

Supporting Information for Swell generation under extra-tropical storms

Momme. C. Hell ¹,

¹University of California, San Diego, Scripps Institution of Oceanography; 9500 Gilman Drive, La Jolla, California 92093

Contents of this file

1. Text T1: Physically informed Optimization of a common Focal Point
2. Figure F1: Comparison of the brute force and Global Optimization
3. Figure F2: Illustration of the Along track averaging.
4. Figure F3: Least square fit to estimate translation velocity
5. Figure F4: Case Study 2 triangulation
6. Figure F5: Case Study 2 characteristic curves
7. Figure F6: Case Study 3 triangulation
8. Figure F7: Simplified time series for one year of fitted initial times uncertainty

estimates

Introduction

Corresponding author: M. C. Hell, University of California, San Diego, Scripps Institution of Oceanography; 9500 Gilman Drive, La Jolla, California 92093 (mhell@ucsd.edu)

June 3, 2021, 6:00pm

This supplementary document contains additional material for *a simple model of swell generation under extra-tropical storms*.

Text T1: Physically informed Optimization of a common Focal Point

T1.1. Wave buoy observations and initial event identification

Each of the chosen wave buoys (CDIP 166, CDIP 179, CDIP 029, CDIP 067, CDIP 106) samples the directional wave spectrum in 30-minute averages. The wave buoy spectrograms and their directional information are retrieved from the CDIP datawell (Behrens et al., 2019).

Local swell maxima are identified in the spectrograms by averaging over the first three frequency bins whose spectral amplitude exceeds a noise threshold of $e^{-1} \text{ m}^2 \text{ Hz}^{-1}$. This results in a time series of the amplitude of the longest swell waves that is sensitive to the amplitude and frequency slope of the swell and. This time series is band-pass filtered for timescales between 18 hours and 7.5 days using a Lanczos filter to retain variability that is mainly related to atmospheric synoptic scales. Examples of the identified swell maxima are shown in main-text figure 1b to e (black dots).

T1.2. Data pre-handling

First, we apply an adaptive directional filter on the wave buoy observations to filter out local wind waves and focus on dispersed swell. The incident directions of the swell forerunners are used to weight the observed spectrograms. The directional component of the wave spectrum $D_\theta(\theta, f, t)$ is used to create a weight for the omni-directional spectral amplitude $D_{amp}(f, t)$ (here, f is the wave frequency and t is time). Frequency bands with wave energy in the same direction as the swell forerunners have a weight of one,

while frequency bands that contain energy from a different direction have a weight close to zero. This selects wave energy in a $\pm 15^\circ$ -angle around the peak direction of the swell forerunners and filters out secondary swell systems or locally generated higher frequency waves if they come from a different direction.

In a second step, the pre-identified wave events are isolated for the optimization procedure. The initial dispersion slope of each swell event is estimated by the difference between the prior identified local maxima (main-text figure 1b to e, black dots) and a local maxima on a frequency band that is 0.01Hz higher compared to the prior identified local maximum. This slope between the two local maxima on different frequency bands is used to select and initialize each wave event following (Hell et al., 2019).

T1.3. Initial model fit

The pre-identified single wave events are then used to fit a model of swell arrival to each case individual. Based on the algorithm in Hell et al. (2019), the two-dimensional model function for the individual events \mathbf{M}^k is defined as

$$\mathbf{M}^k(\tilde{t}, f) = A(m_t, \mu) K(\tilde{t}, \tilde{t}_0, m_t, \sigma_K) S(f, U, f_m, \gamma_{par}), \quad (\text{T1.1})$$

where A describes the amplitude attenuation, $\tilde{K}(\tilde{t}) = K(\tilde{t})/\max(K)$ describes the peak-normalized and time-normalized time component, and $S(f)$ the frequency dependent power spectra. The amplitude of \mathbf{M}^k is defined by the initial spectral power of S and the attenuation A .

The power spectra $S(f)$ is modelled by the standard JONSWAP spectrum

$$\begin{aligned} S(f) &= \frac{\alpha_S g^2}{(2\pi f)^5} \exp \left[-\frac{5}{4} \left(\frac{f}{f_m} \right)^{-4} \right] \gamma^\delta, \\ \alpha_S &= 0.076 \left(\frac{f_m U}{3.5g} \right)^{2/3}, \\ \delta &= \exp \left[-\frac{1}{2} \left(\frac{f - f_m}{\sigma_S f_m} \right)^2 \right], \end{aligned} \quad (\text{T1.2})$$

with f_m as the peak frequency of the spectra, $\sigma_S = 0.07$ for $f \leq f_m$ and $\sigma_S = 0.09$ for $f > f_m$, and γ_{par} as the amplitude of the peak-enhancement factor (Hasselmann et al., 1973).

In time, the model is defined as a form of a χ^2 - or Erlang distribution such that

$$\begin{aligned} K(\tilde{t}, \tilde{t}_m) &= \frac{\tilde{t}_m}{\sigma_K} e^{-\tilde{t}_m}, \\ \tilde{t}_m &= \frac{\tilde{t} - \tilde{t}_0 + f m_t}{\sigma_K} \end{aligned} \quad (\text{T1.3})$$

where \tilde{t} is the non-dimensional time, the relative time of the selected data divided by its time span Δt , \tilde{t}_0 is the non-dimensional initial time, m_t is the slope of the peak frequency in the spectrogram in units of Hz^{-1} , and σ_K a parameter that controls the width of the distribution (Hell et al., 2019).

The swell's attenuation A along the travel path is modeled with a simple exponential decay that does not depend on direction or frequency (Ardhuin et al., 2009). That means the decay only depends on the distance traveled along a great circle path r_0 such that it can be directly related to the spectral slope m_t (Munk, 1947). The attenuation model is defined as

$$A(m_t, \mu) = \exp \left(-\mu \frac{g m_t \Delta t}{4\pi} \right), \quad (\text{T1.4})$$

where $\mu \approx 3.7 \pm 0.210^{-7} \text{ m}^{-1}$ (Jiang et al., 2016). This simple attenuation model allows the spectral power at the origin to be estimated from the observed swell spectrogram alone, assuming that distortions by other processes are small.

The to-be-optimized parameters for each event k are summarized as

$$\mathbf{p}^k = \{m_t, \mu, \tilde{t}_0, U, f_m, \gamma_{par}, \sigma_t\}^T. \quad (\text{T1.5})$$

They are optimized to find the best fit of the model $\mathbf{M}^k(\mathbf{p}^k, \tilde{t}, f)$ to the data \mathbf{D}^k by minimizing the the cost function

$$J^k = \|(\mathbf{D}^k - \mathbf{M}^k)\mathbf{w}^k\|^2 + \left\| \frac{\mathbf{p}_0 - \mathbf{p}}{\mathbf{p}_\sigma} \right\|^2, \quad (\text{T1.6})$$

for a wave event k individually (adapted from (Hell et al., 2019)). The initial guess of the parameters \mathbf{p}_0 was derived from the data, and the priors of the model parameters \mathbf{p}_σ are taken from (Hell et al., 2019). The data weighting function \mathbf{w}^k describes 2D-Gaussian weight around the center of the event such that noise at the corner of the data is excluded (dark shading in main-text figure 7 b to f, also F4 and F6, Hell et al., 2019, sec. 6.d).

The cost function J^k is optimized with three-stage optimization procedure. An initial semi-random ‘basinhopping’ search finds the minimal cost varying only m_f and \tilde{t}_0 to determine the best model slope that goes through the pre-identified forerunner point (Wales & Doye, 1997). In a second step, the cost function is minimized by varying all parameters using the Levenberg-Marquardt Algorithm (LM, damped least-squares, Newville et al., 2014) and finally, a *posteriori* error distribution is derived with a Parallel Tempering Markov-Chain-Monte-Carlo (PTMCMC, Goodman & Weare, 2010; Foreman-Mackey et al., 2013; Earl & Deem, 2005).

This procedure is applied to all pre-identified swell events at five stations between the year 2014 to 2018 resulted in about 77 successfully fitted wave events per station per year. After quality control, only about 56% of these cases can be bundled to sets of observed events with a common source (see next section).

T1.4. Identifying and optimizing common swell source

To derive a common source location we combine the identified wave events from the previous step from the five wave buoys. The initial fitting acts here as a quality control, such that we only use events that provide a reasonable radial distance (> 200 km), a small fractional error ($-w_{err} + 1 < 0.6$, eq. T1.9), and have a $\sigma_K < 0.2$ to sort out short local events.

The matching of events between the five wave buoy stations are done using the fitted initial time and their uncertainty estimates (Figure F7 for the year 2016). Blue lines are two-standard deviation uncertainty ranges around estimated initial times that pass a quality criterion of good model fit (Hell et al., 2019), while light green lines show the initial time uncertainties that do not pass this criterion. Red blocks indicate time ranges where two or more initial time estimates overlap. These events are used to triangulate the source locations in the north pacific (longitude and latitude) from the radial distance estimates of the identified overlapping subset. Figure F7 illustrates that by far not all initial time estimates are good enough and the not all initial time estimates coincide. To account for this, the triangulated location and time from the identified subset of wave buoys are used to re-select data from the not identified wave buoys by forward propagating the model M and estimating the slope and model shape at the buoy location. The now selected data in the additional wave buoys is then again fed in to the parameter estimation described

in sec. T1.3. This results in a data array from each wave buoy and these five data arrays likely contain observations from the same swell event.

Using this procedure, only about 7.5 events per year are well observed at 2 or more wave buoys, while about 50-70 event per year are identified in each wave buoy. That low matching rate by the initial time only is due to a) an insufficient initial detection algorithm based on the forerunners of swell (sec. T1.1), b) noise by local wind swell at buoy locations, c) deflection of waves by currents, and finally d) the fact that not all wave events propagate across the north pacific such that they are detected by multiple wave buoys.

T1.5. Multiple-stations cost function

The identification of a common swell source by their initial time t_0 described in the previous section results in 31 sets of swell events that had a common t_0 . Many other events are distorted by noise at the station or the wave ray refraction on their path through the ocean (Gallet & Young, 2014; Villas Bôas & Young, 2020).

The sets of swell event observations were then used to reassess the model parameters by adding the constraint of a common source. The optimization problem was reformulated in terms of parameters describing a common swell event from a single location

$$\mathbf{p}_m = \{\lambda, \phi, t\}, \quad (\text{T1.7})$$

with the longitude λ , latitude ϕ , and time t define the source location. The slope parameter m_t and attenuation μ at each station k were calculated based on the common source position (Munk, 1947; Barber & Ursell, 1948). Other parameters of the model \mathbf{M}^k were set to the five-station mean of the individual fitted parameters and do not vary during the

multi-station optimization. This reduces the search space of the optimization procedure and allows for faster optimization. Alternatively, the parameter space \mathbf{p}_m (eq. T1.5) could have been extended with parameters that describe the spectral shape as well. However a larger parameter space required larger computational efforts and here we focused on the source location and time, which only requires changes of the dispersion slope, timing, and amplitude. Tests where more parameters are optimized resulted in a lower total fractional error (eq. T1.10), but did not change the results in the optimization of the position. Hence, the reduction of the parameter space leads to larger systematic error in the cost function, but its physical interpretation remains the same.

The parameters for each station $\mathbf{p}^k(\mathbf{p}_m)$ are calculated at each function evaluation of the multi-station optimization. The cost function for optimization over N stations is defined as

$$J_m = \sum_k^N \frac{w_{err}^k}{\sum_i^N w_{err}^i} J^k, \quad (\text{T1.8})$$

where J^k is the regularized cost function for each individual event k (eq.T1.6) and w_{err}^k is the measure of the fit derived from the individual fitting procedure. It is defined for each event at a station k as

$$w_{err}^k = 1 - \frac{\|(\mathbf{D}^k - \mathbf{M}^k)\mathbf{w}^k\|^2}{\|\mathbf{D}^k\mathbf{w}^k\|^2}, \quad (\text{T1.9})$$

where \mathbf{w} is again the geometric weight of the event (dark shading in main-text figure 7 b to f, also F4 and F6, Hell et al., 2019, sec. 6.d). A $w_{err}^k = 1$ expresses a perfect model fit, while a $w_{err}^k = 0$ describes a complete failure of the optimization at the individual station. The weighting emphasizes station data with a high signal-to-noise ratio rather

than uncorrelated noisy data that might appear in some cases. This methodology can be easily extended to incorporate more data from other observations.

The parameters \mathbf{p}_m are not regularized to allow a more unbiased search of the source location. However, the search space is limited to the North Pacific (20°N to 60°N, 140°E to 120° W) and ± 2 days around the 5-station mean of the individual fits.

The error of the model fit for the multi-station cost function is then defined as the sum of the individual weighted cost functions (T1.8) normalized by the sum of the (geometrically) weighted data such that

$$e_f = \frac{J_m}{\sum_k^N (D^k w^k)^2}, \quad (\text{T1.10})$$

for a given set of N stations. The fractional error e_f can be interpreted as a likelihood

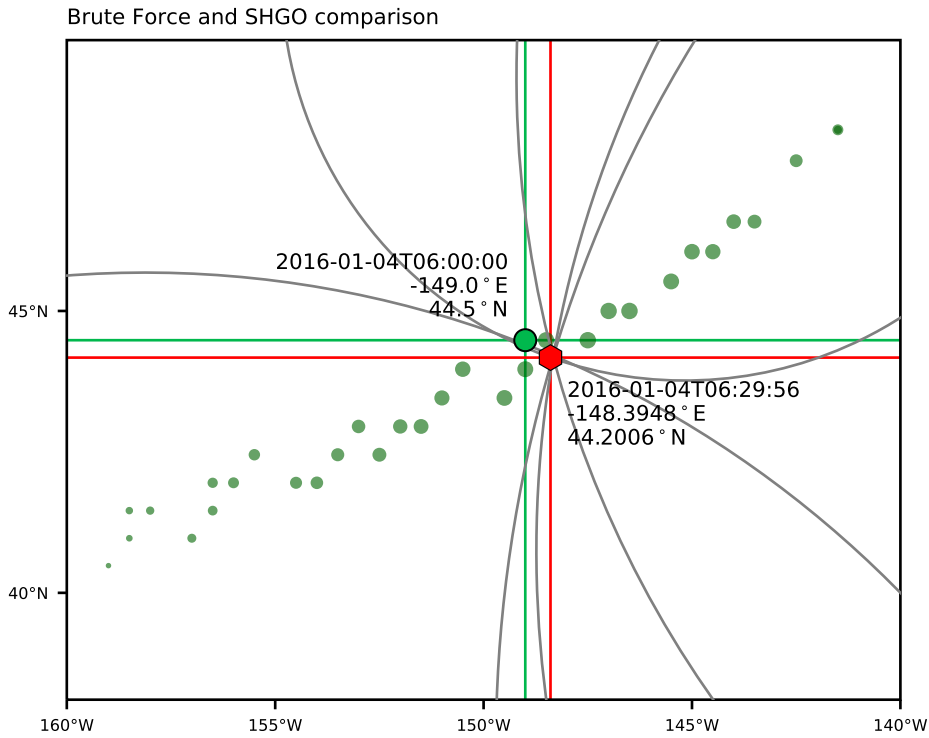
$$L_{e_f} = 1 - e_f = 1 - \frac{J_m}{\sum_k^N (D^k w^k)^2}, \quad (\text{T1.11})$$

such that a perfect match ($e_f = 0$) results in a likelihood of 1 and a failure of the model results in a likelihood of zero.

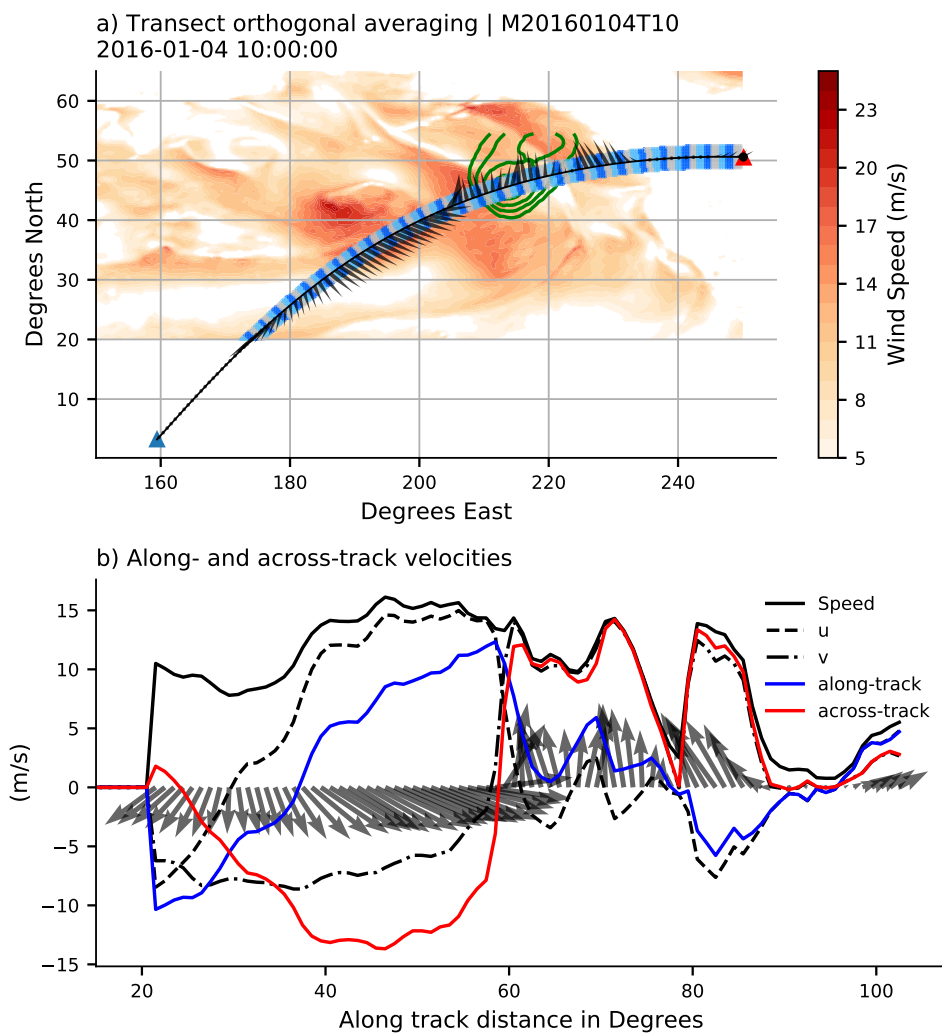
The 31 sets of matched observations are used to explore the multi-station cost function J_m (eq. T1.8) with two different procedures to explore the cost function. The first procedure is a brute-force sampling in the 3-dimensional parameter space of \mathbf{p}_m on the same grid as the wind data was provided (hourly and 25km). This creates a time-varying map of model fit using eq. T1.8 that is transformed to a map of likely wave origins using eq. T1.11 (see main-text section 3.2 and main-text figure 7a, F4 a and F6 a). The second procedure uses a sequence of two gradient decent methods; First simplicial homology global optimization (SHGO, Endres et al., 2018), and then a dual annealing method (DA Tsallis, 1988; Tsallis & Stariolo, 1996; Xiang et al., 1997). Both methods are developed

for fast convergence to a single global optimum of a complex cost-function. Regardless of the method, or procedure all optimizations return a focal points that are the same, as least on scales that are relevant for this study (figure F1).

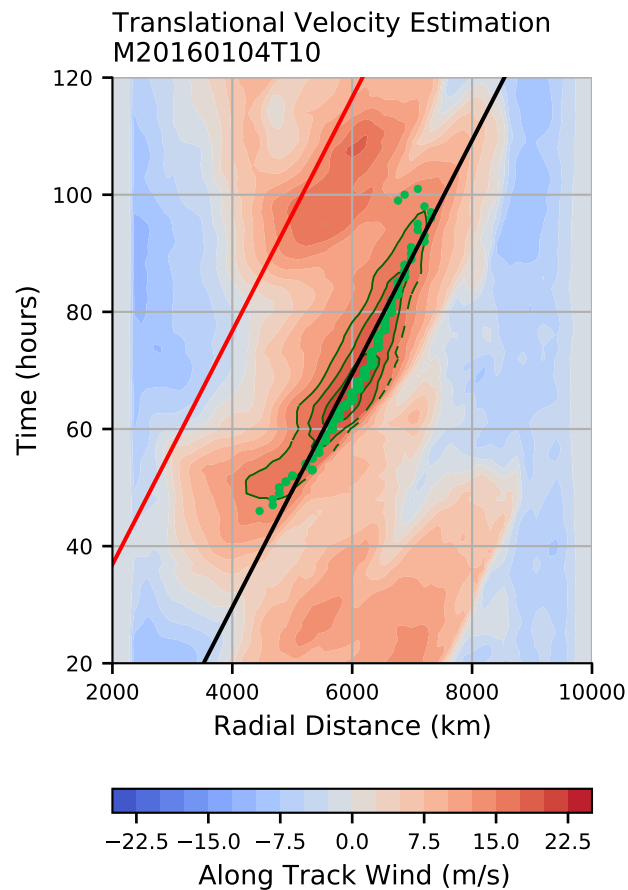
Figures F1 Comparison of the brute force (green dot) and SHGO (red hexagon) optimization, while the optimized location and time are indicated next to them. The dark green dots are positions of minimal fractional error from the brute force method before and after the best smallest fractional error on hourly intervals. The gray circle lines are the great circle distances centered around the stations in main-text figure 1a according to their radial distance through the SHGO point.



Figures F2 Illustration of transects through wind data along the trace of most likely wind origins. The black line in (a) indicates the estimated great-circle line (see section 3.2 in the main text), and the gray, light blue, or dark blue patches are the group grid points used for each transformed wind vector (black thin line). The wind speed is shown as red shading on the likelihood map of wave origin as green lines. The zonal and meridional wind, wind speed, as well as the transformed along- and across-track velocities are shown in panel (b). The vectors again show the zonal and meridional wind direction for each of the along-track averages bins indicated in panel (a).

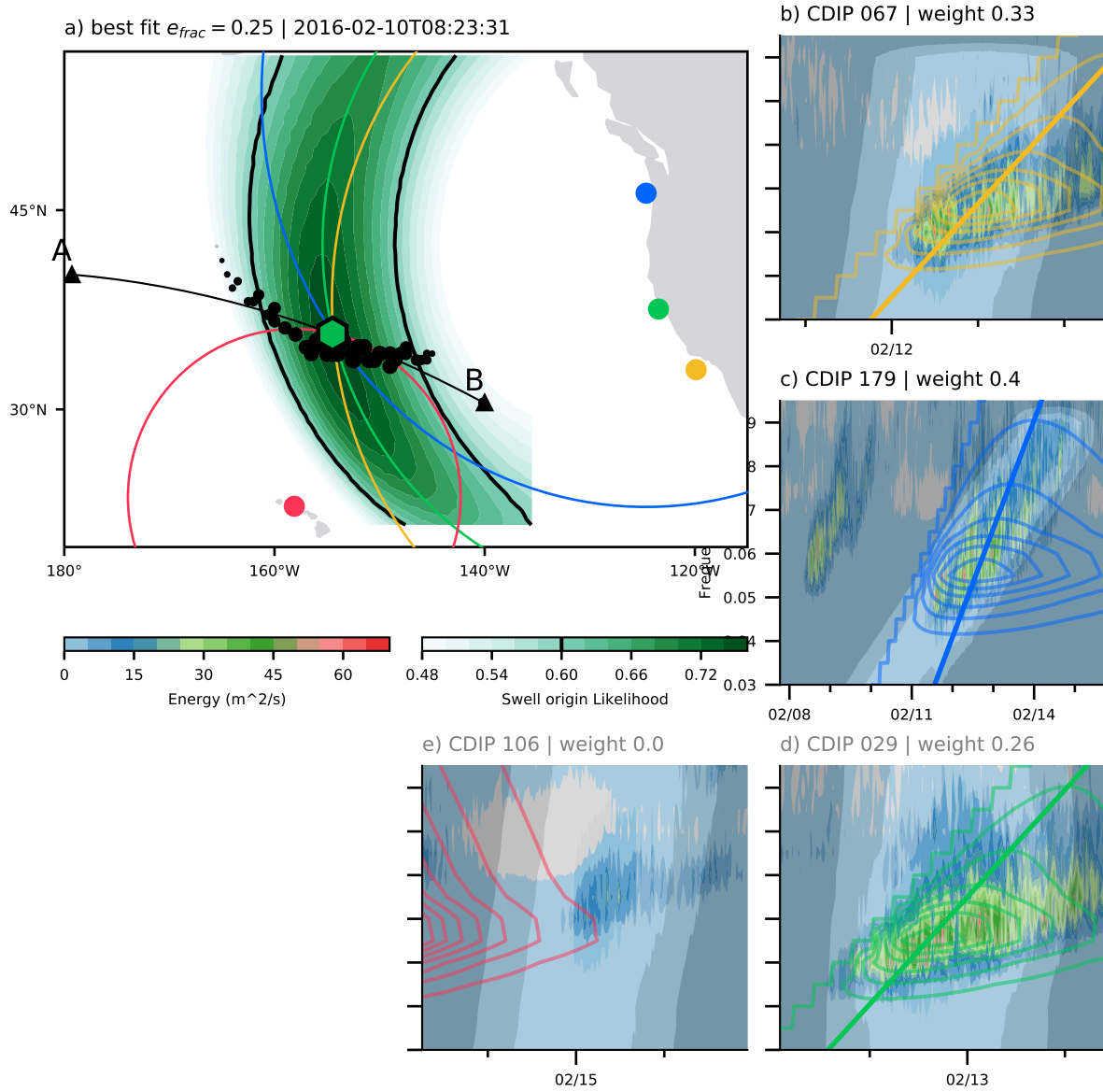


Figures F3 Least square fit (black line) to the points of maximum along track wind at each time step (green dots). These local maxima are determined within the centered wind event (green contours). The along track wind is shown as red and blue shading. The wind data is transformed according to the black line (see section 3.3. in the main text) while the left boundary is defined by the parallel shifted red line.

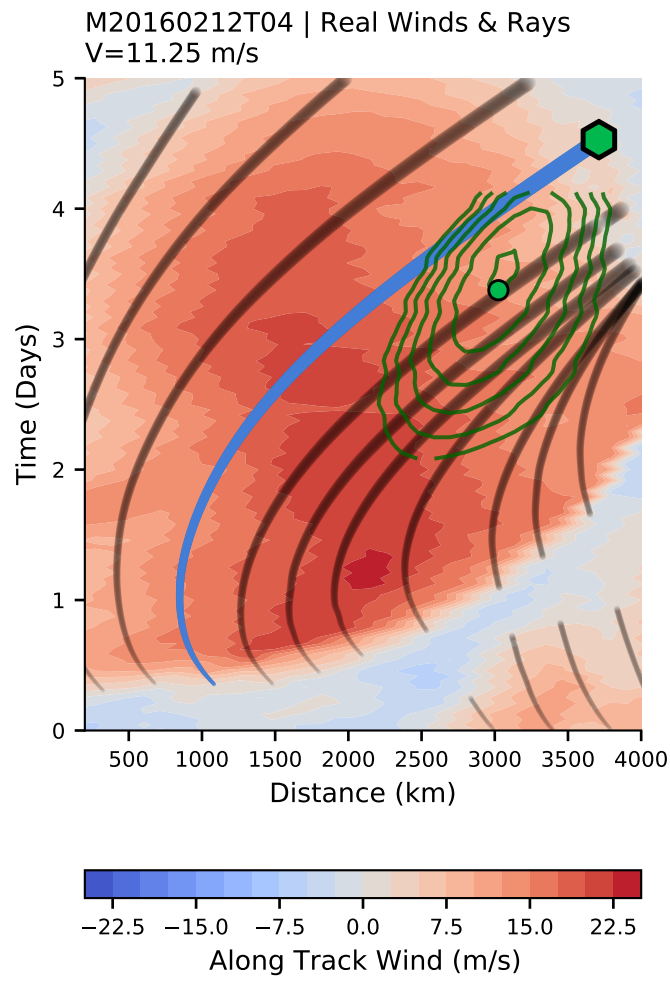


June 3, 2021, 6:00pm

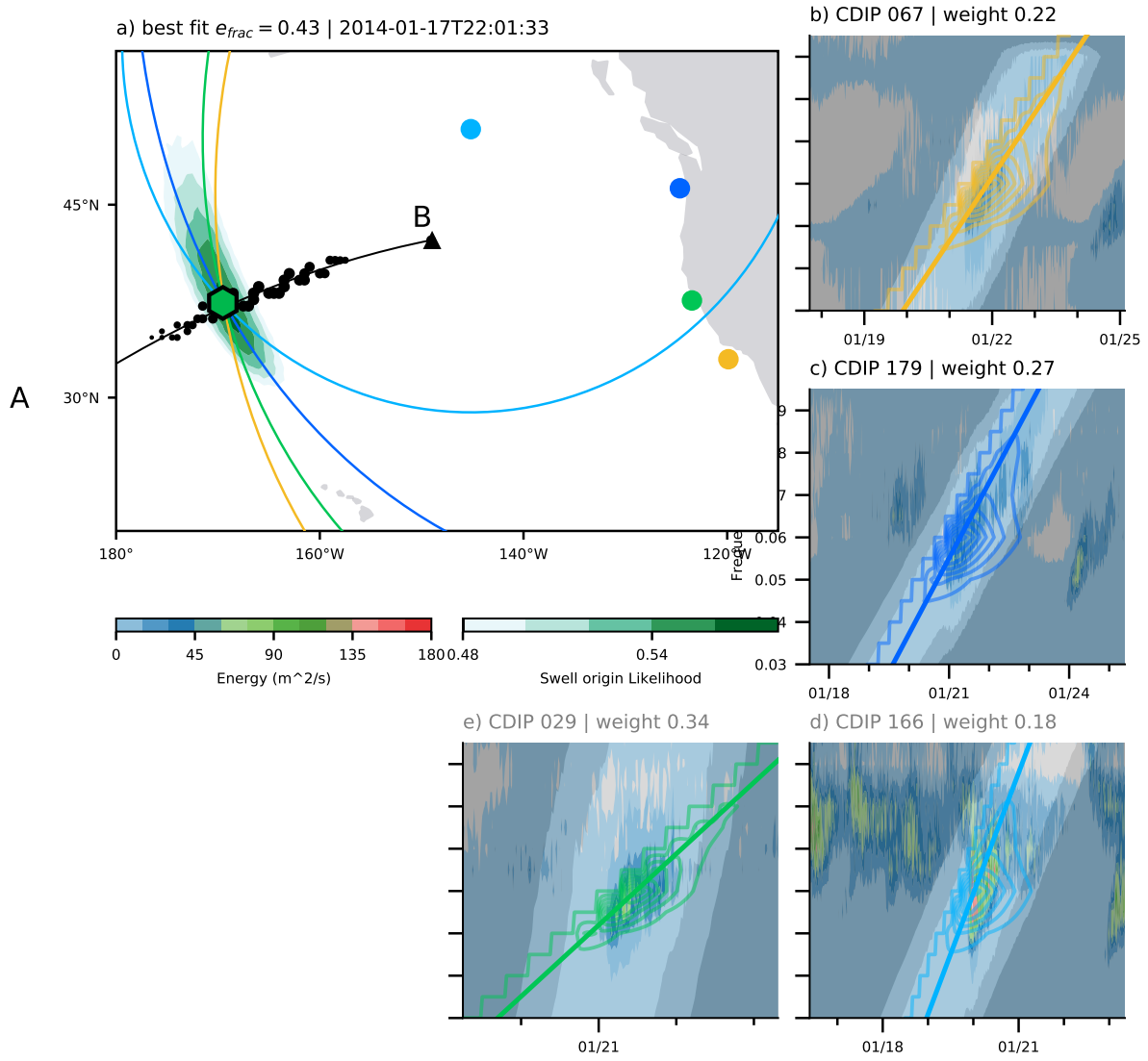
Figures F4 Same as main-text Figure 7 but for a case storm around February 10th 2016.



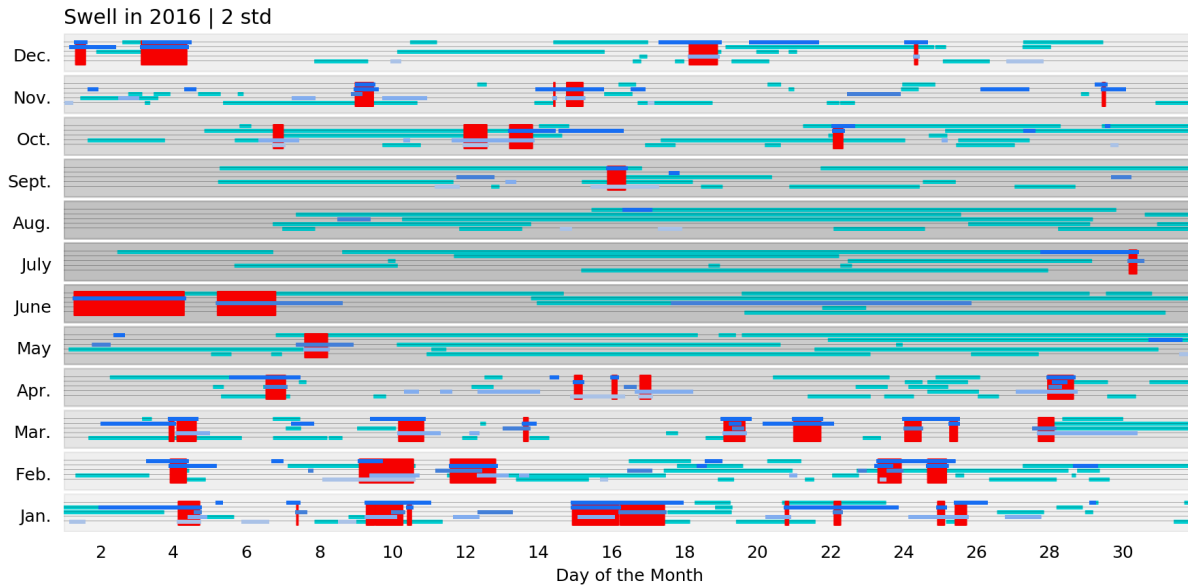
Figures F5 Same as main-text figure 9b, but now for the case in figure F4. See figure caption of figure F4.



Figures F6 Same as main-text Figure 7 but for a case storm around January 17th 2014.



Figures F7 Simplified time series for one year of fitted initial times uncertainty estimates. Each green or dark blue line shows $t_0 \pm 2\sigma_{t_0}$ for events identified at one of the stations. Green bars indicate events that have fractional error $e_f < 0.4$, while blue bars are events with $e_f \geq 0.4$. The red areas show time spans where at least 3 or more events have overlapping estimated initial times.



References

- Ardhuin, F., Chapron, B., & Collard, F. (2009, March). Observation of swell dissipation across oceans. *Geophys. Res. Lett.*, *36*, L06607. doi: 10.1029/2008GL037030
- Barber, N. F., & Ursell, F. (1948, February). The generation and propagation of ocean waves and swell. I. Wave periods and velocities. *Phil. Trans. R. Soc. Lond. A*, *240*, 527–560. doi: 10.1098/rsta.1948.0005
- Behrens, J., Thomas, J., Terrill, E., & Jensen, R. (2019, March). CDIP: Maintaining a Robust and Reliable Ocean Observing Buoy Network. In *2019 IEEE/OES Twelfth Current, Waves and Turbulence Measurement (CWTM)* (pp. 1–5). San Diego, CA, USA: IEEE. doi: 10.1109/CWTM43797.2019.8955166
- Earl, D. J., & Deem, M. W. (2005, November). Parallel tempering: Theory, applications, and new perspectives. *Phys. Chem. Chem. Phys.*, *7*, 3910–3916. doi: 10.1039/B509983H
- Endres, S. C., Sandrock, C., & Focke, W. W. (2018, October). A simplicial homology algorithm for Lipschitz optimisation. *J Glob Optim*, *72*, 181–217. doi: 10.1007/s10898-018-0645-y
- Foreman-Mackey, D., Hogg, D. W., Lang, D., & Goodman, J. (2013, March). Emcee: The MCMC Hammer. *Publ. Astron. Soc. Pac.*, *125*, 306–312. doi: 10.1086/670067
- Gallet, B., & Young, W. R. (2014). Refraction of swell by surface currents. *J. Mar. Res.*, *72*, 105–126. doi: info:doi/10.1357/002224014813758959
- Goodman, J., & Weare, J. (2010, January). Ensemble samplers with affine invariance. *Commun. Appl. Math. Comput. Sci.*, *5*, 65–80. doi: 10.2140/camcos.2010.5.65
- Hasselmann, K., Barnett, T. P., Bouws, E., Carlson, H., Cartwright, D. E., Enke, K., . . .

- Walden, H. (1973). Measurements of wind-wave growth and swell decay during the Joint North Sea Wave Project (JONSWAP). *Ergänzungsheft 8-12*.
- Hell, M. C., Cornuelle, B. D., Gille, S. T., Miller, A. J., & Bromirski, P. D. (2019, October). Identifying Ocean Swell Generation Events from Ross Ice Shelf Seismic Data. *J. Atmos. Oceanic Technol.*, *36*, 2171–2189. doi: 10.1175/JTECH-D-19-0093.1
- Jiang, H., Stopa, J. E., Wang, H., Husson, R., Mouche, A., Chapron, B., & Chen, G. (2016). Tracking the attenuation and nonbreaking dissipation of swells using altimeters. *J. Geophys. Res. Oceans*, *121*, 1446–1458. doi: 10.1002/2015JC011536
- Munk, W. H. (1947, April). Tracking storms by forerunners of swell. *J. Meteor.*, *4*, 45–57. doi: 10.1175/1520-0469(1947)004<0045:TSBFOS>2.0.CO;2
- Newville, M., Stensitzki, T., Allen, D. B., & Ingargiola, A. (2014, September). *LMFIT: Non-Linear Least-Square Minimization and Curve-Fitting for Python*. Zenodo. doi: 10.5281/zenodo.11813
- Tsallis, C. (1988, July). Possible generalization of Boltzmann-Gibbs statistics. *J Stat Phys*, *52*, 479–487. doi: 10.1007/BF01016429
- Tsallis, C., & Stariolo, D. A. (1996, November). Generalized simulated annealing. *Physica A: Statistical Mechanics and its Applications*, *233*, 395–406. doi: 10.1016/S0378-4371(96)00271-3
- Villas Bôas, A. B., & Young, W. R. (2020, May). Directional diffusion of surface gravity wave action by ocean macroturbulence. *J. Fluid Mech.*, *890*. doi: 10.1017/jfm.2020.116
- Wales, D. J., & Doye, J. P. K. (1997, July). Global Optimization by Basin-Hopping

and the Lowest Energy Structures of Lennard-Jones Clusters Containing up to 110 Atoms. *J. Phys. Chem. A*, *101*, 5111–5116. doi: 10.1021/jp970984n

Xiang, Y., Sun, D. Y., Fan, W., & Gong, X. G. (1997, August). Generalized simulated annealing algorithm and its application to the Thomson model. *Physics Letters A*, *233*, 216–220. doi: 10.1016/S0375-9601(97)00474-X