Supplement of Geosci. Model Dev., 19, 933–953, 2026
https://doi.org/10.5194/gmd-19-933-2026-supplement
© Author(s) 2026. CC BY 4.0 License.

Geoscientific Model Development
Open Access
EGU

Supplement of

Identifying sea breezes from atmospheric model output (sea_breeze v1.1)

Andrew Brown et al.

Correspondence to: Andrew Brown (a.brown1@unimelb.edu.au)

The copyright of individual parts of the supplement might differ from the article licence.

Supplementary Material for: "Identifying sea breezes from atmospheric model output (sea_breeze v1.1)"

Contents:

S1. Method for calculating the angle of coastline orientation

Here, a method to calculate coastline orientations from a model land-sea mask is presented. This is used to calculate onshore and offshore wind directions at all grid points from each model dataset (ERA5, BARRA-R, BARRA-C and AUS2200, see Section 2 of the main text). Sea breeze objects are filtered based on their orientation relative to the orientation of the coast, as well as their onshore wind speed. The process is summarised as follows:

- From a binary land-sea mask, define coastline points as the boundary between land and sea.
- For each grid point, calculate the distance and angle between the point and the set of all coastline points.
- For each grid point, calculate the average coastline angle using all coastline points, weighted by the inverse distance.
- Use the average angles for the orientation and onshore wind filters described in the main text (Section 3.3).

These steps are now described in more detail in the following paragraphs.

Firstly, the model coastline is defined as the boundary between land and sea points using the scikit-image Python package (Van Der Walt et al., 2014). If the land-sea fraction is provided rather than a land-sea mask, then a mask is created using a threshold fraction of 0.5. For ERA5, inland lakes are counted as sea points, and we convert these to land points to focus on coastlines facing open ocean.

Then, the distance and angle between each pair of grid points and coastline points are calculated based on standard great-circle computations (Snow et al., 2023). At each grid point, the set of angles and distances to each coastline point is transformed into the complex number space, and an inverse distance weighted average is taken to define the average direction towards the coast, using a weighting function, $W(x)$.

$$W(x) = \begin{cases} 0.75 - 0.50(\frac{x}{R_1})^2 & x \leq R_1 \\ 0.25(\frac{x}{R_1})^{-4} & R_1 < x < R_2 \\ 0 & x \geq R_2 \end{cases} \tag{S1}$$

where x is distance from the coast, in km. The weighting of the coastline angles decreases as a quadratic function, until a manually defined distance of R_1. After the distance of R_1 is reached, the weights decrease as a function of $1/x^4$, until a distance of R_2 where the weights are set to zero. The weighting functions for each of the model datasets are shown in Figure S1, with different choices of R_1 for each model to be discussed shortly.

The piecewise weighting function ensures that the influence of individual coastline points decreases relatively slowly with distance near the coast, and relatively fast offshore, with the transition between these two regimes controlled by the value of R_1. The value of R_1 therefore influences the smoothness of the resulting weighted average angles near the coast for a given land-sea mask, with higher values increasing smoothness. Given that the sea breeze is a mesoscale phenomenon, it is assumed that coastline features of O(10 km) or larger are of importance. Based on this, and manual inspection of average coastline angles, an R_1 value of $2 \times dx$ is chosen, where dx is the model grid spacing (4 km for AUS2200, 8 km for BARRA-C, 24 km for BARRA-R, and 50 km for ERA5). A large value of R_2 =10,000 km is chosen. This upper limit is much larger than the distance

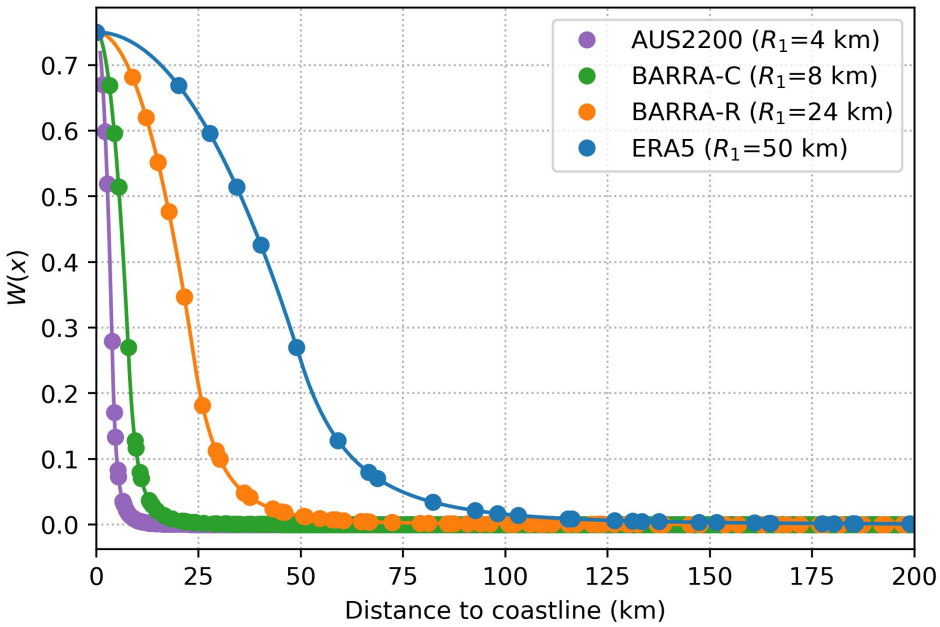

Figure S1: Weighting functions (Eqn. S1) for averaging the angles between a model grid point and coastline points. Shown separately for different R_1 values chosen for each model dataset. For each weighting function, the distance between a representative model grid point at the coast and other coastline points are indicated by circle markers.

at which a coastline might influence an individual point, but is included to avoid overflow errors during the computation of $W(x)$.

After the weighted average coastline angles are calculated for each grid point, three final steps are required. Firstly, because average coastline directions are opposing by 180° on either side of the coast (that is, the angles point in the offshore direction over the land and the onshore direction over the ocean), angles over the land are reversed in direction for continuity across the coastline. The average angles are also converted from a coastline direction to a coastline orientation, by rotating them 90° clockwise, giving the average coastline orientation associated with each point (θ). Finally, because average angles are undefined at coastline points based on a distance of zero, a linear interpolation is used to define coastline orientation at those points.

In addition to the average coastline orientation associated with each point (θ), the associated weighted variance in coastline angles can also be calculated. This can be used to evaluate uncertainty in coastline orientation, relevant for points where there are influences from multiple coastlines of different orientations. The weighted coastline variance uses the same weighting function and method as described above for calculating the weighted average. The amount of variance is somewhat dependent on the value of R_1 chosen, with larger values of R_1 for a given land-sea mask resulting in higher variance near the coast. Example maps of θ and the weighted coastline variance can be seen in Section 3.1 of the main text.

S2. Model evaluation using weather station data and the "hourly rate of change" (H) diagnostic

S2.1 Evaluating the spatial distribution of H over the entire 6-month period

Weather station observations are used to evaluate the ability of each model dataset to represent hourly local changes in temperature, moisture, and onshore wind speed, as summarized by the hourly rate of change (H) diagnostic. Observed H is computed using data from the network of Automatic Weather Stations (AWS) managed by the Australian Bureau of Meteorology, over the 6-month study period, with the process for calculating H described in Section 3.2.3 of the main text. The H diagnostic is calculated using hourly data from 712 stations around Australia, and exceedances of the 99.5^{th} percentile are used to define sea breeze occurrences. To define an onshore direction for the computation of H from AWS data, coastline angles based on the AUS2200 land-sea mask are used, given that it is the highest resolution out of the model land-sea masks considered here. It is not expected that this will have significant impacts on the results, given that the rate of change in onshore wind is of interest, rather than the magnitude or direction of the onshore wind.

Because H is calculated from time series data, it is the only diagnostic that can be compared directly with observations, noting that there are a lack of observational datasets with horizontal and vertical coverage for computing F and the SBI (see Section 3.2 of the main text for a description of these diagnostics). We are unable to filter the observed H diagnostic for sea breeze objects, as this requires gridded spatial data, so the analysis here serves as an evaluation of local changes in general, rather than sea breezes specifically. However, we will still refer to instances of high H values as "sea breeze occurrences", for simplicity and consistency with other sections.

Figure S2 shows the number of days with a sea breeze occurrence as calculated from AWS observations, based on the H diagnostic. For comparison with the observed H diagnostic, model sea breeze occurrences are counted using exceedances of the 99.5^{th} percentile of H at the closest pixels to the 712 station locations, as shown in Figure S2. Figure S3 shows the relationship between the observed and modeled number of (unfiltered) sea breeze days at each AWS location, with the Spearman correlation coefficient (r) indicated above each figure panel. Figure S3 shows that there is a high spatial correlation between each model and the observations ($r > 0.8$), suggesting that each model dataset can represent the spatial distribution of hourly local changes in temperature, moisture, and onshore wind to some extent, with some of these changes likely related to sea breeze occurrences. There is some variability in the strength of the correlation coefficient between models, with increasing r values with increasing resolution.

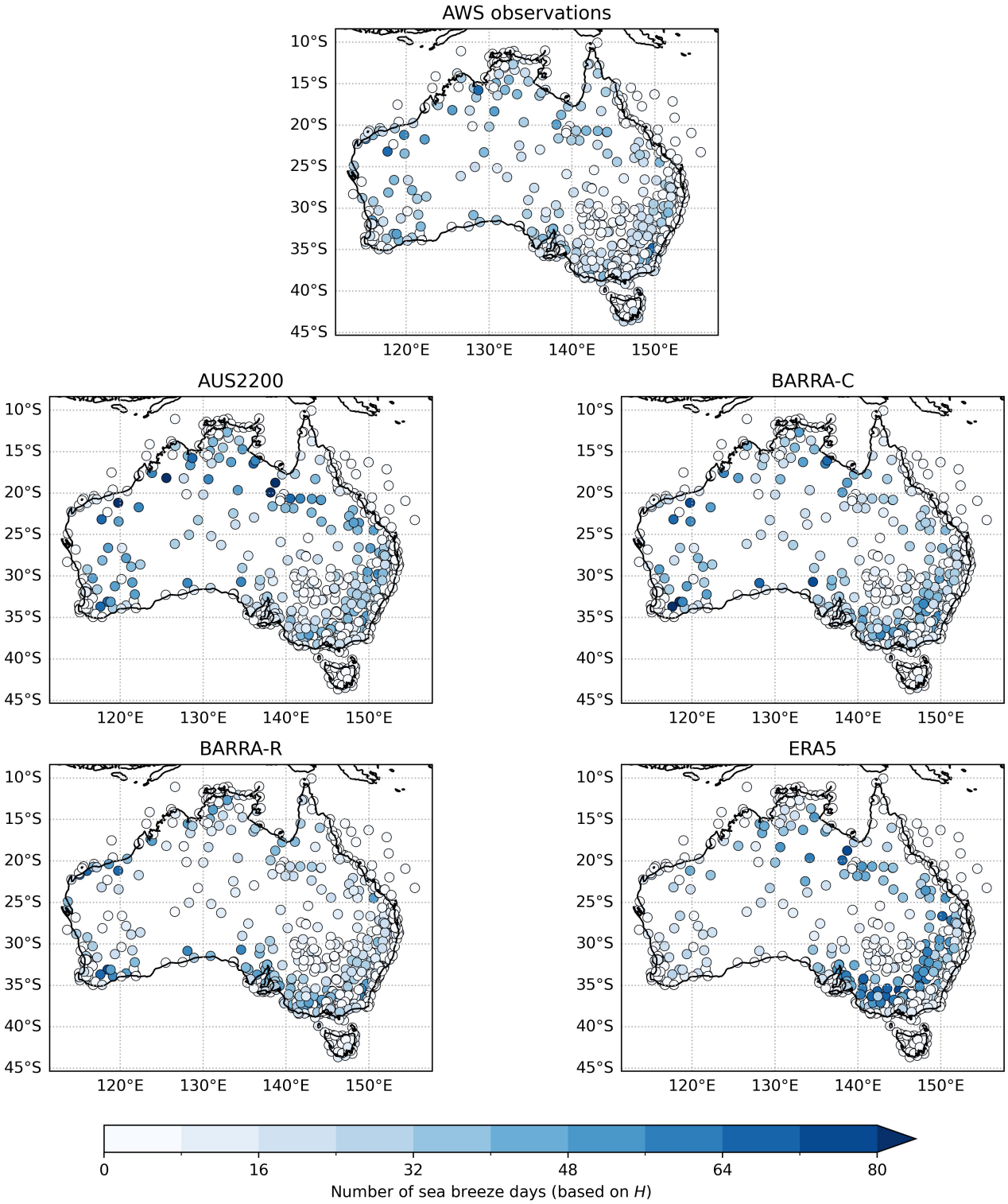

Figure S2: The number of days with a sea breeze occurrence over the 6-month study period, for all AWS locations, defined using exceedances of the 99.5^{th} percentile of H. Shown for AWS observations, as well as each model dataset at the closest model data point to each station location.

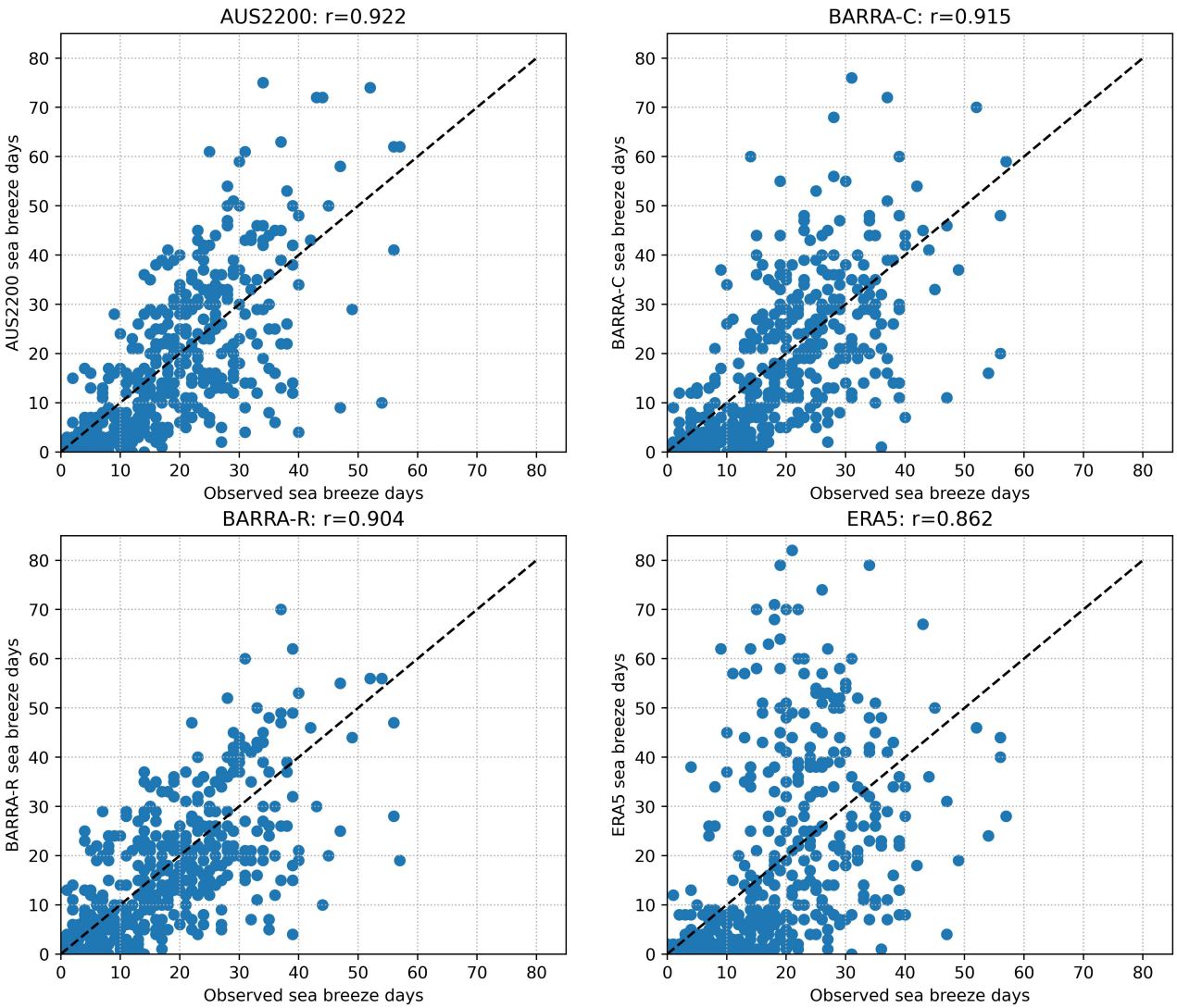

Figure S3: The relationship between observed and modeled sea breeze occurrences, defined using exceedances of the 99.5^{th} percentile of H, across all AWS locations. Shown for each different model dataset. The spatial (Spearman's ranked) correlation coefficient between the modeled and observed occurrences is shown at the top of each panel.

S2.2 Evaluating occurrences of high values of H for individual cases

Occurrences of high H values for each of the example cases in the main text (see Section 4) are compared spatially between observations and each model dataset (Figure S4). As in the previous section, high values of H are defined by exceedances of the 99.5^{th} percentile over the 6-month period. For each case, daily maximum H values are considered, to account for small differences in the timing and/or location of local changes in wind, moisture and temperature between the models and observations.

Figure S4 indicates that the spatial patterns of high H occurrences are broadly similar between the models and the observations, for most cases. This includes the number of stations with high H values as well as their location. Exceptions to this are the first of the two Darwin cases, where there is only one location observed with high H, compared with 4–8 between each of the models. Also, for the first of the two Southwest Australia cases, only AUS2200 appears to capture the correct distribution of high H occurrences.

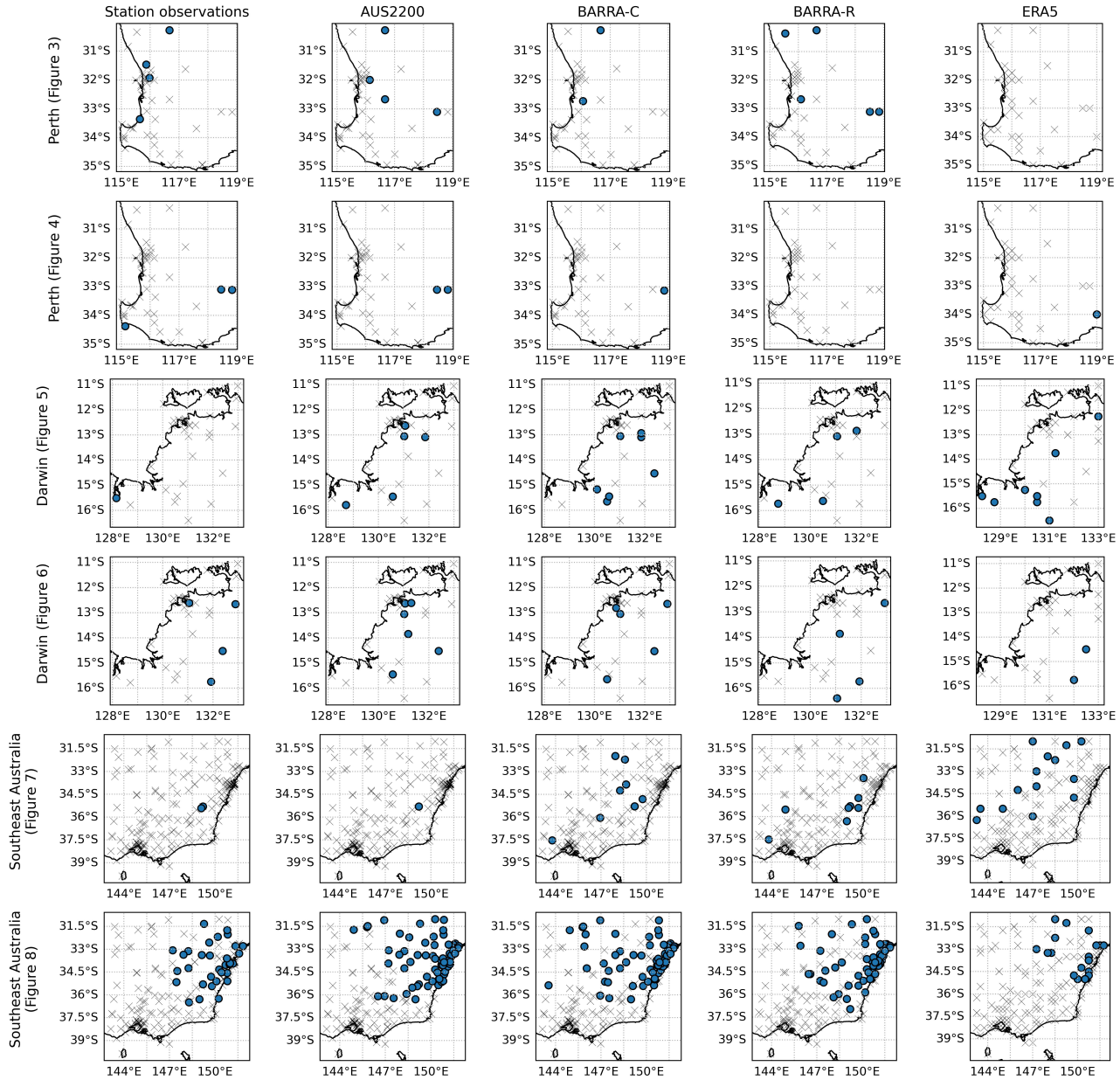

Figure S4: Maps of occurrences of high values of H, defined by exceedances of the 99.5th percentile over the 6-month period, for (from left to right columns) station observations, AUS2200, BARRA-C, BARRA-R, and ERA5, for six case study days analysed in the main text (with the corresponding figure referenced here alongside each row). For each case, the daily maximum H value at each station location is considered. Stations where the 99.5th percentile of H is exceeded are shown with blue dots, with other stations shown with black crosses.

S3. Sensitivity tests

S3.1 Hourly rate of change diagnostic (H) variables

The sensitivity of sea breeze objects to the formulation of the hourly rate of change (H) diagnostic is investigated here (see Section 3.2.3 of the main text for details on the formulation of H). This is done by performing three tests, each time removing one of the rate of change variables (onshore wind speed, temperature, and moisture) from the diagnostic, and calculating sea breeze objects over the 6-month period from AUS2200 (following the procedure described in Section 3.3 of the main text). We then analyse the diurnal cycle of the resulting objects as a function of distance from the coast, consistent with the analysis in Section 5.2 of the main text. This analysis was chosen based on summarizing the spatial and temporal occurrences of sea breeze objects.

Figure S5 shows that the overall timing of the sea objects using H has very little dependence on the variables chosen. However, including temperature and/or moisture results in more objects, as evidenced by less objects when removing those variables. Removing moisture from the diagnostic appears to result in less realistic inland propagation, with objects instead occurring at a wide range of distances at around 18:00 local time.

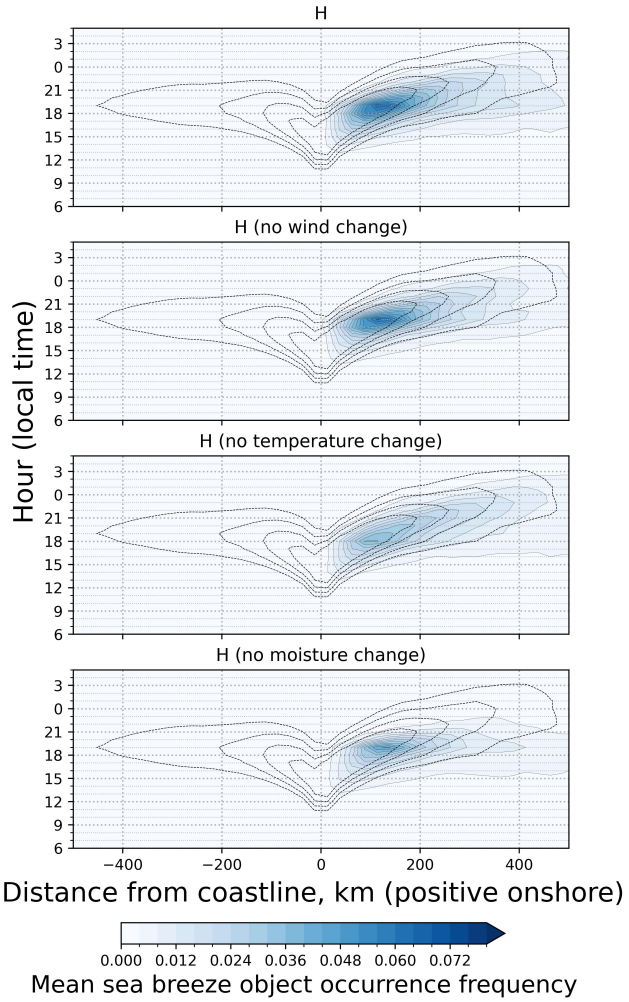

Figure S5: The mean occurrence frequency of sea breeze objects at each hour of the day in local time (vertical axis) and at different distances from the coast (horizontal axis), for objects derived from the hourly rate of change (H) diagnostic from AUS2200, as well as for three different sensitivity tests that change the formulation of H. These tests include removing terms for the rate of change in onshore wind speed ("no wind change"), temperature ("no temperature change") and moisture ("no moisture change"). Occurrence frequencies are averaged over bins spaced at 25 km intervals from the coast, with positive values representing onshore locations. The mean onshore wind (u') perturbation from the daily mean is shown with a dashed contour, for comparison, with contours of 0.25, 0.50, 0.75, and 1.0 m/s.

S3.2 Filtering settings

A range of sensitivity tests are performed here to evaluate the impact of filtering settings and the smoothing of input fields on sea breeze identification, using the F diagnostic applied to the AUS2200 model (see Sections 2.3 and 3.2.2 of the main text for a description of this model dataset and diagnostic). It is expected that these tests will generalize to other diagnostics and methods. 12 different sensitivity tests are performed as summarized and named in Table S1. These include adjusting the amount of smoothing of input fields (Smooth 2 and Smooth 6), the coastline orientation filter (Orientation 22.5 and Orientation Off), the aspect ratio filter (Aspect 4 and Aspect Off), the area filter (Area 6 and Area 24), and the percentile for creating a binary sea breeze mask from the F diagnostic (Percentile 99.0 and Percentile 99.9). We also tested turning off the land-sea temperature filter (Land-sea Off) and onshore wind speed filter (Onshore Off). The default conditions of each filter can be seen in Table 2 of the main text, along with discussions of smoothing and percentile thresholds in Section 3.3.

Figure S6 shows the impact of each filtering test on the diurnal cycle of sea breeze object occurrences with distance from the coast. This can be compared with the equivalent figure using the default filtering settings in Figure 13a of the main text. Again, this analysis was chosen based on summarizing the spatial and temporal occurrences of sea breeze objects. For each test, all other settings remain as default, as described in Table 2 of the main text. Figure S6 demonstrates that the diurnal and spatial structure of object occurrences are similar between each sensitivity test, including the inland propagation and timing of object occurrences. The main difference between each test is the number of objects produced. For example, more objects are produced by turning off the orientation filter (Orientation Off), and less objects are produced by increasing the percentile used for the threshold (Percentile 99.9). Turning off the onshore wind speed filter, however (Onshore Off), results in the identification of objects on the coastline for a longer portion of the day. This may reflect misclassifications due to coastline convergence, and therefore this filter should be left on. Similarly, lowering the percentile threshold to 99.0 results in a greater number of coastline objects throughout the day, and is likely therefore too low of a threshold.

Table S1: Sensitivity tests for sea breeze identification methods, applied to the F diagnostic from AUS2200.

Sensitivity test	Name
Lower standard deviation parameter for Gaussian smoothing of input fields ($\sigma = 2$)	Smooth 2
Higher standard deviation parameter for Gaussian smoothing of input fields ($\sigma = 6$)	Smooth 6
Stricter tolerance for orientation (relative to coastline) filter, of 22.5°	Orientation 22.5
No orientation (relative to coastline) filter	Orientation Off
A stricter aspect ratio filter of 4:1	Aspect 4
No aspect ratio filter	Aspect Off
Smaller area filter of 6 pixels	Area 6
Larger area filter of 24 pixels	Area 24
Lower threshold for creating binary mask from diagnostic (99.0^{th} percentile)	Percentile 99.0
Higher threshold for creating binary mask from diagnostic (99.9^{th} percentile)	Percentile 99.9
No land-sea temperature difference filter	Land-sea Off
No onshore wind speed filter	Onshore Off

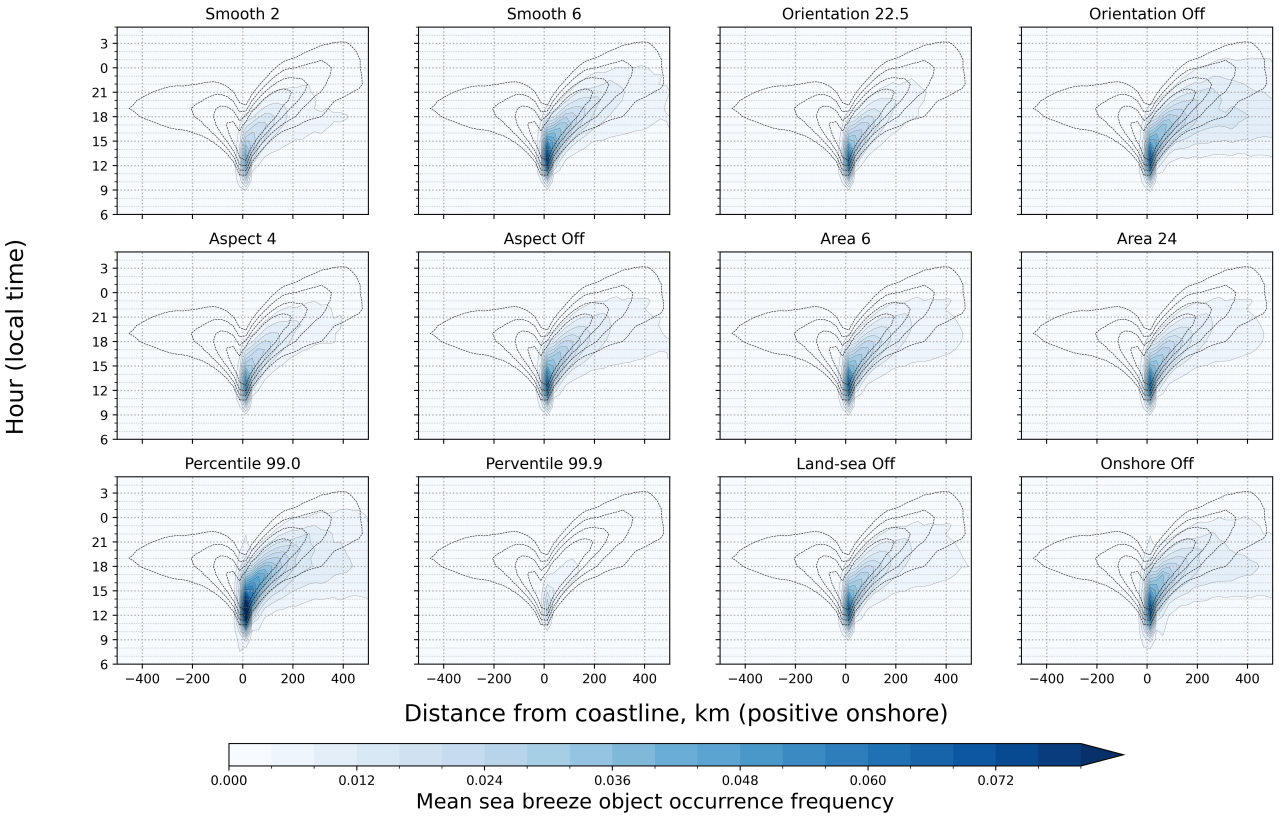

Figure S6: The mean occurrence frequency of sea breeze objects at each hour of the day in local time (vertical axis) and at different distances from the coast (horizontal axis) for a range of sensitivity tests. The name of each sensitivity test is shown above each figure panel, with explanations in Table S1. In each panel, sea breeze objects are identified using the F diagnostic applied to AUS2200 data. Occurrence frequencies are averaged over bins spaced at 25 km intervals from the coast, with positive values representing onshore locations. The mean onshore wind (u') perturbation from the daily mean is shown with a dashed contour, for comparison, with contours of 0.25, 0.50, 0.75, and 1.0 m/s.

References

Snow, A. D., Whitaker, J., Cochran, M., Miara, I., den Bossche, J. V., Mayo, C., Cochrane, P., Lucas, G., de Kloe, J., Karney, C., Filipe, Couwenberg, B., Lostis, G., Dearing, J., Ouzounoudis, G., Jurd, B., Gohlke, C., McDonald, D., Hoese, D., Itkin, M., May, R., Little, B., Heitor, Shadchin, A., Wiedemann, B. M., Barker, C., Willoughby, C., Schneck, C., and DWesl: pyproj4/pyproj: 3.6.1 Release, https://doi.org/10.5281/zenodo.8365173, 2023.

Van Der Walt, S., Schönberger, J. L., Nunez-Iglesias, J., Boulogne, F., Warner, J. D., Yager, N., Gouillart, E., and Yu, T.: Scikit-image: Image processing in python, PeerJ, 2014, https://doi.org/10.7717/peerj.453, 2014.