K_1 = -1.6 \times 10^{-11}\,\mathrm{K\,m^{-2}}$ and $K_2 = 7.5 \times 10^{-6}\,\mathrm{K\,m^{-1}}$. Thin black lines indicate contours of wet-bulb potential temperature θ_W. Thick blue lines indicate cold fronts, thick red lines indicate warm fronts and thick black lines indicate quasi-stationary fronts. All fronts were classified using a threshold of $K_3 = 1.5\,\mathrm{m\,s^{-1}}$

that can be seen is the warm front directly to the southeast of the cold front that crosses the south of the UK. This is a common feature and tends to relate to pre-frontal troughs that are often analysed on synoptic charts.

Figure 4 compares the front frequency climatologies from three different implementations of the Hewson (1998) algorithm applied to ERA-Interim with identical parameters (i.e. the same number of smoothing passes, and thresholds K_1 and K_2). Specifically, the implementation of Berry et al. (2011b) used by Dowdy and Catto (2017) in Figure 4(a), a version incorporating our numerical updates but using the original mask-then-join approach in Figure 4(b), and our final version using the contour-then-mask approach in Figure 4(c). Figure 4(d) shows that our numerical updates result in slightly greater numbers of fronts identified in almost all regions, with the exception of areas of very high orography and the far Southern Ocean. Figures 4(e) and (f) compare our final version with the implementation of Berry et al. (2011b) and the version incorporating only the numerical updates. The greatest increases in the number of fronts identified are seen in or adjacent to regions where fronts were already common, highlighting the effectiveness of the contour-then-mask approach at reducing the breaks in the fronts. In the densest region of the North Atlantic storm track, the number of fronts identified increases by almost $100\,\%$.
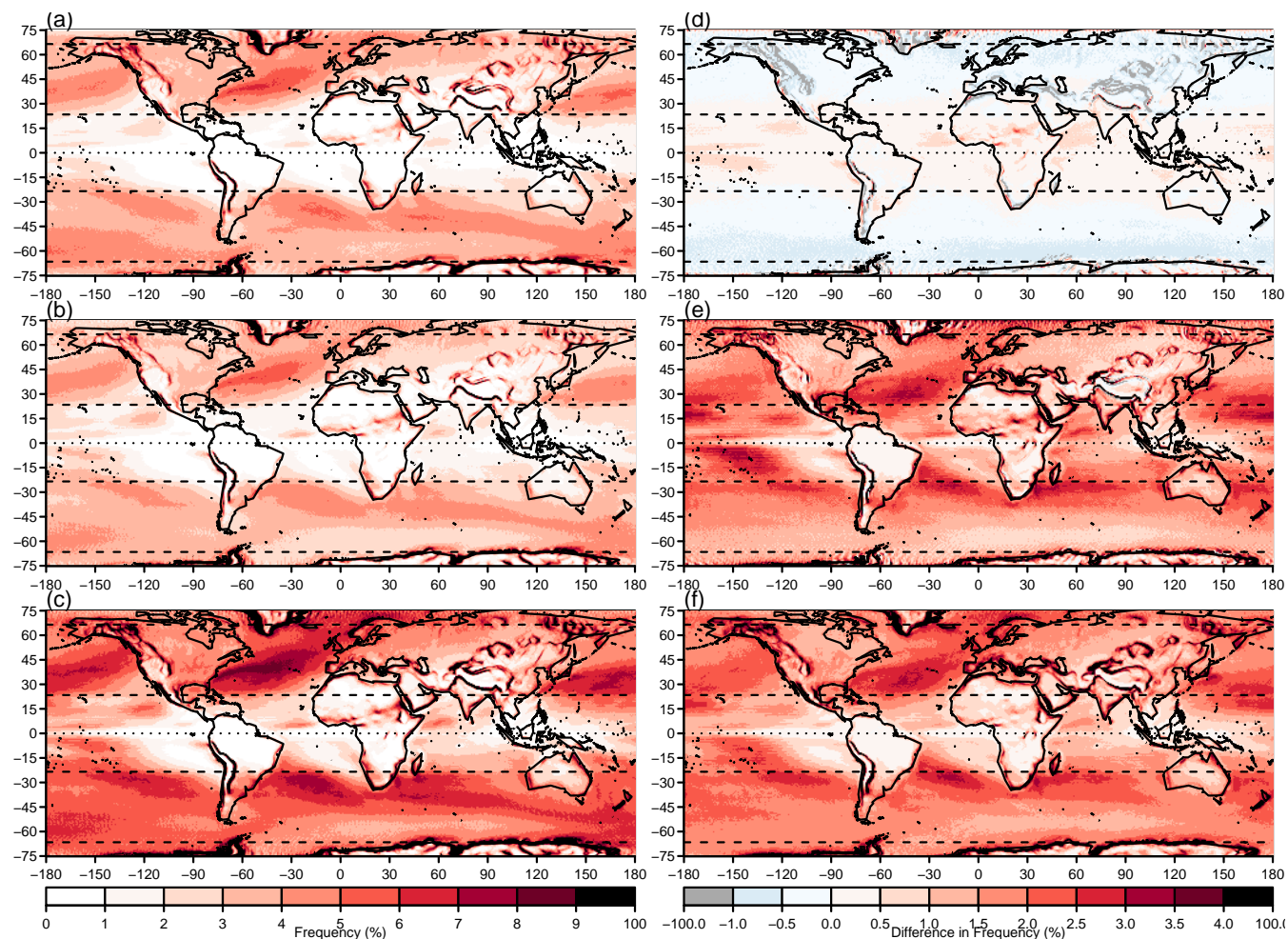
Figure 4. Comparison of global climatologies of front frequency as a percentage of 6-hourly times. (a) Dowdy & Catto (2017), (b) updated implementation using the mask-then-join approach (c) updated implementation using the contour-then-mask approach, (d) difference between mask-then-join and Dowdy & Catto (2017), (e) difference between contour-then-mask and Dowdy & Catto (2017), and (f) difference between contour-then-mask and mask-then-join. All climatologies were computed with $n = 2$ smoothing cycles, $K_1 = -5 \times 10^{-11} \, \text{K m}^{-2}$ and $K_2 = 0 \, \text{K m}^{-1}$

While changing the implementation of the front identification leads to an increase in the number of fronts identified, as shown in Figure 4, the next aspect of the updated method is a change to the parameters used. Figure 5 compares the climatology of front frequency of our final version with updated parameters (i.e. smoothing passes, and thresholds K_1 and K_2) applied to ERA-Interim against the implementation by Berry et al. (2011b) with the original parameters. The updated parameters result in slightly fewer fronts identified in almost all regions, due to the increased smoothing, making the climatology more similar to earlier estimates, but with the smoother individual fronts given by the contour-then-mask method. The greatest decreases are seen on the edge of the tropics, adjacent to regions with high front activity. This pattern is to be expected due to the rapid drop-off in the climatology of the masking parameters in the tropics in Figure 2.
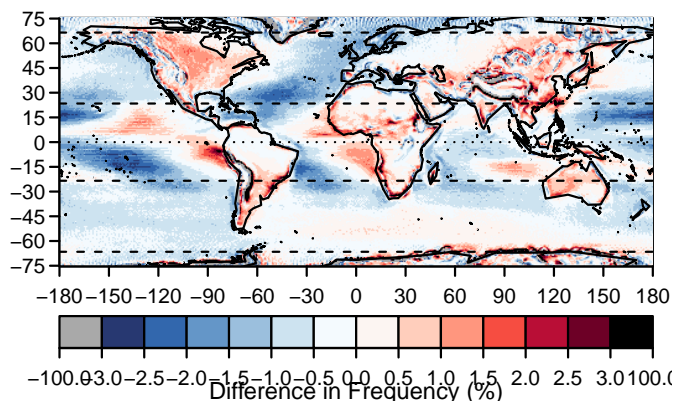
Figure 5. Updated parameters. Difference between ERA-Interim climatology of front frequency computed using updated parameters $n = 8$, $K_1 = -1.6 \times 10^{-11} \, \mathrm{K\,m^{-2}}$ and $K_2 = 7.5 \times 10^{-6} \, \mathrm{K\,m^{-1}}$, and the original parameters $n = 2$, $K_1 = -5 \times 10^{-11} \, \mathrm{K\,m^{-2}}$ and $K_2 = 0 \, \mathrm{K\,m^{-1}}$.
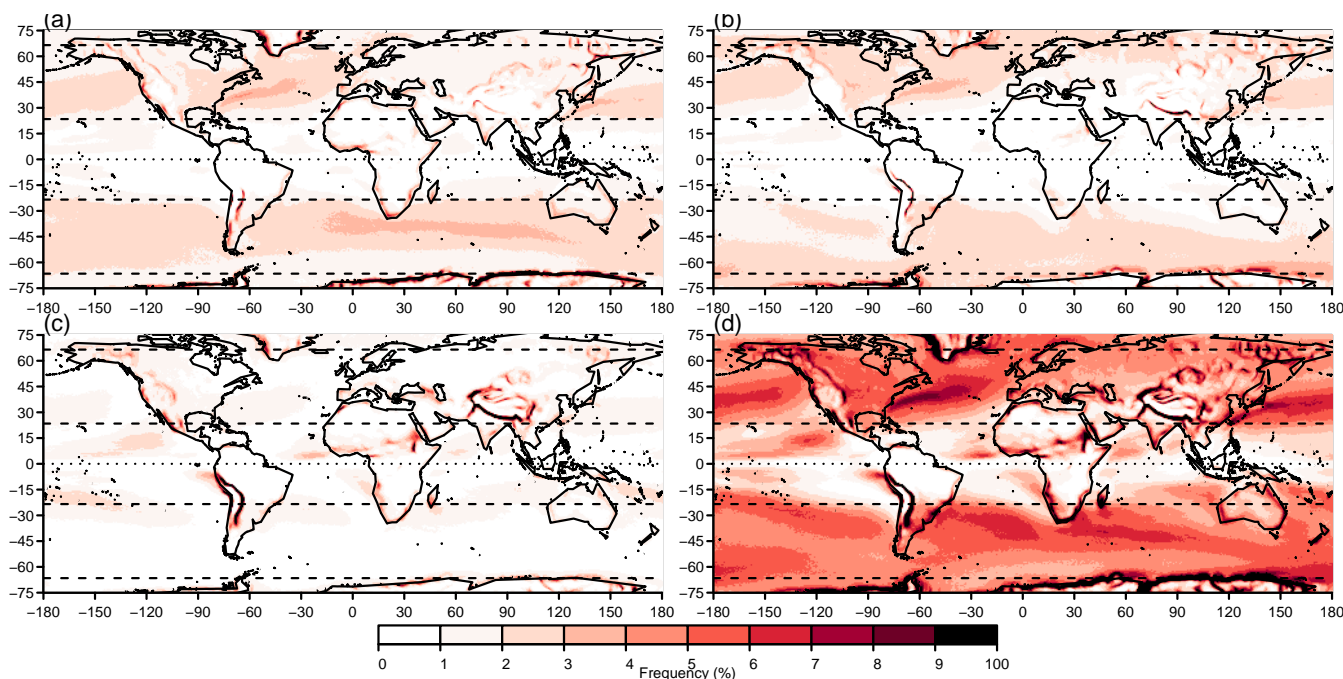
Figure 6. Updated global climatologies of front frequency as a percentage of times. (a) Cold fronts, (b) warm fronts, (c) quasi-stationary fronts, (d) all fronts. All climatologies were computed with $n = 8$ smoothing cycles, $K_1 = -1.6 \times 10^{-11} \, \mathrm{K\,m^{-2}}$ and $K_2 = 7.5 \times 10^{-6} \, \mathrm{K\,m^{-1}}$ and $K_3 = 1.5 \, \mathrm{m\,s^{-1}}$.

4.2 Fronts identified in moderate resolution reanalysis data

Figure 6 shows the climatology of front frequency of our final version with updated parameters applied to ERA-Interim, including the breakdown into cold, warm and quasi-stationary fronts. Figure 6(d) allows a direct comparison with the climatologies

in Figure 4, showing that while the updated parameters reduce the number of fronts identified compared to the updated numerical implementation only, overall more fronts are still identified in almost all regions than in earlier versions. Figure 6(a) and (b) show that warm and cold fronts occur with similar frequencies in most extra-tropical regions. Figure 6(c) shows that quasi-stationary fronts occur most often in the tropics where gradients are weaker, and adjacent to very high orography, as expected.

Figure 7 breaks the classification of fronts down still further, reporting cold and warm fronts by season. Unsurprisingly, cold fronts in the Northern Hemisphere are most common at the beginning of the storm track regions of both the Atlantic and Pacific oceans in northern winter (DJF) in Figure 7(a). In contrast, warm fronts in northern summer in in Figure 7(g) tend to outnumber cold fronts in in Figure 7(c). In agreement with Berry et al. (2011b), the seasonal distribution of fronts in the Southern Hemisphere is much more stable. Somewhat surprisingly, cold fronts are slightly more common though less widely distributed in the Southern Hemisphere during southern summer (DJF, Figure 7(a)) than in southern winter (JJA, Figure 7(c)). When the storm track moves poleward in Southern winter there are larger numbers of warm fronts near Antarctica in Figure 7(g).

4.3   Fronts identified in high-resolution reanalysis data

The ERA5 reanalysis has a higher resolution than ERA-Interim, with grid spacing of $0.25° \times 0.25°$ compared to $0.75° \times 0.75°$ for ERA-Interim. For ERA5, a total of $n = 96$ smoothing cycles were required to make the climatologies of the TFP and gradient similar to ERA-Interim. Figure 3(d) illustrates fronts identified over Europe and the North Atlantic at 00:00 on 2001-01-01. As expected, the features are very similar to those identified in ERA-Interim against which it was calibrated in Figure 3(c). Figure 8 compares the frequency of fronts identified in ERA5 with ERA-Interim when fronts are identified in ERA5 at $0.25° \times 0.25°$ grid spacing with $n = 96$ smoothing cycles but identical thresholds then aggregated to $0.75° \times 0.75°$ grid spacing for comparison with ERA-Interim. When aggregated to the same resolution, more fronts are identified almost everywhere in ERA5 than in ERA-Interim due to the increased ability to resolve the required derivatives. The only decreases in frequency are seen around areas of high orography. The pattern of increase broadly follows the general distribution of fronts, with more fronts seen where there fronts were already common. Some of the greatest increases are seen in the tropics where large numbers of fronts are identified as a result of the increased resolution.

Figure 9 shows the climatology of fronts by type identified in ERA5 at it's native $0.25° \times 0.25°$ resolution. Due to the smaller grid boxes, the frequency is necessarily lower than for ERA-Interim in Figure 6 and the aggregated data in 8. Figure 9(c) suggests that the large increases in the frequency of tropical fronts seen in Figure 8(b) are largely due to quasi-stationary fronts, as might be expected given the more stable atmosphere of the region.

Figure 10 shows the seasonal breakdown of cold and warm fronts in ERA5, which is provided to be able to compare the most up-to-date climatology from ERA5 with previous studies. The increase in resolution provides a much smoother picture of the variation in front frequency across the globe. In general the maximum warm front frequency occurs at higher latitudes than the maximum cold front frequency, associated with the structure of extratropical cyclones and the poleward transport of warm air. During DJF especially, the sea surface temperature (SST) front associated with the Gulf Stream is clearly visible in
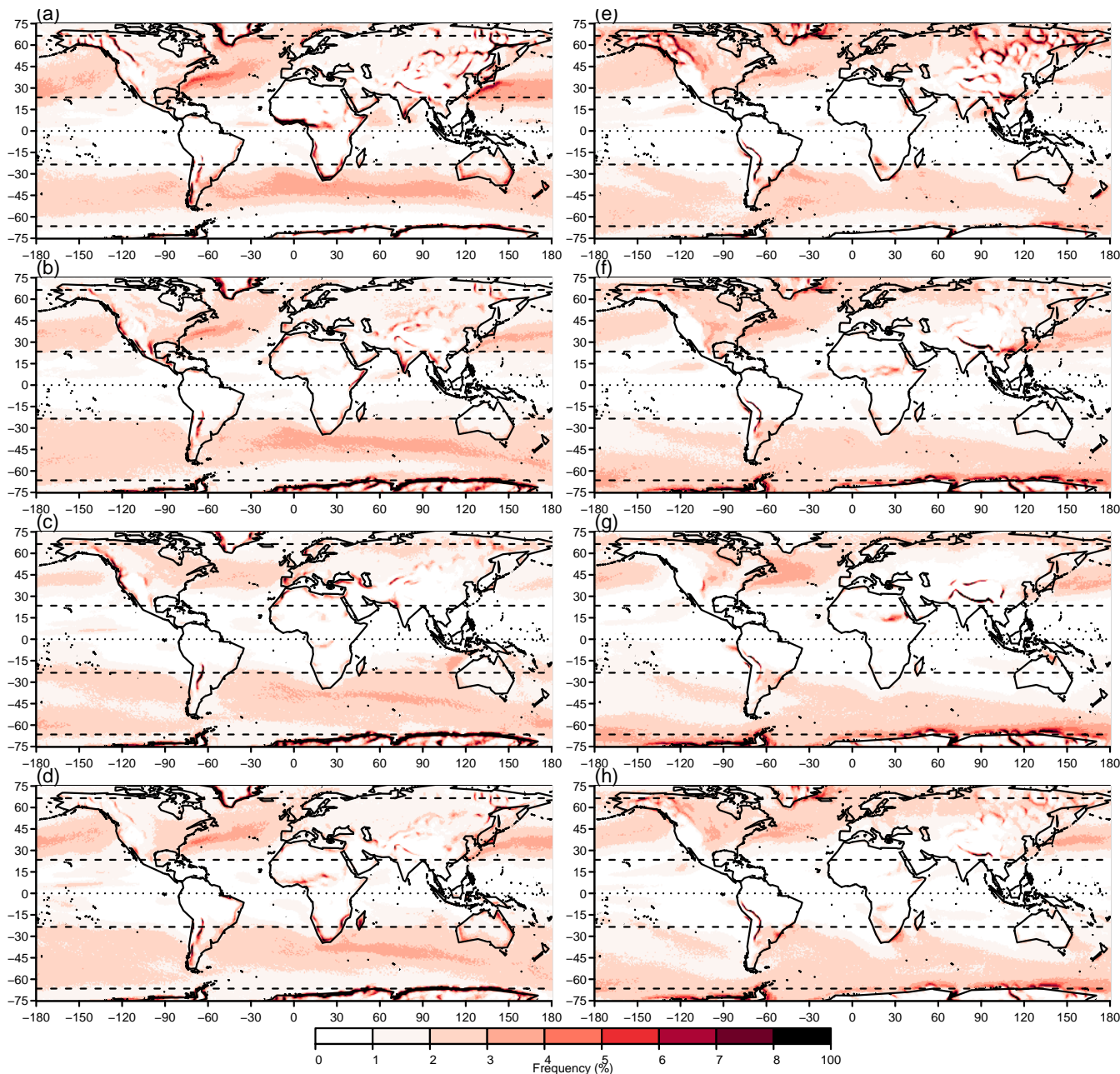
Figure 7. Updated seasonal climatologies of front frequency. (a) Cold fronts in DJF, (b) Cold fronts in MAM, (c) Cold fronts in JJA, (d) Cold fronts in SON, (e) Warm fronts in DJF, (f) Warm fronts in MAM, (g) Warm fronts in JJA, (h) Warm fronts in SON, All climatologies were computed with $n = 8$ smoothing cycles, $K_1 = -1.6 \times 10^{-11}\,\mathrm{K\,m^{-2}}$ and $K_2 = 7.5 \times 10^{-6}\,\mathrm{K\,m^{-1}}$ and $K_3 = 1.5\,\mathrm{m\,s^{-1}}$.

the North Atlantic. The influence of the SST on the atmosphere is more marked for higher resolution ocean and atmosphere (Parfitt et al., 2016, 2017a).
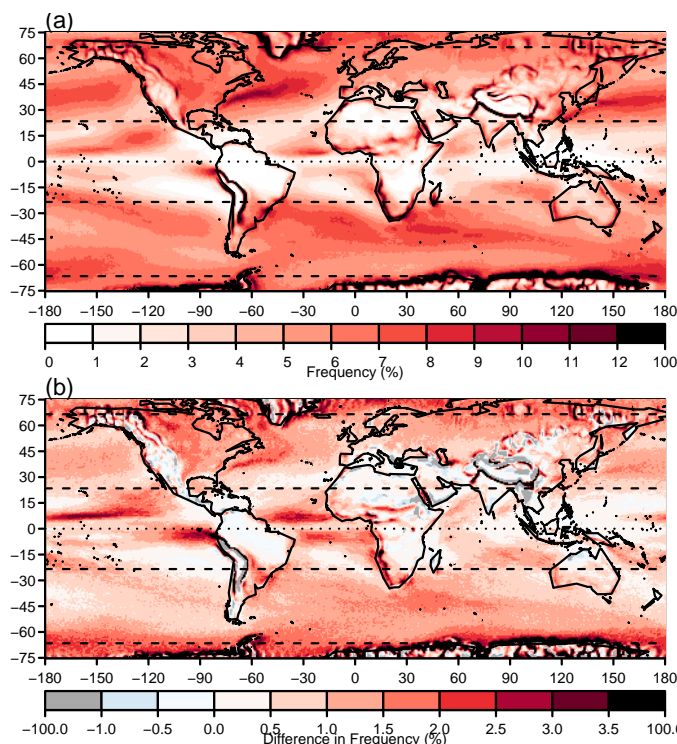
Figure 8. ERA5 compared to ERA-Interim. (a) ERA5 climatology of all fronts at $0.75° \times 0.75°$, and (b) difference between ERA5 and ERA-Interim. The ERA5 climatology were computed with $n = 96$ smoothing cycles, $K_1 = -1.6 \times 10^{-11} \, \mathrm{K\,m^{-2}}$ and $K_2 = 7.5 \times 10^{-6} \, \mathrm{K\,m^{-1}}$ and $K_3 = 1.5 \, \mathrm{m\,s^{-1}}$. ERA5 fronts were identified at $0.25° \times 0.25°$ then regridded to $0.75° \times 0.75°$ for comparison with ERA-Interim.

5 Discussion

In this paper, we have presented an updated implementation of the automatic front identification method of Berry et al. (2011b),

255   based on Hewson (1998). The updated implementation was designed specifically to scale to modern high resolution data sets. It

is open source and does not require compilation, making it extremely portable. Despite not requiring compilation, performance

is improved over earlier versions. One month of global ERA-Interim data at 6-hourly intervals and $0.75° \times 0.75°$ resolution

can be processed using a single core of a 2 year old laptop in around 6 minutes. The same amount of global ERA5 data at

$0.25° \times 0.25°$ can be processed in around 1 hour. Memory requirements are minimal since only one time step is processed at

260   once. The improved scalability enables us to present high resolution climatologies of cold, warm and quasi-stationary fronts

for all seasons from the ERA5 reanalysis.

In addition to several numerical improvements, we present one key methodological update in the contour-then-mask ap-

proach. We have demonstrated that moving from a mask-then-join to a contour-then-mask approach offers clear benefits.

Gaps in what should be continuous fronts are dramatically reduced in moderate resolution data sets such as ERA-Interim, and

265   improvements are even greater when applied to general circulation models. This improvement will be useful when linking

frontal features to precipitation or winds (or compound extreme events) as in Catto and Dowdy (2021), or when using more
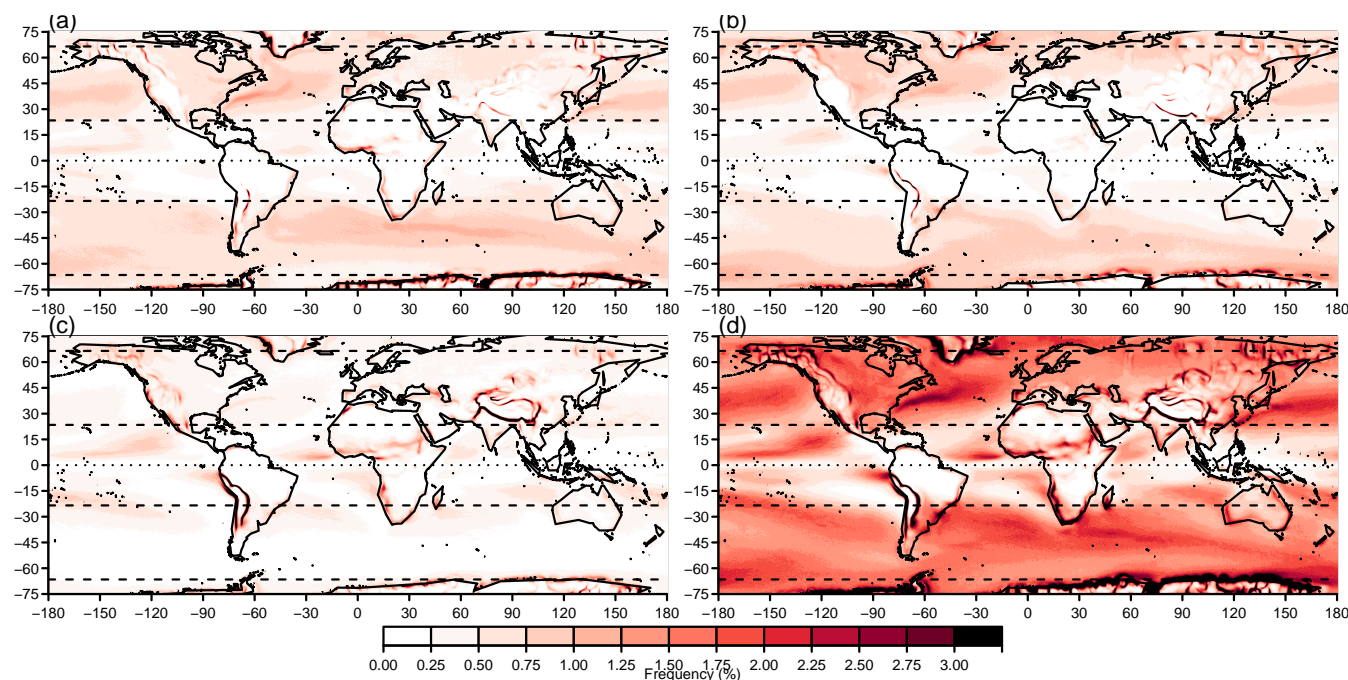
Figure 9. ERA5 global climatologies. (a) Cold fronts, (b) warm fronts, (c) quasi-stationary fronts, (d) all fronts. All climatologies were computed with $n = 96$ smoothing cycles, $K_1 = -1.6 \times 10^{-11}\,\mathrm{K\,m^{-2}}$ and $K_2 = 7.5 \times 10^{-6}\,\mathrm{K\,m^{-1}}$ and $K_3 = 1.5\,\mathrm{m\,s^{-1}}$.

object-based connections such as Papritz et al. (2014). For high resolution data such as ERA5, performance is improved since contouring is computationally cheaper than line joining.

Most automatic feature detection algorithms involve a calibration or training step involving comparison to analyses by a
270 human meteorologist. While this step should not be wholly neglected, we offer an objective quantile based method of setting thresholds for front identification. This quantile based method can also be used to compare analyses from different models at different resolutions while allowing feature identification to be conducted at the native resolution of each model. By considering the climatological distributions of the masking variables, we have demonstrated for the first time the regional and seasonal variation of the TFP and gradient fields. Subsequent analyses may wish to adopt latitudinally varying thresholds, although
275 seasonal variation is less strong outside of the polar regions.

In addition to the various numerical and methodological improvements presented in this study, further improvements or alternative choices are possible. As well as improving the accuracy of the finite difference approximations of the second derivative fields to second order, more accurate finite difference schemes could be used for both the first and second order derivatives. In testing, moving to fourth order accuracy for both first and second derivative fields did yield modest performance
280 increases in front identification for minimal computational expense. However, the implied accuracy is difficult to justify given the coarse resolution of the fields in question. Following Berry et al. (2011b), we identify fronts as zero contours in the field defined by Equation 5 of Hewson (1998), effectively the third derivative of the wet-bulb potential temperature field at 850 hPa. Firstly other meteorological fields than wet-bulb potential temperature could be considered, see Hewson (1998) for a list
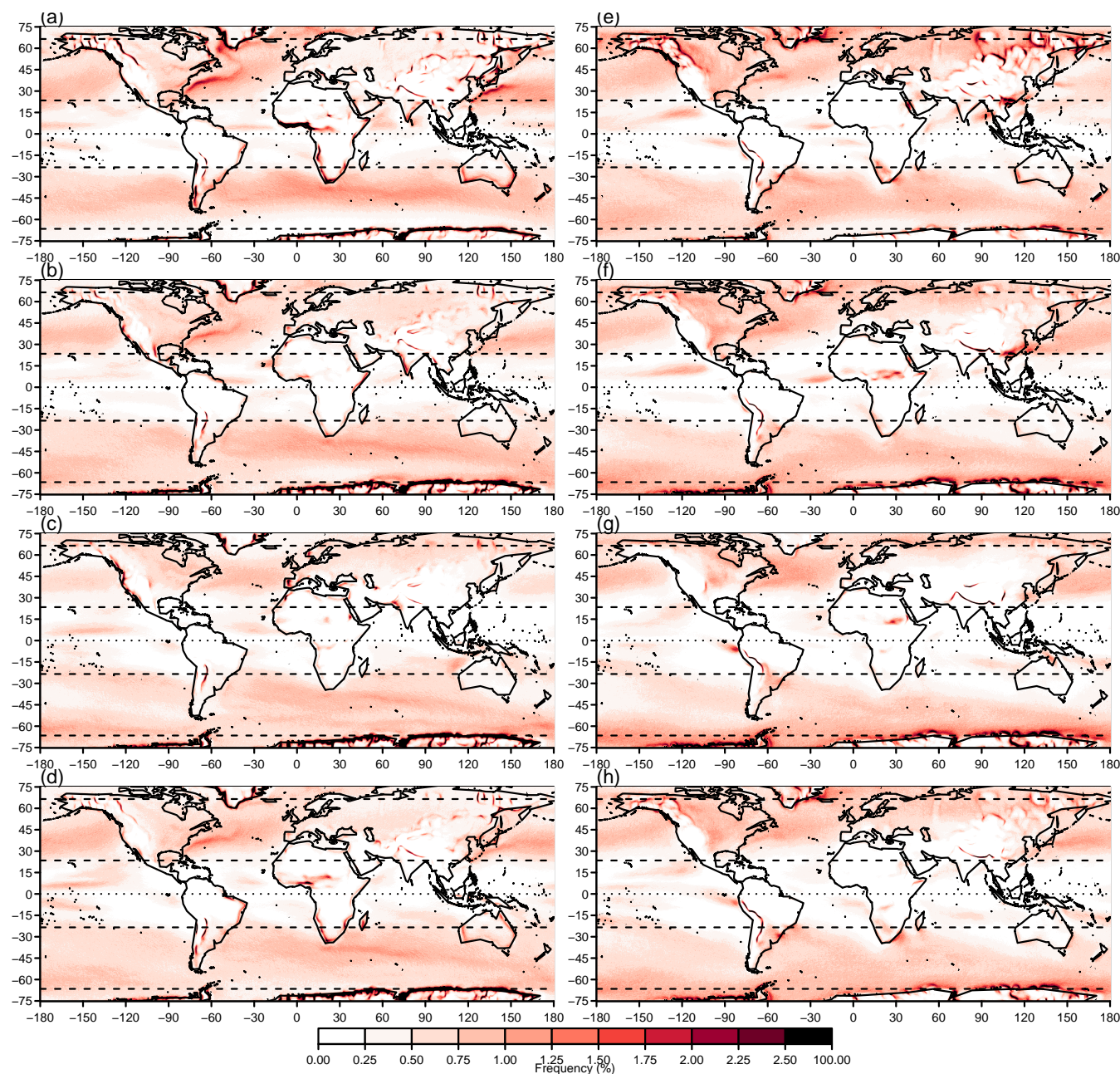
Figure 10. ERA5 seasonal climatologies of front frequency. (a) Cold fronts in DJF, (b) Cold fronts in MAM, (c) Cold fronts in JJA, (d) Cold fronts in SON, (e) Warm fronts in DJF, (f) Warm fronts in MAM, (g) Warm fronts in JJA, (h) Warm fronts in SON, All climatologies were computed with $n = 96$ smoothing cycles, $K_1 = -1.6 \times 10^{-11}\,\mathrm{K\,m^{-2}}$ and $K_2 = 7.5 \times 10^{-6}\,\mathrm{K\,m^{-1}}$ and $K_3 = 1.5\,\mathrm{m\,s^{-1}}$.

of previously considered fields. Secondly, Hewson (1998) identified three alternative expressions for the front locator field
(Equations 6, 7, & 8 of that paper), two of which they dismissed as poorly performing. The third, the so called "mean axis"

approach preferred by Hewson (1998) was considered for implementation here, but no performance benefit was found. As an addition to the methodology of Berry et al. (2011b) and Hewson (1998), the distinction between local and synoptic fronts suggested by Jenkner et al. (2010) could be implemented. While all distance calculations are carried out on a sphere in the updated implementation, contouring and interpolation still take place on a regular longitude-latitude grid. Greater accuracy could be achieved at high latitudes by also carrying out these operations on a sphere.

While cyclone identification algorithms routinely include the ability to track cyclonic features over subsequent time steps, similar feature tracking algorithms are almost absent for fronts. To the authors' knowledge, only Rüdisühli et al. (2020) claim a front tracking algorithm. Front tracking is inherently more complex than cyclone tracking since fronts are complex line objects whereas cyclones can be reduced to simple point objects or point objects with an associated area. An openly available front tracking algorithm would offer new possibilities in terms of attributing and analysing impacts of individual fronts, e.g., precipitation or wind events, or understanding biases in weather and climate models.

Code availability. Code for revised method detailed in this paper are available from https://doi.org/10.5281/zenodo.7278068 and future developments will be available at https://github.com/phil-sansom/front_id.

Author contributions. PGS developed and tested the software, produced the results, and wrote the paper. JLC led the project, interpreted results, and wrote the paper.

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

Acknowledgements. This research was supported by Natural Environment Research Council (NERC) grant NE/V004166/1. The authors thank Dr. Duncan Ackerley for comments on the manuscript.

References

Berry, G., Jakob, C., and Reeder, M.: Recent global trends in atmospheric fronts, Geophysical Research Letters, 38, 1–6, https://doi.org/10.1029/2011GL049481, 2011a.

Berry, G., Reeder, M. J., and Jakob, C.: A global climatology of atmospheric fronts, Geophysical Research Letters, 38, 1–5, https://doi.org/10.1029/2010GL046451, 2011b.

Bitsa, E., Flocas, H. A., Kouroutzoglou, J., Galanis, G., Hatzaki, M., Latsas, G., Rudeva, I., and Simmonds, I.: A Mediterranean cold front identification scheme combining wind and thermal criteria, International Journal of Climatology, 41, 6497–6510, https://doi.org/https://doi.org/10.1002/joc.7208, 2021.

Browning, K. A.: The sting at the end of the tail: Damaging winds associated with extratropical cyclones, Quarterly Journal of the Royal Meteorological Society, 130, 375–399, https://doi.org/10.1256/qj.02.143, 2004.

Catto, J. L. and Dowdy, A. J.: Understanding compound hazards from a weather system perspective, Weather and Climate Extremes, 32, https://doi.org/10.1016/j.wace.2021.100313, 2021.

Catto, J. L. and Pfahl, S.: The importance of fronts for extreme precipitation, Journal of Geophysical Research Atmospheres, 118, 10,791–10,801, https://doi.org/10.1002/jgrd.50852, 2013.

Catto, J. L., Jakob, C., Berry, G., and Nicholls, N.: Relating global precipitation to atmospheric fronts, Geophysical Research Letters, 39, 1–6, https://doi.org/10.1029/2012GL051736, 2012.

Catto, J. L., Nicholls, N., Jakob, C., and Shelton, K. L.: Atmospheric fronts in current and future climates, Geophysical Research Letters, 41, 7642–7650, https://doi.org/10.1002/2014GL061943, 2014.

Catto, J. L., Ackerley, D., Booth, J. F., Champion, A. J., Colle, B. A., Pfahl, S., Pinto, J. G., Quinting, J. F., and Seiler, C.: The Future of Midlatitude Cyclones, Current Climate Change Reports, 5, 407–420, https://doi.org/10.1007/s40641-019-00149-4, 2019.

Davies-Jones, R.: An efficient and accurate method for computing the wet-bulb temperature along pseudoadiabats, Monthly Weather Review, 136, 2764–2785, https://doi.org/10.1175/2007MWR2224.1, 2008.

Dee, D. P., Uppala, S. M., Simmons, A. J., Berrisford, P., Poli, P., Kobayashi, S., Andrae, U., Balmaseda, M. A., Balsamo, G., Bauer, P., Bechtold, P., Beljaars, A. C. M., van de Berg, L., Bidlot, J., Bormann, N., Delsol, C., Dragani, R., Fuentes, M., Geer, A. J., Haimberger, L., Healy, S. B., Hersbach, H., Hólm, E. V., Isaksen, L., Kållberg, P., Köhler, M., Matricardi, M., McNally, A. P., Monge-Sanz, B. M., Morcrette, J. J., Park, B. K., Peubey, C., de Rosnay, P., Tavolato, C., Thépaut, J. N., and Vitart, F.: The ERA-Interim reanalysis: Configuration and performance of the data assimilation system, Quarterly Journal of the Royal Meteorological Society, 137, 553–597, https://doi.org/10.1002/qj.828, 2011.

Dowdy, A. J. and Catto, J. L.: Extreme weather caused by concurrent cyclone, front and thunderstorm occurrences, Scientific Reports, 7, 1–8, https://doi.org/10.1038/srep40359, 2017.

ECMWF: Part IV: Physical processes, in: IFS Documentation CY47R3, ECMWF, https://doi.org/10.21957/eyrpir4vj, 2021.

Hersbach, H., Bell, B., Berrisford, P., Hirahara, S., Horányi, A., Muñoz-Sabater, J., Nicolas, J., Peubey, C., Radu, R., Schepers, D., Simmons, A., Soci, C., Abdalla, S., Abellan, X., Balsamo, G., Bechtold, P., Biavati, G., Bidlot, J., Bonavita, M., De Chiara, G., Dahlgren, P., Dee, D., Diamantakis, M., Dragani, R., Flemming, J., Forbes, R., Fuentes, M., Geer, A., Haimberger, L., Healy, S., Hogan, R. J., Hólm, E., Janisková, M., Keeley, S., Laloyaux, P., Lopez, P., Lupu, C., Radnoti, G., de Rosnay, P., Rozum, I., Vamborg, F., Villaume, S., and Thépaut, J. N.: The ERA5 global reanalysis, Quarterly Journal of the Royal Meteorological Society, 146, 1999–2049, https://doi.org/10.1002/qj.3803, 2020.

Hewson, T. D.: Objective fronts, Meteorological Applications, 5, 37–65, https://doi.org/10.1017/S1350482798000553, 1998.

Hope, P., Keay, K., Pook, M., Catto, J. L., Simmonds, I., Mills, G., Mcintosh, P., Risbey, J., and Berry, G.: A comparison of automated methods of front recognition for climate studies: A case study in southwest Western Australia, Monthly Weather Review, 142, 343–363, https://doi.org/10.1175/MWR-D-12-00252.1, 2014.

345   Jenkner, J., Sprenger, M., Schwenk, I., Schwierz, C., Dierer, S., and Leuenberger, D.: Detection and climatology of fronts in a high-resolution model reanalysis over the Alps, Meteorological Applications, 17, 1–18, https://doi.org/10.1002/met.142, 2010.

Leung, L. R., Boos, W. R., Catto, J. L., A. DeMott, C., Martin, G. M., Neelin, J. D., O'Brien, T. A., Xie, S., Feng, Z., Klingaman, N. P., Kuo, Y., Lee, R. W., Martinez-Villalobos, C., Vishnu, S., Priestley, M. D. K., Tao, C., and Zhou, Y.: Exploratory Precipitation Metrics: Spatiotemporal Characteristics, Process-Oriented, and Phenomena-Based, Journal of Climate, 35, 3659–3686, https://doi.org/10.1175/JCLI-D-21-0590.1, 2022.

350

Papritz, L., Pfahl, S., Rudeva, I., Simmonds, I., Sodemann, H., and Wernli, H.: The Role of Extratropical Cyclones and Fronts for Southern Ocean Freshwater Fluxes, Journal of Climate, 27, 6205–6224, https://doi.org/10.1175/JCLI-D-13-00409.1, 2014.

Parfitt, R., Czaja, A., Minobe, S., and Kuwano-Yoshida, A.: The atmospheric frontal response to SST perturbations in the Gulf Stream region, Geophysical Research Letters, 43, 2299–2306, https://doi.org/https://doi.org/10.1002/2016GL067723, 2016.

355   Parfitt, R., Czaja, A., and Kwon, Y.-O.: The impact of SST resolution change in the ERA-Interim reanalysis on wintertime Gulf Stream frontal air-sea interaction, Geophysical Research Letters, 44, 3246–3254, https://doi.org/https://doi.org/10.1002/2017GL073028, 2017a.

Parfitt, R., Czaja, A., and Seo, H.: A simple diagnostic for the detection of atmospheric fronts, Geophysical Research Letters, 44, 4351–4358, https://doi.org/10.1002/2017GL073662, 2017b.

R Core Team: R: A Language and Environment for Statistical Computing, https://www.r-project.org, 2021.

360   Raveh-Rubin, S. and Catto, J. L.: Climatology and dynamics of the link between dry intrusions and cold fronts during winter, Part II: Front-centred perspective, Climate Dynamics, 53, 1893–1909, https://doi.org/10.1007/s00382-019-04793-2, 2019.

Renard, R. J. and Clarke, L. C.: Experiments in Numerical Objective Frontal Analysis 1, Monthly Weather Review, 93, 547–556, https://doi.org/10.1175/1520-0493(1965)093<0547:einofa>2.3.co;2, 1965.

Rüdisühli, S., Sprenger, M., Leutwyler, D., Schär, C., and Wernli, H.: Attribution of precipitation to cyclones and fronts over Europe in a
365   kilometer-scale regional climate simulation, Weather and Climate Dynamics, 1, 675–699, https://doi.org/10.5194/wcd-1-675-2020, 2020.

Schemm, S., Rudeva, I., and Simmonds, I.: Extratropical fronts in the lower troposphere-global perspectives obtained from two automated methods, Quarterly Journal of the Royal Meteorological Society, 141, 1686–1698, https://doi.org/10.1002/qj.2471, 2015.

Schemm, S., Sprenger, M., Martius, O., Wernli, H., and Zimmer, M.: Increase in the number of extremely strong fronts over Europe? A study based on ERA-Interim reanalysis (1979–2014), Geophysical Research Letters, 44, 553–561, https://doi.org/10.1002/2016GL071451,
370   2017.

Simmonds, I., Keay, K., and Bye, J. A. T.: Identification and climatology of Southern Hemisphere mobile fronts in a modern reanalysis, Journal of Climate, 25, 1945–1962, https://doi.org/10.1175/JCLI-D-11-00100.1, 2012.

Soster, F. and Parfitt, R.: On Objective Identification of Atmospheric Fronts and Frontal Precipitation in Reanalysis Datasets, Journal of Climate, 35, 4513–4534, https://doi.org/10.1175/JCLI-D-21-0596.1, 2022.

375   Thomas, C. and Schultz, D.: What are the Best Thermodynamic Quantity and Function to Define a Front in Gridded Model Output?, Bulletin of the American Meteorological Society, 100, https://doi.org/10.1175/BAMS-D-18-0137.1, 2019.