Supplementary materials to:
The motion of trees in the wind: a data synthesis

Toby D Jackson[1], Sarab Sethi[2], Ebba Dellwik[3], Nikolas Angleou[3], Amanda Bunce[4], Tim van Emmerik[5], Marine Duperat[6], Jean-Claude Ruel[6], Axel Wellpott[7], Skip Van Bloem[8], Alexis Achim[9], Brian Kane[10], Dominick M Ciruzzi[11], Steven P Loheide II[11], Ken James[12], Daniel Burcham[13], John Moore[14], Dirk Schindler[15], Sven Kolbe[15], Kilian Wiegmann[16], Mark Rudnicki[17], Victor J Lieffers[18], John Selker[19], Andrew V Gougherty[20], Tim Newson[21], Andrew Koeser[22, 23], Jason Miesbauer[24], Roger Samelson[25], Jim Wagner[26], David Coomes[1], Barry Gardiner[27]

[1] Plant Sciences, University of Cambridge, CB2 3EA, UK
[2] Department of Mathematics, Imperial College London, UK
[3] Department of Wind Energy, Technical University of Denmark, Frederiksborgvej 399, Roskilde, 4000, Denmark
[4] Department of Natural Resources, University of Connecticut, Mansfield, CT 06269, USA
[5] Hydrology and Quantitative Water Management Group, Wageningen University, Wageningen, The Netherlands
[6] Department of Wood and Forest Sciences, Laval University, Quebec, G1V 0A6, Canada
[7] Bavarian State Institute of Forestry (LWF), Hans-Carl-von-Carlowitz-Platz 1, D-85354 Freising
[8] Baruch Institute of Coastal Ecology and Forest Science, Clemson University, PO Box 596, Georgetown, SC 29442, USA
[9] Centre de recherche sur les matériaux renouvelables, Département des sciences du bois et de la forêt, Université Laval, Québec, QC G1V 0A6, Canada
[10] Department of Environmental Conservation, University of Massachusetts, Amherst, MA 01003, USA
[11] Civil and Environmental Engineering, University of Wisconsin Madison, Madison, WI
[12] School of Ecosystem and Forest Sciences, Faculty of Science, University of Melbourne, Melbourne, Australia
[13] Centre for Urban Greenery and Ecology, National Parks Board, 259569 Singapore
[14] Timberlands Ltd., Rotorua 3010, New Zealand
[15] Environmental Meteorology, University of Freiburg, Germany
[16] Argus Electronics gmbh, Erich-Schlesinger-Str. 49d, 18059 Rostock
[17] College of Forest Resources and Environmental Science, Michigan Technological University, Houghton, MI 49931 USA
[18] Renewable Resources Dept, University of Alberta, USA
[19] Oregon State University, Corvallis, OR 97331, USA
[20] Department of Botany, University of British Columbia, Canada
[21] Department of Civil and Environmental Engineering, Western University, Canada
[22] Department of Environmental Horticulture, IFAS, University of Florida
[23] Gulf Coast Research and Education Center, 14625 County Road 672, Wimauma, FL 33598, United States
[24] The Morton Arboretum, Lisle, IL 60532, USA
[25] College of Earth, Ocean, and Atmospheric Sciences, Oregon State University, Corvallis OR 97331 USA
[26] Oregon Research Electronics, Tangent, OR 97389, USA
[27] Institut Européen de la Forêt Cultivée, 69 route d'Arcachon, 33612, Cestas, France

Correspondence to: Toby Jackson (tobydjackson@gmail.com)

S1. Calculating the slope of the power spectrum

In this paper we found that the slope of the power spectrum, S_{freq}, was a key feature of tree motion. Here we outline how it was calculated in detail and how sensitive the results are to the calculation method.

Each tree has two horizontal axes of motion and we applied a 10-minute high-pass Butterworth filter to remove offsets from each channel separately, and then calculated the resultant tree motion. We resampled each one-hour time-series to 4 Hz and calculated the power spectral density using Welch's method (pwelch, in Matlab). We then plot the log-transformed power spectral density (y-axis) against the log-transformed frequency (x-axis) and fit a linear model to the data across the specified frequency range to obtain the slope of the power spectrum, S_{freq}. The fitting method and frequency range are somewhat arbitrary, and we test multiple options below. We noted that the high frequency range showed the most variation and the frequency ranges up to 2 Hz were more significant in the analyses, we therefore test this frequency range in more detail.

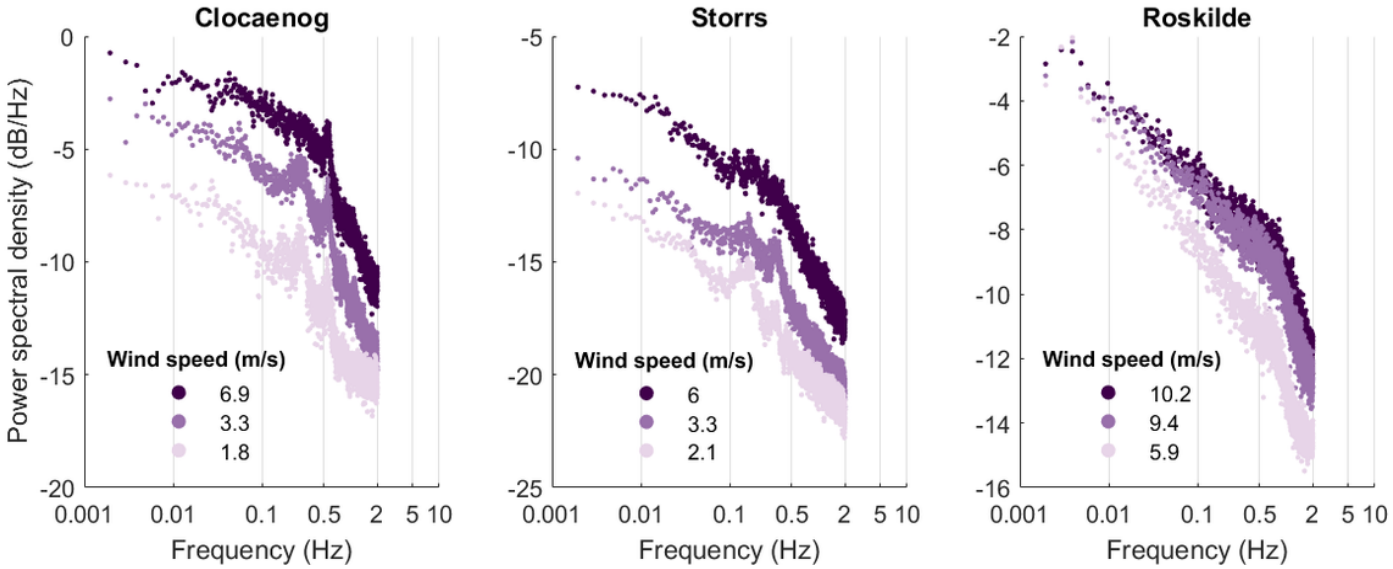

Figure S1 - Power spectra for example forest conifer (left), forest broadleaf (middle) and open-grown broadleaf (right) in high medium and low wind speeds (colours).

We tested two methods to fit linear models to the power spectra: (1) using the output of *pwelch* directly and (2) logarithmically re-sampling the output to give evenly distributed log-transformed data. We found that this re-sampling altered the absolute values of the slope slightly, but did not alter the observed trends.

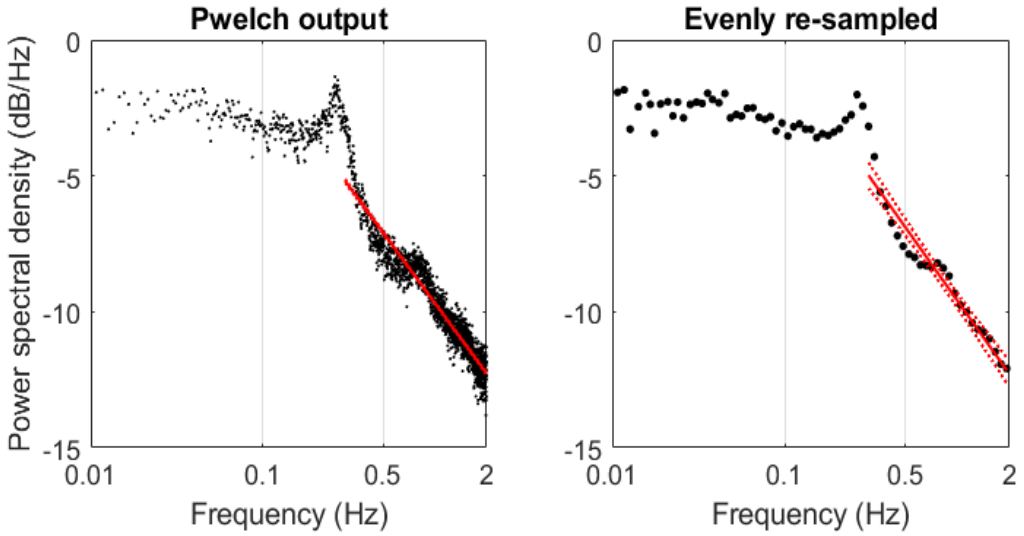

Figure S2 - Linear models of the power spectra to calculate slope using output of pwelch directly (left) and logarithmically re-sampled output (right). The frequency interval used in the above is 0.3 - 2 Hz.

We tested three frequency intervals over which to fit the linear models, 0.05-2, 0.3-2 and 1-2 Hz. The first two intervals produced similar trends with slightly different absolute values. The shortest interval had a similar trend but was partially obscured by an increased level of noise.

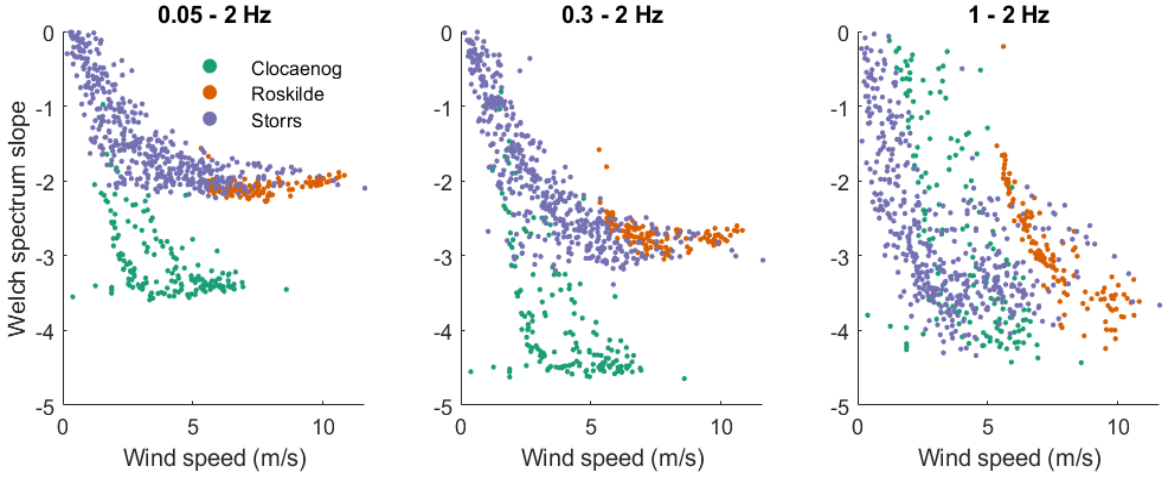

Figure S3 - Slope of the power spectrum against wind speed for three example trees (same as Figure S1) for three different frequency ranges.

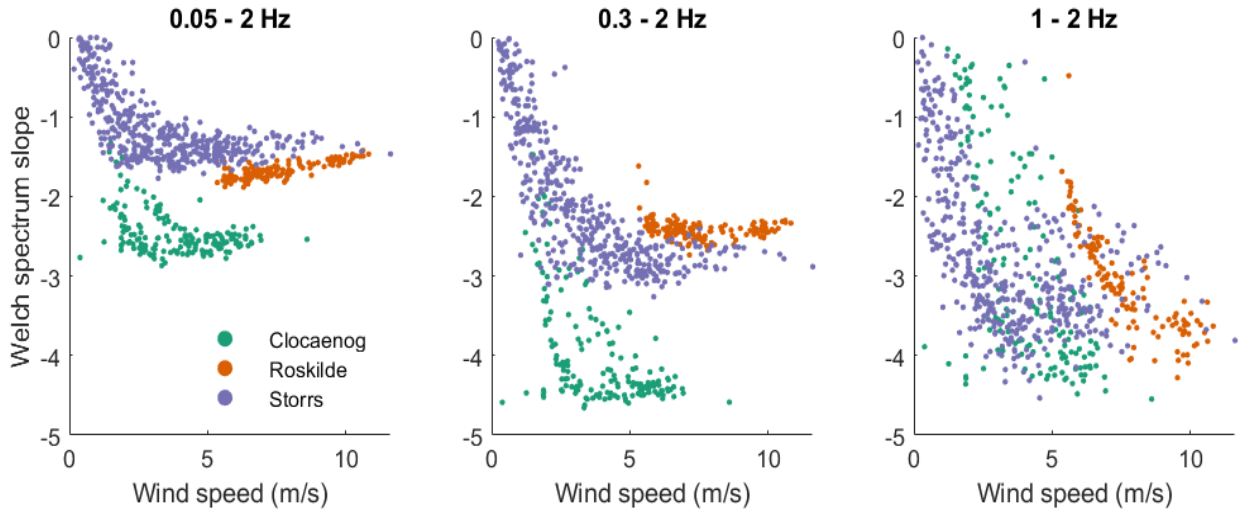

Figure S4 – Same as figure S3 but calculated using the logarithmically re-sampled power spectrum output.

Overall, we find that the trend described in Figure 3d of the main text is robust to the different frequency ranges and fitting methods.

We also tested the effect of the Butterworth high-pass filter on the power spectra. The purpose of this filter is to remove offsets in the tree motion data. These offsets vary slowly so we chose a 10-minute high-pass filter. We found this had no significant effect on the power spectrum in the region of interest (0.05-2 Hz).

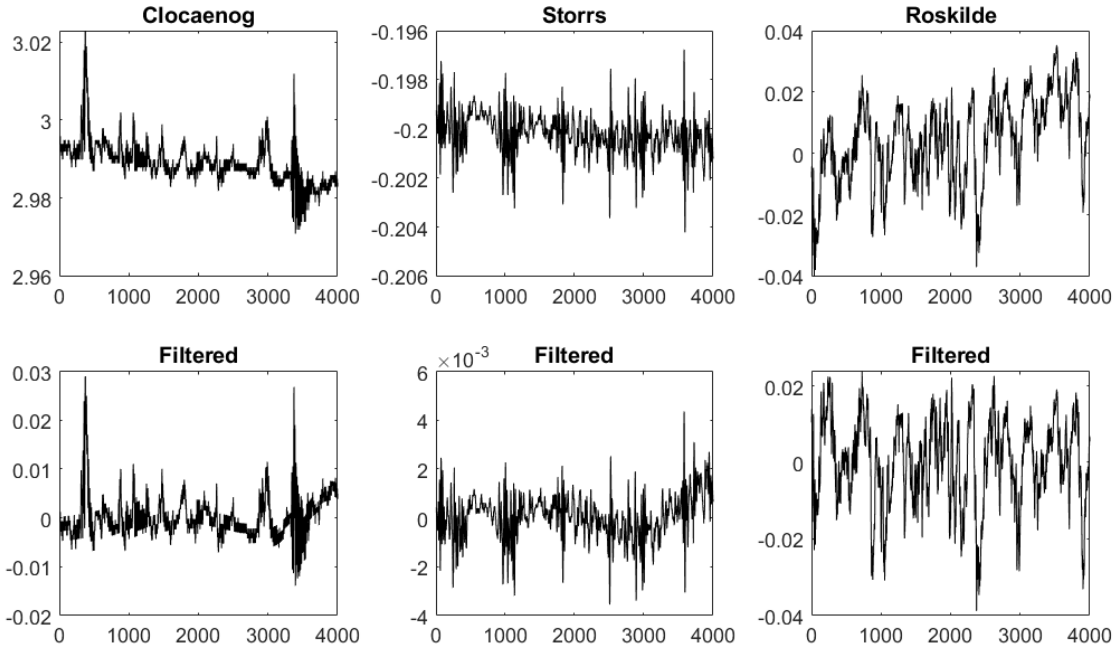

Figure S5 – Time-series for the same three trees as figures S1, 3 and 4 with and without the Butterworth high-pass filter.

S2. Example power spectra at original resolution

In the following we provide power spectra of tree motion (colours) and locally measured wind speed time-series (black) for the sites in which we have sufficient data. These are the forest broadleaf trees (figure S5) and forest conifers (figure S6). These power spectra were calculated at the original sampling frequency (all analysis in the main text was re-sampled to 4Hz) and we therefore pre-multiplied the y-axes by the frequency to allow a direct comparison between sites.

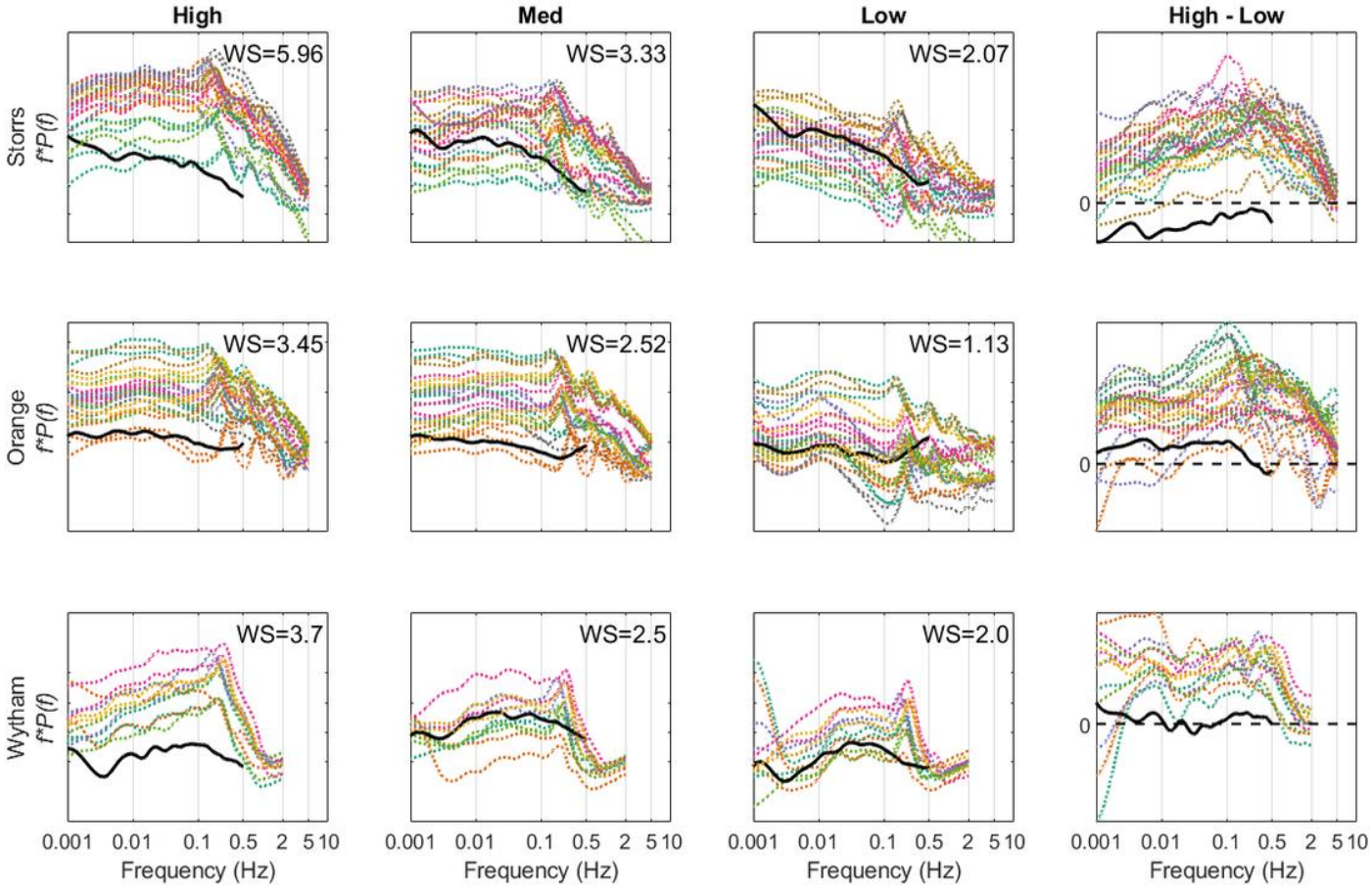

Figure S6 – power spectra for hourly samples of tree motion data (coloured lines) and wind speeds (black lines) at high, medium and low wind speed. Y-axis labels correspond to site names. The y-axis can be thought of as a measure of the relative energy content, in arbitrary units. The right-hand panels shows the difference between high and low wind speeds and the horizontal dashed line represents 0 change. Numbers in the top right-hand corners show the mean hourly wind speeds for the data sample. All of these sites are forest broadleaf trees.

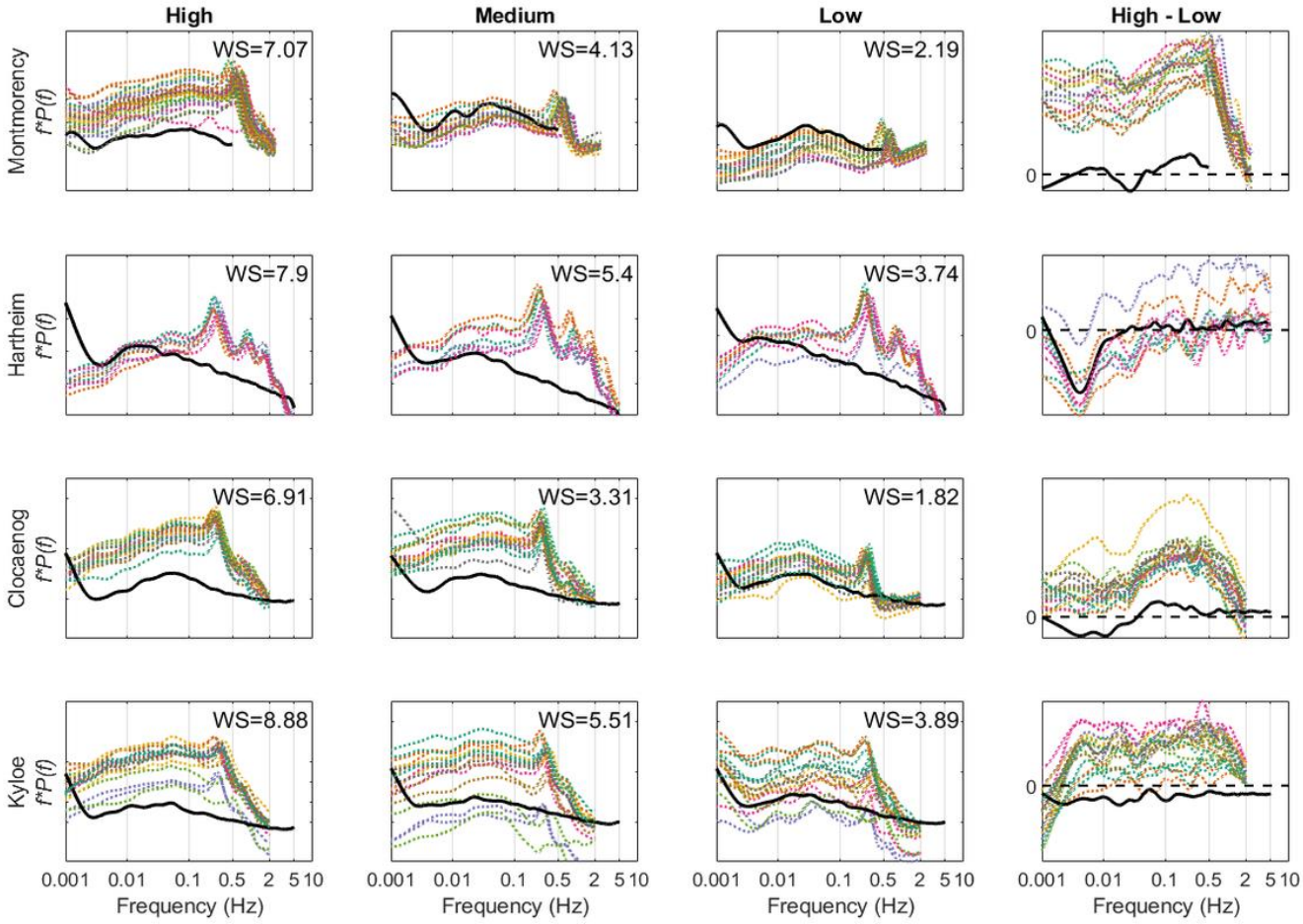

Figure S7 - Same as figure S6 but for sites in conifer forests.

S3. Correlation of all tree motion features

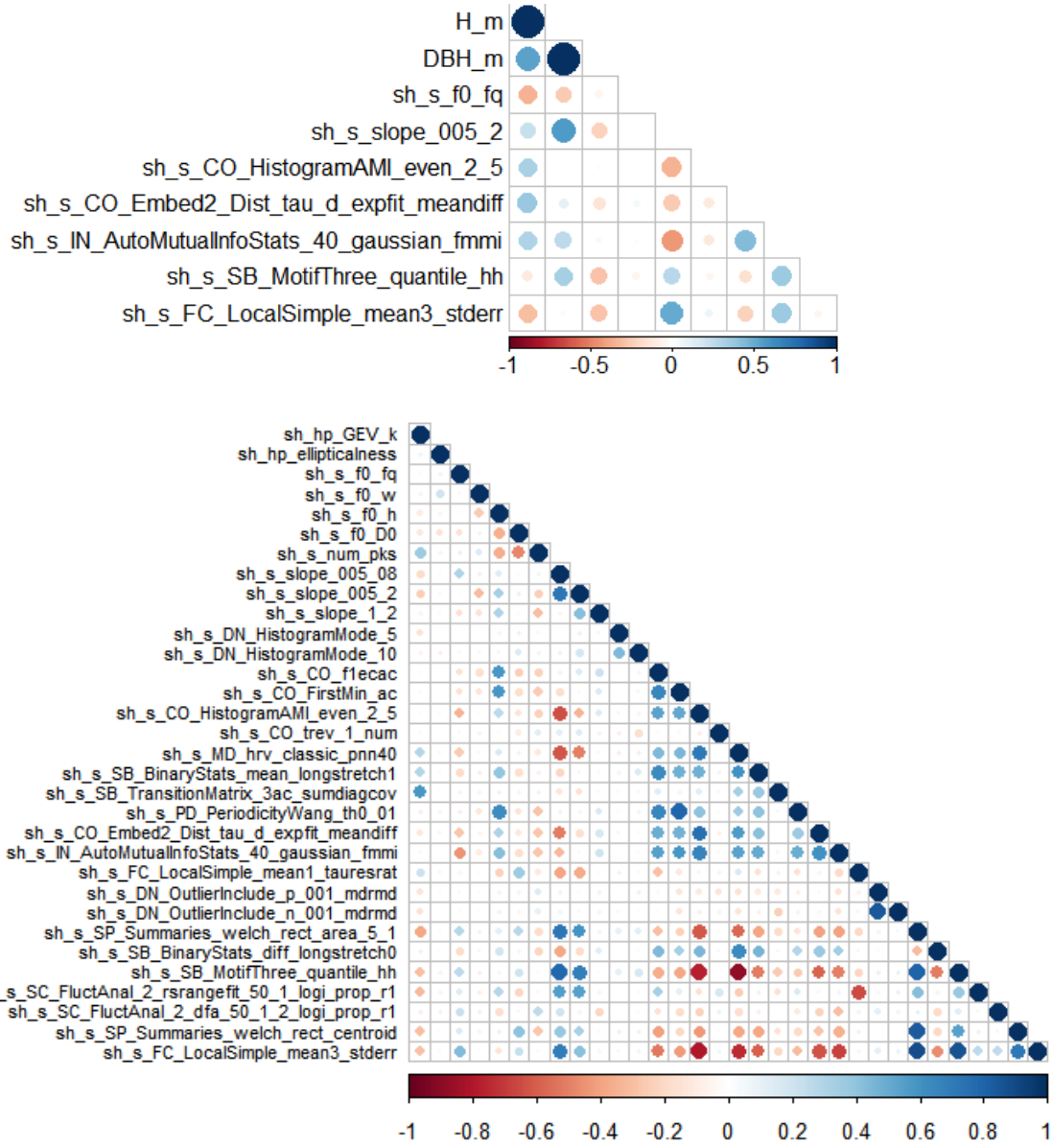

Figure S8 – Correlation plot for the tree motion features. Top panel shows tree height and diameter and only those features which are strongly correlated with them (R^2>0.3). Bottom panels show all features used in this study. Colours represent strength of correlation as determined by the coefficient of determination (R^2)

S4. Correlation between tree size and tree motion features

We considered a subset of trees (N=168, 86 forest broadleaves, 54 forest conifers and 28 open-grown broadleaves) for which height, *dbh* and tree motion data were available. In order to test which features were closely related to tree size, while accounting for tree types, we predicted tree height and diameter from the tree motion features using a multiple linear regression including tree type in the model as a factor (Table 3). The best single predictor of *dbh* was S_{freq} while f_0 was the second best predictor of tree height after a *catch22* feature (CO_Embed2_Dist_tau_d_expfit_meandiff). The factor "tree type" was the 9th most explanatory feature in the model of height and 6th in the model of *dbh*. This demonstrates that tree size is more strongly related to tree motion features than it is to tree type. Therefore, the relationship between tree type and tree motion features is unlikely to be confounded by differences in tree size, and hence the results of our classification analyses are valid.

Linear model	R^2	AIC
$DBH \sim S_{freq}$	0.312	173
$DBH \sim S_{freq} + C_{22}m$	0.357	162
$DBH \sim S_{freq} + C_{22}m + C_{22}g$	0.400	151
$H \sim C_{22}r$	0.132	1240
$H \sim C_{22}r + f_0$	0.196	1228
$H \sim C_{22}r + f_0 + C_{22}u$	0.239	1219

Table S1 - Summary statistics from the most parsimonious multiple linear models relating tree height and dbh to tree motion features. Each feature is added to the model sequentially in order of the largest decrease in AIC. A maximum of 3 features are included here for clarity.$C_{22}m$ - IN_AutoMutualInfoStats_40_gaussian_fmmi, $C_{22}g$ - CO_FirstMin_ac, $C_{22}r$ - CO_Embed2_Dist_tau_d_expfit_meandiff, $C_{22}u$ - SB_TransitionMatrix_3ac_sumdiagcov. A brief description of the catch22 features can be found in the supplementary materials (S4) and a more detailed description in the associated publication (Lubba et al 2019).

S5. Catch22 features table

	Name	Description
a	DN_HistogramMode_5	Mode of z-scored distribution (5-bin histogram)
b	DN_HistogramMode_10	Mode of z-scored distribution (10-bin histogram)
c	SB_BinaryStats_mean_longstretch1	Longest period of consecutive values above the mean
d	DN_OutlierInclude_p_001_mdrmd	Time intervals between successive extreme events above the mean
e	DN_OutlierInclude_n_001_mdrmd	Time intervals between successive extreme events below the mean
f	CO_f1ecac	First 1/e crossing of autocorrelation function
g	CO_FirstMin_ac	First minimum of autocorrelation function
h	SP_Summaries_welch_rect_area_5_1	Total power in lowest fifth of frequencies in the Fourier power spectrum
i	SP_Summaries_welch_rect_centroid	Centroid of the Fourier power spectrum
j	FC_LocalSimple_mean3_stderr	Mean error from a rolling 3-sample mean forecasting
k	CO_trev_1_num	Time-reversibility statistic, h(xt+1 − xt)3it
l	CO_HistogramAMI_even_2_5	Automutual information, m = 2, = 5
m	IN_AutoMutualInfoStats_40_gaussian_fmmi	First minimum of the automutual information function
n	MD_hrv_classic_pnn40	Proportion of successive differences exceeding 0.04
o	SB_BinaryStats_diff_longstretch0	Longest period of successive incremental decreases
p	SB_MotifThree_quantile_hh	Shannon entropy of two successive letters in equiprobable 3-letter symbolization
q	FC_LocalSimple_mean1_tauresrat	Change in correlation length after iterative differencing
r	CO_Embed2_Dist_tau_d_expfit_meandiff	Exponential fit to successive distances in 2-d embedding space
s	SC_FluctAnal_2_dfa_50_1_2_logi_prop_r1	Proportion of slower timescale fluctuations that scale with DFA (50% sampling)
t	SC_FluctAnal_2_rsrangefit_50_1_logi_prop_r1	Proportion of slower timescale fluctuations that scale with linearly rescaled range fits
u	SB_TransitionMatrix_3ac_sumdiagcov	Trace of covariance of transition matrix between symbols in 3-letter alphabet
v	PD_PeriodicityWang_th0_01	Periodicity measure of