Measuring Tree Sway Frequency with Videos for Ecohydrologic Applications: Assessing the Efficacy of Eulerian Processing Algorithms

Joseph H. Ammatelli[1], Ethan Gutmann[2], Sidney A. Bush[3], Holly Barnard[3], Dominick M. Ciruzzi[4,5,7], Steven P. Loheide II[4,5], Mark S. Raleigh[6], Jessica D. Lundquist[1]

[1]Department of Civil and Environmental Engineering, University of Washington, Seattle, Washington, USA. [2]National Center for Atmospheric Research, Boulder, CO, USA. [3]Department of Geography, Institute of Arctic and Alpine Research, University of Colorado, Boulder, Colorado, USA. [4]Department of Civil & Environmental Engineering, University of Wisconsin-Madison, Madison, WI, USA. [5]Geological Engineering, University of Wisconsin-Madison, Madison, WI, USA. [6]College of Earth, Ocean, and Atmospheric Sciences, Oregon State University, Corvallis, OR, USA. [7]Department of Geology, William & Mary, Williamsburg, VA, USA.

Corresponding authors: Joseph Ammatelli (jamma@uw.edu), Jessica Lundquist (jdlund@uw.edu)

Key Points:

- Measurements of tree sway frequency, typically obtained from accelerometers, can be used to quantify important ecohydrologic processes.
- Two video processing algorithms can measure tree sway with comparable accuracy to accelerometers while improving the spatial scalability.
- For best results, long videos of trees with exposed portions of trunk and strong contrast with the background should be used for analysis.

Abstract

Measurements of biophysical tree properties and hydrologic fluxes are necessary for improving models and monitoring the impact of disturbances. Prior research has demonstrated that measurements of tree sway frequency can be used to quantify important ecohydrologic processes, such as drought stress and snow interception, that otherwise require expensive measurement techniques. However, existing instruments used to measure tree sway lack spatial scalability. We investigate whether the virtual vision sensor and multilevel binary thresholding video processing algorithms can be used to accurately extract tree sway frequency at multiple points in a video camera field of view and enable scalable measurements of ecohydrologic processes. Comparing sway frequencies extracted from video and accelerometer data at two sites, we show that for 30-60 s videos, the video processing algorithms can reproduce accelerometer sway frequencies with ±0.03 Hz accuracy. The results suggest that video processing algorithms may be suitable for applications where changes in sway frequency are on the order of tenths of hertz or larger, for example the measurement of snow in trees. Further work is needed to clarify the accuracy of the algorithms when applied to longer videos, which may be required to monitor processes with more subtle changes in sway frequency, such as diurnal changes in tree water content.

Plain Language Summary

Wind causes trees to sway back and forth at a particular frequency, much like a pendulum. Measurements of the resultant tree sway frequency can be used to quantify changes in tree properties, such as mass and stiffness, that are related to tree-water interactions. Instruments commonly used to measure tree sway frequency can be difficult to install and only collect data for one point on a single tree. Therefore, it is challenging to measure the sway frequency of multiple trees. We show that two video processing techniques can be used to extract the sway frequency of multiple trees from videos recorded with inexpensive cameras. The findings suggest video processing techniques have the potential to simplify tree sway data collection, expand the number of trees monitored, and ultimately improve our understanding of how different trees are responding to various weather/climate events.

1 Introduction

Understanding the timing and magnitude of ecohydrologic processes, including evapotranspiration and precipitation interception, at tree and stand scales is critical for improving models (e.g., Lundquist et al., 2021) and monitoring the impact of disturbances such as drought (e.g., Clark et al., 2016).

Despite its importance to modeling and monitoring efforts, measuring tree water status and precipitation interception across space and time is challenging. Existing measurement techniques exhibit tradeoffs between measurement type (point vs. aggregated), spatial resolution/scalability, and time resolution (episodic vs. continuous). In situ point measurements are often destructive (e.g., sap flow), lack spatial scalability (e.g., weighing a severed tree (Stork et al., 2002)), episodic (e.g., leaf water potential), and/or hampered by spatial variability (e.g., precipitation interception). Airborne remote sensing products have varying space and time resolutions that are often too coarse for examining tree scale processes.

In recent decades, measurements of natural tree sway frequency have been used to infer difficult-to-measure tree processes (e.g., Ciruzzi & Loheide II, 2019; Raleigh et al., 2022; Van Emmerick et al., 2017). When excited by the wind, trees are compelled into oscillatory motion. The resulting vibration has a dominant frequency defined as the natural frequency, henceforward referred to as the sway frequency. Trees can be modeled as a damped harmonic oscillator, and the sway frequency can be related to biophysical properties including mass and stiffness (Jackson et al., 2019). Observing changes in these biophysical properties over time enables the quantification of ecohydrologic processes. Previous studies have used tree sway to classify tree health (Baker, 1997), identify changes in phenology (Gougherty et al., 2018; Jaeger et al., 2022), monitor plant water content (Ciruzzi & Loheide II, 2019; Kooreman, 2013), detect precipitation interception events (Ciruzzi & Loheide II, 2021b; Raleigh et al., 2022; Van Emmerick et al., 2017), and estimate changes in canopy mass (Raleigh et al., 2022; Selker et al., 2011).

Accelerometers have become a popular tool for measuring tree sway given their low cost and performance in challenging environments (Jackson et al., 2021). Using an accelerometer to obtain a vibration signal with a high signal-to-noise ratio (SNR) typically requires placing the sensor several meters above the ground (Ciruzzi & Loheide II 2019; Raleigh et al., 2022; Van Emmerick et al., 2017). Thus, accelerometers can provide useful data for a small number of trees, but it can be expensive and dangerous to scale measurements to community or stand scales.

Recent work has demonstrated the potential of using a video camera to simultaneously observe mechanical vibrations at different points in space associated with groups of pixels (Chen et al., 2017; Ferrer et al., 2013; Schumacher & Shariati, 2013; Wang et al., 2022). Videos encode changes in brightness across space and time for the entire camera field of view (FOV), effectively providing an array of motion sensors. Temporal variations in brightness correspond with local motion, presuming changes in lighting, changes to the visual properties of objects in the FOV, camera motion, and camera noise are all negligible. Vibration signals can be extracted from videos using Lagrangian or Eulerian frameworks. In the Langrangian framework, a feature is tracked across space and time. Despite demonstrated success measuring the motion of tree features (Wang et al., 2022), Lagrangian vibration analysis can only resolve pixel scale motion for individual features and suffers from issues such as target occlusion. In the Eulerian framework, the evolution of the brightness of a fixed pixel (or cluster of pixels) is analyzed over time, representing the movement of objects into and out of a region. Eulerian video processing methods can resolve sub-pixel motion across large spatial extents (Chen et al., 2017; Ferrer et al., 2013; Schumacher & Shariati, 2013), showing promise for measuring sway frequency for multiple trees in a camera FOV. Here, we develop and demonstrate two simple Eulerian video processing algorithms, based on the methods of Schumacher and Shariati (2013) and Ferrer et al. (2013), that can be used to capture the frequencies of swaying trees and enable analysis of ecohydrologic processes over greater spatial extents.

2 Materials and Methods

2.1 Video and Accelerometer Datasets

We compiled videos of swaying trees (input data) and co-located accelerometer data (validation data) from two existing studies. Video and accelerometer data collected by Ciruzzi and Loheide II (2019) for a red oak tree (Quercus rubra) in the Trout Lake Watershed in northern

Wisconsin were used for a simple single-tree validation. A video camera (Bushnell TrophyCam, 30 fps, 1080p resolution, 45° FOV) was tilted upwards at a slightly oblique angle towards the target tree and fastened with straps to the base of an adjacent tree so camera motion could be assumed negligible. Five, 60 s videos were recorded over a two hour period on 15 August 2019. A 3-axis accelerometer (Gulf Coast Data Concepts 2g MEL-X2,16 Hz continuous sampling) was positioned beneath the main branching of the target tree at ~8 m (total tree height ~22 m).

For multitree validation, we processed video and accelerometer data recorded by Bush (2022) at the Manitou Experimental Forest in Colorado (henceforward referred to as Manitou). Bush (2022) equipped six ponderosa pine trees (Pinus ponderosa, PIPO) with 3-axis accelerometers (Gulf Coast Data Concepts 2g MEL-X2,16 Hz continuous sampling) and mounted a GoPro camera (30 fps, 1080p resolution, 155° FOV) at the top of a nearby tower to capture the motion of the study trees. Accelerometers were mounted ~6-8 m above the ground (total tree height ~8-10 m). Videos were recorded every 15-30 min for 30 s between May and September 2020. We evaluated two of the metered trees.

2.2 Video Data Processing

We decomposed the video processing logic into three generalized steps: translating a video into vibration signals, estimating the power spectral density (PSD) of each signal, and aggregating the frequency content across all vibration spectra. Our analysis includes two adapted end-to-end algorithms: the virtual vision sensor (VVS) algorithm from Schumacher and Schariati (2013) and the multilevel binary thresholding (MBT) algorithm from Ferrer et al. (2013), which are illustrated in Figure 1 and detailed below. The algorithms differ primarily in how vibration signals are extracted from videos. In the VVS algorithm, pixel brightness time series are treated as vibration signals. Meanwhile, the MBT algorithm constructs vibration signals by counting the number of pixels in the region of interest (ROI) that are below a brightness threshold at each time step.

2.2.1 Selecting a Region of Interest

For all algorithms, regions of interest (ROIs) were manually selected following Schumacher and Shariati (2013) and Ferrer et al. (2013), who emphasize the importance of selecting areas of high contrast. To maximize brightness fluctuations and evaluate pixels most representative of the tree motion, we visually inspected the FOV for areas of exposed trunk with high-contrast backgrounds (e.g., the sky). The vertical range of the bounding box was chosen so it spanned the longest vertical portion of the exposed trunk with a high-contrast background. The horizontal range was chosen so it included one edge of the trunk, approximately equal parts of trunk and background at equilibrium, and enough background pixels to maintain a baseline signal for MBT when the trunk reaches its maximum displacement. ROIs high on the trunk, where larger sway displacements result in greater changes in pixel brightnesses, were preferred to ROIs lower on the trunk. When the trunk was difficult to identify, or the trunk sway was observed to be insignificant, ROIs on the crown with high contrast were evaluated.

2.2.2 Virtual Vision Sensor

The Virtual Vision Sensor (VVS) method introduced by Schumacher and Shariati (2013) treats each pixel as a “virtual vision sensor,” whose brightness time series captures the temporal

evolution of local motion. Here, we extracted vibration signals by converting the video to grayscale and loading the brightness time series of each pixel in the ROI (Figure 1a).

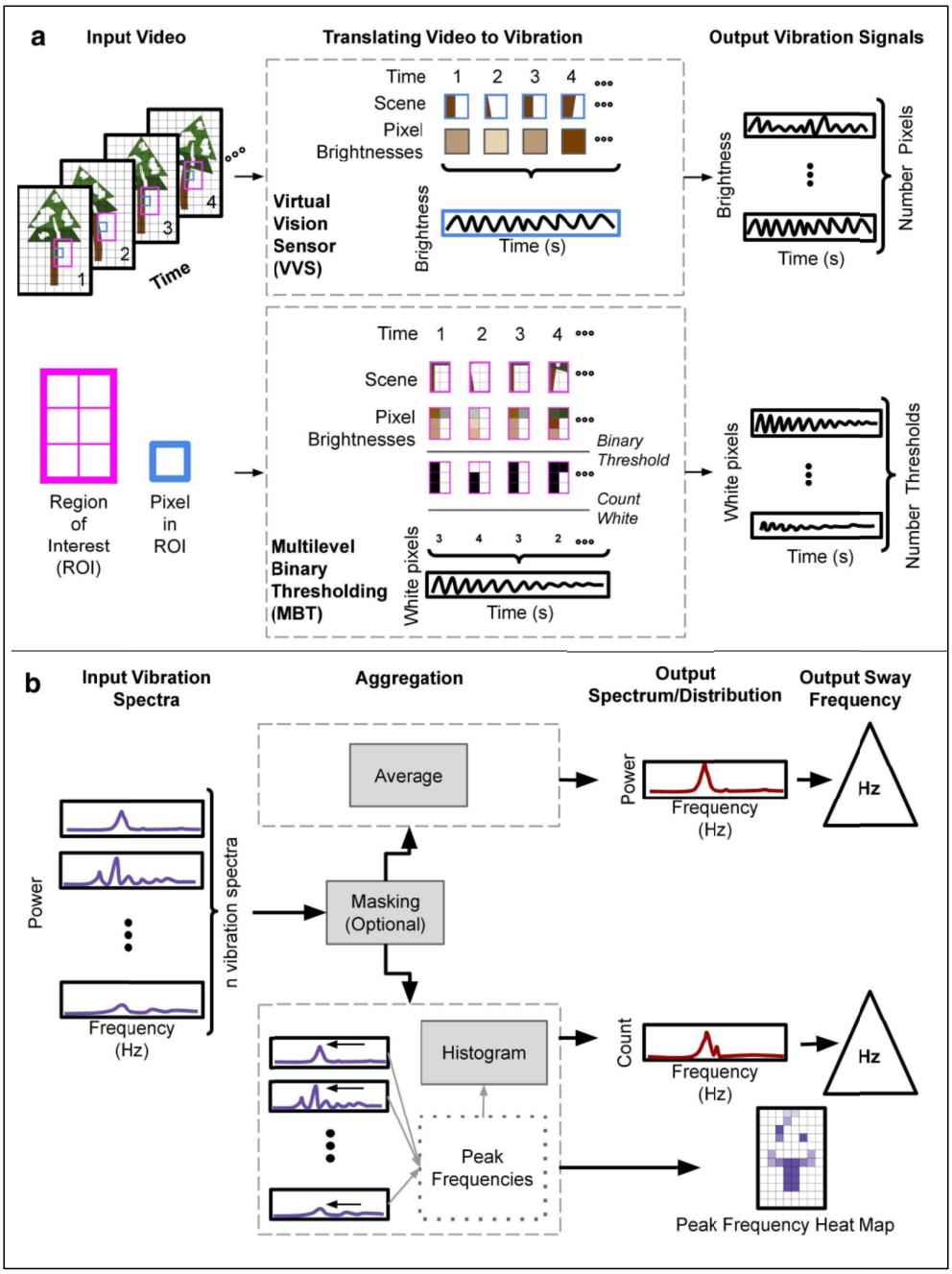

Figure 1: **(a)** Illustration of how vibration signals are extracted from a video using two methods: VVS (top) and MBT (bottom). **(b)** Illustration of the two methods used for aggregating power spectra: average spectrum (top) and peak frequency histogram (bottom).

157 The PSD of each pixel in the ROI was estimated using the periodogram technique
158 (Oppenheim & Schafer, 2010). For a given input signal, prior to computing the PSD, we
159 removed any linear trend, subtracted the mean, applied the Hann windowing function to reduce
160 spectral leakage, and zero padded the signal to 8x the nearest power of two for high-quality
161 interpolation between frequency samples. Spectra from the Trout Lake videos had a frequency
162 resolution of 0.016 Hz, while Manitou spectra had a frequency resolution of 0.033 Hz. The
163 frequency resolution, given by the inverse of the video duration in seconds, equivalently
164 describes the uncertainty in peak estimates and the magnitude of frequency fluctuations that can
165 actually be resolved. As dictated by the Nyquist-Shannon sampling theorem, the spectra initially
166 described frequency content from 0-15 Hz. To simplify peak selection and reduce the memory
167 burden, we trimmed the PSDs to a range of plausible candidate frequencies: 0.15-0.5 Hz.

168 Whereas Schumacher and Shariati (2013) analyzed individual pixels, we aggregated
169 frequency content from a subset of pixels in the ROI using two methods: (1) the average
170 spectrum and (2) the peak frequency histogram (Figure 1b). We defined the peak frequency of a
171 spectrum to be the frequency of the local maximum with the greatest magnitude. For both
172 aggregation techniques, we first masked any spectra corresponding to background pixels or
173 vibration signals with low SNRs. In particular, we only considered spectra whose frequency peak
174 prominence, the vertical distance between the frequency peak and its lowest contour line, was
175 above the 75th percentile of all peak prominences in the ROI. For the first aggregation technique,
176 we computed the unweighted average of the pixel spectra and determined the peak frequency of
177 the resulting spectrum. Compared to the raw pixel spectra, which each have 2 degrees of
178 freedom, the average spectrum typically had over 50 degrees of freedom (greater statistical
179 quality). For the second method, we found the peak frequency of each vibration spectrum and
180 then counted the number of times each peak frequency occurred, building a histogram for spectra
181 peak frequencies, where each bin represents a frequency in the PSD. The mode peak frequency
182 was chosen as the dominant sway frequency for the ROI. Because the VVS spectra were
183 associated with spatial coordinates, the peak frequencies were also used to create a heat map
184 (Fig. 1b) to assess the spatial distribution of peak frequencies.

185 2.2.3 Multilevel Binary Thresholding

186 The multilevel binary thresholding (MBT) scheme developed by Ferrer et al. (2013)
187 generates vibration signals by counting the number of pixels in the ROI whose grayscale
188 brightness is below a particular threshold at each time step (Figure 1a). Each threshold emits one
189 vibration signal, so multiple vibration signals are generated using a suite of thresholds. The
190 resultant signals represent how much an object of a particular brightness occupies the ROI at
191 each time step. MBT is predicated upon the motion of some high-contrast boundary. Therefore,
192 unlike VVS, MBT requires a distinct boundary between the vibrating object and the
193 background.

194 We generated eight vibration signals using MBT video to vibration translation with eight
195 thresholds evenly spaced between the minimum and maximum brightness in the ROI across all
196 frames. Vibration spectra were computed in the same manner described for VVS. We aggregated
197 the resultant spectra by finding the peak frequency of the average spectrum (Figure 1b). All
198 spectra were included in the average.

2.3 Accelerometer Data Processing

The aggregated spectra and sway frequencies generated by each video processing algorithm were compared against spectra and sway frequencies extracted from two lengths of accelerometer data. First, a short segment of accelerometer data with the same start time and duration as the video was used to assess accuracy. Here, the accelerometer PSD was generated by averaging the spectra of the two horizontal axes, which were computed using the same periodogram method and parameters applied to the video vibration signals. The resultant PSD had four degrees of freedom and the same frequency resolution as the video. Second, a 30-min accelerometer segment centered on the video start time was used to compare the video processing output to sway frequencies with greater statistical quality (i.e., more degrees of freedom and a higher frequency resolution), analogous to those used in practice. As before, the spectra from the two horizontal accelerometer axes were averaged together. This time, however, the PSD of each horizontal axis was estimated using Welch's method (Welch, 1967), with 50% overlapping 5-min segments, to reduce the output spectrum noise. The PSD of each segment was computed using the same method and parameters applied to the video vibration signals. The 30-min accelerometer spectra had a frequency resolution of 0.003 Hz and approximately 44 degrees of freedom.

2.4 Video Processing Validation

We validated the video processing methods using five videos from the Trout Lake dataset and three videos from the Manitou dataset. For the Trout Lake dataset, we focused our analysis on a 30x30 pixel ROI spanning the sky-trunk boundary and applied three video processing methods: VVS with average spectrum aggregation, VVS with modal aggregation of spectral peaks, and MBT with average spectrum aggregation. For the Manitou dataset, we extracted sway frequencies for two target trees, Tree 1 and Tree 2, using 30x80 and 70x175 pixel ROIs, respectively. Since ROIs enclosing the target trees included mostly crown features that lacked distinct boundaries, we only applied the two VVS video processing methods. For all video and accelerometer data, we assumed tree sway did not change substantially over a 30-min interval. However, this assumption is invalid during events (e.g, precipitation interception/unloading) when tree mass can change substantially in less than 30 min.

3 Results

Sway frequencies extracted from video and accelerometer data for both sites are compiled in Table 1. Values have been rounded to two decimal places to be consistent with the frequency resolution associated with the video data of each site. Across all trials and video processing methods, the median and maximum observed differences between video and short-segment accelerometer sway frequencies were 0.03 Hz and 0.26 Hz, respectively. Meanwhile, the median and maximum observed differences between video and 30-min accelerometer sway frequencies were 0.00 Hz and 0.02 Hz, respectively.

3.1 Single Tree Validation - Trout Lake

For the Trout Lake dataset, the maximum differences between video and 30-min accelerometer sway frequencies was 0.02 Hz. Meanwhile, the maximum differences between video and 60-s accelerometer sway frequencies was 0.26 Hz. Even when the accelerometer

signals yielded spectra with no obvious dominant peak (e.g. the 60 s spectrum in Figure 3e), the video processing methods yielded spectra with a distinct peak.

Video identification		Video processing peak frequency (Hz)			Accelerometer peak frequency (Hz)	
Video ID	Video timestamp (local time)	VVS average spectrum	VVS frequency histogram	MBT average spectrum	Short accelerometer segement (same length as video)	Long accelerometer segment (30 min)
Trout Lake (60 s videos)						
Trout-A	2019-8-15 17:07:18	0.21	0.21	0.21	0.26	0.23
Trout-B	2019-8-15 17:30:01	0.24	0.23	0.25	0.29	0.25
Trout-C	2019-08-15 17:32:34	0.23	0.24	0.23	0.49	0.24
Trout-D	2019-08-15 18:00:01	0.24	0.24	0.24	0.31	0.24
Trout-E	2019-08-15 18:09:06	0.24	0.24	0.24	0.23	0.24
Manitou, Tree 1 (30 s videos)						
Manitou-A1	2020-08-15 12:06:00	0.29	0.29	-	0.29	0.29
Manitou-B1	2020-08-20 17:28:00	0.30	0.30	-	0.27	0.29
Manitou-C1	2020-08-31 11:50:00	0.30	0.31	-	0.29	0.30
Manitou, Tree 2 (30 s videos)						
Manitou-A2	2020-08-15 12:06:00	0.38	0.39	-	0.37	0.37
Manitou-B2	2020-08-20 17:28:00	0.37	0.37	-	0.39	0.38
Manitou-C2	2020-08-31 11:50:00	0.38	0.38	-	0.37	0.38

Table 1. Video and Accelerometer Sway Frequency by Site and Tree

Considering video sample Trout-D as an example (Figure 2), the VVS peak frequency heat maps show that pixels with relatively prominent frequency peaks were concentrated around the trunk-sky boundary. Masking removed the pixels that overlapped with the sky and regions of the trunk with poor contrast. The VVS average spectrum, VVS frequency distribution, and MBT average spectrum all yielded the same peak frequency of 0.24 Hz. The 60-s and 30-min accelerometer segments yielded sway frequencies of 0.31 Hz and 0.24 Hz, respectively.

3.2 Multitree Validation – Manitou

For the Manitou dataset, we examined video output for a stand of trees, as illustrated for the Manitou-B video sample (Figure 3). Sway frequencies extracted for Tree 1 and Tree 2 (marked in Figure 3) from three video samples differed from accelerometer sway frequencies on the order of hundredths of Hertz. The maximum difference between video and 30-s accelerometer sway frequencies for Tree 1 and Tree 2 were 0.03 Hz and 0.02 Hz, respectively. The maximum difference between video and 30-min accelerometer sway frequencies for Tree 1 and Tree 2 were 0.01 Hz and 0.02 Hz, respectively. Unmetered trees adjacent to the target trees were observed to have distinct sway frequencies (Figure 3b).

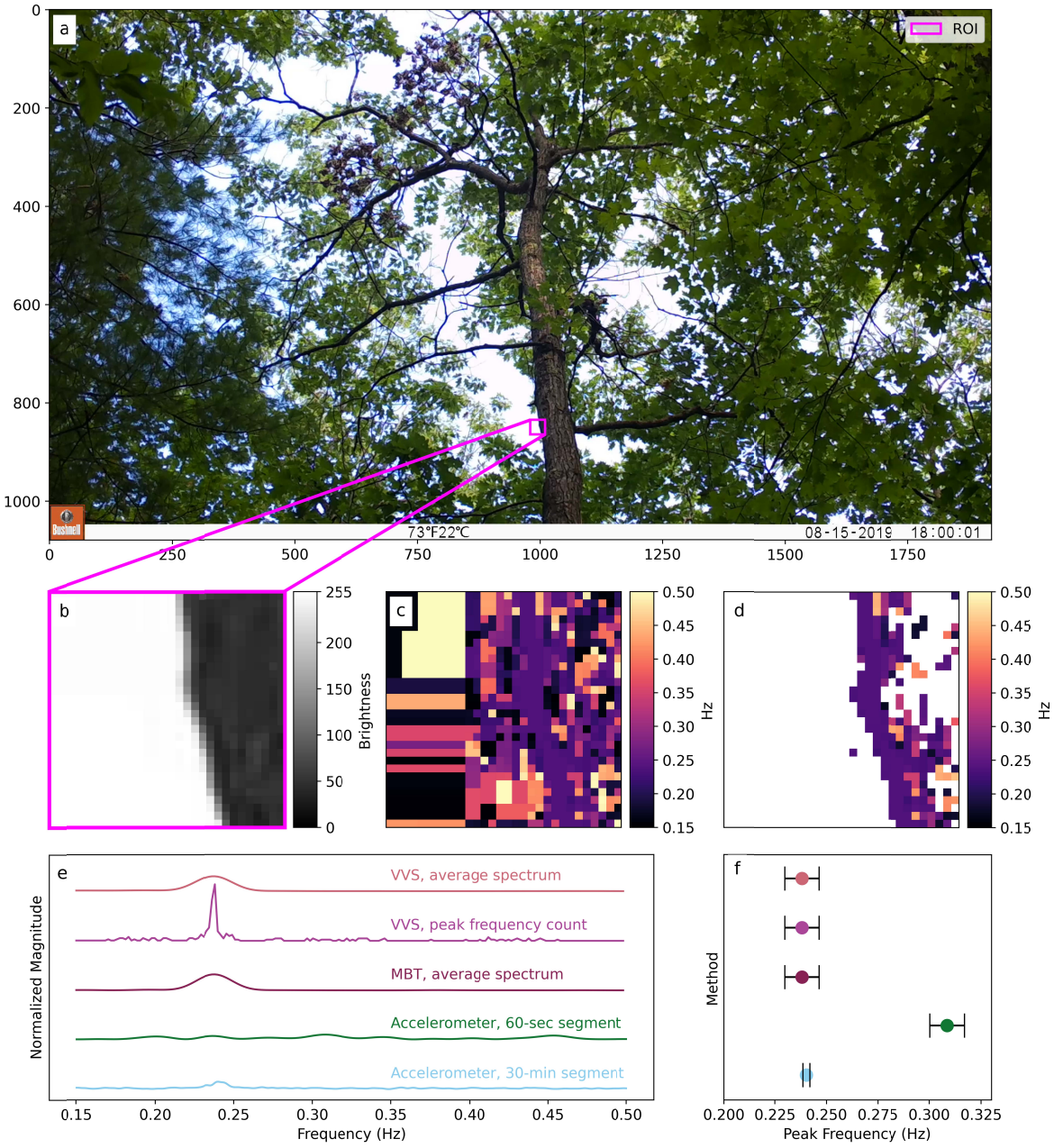

Figure 2. Output for Trout-D. All spectra/distributions have been normalized for comparison. **(a)** Camera FOV. **(b)** Grayscale region of interest. **(c)** Unmasked frequency heat mask. **(d)** Masked frequency heat map where only pixels whose peak prominence is above the 75th percentile of all prominences are shown. **(e)** Video and accelerometer spectra. **(f)** Peak frequency and corresponding frequency resolution by method.

4 Discussion

4.1 Video Processing Performance

Despite inconsistent agreement with short-segment accelerometer sway frequencies, the video processing output agreed well with 30-min accelerometer sway frequencies in all cases

(maximum difference of 0.03 Hz). If we assume stationarity of sway frequency over a 30-min period, the 30-min accelerometer sway frequencies provide a robust estimate of the tree's sway frequency. Thus, the results suggest that when applied to 30-60 s videos, the video processing methods can capture tree sway for one or more trees with approximately 0.03 Hz accuracy.

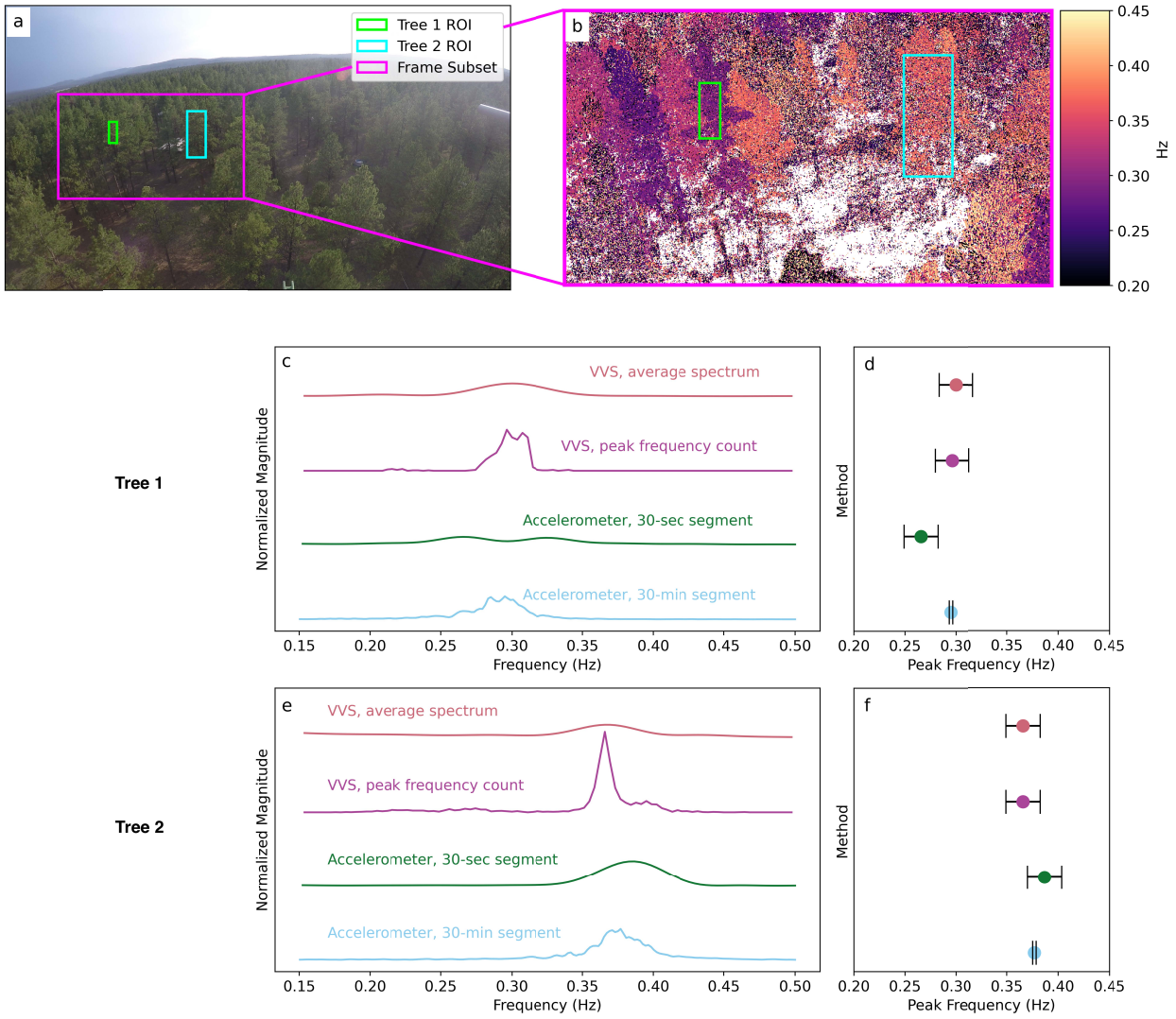

Figure 3. Output for Manitou-B. Only the VVS methods are reported since MBT was not suitable for the ROI enclosing the crown. All spectra/distributions have been normalized for comparison. **(a)** Camera FOV and ROIs. **(b)** Masked frequency heat map for a subset of the FOV **(c)** Video and accelerometer spectra for Tree 1. **(d)** Peak frequency and frequency resolution by method for Tree 1. **(e)** Video and accelerometer spectra for Tree 2. **(f)** Peak frequency and frequency resolution by method for Tree 2.

Differences between sway frequencies extracted from videos and analogous accelerometer segments (same start time, duration, and processing parameters as the videos) can be attributed to several factors. For the Trout Lake dataset, the accelerometer vibration signals were observed to have poor SNRs, leading to spectra with no single dominant peak and the

extraction of spurious sway frequencies (e.g. Trout-C). Poor accelerometer SNRs may have been due to insufficient wind forcing, chaotic motion before the natural sway oscillation, or the trees's broadleaf canopy architecture (Jackson et al., 2019). Despite weak accelerometer signals, the video processing methods were still able to generate spectra with prominent peaks, highlighting the benefit of spatial aggregation. For the Manitou dataset, the maximum error (0.03 Hz) between video and 30-s accelerometer sway frequencies can be partially explained by the frequency resolution. However, the video vibration signals also tended to be noisier and were likely impacted by differences in ROI and accelerometer positions, branch interference, and camera motion.

All three video processing algorithms performed similarly, despite constructing and/or aggregating vibration signals differently. However, each algorithm has practical tradeoffs that make it better suited for certain applications. For general use, we recommend using VVS with average spectrum aggregation. Unlike MBT, VVS can be used to generate a frequency heat map and resolve frequencies from ROIs with no distinct boundary motion (e.g., the tree crowns at Manitou). Additionally, the average spectrum aggregation yields a PSD that is more statistically robust (more degrees of freedom). The MBT method may be preferable when computational resources are a concern, the ROI has a distinct boundary (e.g., Trout Lake), and/or there are significant lighting changes in the video.

4.2 Ecohydrology Capabilities and Limitations

The applicability of the video processing methods to ecohydrologic studies depends on the expected magnitude of sway frequency changes and the length of videos that can be recorded. The existing videos were limited to 30 s and 60 s segments, which meant we could only achieve frequency resolutions of 0.033 Hz and 0.016 Hz, respectively. We used zero-padding to interpolate to a finer resolution and improve our peak estimates, but we were unable to reliably measure frequency with 0.01 Hz precision. The results indicate that using the video processing algorithms with 30-60 s videos provides enough accuracy to resolve changes in sway frequency on the order of 0.1 Hz. For some trees, this is sufficient for observing drought stress (Ciruzzi & Loheide II, 2019), estimating the timing of major phenological changes (Jaeger et al., 2022), and quantifying the timing and magnitude of snow interception (Raleigh et al., 2022). Currently, few methods exist to observe these processes at the stand scale (e.g., Friesen et al., 2014). To observe changes in sway frequency an order of magnitude smaller, longer videos are needed (discussed more below). This may be necessary, for example, to characterize diurnal variations in tree water content (Ciruzzi & Loheide II, 2019), phenology changes (Gougherty et al., 2018), and rain interception (Van Emmerick et al., 2017).

Results from the Manitou dataset suggest the video processing methods scale well to multiple trees and provide a spatial advantage over accelerometers. For example, in Figure 3b, we can observe that Tree 1 has a lower sway frequency than its neighbors. The identification of distinct sway frequencies in the adjacent unmetered trees coupled with the accuracy of extracted sway frequencies demonstrates the promise of using videos to measure sway frequency across a stand. However, the video processing methods have several tradeoffs. First, visible light video data is only useful during daylight hours and in clear conditions (e.g., no fog). Second, given the memory limitations of collecting video data, obtaining sway frequency measurements with high temporal resolution requires considerable storage and/or real time processing. Third, when no suitable trunk ROI is available, it can be difficult to disentangle crown signals and interpret the

output frequency. Accelerometers may be preferable when measurement accuracy and density are a priority.

4.3 Deployment Considerations

The capacity to extract meaningful sway frequency values from video data depends heavily on how the video data are collected. Simultaneously, field site constraints and the memory burden associated with video data impose considerable practical limitations. To reduce the storage and processing overhead while preserving the desired signal processing properties, special care must be taken when choosing the video parameters, camera position, and target tree(s).

To reliably measure frequency with 0.01 Hz precision, which is needed for many tree sway applications, videos need to be at least 100 s long. However, we recommend recording 10-min videos to improve the SNR and enable use of more robust spectral methods (e.g., Welch's method). For studies requiring fine temporal resolution over long periods of time, 10-min videos may result in hundreds of gigabytes of data. When memory is a constraint, we recommend prioritizing video length over video quantity. In most cases, the memory burden can also be reduced by choosing a sufficiently small framerate, recording in grayscale, and eliminating audio. Following the Shannon-Nyquist sampling theorem, we advise choosing a minimum framerate around 8-10 fps.

Frame resolution, lens type, distance between the camera and target trees, and tree features control the size of observed sway and corresponding signal strength. We suggest placing the camera as close to the target trees as possible and using the largest frame resolution that memory constraints allow, after first prioritizing frequency resolution and collection interval. It is unclear how the lens FOV angle affects sway measurements. Therefore, we encourage avoiding wide-angle lenses. We suggest positioning the camera so contrast between the sky and trees is maximized. When the camera is at a high point looking down (e.g., at Manitou) and/or when there is a dense forest canopy, it can be difficult to identify individual trees and choose ROIs with suitable contrast. We recommend limiting video processing analysis to tall, slender trees that are well represented by the cantilever beam model and whose trunk sway dominates branch sway (Jackson et al., 2019). We also recommend choosing trees with significant portions of unobstructed trunk. Unlike sway signals from the crown, sway signals from the trunk-background boundary are straightforward to interpret and usually less noisy.

4.4 Directions for Future Work

Future work should apply the video processing methods to longer videos to determine how well they can measure small changes in sway frequency. Further analysis, particularly of video time series over large time scales (weeks, months), is needed to clarify when and where video processing can infer tree properties. Future work could begin by comparing sway frequency time series extracted from video and accelerometer data when there are large expected changes in sway frequency (e.g., snow interception, leaf drop). Another major advantage of video analysis is the ability to measure sway frequencies of individual branches independent of the main stem, but more work is needed to quantify when and where this is possible. Experimentation with different camera systems could also be studied. Cameras recording at the near infrared may provide high-contrast ROIs through use of the Normalized Difference Vegetation Index (NDVI). Joint camera-accelerometer systems could be developed to leverage

the advantages of both sensors. Significant improvements in the processing algorithm are also likely to improve the frequency resolution and refine the aggregation across pixels in a tree.

5 Conclusions

We demonstrate that using 30-60s videos, VVS and MBT Eulerian video processing algorithms can measure tree sway frequency to within approximately 0.03 Hz of accelerometer sway frequencies at multiple points in a camera field of view. The results indicate video processing can be used to collect tree sway data over greater spatial extents, improving upon a key limitation of accelerometers. Our analysis suggests that when applied to 30-60 s videos, the technology is suitable for detecting changes in sway frequency on the order of 0.1 Hz, which is sufficient to resolve changes due to snow interception and phenology variations but not diurnal drought stress, as based on currently published values (Raleigh et al., 2022; Jaeger et al, 2022; Ciruzzi & Loheide II, 2019). For best results, we recommend recording long videos (~10 min), choosing trees with unobstructed segments of trunk, and positioning the camera so contrast between the trunk and background is maximized. Further applications in the field should test the extent and effectiveness of video processing to infer tree properties.

Acknowledgments

Joseph Ammatelli and Jessica Lundquist received funding support from The National Aeronautics and Space Administration (NASA Grant NNX17AL59G). For Ethan Gutmann, this material is based upon work supported by the National Center for Atmospheric Research, which is a major facility sponsored by the National Science Foundation under Cooperative Agreement No. 1852977. Holly Barnard and Sidney Bush received funding support from the National Science Foundation (NSF 2012669) and the University of Colorado – Boulder Research and Innovation Office. Contributions from Dominick Ciruzzi and Steven Loheide II were sponsored by the National Science Foundation Division of Earth Sciences (NSF EAR-1700983). Mark Raleigh was supported by the National Science Foundation (NSF EAR Award 1761441).

Open Research

Data Availability Statement

All datasets used for this study are freely available in public repositories. Video samples, accelerometer data, and extracted sway frequencies (Table 1) for the oak tree in Trout Lake,

Wisconsin (Ammatelli, Ciruzzi, Loheide II, & Lundquist, 2023) are available at https://doi.org/10.5281/zenodo.7988211 . More accelerometer data for the Trout Lake target tree (Ciruzzi & Loheide II, 2021a) can be found using the TL_Tree1 identifier in the repository https://www.hydroshare.org/resource/38ae9d9fb88d49f9ad2eed1ee07475c0/ . Video samples, accelerometer data, and extracted sway frequencies (Table 1) for the two target ponderosa pine trees at the Manitou Experimental Forest, Colorado (Ammatelli, Bush, Barnard, & Lundquist, 2023) are available at https://doi.org/10.5281/zenodo.7988227 .

Software Availability Statement

All python code used for tree sway analysis and the development of subsequent plots (Ammatelli, 2023) is hosted at https://github.com/j-amma/swayfreq and is preserved at https://doi.org/10.5281/zenodo.7992823 .

References

Ammatelli, J. H. (2023). j-amma/swayfreq: v1.0.0 (v1.0.0) [software]. Zenodo. https://doi.org/10.5281/zenodo.7992823

Ammatelli, J. H., Bush, S. A., Barnard, H. R., & Lundquist, J. D. (2023). Video and accelerometer tree sway data for two ponderosa pine trees in the Manitou Experimental Forest, Colorado (Version 1) [Dataset]. Zenodo. https://doi.org/10.5281/zenodo.7988227

Ammatelli, J. H., Ciruzzi, D. M., & Loheide II, S. P, & Lundquist, J. D. (2023). Video and accelerometer tree sway data for an oak tree in Trout Lake, Wisconsin (Version 1) [Dataset]. Zenodo. https://doi.org/10.5281/zenodo.7988211

Baker, C. J. (1997). Measurements of the natural frequencies of trees. *Journal of Experimental Botany*, 48(5), 1125–1132. https://doi.org/10.1093/jxb/48.5.1125

425 Bush, S. A. (2022). Ecohydrologic Processes in the Montane Headwaters of the Upper South
426 Platte River (Doctoral dissertation). Retrieved from ProQuest Dissertations Publishing.
427 (cub.b12869881). Boulder, CO: University of Colorado at Boulder.
428 Chen, J. G., Davis, A., Wadhwa, N., Durand, F., Freeman, W. T., & Büyüköztürk, O. (2017).
429 Video Camera–Based Vibration Measurement for Civil Infrastructure Applications.
430 *Journal of Infrastructure Systems*, 23(3), B4016013.
431 https://doi.org/10.1061/(ASCE)IS.1943-555X.0000348
432 Ciruzzi, D. M., & Loheide II, S. P. (2019). Monitoring Tree Sway as an Indicator of Water
433 Stress. *Geophysical Research Letters*, 46(21), 12021–12029.
434 https://doi.org/10.1029/2019GL084122
435 Ciruzzi, D.M., Loheide II, S. P. (2021a). Continuous acceleration time series for tree sway
436 monitoring in a temperate, humid environment [Dataset]. HydroShare.
437 http://www.hydroshare.org/resource/38ae9d9fb88d49f9ad2eed1ee07475c0
438 Ciruzzi, D. M., & Loheide II, S. P. (2021b). Monitoring Tree Sway as an Indicator of
439 Interception Dynamics Before, During, and Following a Storm. *Geophysical Research*
440 *Letters*, 48(20), e2021GL094980. https://doi.org/10.1029/2021GL094980
441 Clark, J. S., Iverson, L., Woodall, C. W., Allen, C. D., Bell, D. M., Bragg, D. C., D'Amato, A.
442 W., Davis, F. W., & et al. (2016). The impacts of increasing drought on forest dynamics,
443 structure, and biodiversity in the United States. *Global Change Biology*, 22(7), 2329–
444 2352. https://doi.org/10.1111/gcb.13160
445 Ferrer, B., Espinosa, J., Roig, A. B., Perez, J., & Mas, D. (2013). Vibration frequency
446 measurement using a local multithreshold technique. *Optics Express*, 21(22), 26198–
447 26208. https://doi.org/10.1364/OE.21.026198

448 Gougherty, A. V., Keller, S. R., Kruger, A., Stylinski, C. D., Elmore, A. J., & Fitzpatrick, M. C.
449 (2018). Estimating tree phenology from high frequency tree movement data. *Agricultural*
450 *and Forest Meteorology*, 263, 217–224. https://doi.org/10.1016/j.agrformet.2018.08.020
451 Jackson, T. D., Sethi, S., Dellwik, E., Angelou, N., Bunce, A., van Emmerik, T., Duperat, M.,
452 Ruel, J.-C., & et al. (2020). The motion of trees in the wind: A data synthesis [Preprint].
453 Biodiversity and Ecosystem Function: Terrestrial. https://doi.org/10.5194/bg-2020-427
454 Jackson, T., Shenkin, A., Moore, J., Bunce, A., van Emmerik, T., Kane, B., Burcham, D.,
455 James, K., & et al. (2019). An architectural understanding of natural sway frequencies in
456 trees. *Journal of The Royal Society Interface*, 16(155), 20190116.
457 https://doi.org/10.1098/rsif.2019.0116
458 Jaeger, D. M., Looze, A. C. M., Raleigh, M. S., Miller, B. W., Friedman, J. M., & Wessman, C.
459 A. (2022). From flowering to foliage: Accelerometers track tree sway to provide high-
460 resolution insights into tree phenology. *Agricultural and Forest Meteorology*, 318,
461 108900. https://doi.org/10.1016/j.agrformet.2022.108900
462 Kooreman, B. (2013). Measuring weight fluctuations in trees based on natural frequency.
463 Retrieved from TU Delft Repositories.
464 (https://repository.tudelft.nl/islandora/object/uuid%3Ab5c5466b-8165-4871-81a3-
465 db8574cf1fd1). Delft, Netherlands: Delft University of Technology.
466 Lundquist, J. D., Dickerson-Lange, S., Gutmann, E., Jonas, T., Lumbrazo, C., & Reynolds, D.
467 (2021). Snow interception modelling: Isolated observations have led to many land surface
468 models lacking appropriate temperature sensitivities. *Hydrological Processes*, 35(7),
469 e14274. https://doi.org/10.1002/hyp.14274
470 Oppenheim, A. V., & Schafer, R. W. (2010). Discrete-time Signal Processing (Third edition.).

Upper Saddle River, NJ : Pearson.

Raleigh, M. S., Gutmann, E. D., Van Stan II, J. T., Burns, S. P., Blanken, P. D., & Small, E. E. (2021). Challenges and Capabilities in Estimating Snow Mass Intercepted in Conifer Canopies with Tree Sway Monitoring. *Water Resources Research*, 58, e2021WR030972. https://doi.org/10.1029/2021WR030972

Schumacher, T., & Shariati, A. (2013). Monitoring of Structures and Mechanical Systems Using Virtual Visual Sensors for Video Analysis: Fundamental Concept and Proof of Feasibility. *Sensors*, 13(12), Article 12. https://doi.org/10.3390/s131216551

Selker, J. S., Lane, J. W., Rupp, D. E., Hut, R., Abou Najm, M. R., Stewart, R. D., Van De Giesen, N., & Selker, F. (2011). The answer is blowing in the wind: Using wind induced resonance of trees to measure time varying canopy mass, including interception. 2011, H11G-1155.

Storck, P., Lettenmaier, D. P., & Bolton, S. M. (2002). Measurement of snow interception and canopy effects on snow accumulation and melt in a mountainous maritime climate, Oregon, United States. *Water Resources Research*, 38(11), 5-1-5–16. https://doi.org/10.1029/2002WR001281

Van Emmerik, T., Steele-Dunne, S., Hut, R., Gentine, P., Guerin, M., Oliveira, R. S., Wagner, J., Selker, J., & Van de Giesen, N. (2017). Measuring Tree Properties and Responses Using Low-Cost Accelerometers. *Sensors*, 17(5), Article 5. https://doi.org/10.3390/s17051098

Wang, A., Yang, X., & Xin, D. (2022). The Tracking and Frequency Measurement of the Sway of Leafless Deciduous Trees by Adaptive Tracking Window Based on MOSSE. *Forests*, 13(1), Article 1. https://doi.org/10.3390/f13010081

494 Welch, P. (1967). The use of fast Fourier transform for the estimation of power spectra: A
495 method based on time averaging over short, modified periodograms. *IEEE Transactions*
496 *on Audio and Electroacoustics*, 15(2), 70–73. https://doi.org/10.1109/TAU.1967.1161901