Supplement for: Probing Environmental and Tectonic changes underneath Mexico City with the urban seismic field

Laura Ermert[1, **], Enrique Cabral-Cano[2, *], Estelle Chaussard[3, *], Darío Solano Rojas[4, *], Luis Quintanar[2, *], Diana Morales Padilla[4], Enrique A. Fernández-Torres[2], and Marine A. Denolle[1]

[1]Department of Earth and Space Sciences, University of Washington, Seattle WA, USA
[**]Now at: Swiss Seismological Service, ETH Zürich, Zürich, Switzerland
[3]Independent Researcher
[2]Instituto de Geofísica, Universidad Nacional Autónoma de México, CDMX 04510, México
[4]Facultad de Ingeniería, Universidad Nacional Autónoma de México, CDMX 04510, México
[*]Who contributed equally to this work

Correspondence: Laura Ermert (laura.ermert@sed.ethz.ch)

1 Content

This supplementary document provides additional information about the methodology used in the study, in particular the waveform clustering technique, the variable reference for the monitoring approach, and the Bayesian inversion. We also provide the following additional figure to support the main manuscript statements:

- Figure S1: Comparison of waveform clustering results for phase auto-correlation and raw auto-correlation
- Figure S2: Rule-base 1-D velocity model, examples
- Figure S3: Surface wave phase velocity sensitivity kernels with respect to isotropic shear wave velocity
- Figure S4: Effect of hydraulic diffusivity and Poisson ratio on estimated pore pressure changes
- Figure S5: Results of Monte Carlo inversion (RSVM network)
- Figure S6: Results of Monte Carlo inversion (temporary TO deployment)
- Figure S7: Elliptical particle motion

2 Auto-correlation waveform clustering

We adopt the correlation waveform clustering strategy by Viens and Iwata (2020) based on Gaussian Mixture Models, principal component analysis (PCA) for dimensionality reduction, and the Bayesian Information Criterion (BIC) for determining the optimal number of clusters. We introduce the following modifications, which adapt the clustering to urban seismic noise: First, we apply the clustering separately to each octave bands between 0.5 and 8 Hz: 0.5-1 Hz, 1-2Hz, 2-4Hz, and 4-8 Hz. We do

this because anthropogenic noise tends to contain narrow spectral peaks (see Figure 2 in the main text). Second, we perform the PCA on a randomly selected subset of the raw correlations for each station, and subsequently expand all correlations in the principal component basis to proceed with the clustering. This enables us to process much larger data volumes, e.g. 26 years worth of data at the site G.UNM. Third, we normalize each waveform to unity (by peak absolute amplitude) rather than applying a Standard (feature) normalization. Consequently, the clustering is likely dominated by early, high relative amplitude arrivals, which are expected to be particularly sensitive to the incident ambient noise compared to later, scattered arrivals (e.g. Takano et al., 2020). Fourth, we run the clustering twice, first determining the number of clusters for each station which would be picked by the BIC, and then applying the clustering a second time with a fixed, constant number of 4 clusters based on the average optimal number of clusters found. This results in a consistent classification for all stations, observation times, and frequency bands. Finally, instead of using the principal components characteristics to select the "best" cluster, we re-label the clusters according to local time. Contrary to a natural noise setting, the signature of urban activity appeared clearly in our clustering results. This is illustrated in Figure 2 of the main text, bottom panel: Clustering of the same data displayed in the spectrogram above results in a separation of day- and nighttime urban noise correlations. To re-label, we determine the distribution of local time in each cluster, compute its mode (the most frequently occurring hour of the day in this cluster), then assign the label "night" to the cluster with the mode that is closest to 2 am local time. Of the remaining clusters, the largest one is labelled "day" and the smallest one "noise", with an additional unused "other" cluster.

3 Monitoring with a variable reference

When waveform similarity systematically varies over time or strongly decreases after events like earthquakes, a single long-term reference stack may not be representative of the change, and waveform decoherence may prevent the use of seismic interferometry to measure the changes. For this scenario, two solution strategies have emerged: Defining and iteratively improving a long-term reference (Richter et al., 2014) or using multiple references (e.g. Sens-Schönfelder et al., 2014). Brenguier et al. (2014) take the multiple-reference approach to the limit by measuring the relative changes between all possible pairs of short-term stacks, and inverting them for a $\frac{dv}{v}$ timeseries assuming linear change in between observation times, as well as a Gaussian covariance.

In our data, the 2017 Puebla Earthquake causes a marked drop in waveform similarity at several stations. After preliminary testing, we decided to use a multiple-reference approach, which is conceptually similar to the approach of Sens-Schönfelder et al. (2014). We select several time periods to construct references; concretely, we use 1-year stacks for each year. We construct a separate time series of $\frac{dv}{v}$ for each reference period. To concatenate the resulting time series, we consider the $\frac{dv}{v}$ time series valid during the observation period corresponding to its reference, and minimize the offset with respect to the preceding period considering a 10-point overlap. Finally, the mean $\frac{dv}{v}$ is removed.

4 Modeling of the $\frac{dv}{v}$ time series: MCMC inversion setup and convergence

We initialize the runs starting from the maximum likelihood model determined using `scipy.optimize`, and perturbed by small random values. Then, we run 25 000 iterations for 20 chains, of which we discard 10 000 as burn-in. If the auto-correlation period of these chains is no longer than 2 $\%$ of the post-burn-in samples, we assume that a sufficient number of independent samples of the posterior distribution was generated, and terminate the inversion. If it is longer, we continue iterating, adding 25 000 samples. After a maximum of 75 000 samples, we interrupt the inversion, and mark all inversions that still do not have a satisfying result as "not converged".

We run this inversion separately for all three components, four frequency bands, two time windows (earlier / later lags) and 18 stations in order to obtain medians and ranges for parameters s_t, p_0, Δc, τ_{max}, a and b. The diffusivities κ_t, κ_h are not inverted for (see section 4.2.1 in the main manuscript). Instead, we repeat the inversion for 3 (2) trial values for κ_t (κ_h), and select the optimal values at each station.

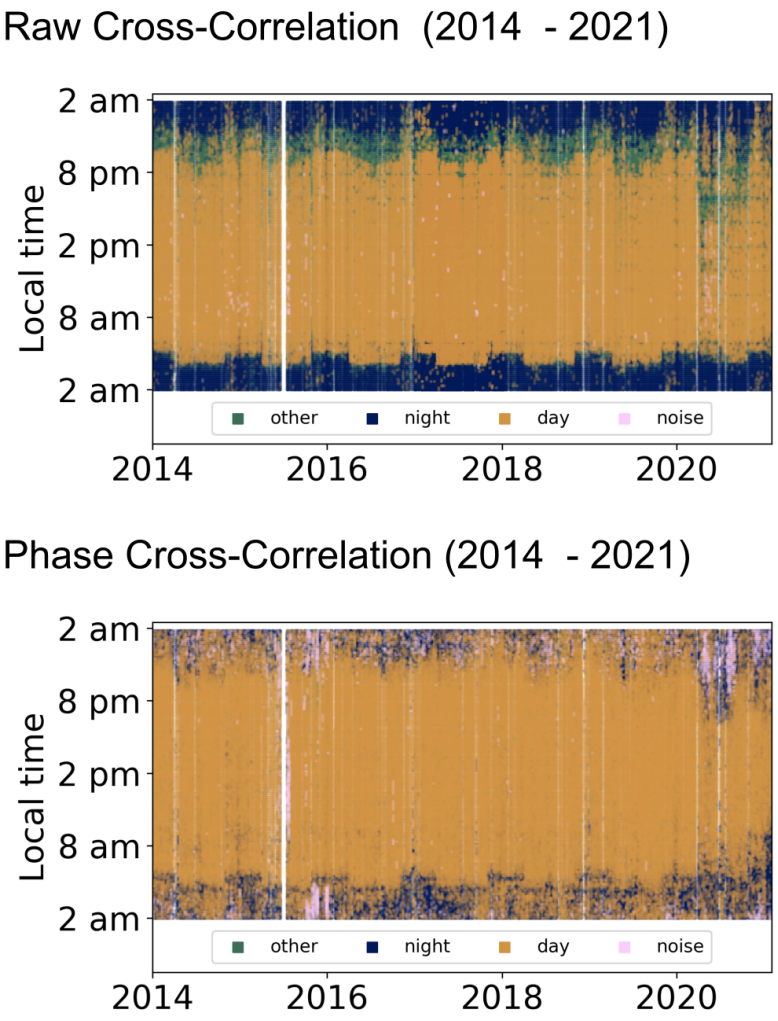

Figure S 1. Comparison of waveform clustering applied to classical auto-correlations (top) and phase auto-correlations (Schimmel et al., 2011, bottom). Phase correlations are amplitude-unbiased so that the influence of ambient noise sources on them is reduced. However, we still observe a diurnal pattern and daylight saving time change in the resulting clusters. Because of this, we expect that the waveform clustering will be effective at reducing source variability, while amplitude and spectral normalizations may be less effective.

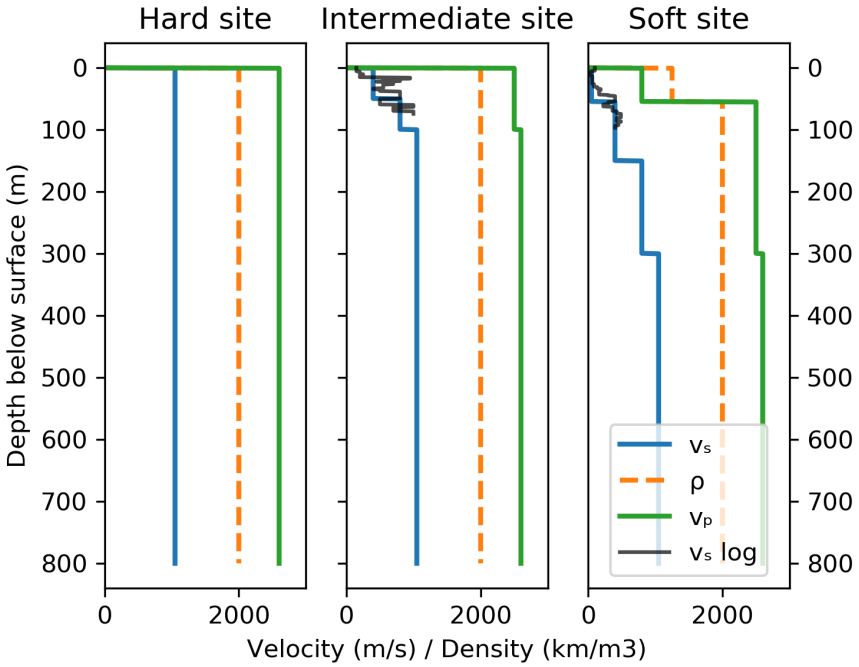

Figure S 2. Rule-based velocity model for hard, intermediate and soft sites. In the case of hard sites, we assume that there is no sediment layer. In the case of intermediate sites, we assume that there are two sediment layer of moderate and increasing velocity close to the surface. For soft sites, we include an additional very low-velocity layer near the surface, and deeper moderate-velocity sediments. The thickness of the shallowest layer varies and is determined by the aquitard model of the basin. The red dashed-dotted line shows redigitized shear wave velocity profiles at intermediate and soft sites from Singh et al. (1995, 1997).

References

Brenguier, F., Campillo, M., Takeda, T., Aoki, Y., Shapiro, N., Briand, X., Emoto, K., and Miyake, H.: Mapping pressurized volcanic fluids from induced crustal seismic velocity drops, Science, 345, 80–82, 2014.

Fichtner, A.: surf, https://github.com/afichtner/surf, 2020.

Richter, T., Sens-Schönfelder, C., Kind, R., and Asch, G.: Comprehensive observation and modeling of earthquake and temperature-related seismic velocity changes in northern Chile with passive image interferometry, J. Geophys. Res., 119, 4747–4765, 2014.

Schimmel, M., Stutzmann, E., and Gallart, J.: Using instantaneous phase coherence for signal extraction from ambient noise data at a local to a global scale, Geophys. J. Int., 184, 494–506, 2011.

Sens-Schönfelder, C., Pomponi, E., and Peltier, A.: Dynamics of Piton de la Fournaise volcano observed by passive image interferometry with multiple references, J. Geophys. Res., 276, 32–45, 2014.

Singh, S., Quaas, R., Ordaz, M., Mooser, F., Almora, D., Torres, M., and Vasquez, R.: Is there truly a “hard” rock site in the Valley of Mexico?, Geophys. Res. Lett., 22, 481–484, 1995.

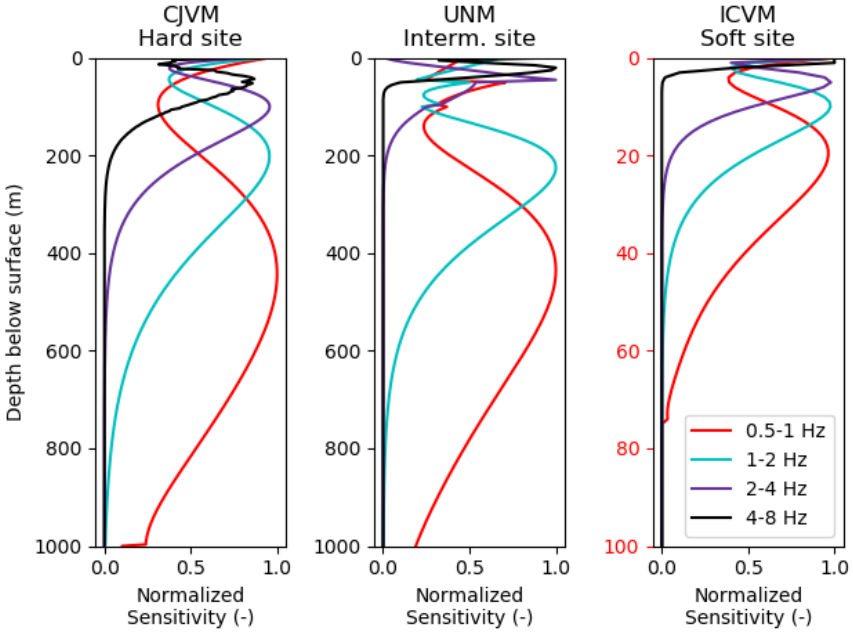

Figure S 3. Rayleigh wave phase velocity sensitivity kernels with respect to isotropic shear wave velocity computed using the surf tool (Fichtner, 2020) for a hard, intermediate and soft site (from left to right). Sensitivity kernels at the hard site have a "classic" shape, while those at the intermediate site include considerable shallow sensitivity in the sedimentary layers, with the sensitivity at higher frequencies concentrated near the surface. Finally, sensitivity at the soft site is entirely concentrated in the upper slow sedimentary layers (note the different depth scale).

Singh, S., Santoyo, M., Bodin, P., and Gomberg, J.: Dynamic deformations of shallow sediments in the Valley of Mexico, Part II: Single-station estimates, Bull. Seismol. Soc. Am., 87, 540–550, 1997.

Takano, T., Brenguier, F., Campillo, M., Peltier, A., and Nishimura, T.: Noise-based passive ballistic wave seismic monitoring on an active volcano, Geophys. J. Int., 220, 501–507, 2020.

Viens, L. and Iwata, T.: Improving the retrieval of offshore-onshore correlation functions with machine learning, J. Geophys. Res., 125, e2020JB019 730, 2020.

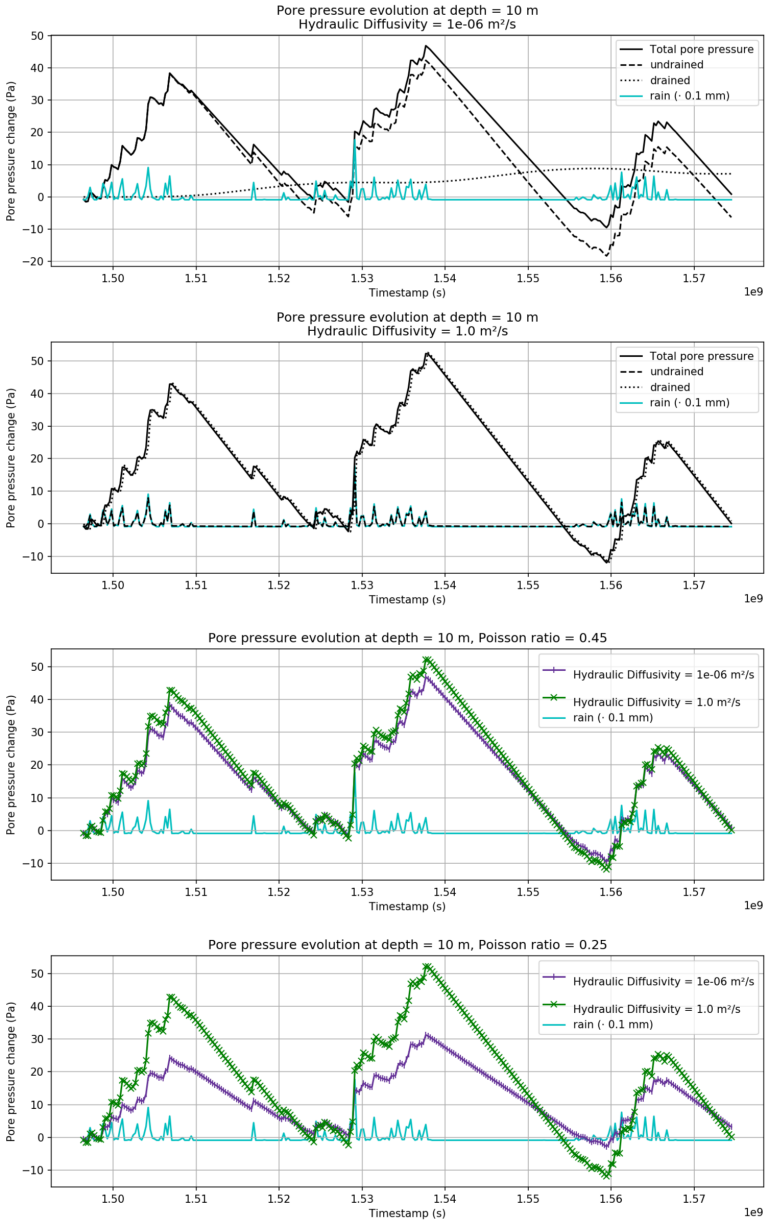

Figure S 4. Illustration of the behaviour of Roeloff's 1-D model for poroelastic pore pressure variation. Top panel: When hydraulic diffusivity is low, the undrained term dominates, while the drained term is strongly smoothed and delayed during diffusion. Second panel: When hydraulic diffusivity is higher, the drained term dominates, but has very little delay and smoothing because diffusion happens very fast. Therefore, the depth-integrated effect of both cases is very similar (panel 3 from top), *unless* Poisson's ratio of the material is low to moderate (bottom panel). Poisson's ratio of the clay-rich sediments in the lake zone is high, resulting in a scenario similar to the 3rd panel. tl, dr: Our $\frac{dv}{v}$ model has low sensitivity to hydraulic diffusivity.

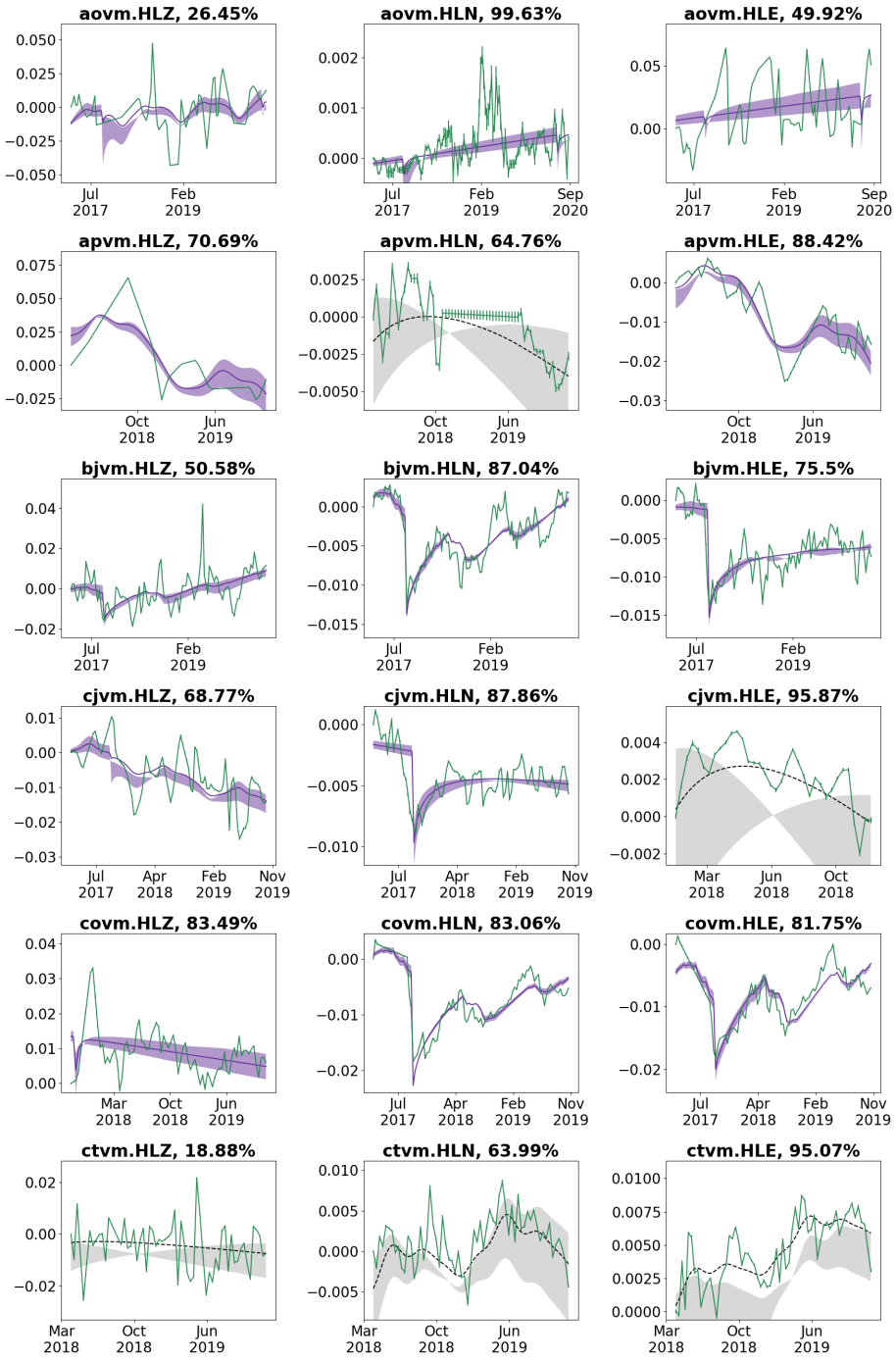

Figure S 5. Results of the Monte Carlo inversion for the RSVM network , stations AOVM-CTVM: Data (green), model (purple line) and range of the 16th–84th percentile models (shaded purple area). Models where the inversion failed to converge are shown in red. Missing panels indicate time series with an insufficient number of high-quality $\frac{dv}{v}$ measurements. This page: 0.5 – 1 Hz, early coda

.

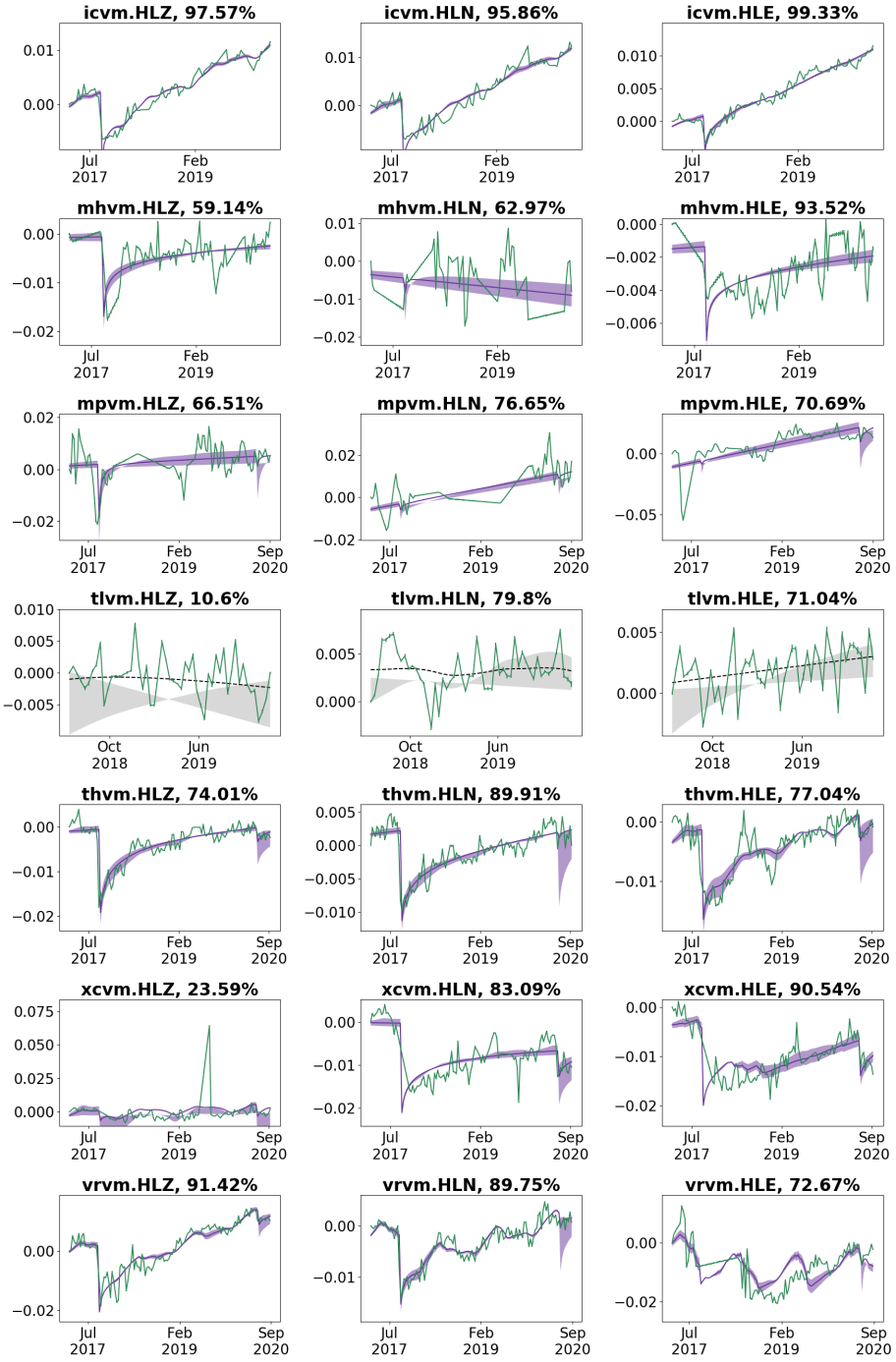

Figure S 5. (continued) RSVM data and models for 0.5 – 1 Hz, stations ICVM-VRVM, early coda

; see above for description.

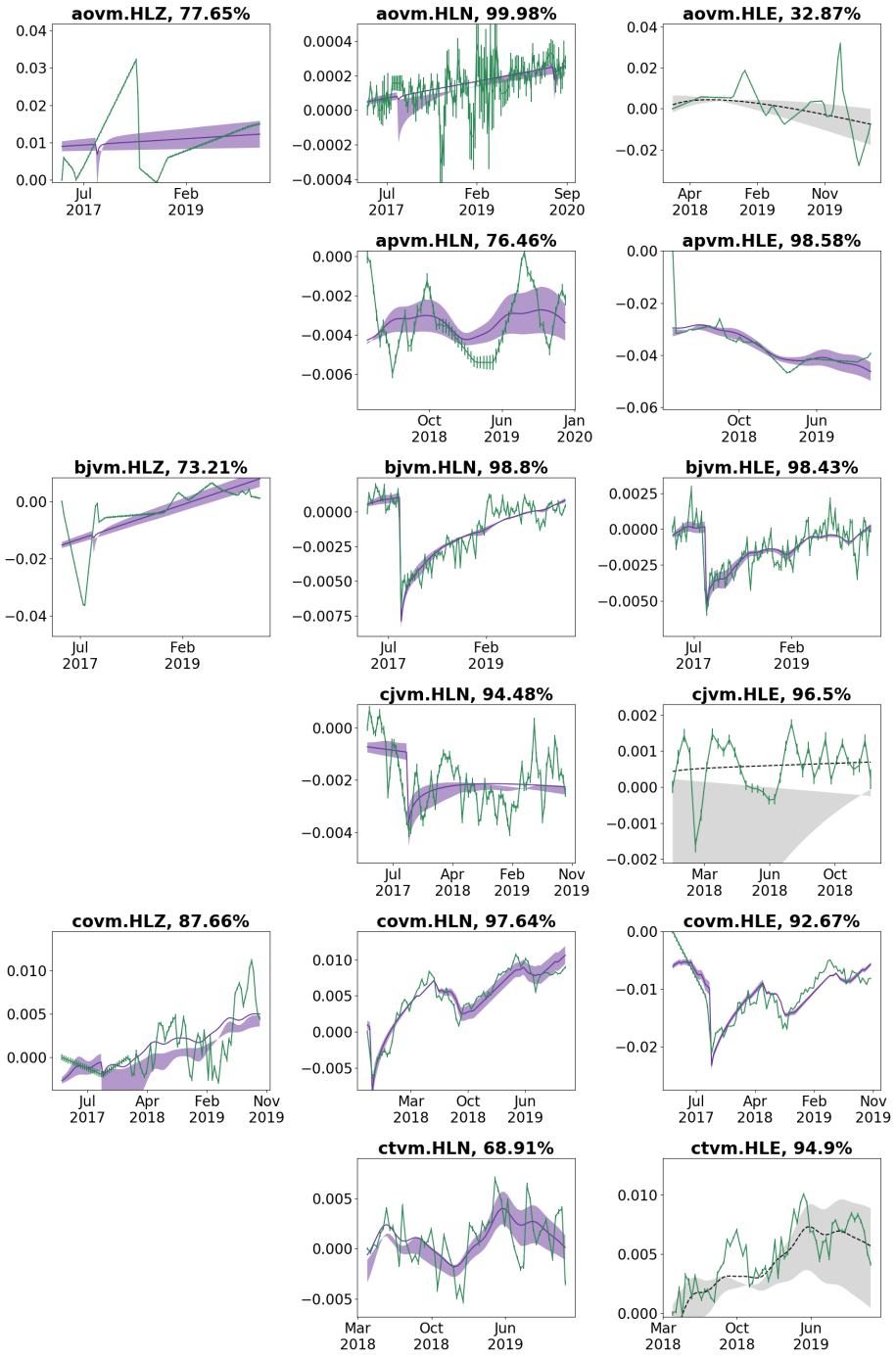

Figure S 5. (continued) RSVM data and models for 0.5 – 1 Hz, stations AOVM-CTVM, late coda; see above for description.

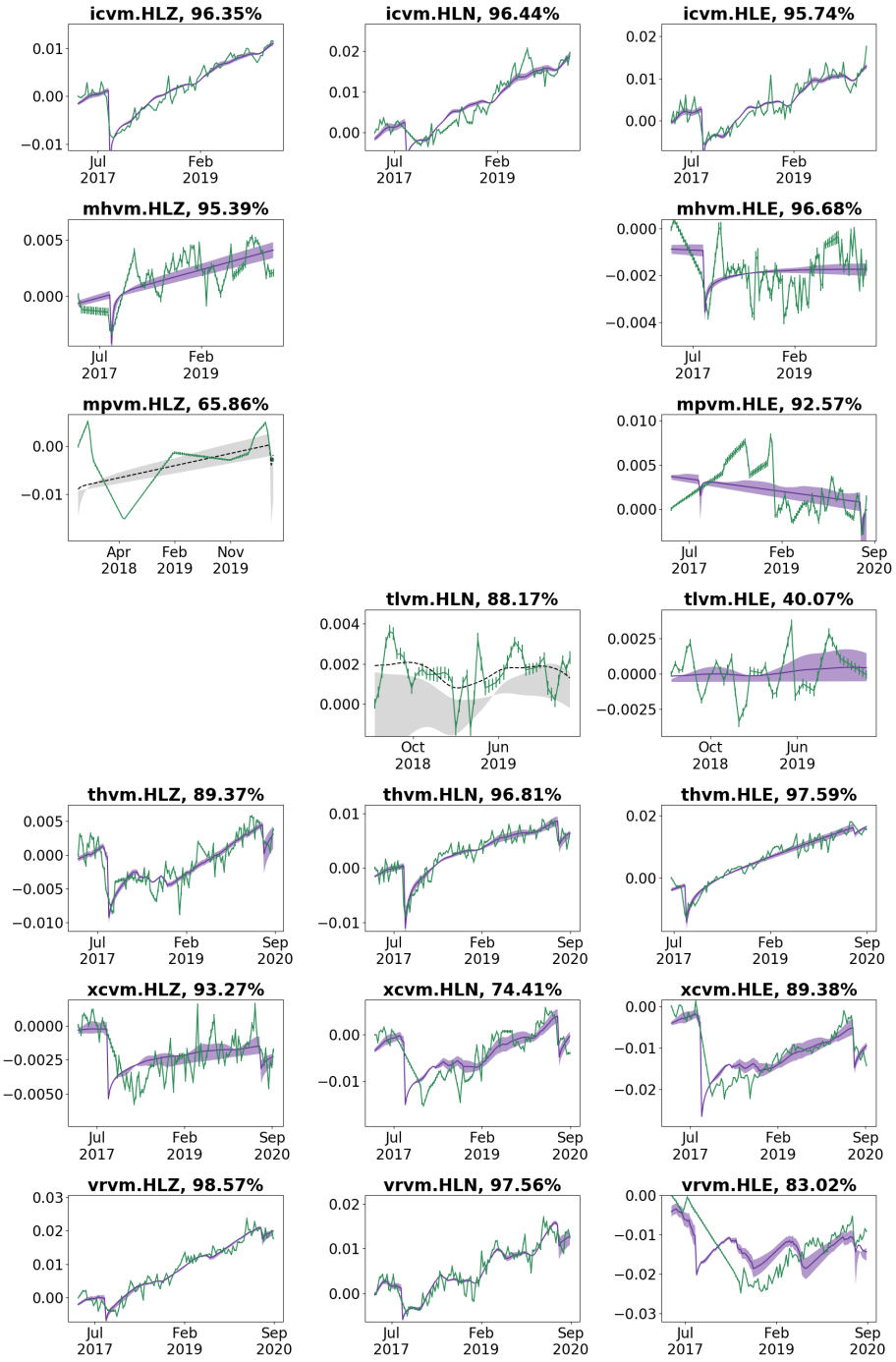

Figure S 5. (continued) RSVM data and models for 0.5 – 1 Hz, stations ICVM-VRVM, late coda; see above for description.

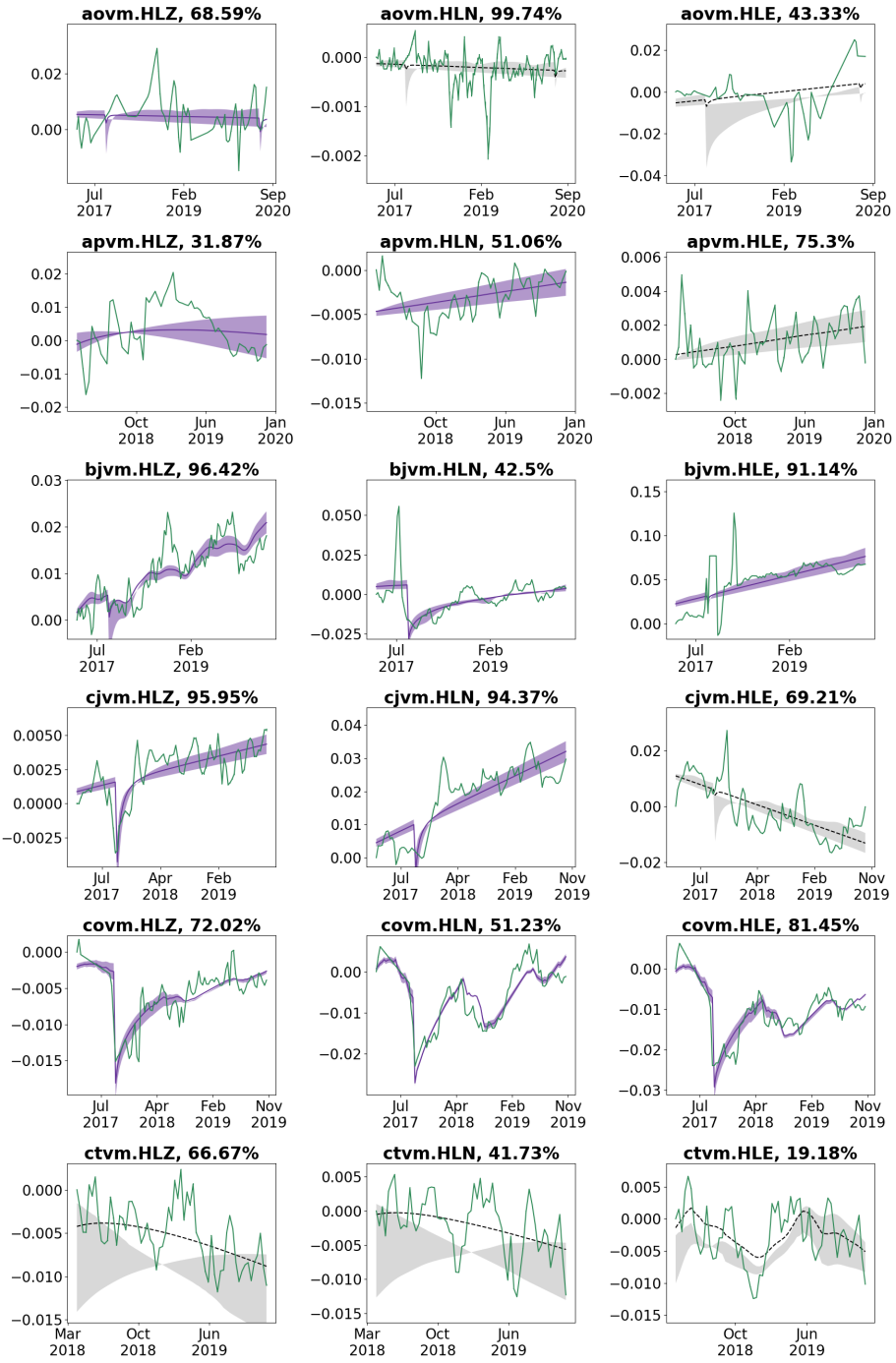

Figure S 5. (continued) RSVM data and models for 1.0 – 2.0 Hz, stations AOVM-CTVM, early coda; see above for description.

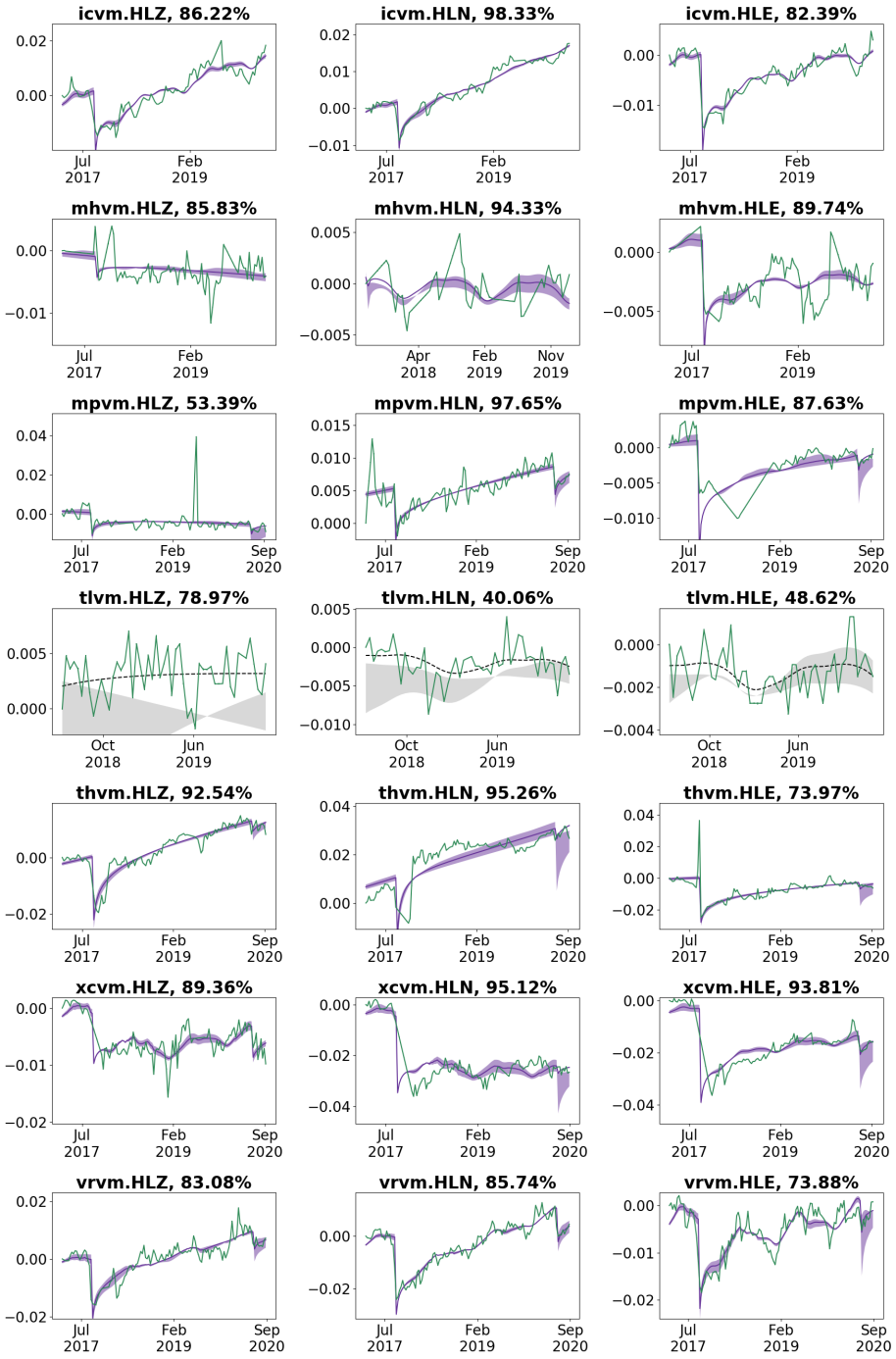

Figure S 5. (continued) RSVM data and models for 1.0 – 2.0 Hz, stations ICVM-VRVM, early coda; see above for description.

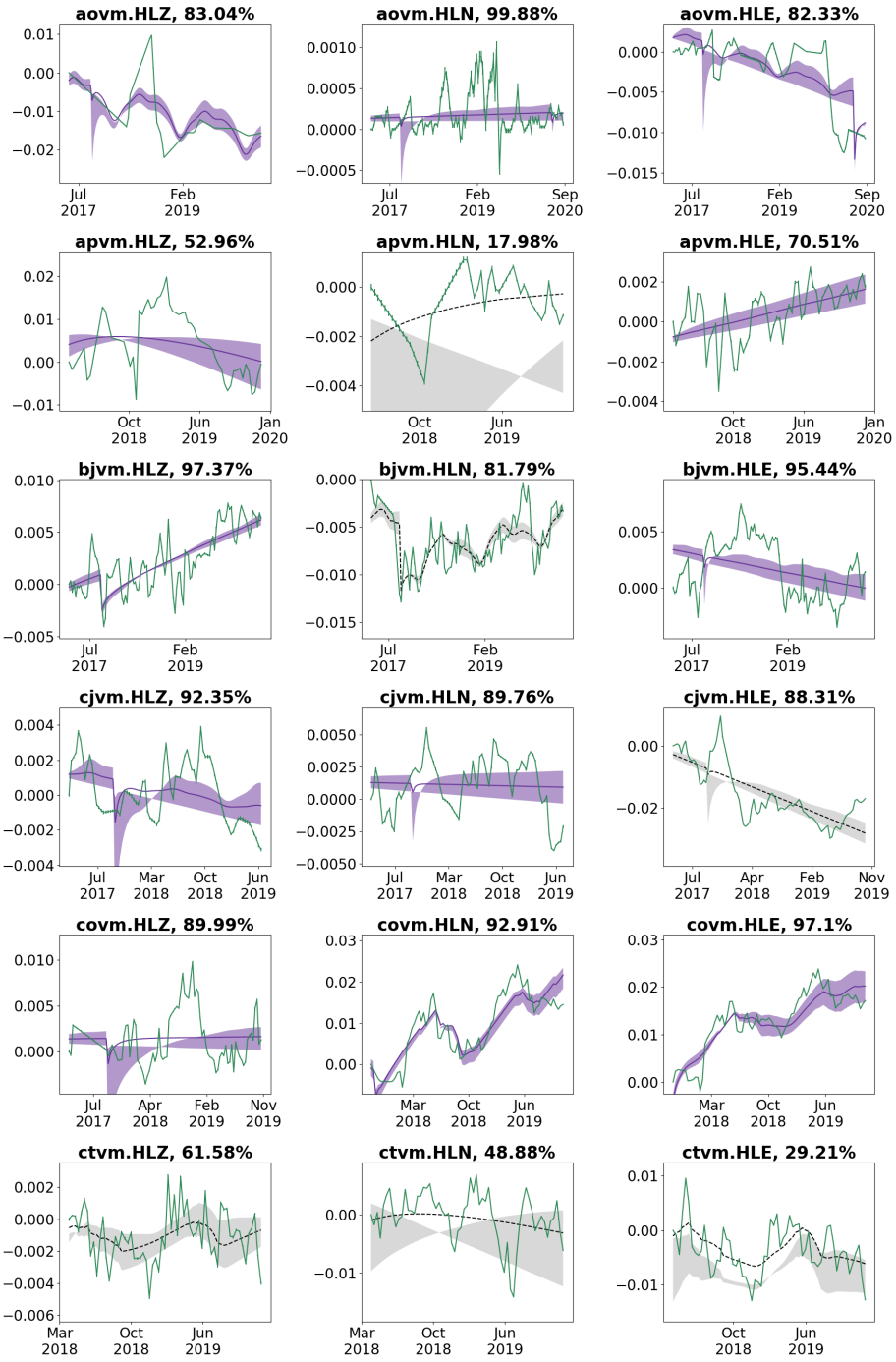

Figure S 5. (continued) RSVM data and models for 1.0 – 2.0 Hz, stations AOVM-CTVM, late coda; see above for description.

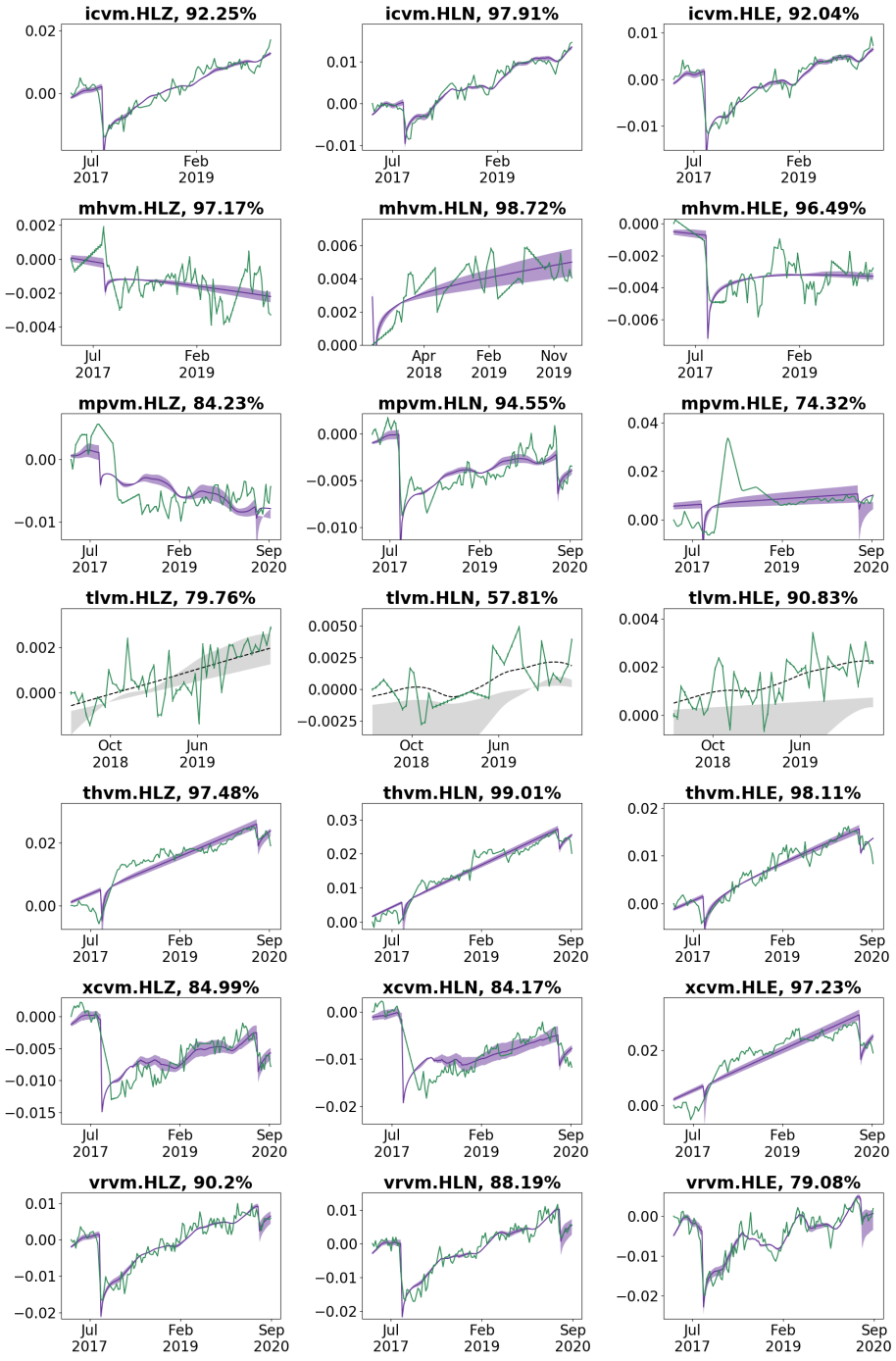

Figure S 5. (continued) RSVM data and models for 1.0 – 2.0 Hz, stations ICVM-VRVM, late coda; see above for description.

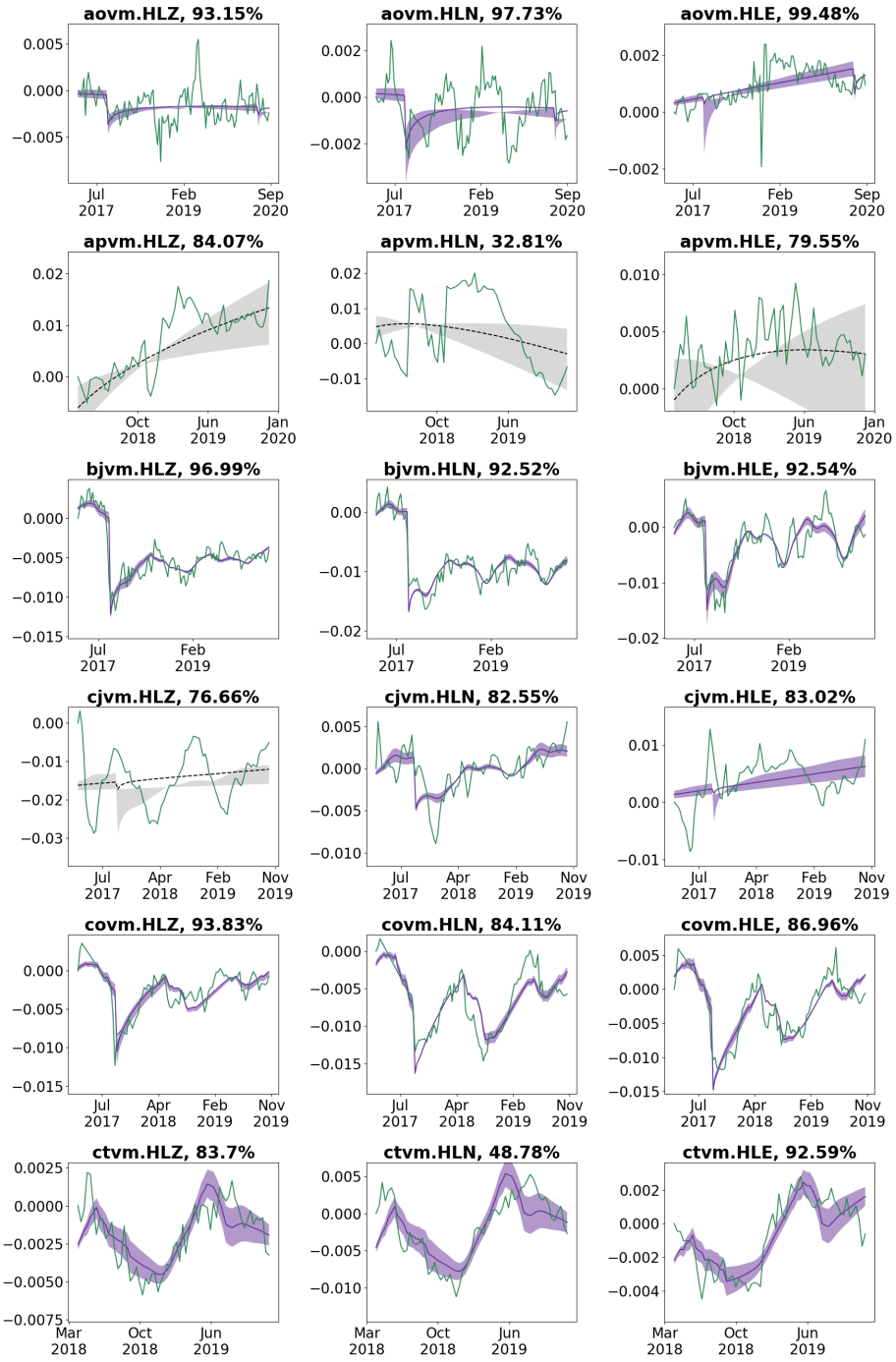

Figure S 5. (continued) RSVM data and models for 2.0 – 4.0 Hz, stations AOVM-CTVM, early coda; see above for description.

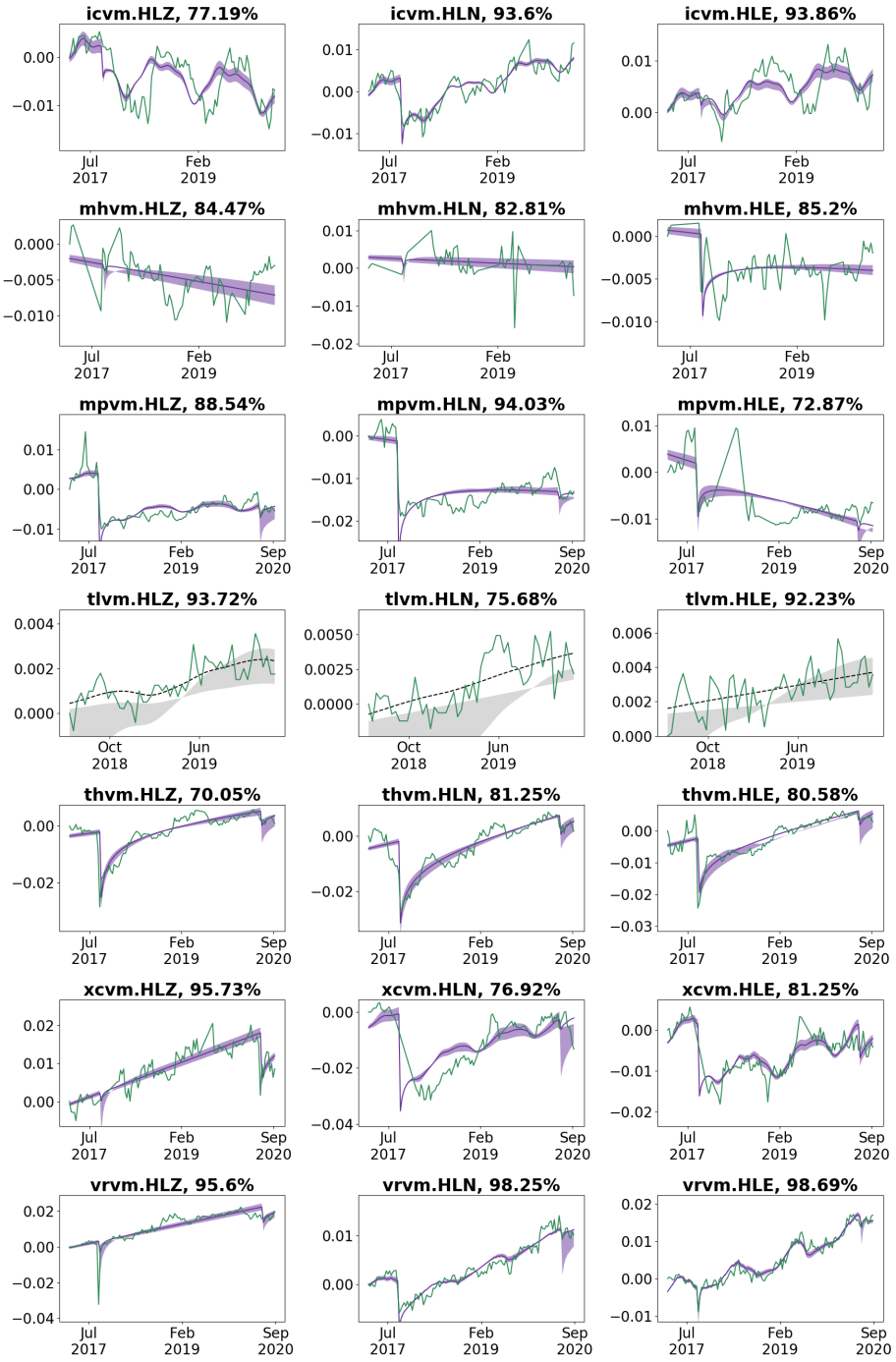

Figure S 5. (continued) RSVM data and models for 2.0 – 4.0 Hz, stations ICVM-VRVM, early coda; see above for description.

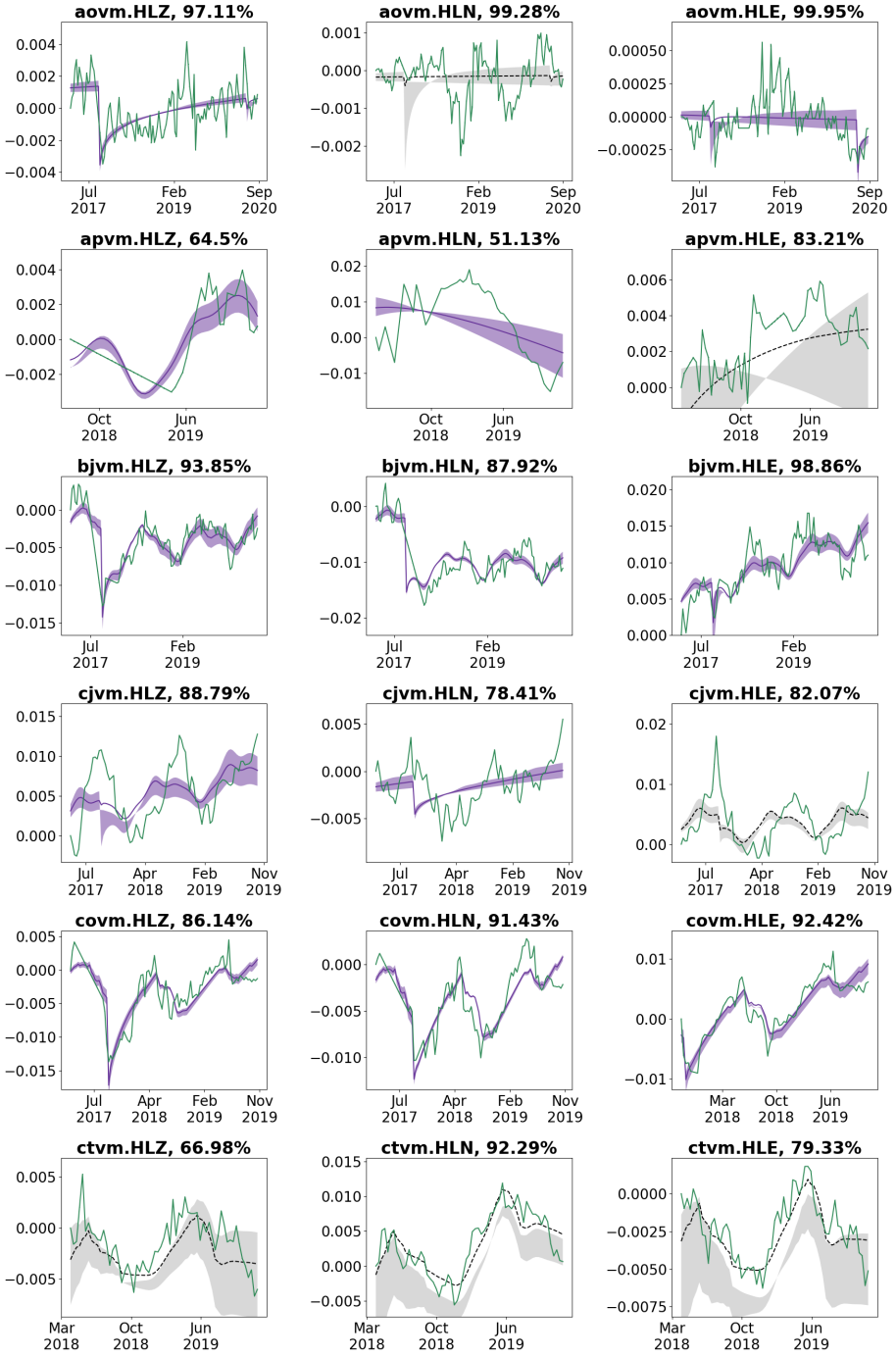

Figure S 5. (continued) RSVM data and models for 2.0 – 4.0 Hz, stations AOVM-CTVM, late coda; see above for description.

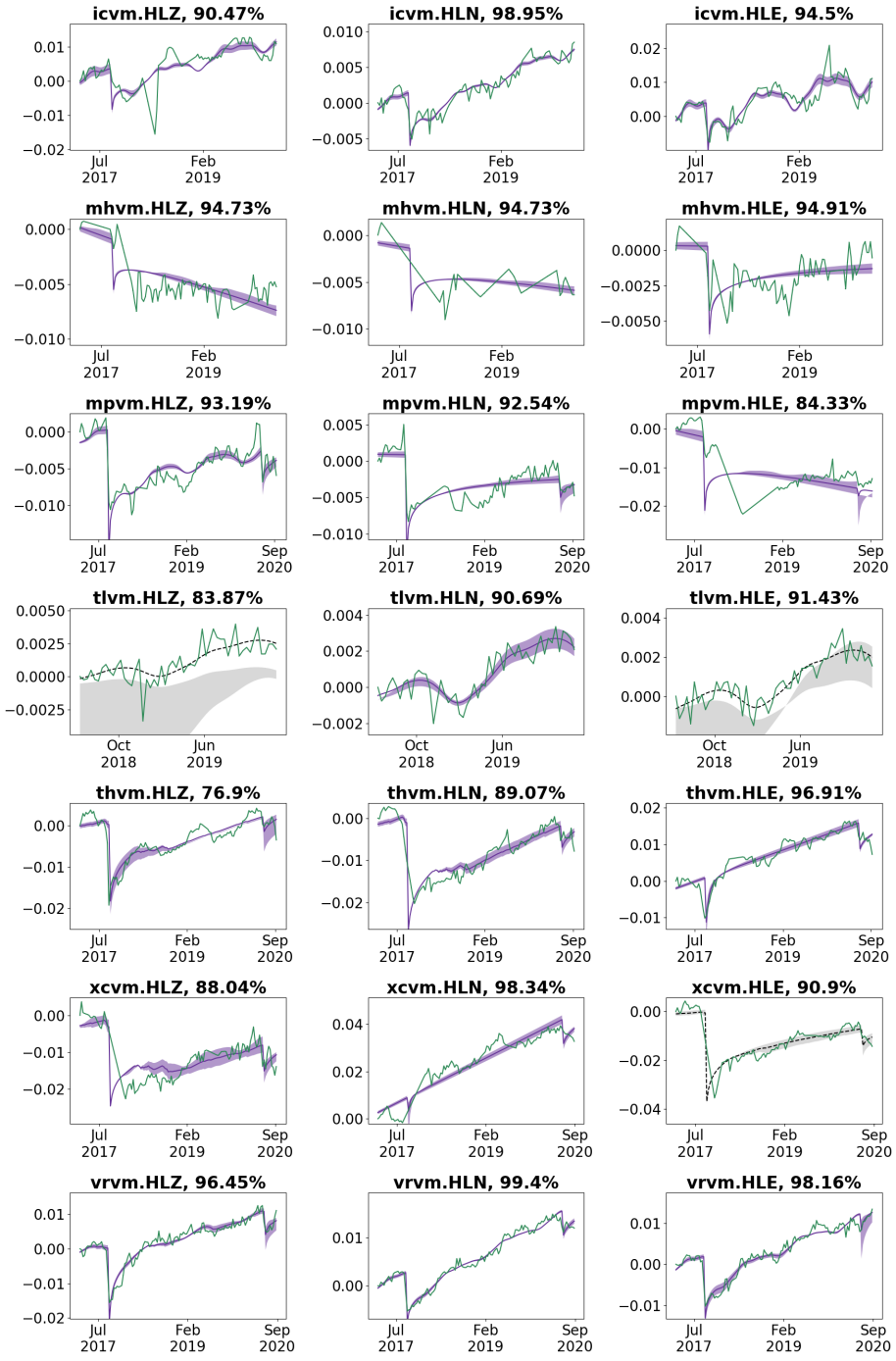

Figure S 5. (continued) RSVM data and models for 2.0 – 4.0 Hz, stations ICVM-VRVM, late coda; see above for description.

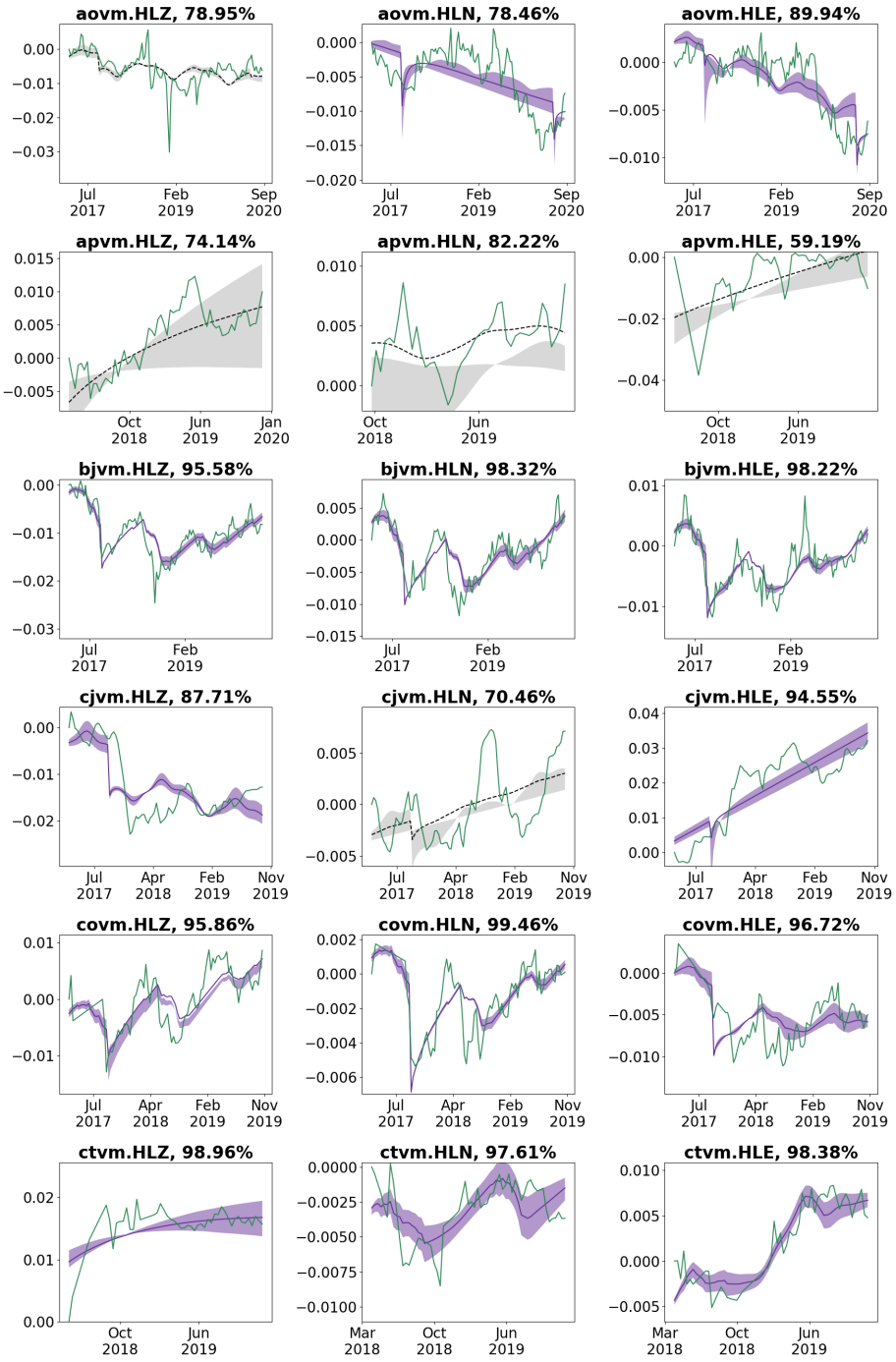

Figure S 5. (continued) RSVM data and models for 4.0 – 8.0 Hz, stations AOVM-CTVM, early coda; see above for description.

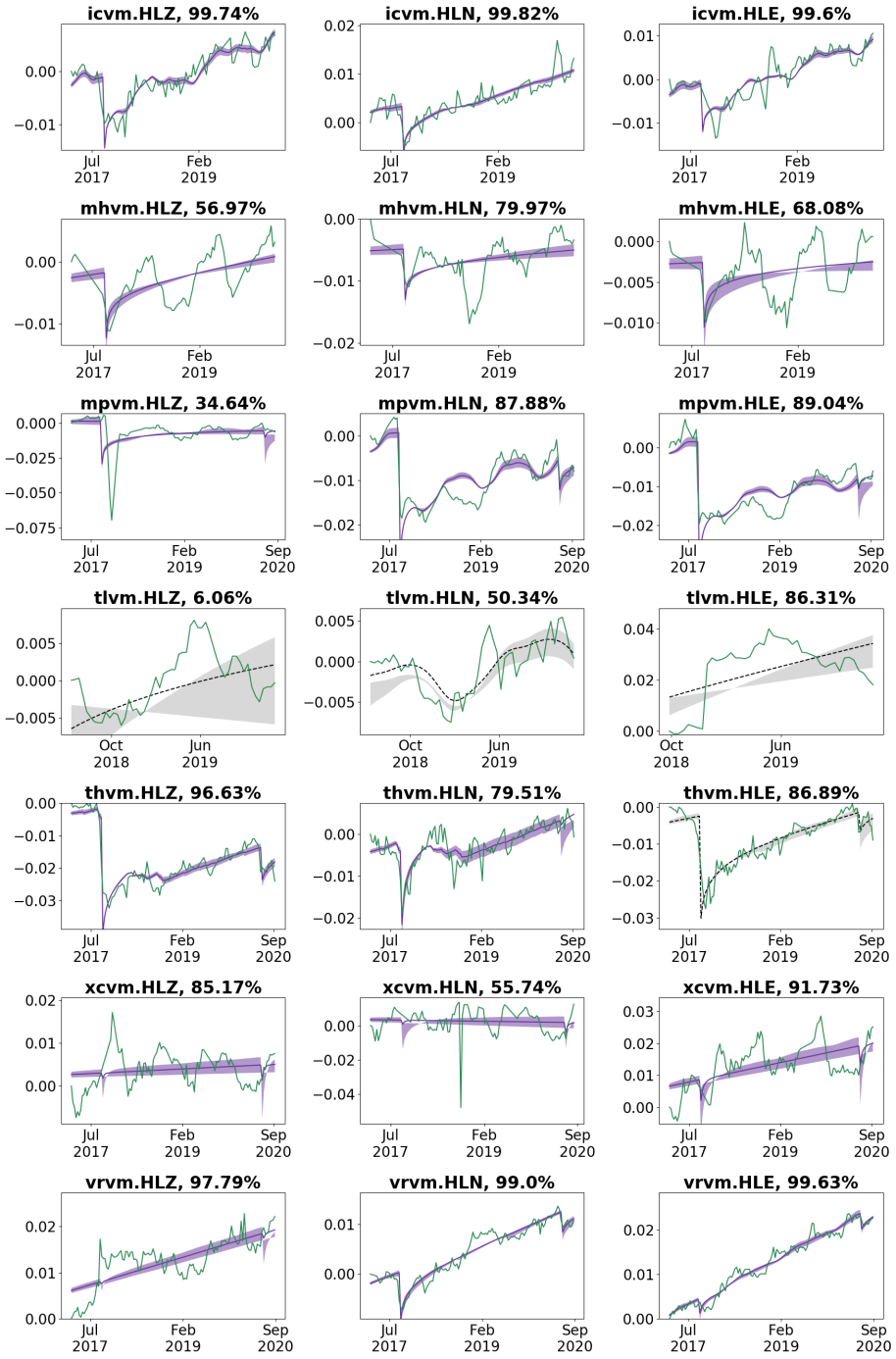

Figure S 5. (continued) RSVM data and models for 4.0 – 8.0 Hz, stations ICVM-VRVM, early coda; see above for description.

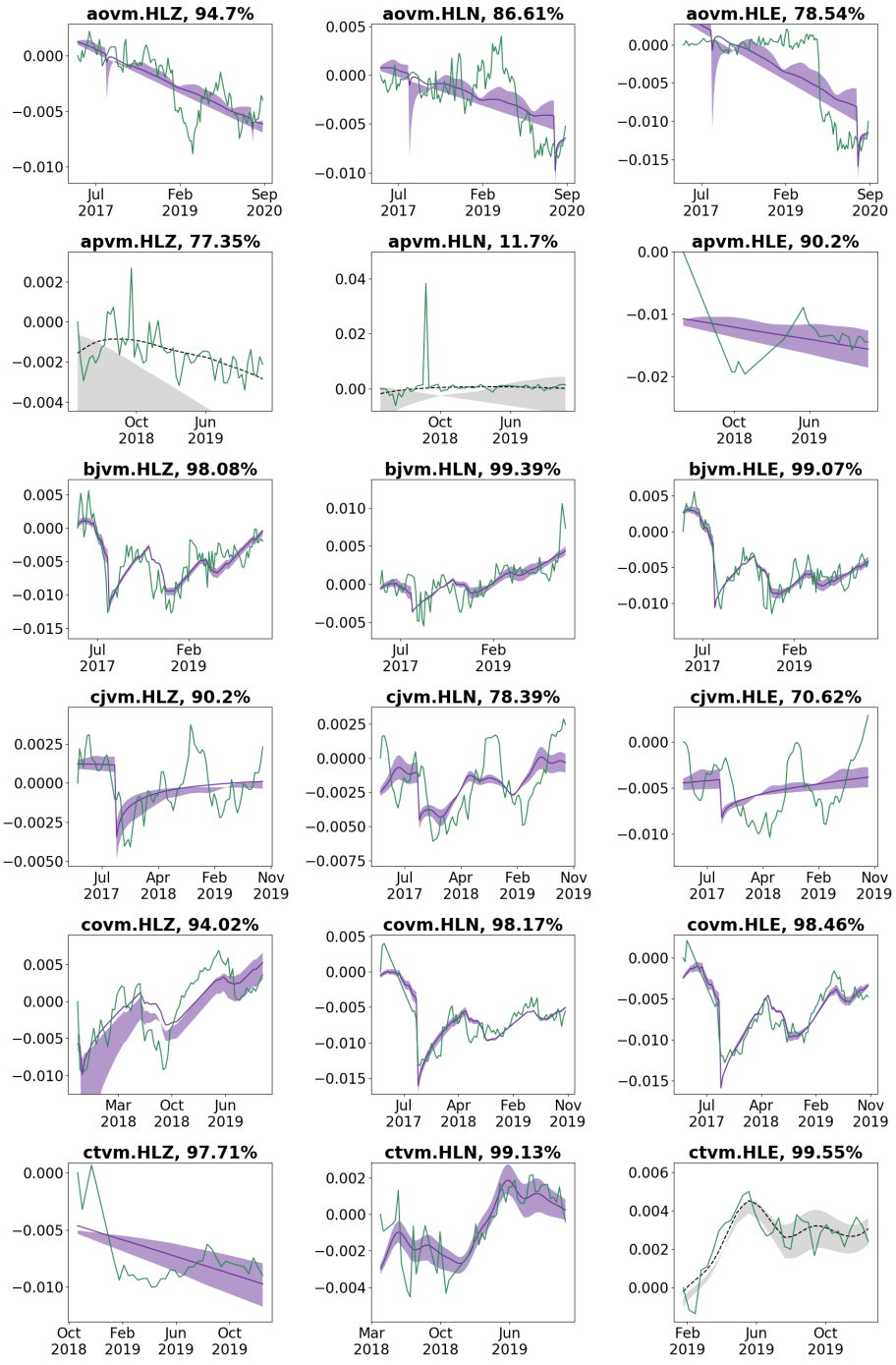

Figure S 5. (continued) RSVM data and models for 4.0 – 8.0 Hz, stations AOVM-CTVM, late coda; see above for description.

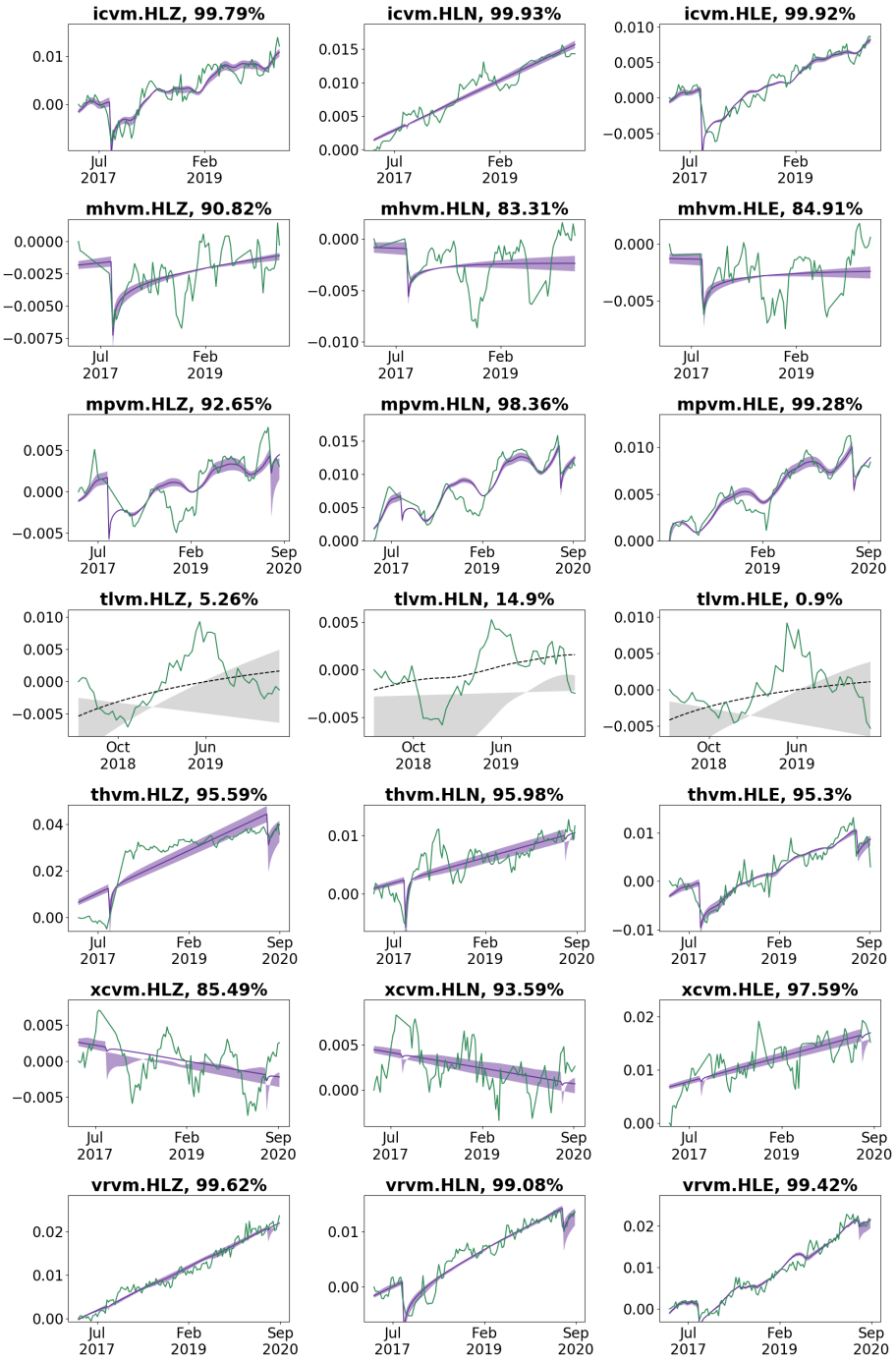

Figure S 5. (continued) RSVM data and models for 4.0 – 8.0 Hz, stations ICVM-VRVM, late coda; see above for description.

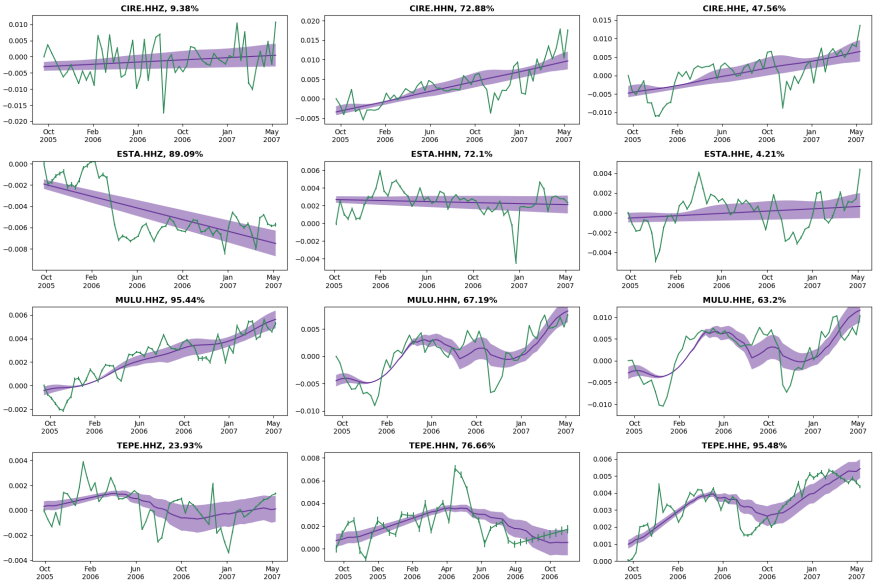

Figure S 6. Results of the Monte Carlo inversion for the TO temporary deployment: Data (green), model (purple line) and range of the 16th–84th percentile models (shaded purple area). Note that this model does not include the slow dynamics term, as strong ground shaking occurred in Mexico City during the deployment. Missing panels indicate time series with an insufficient number of high-quality $\frac{dv}{v}$ measurements. This page: 0.5 – 1 Hz, early coda.

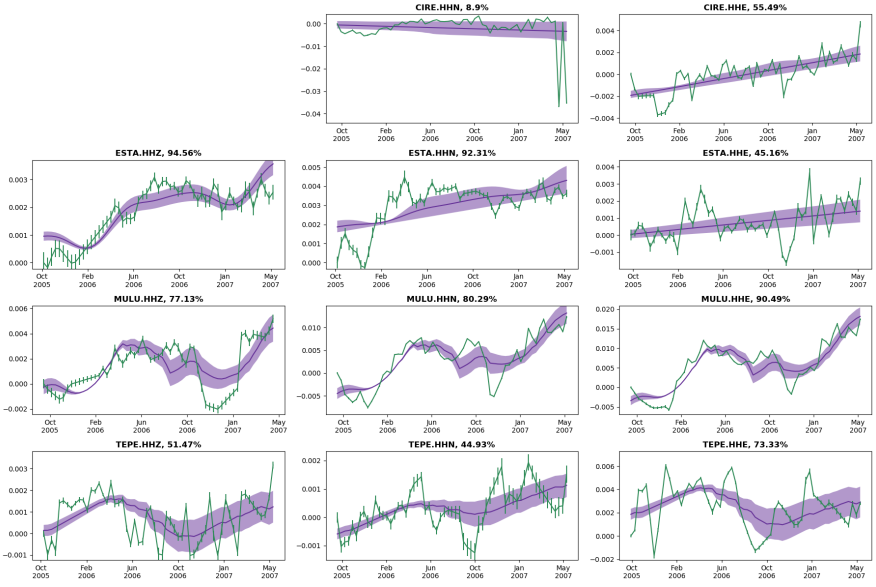

Figure S 6. (continued) TO data and models for 0.5 – 1 Hz, late coda; see above for description.

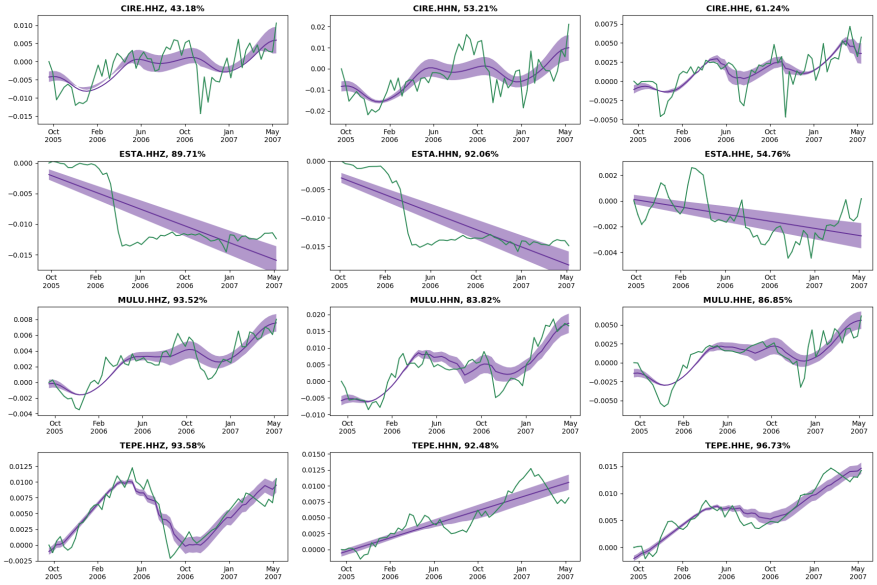

Figure S 6. (continued) TO data and models for 1 – 2 Hz, early coda; see above for description.

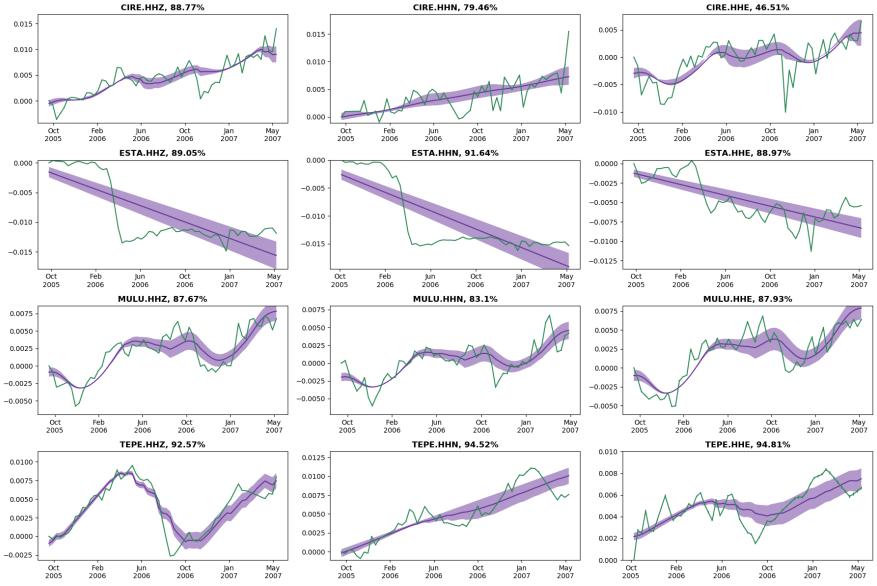

Figure S 6. (continued) TO data and models for 1 – 2 Hz, late coda; see above for description.

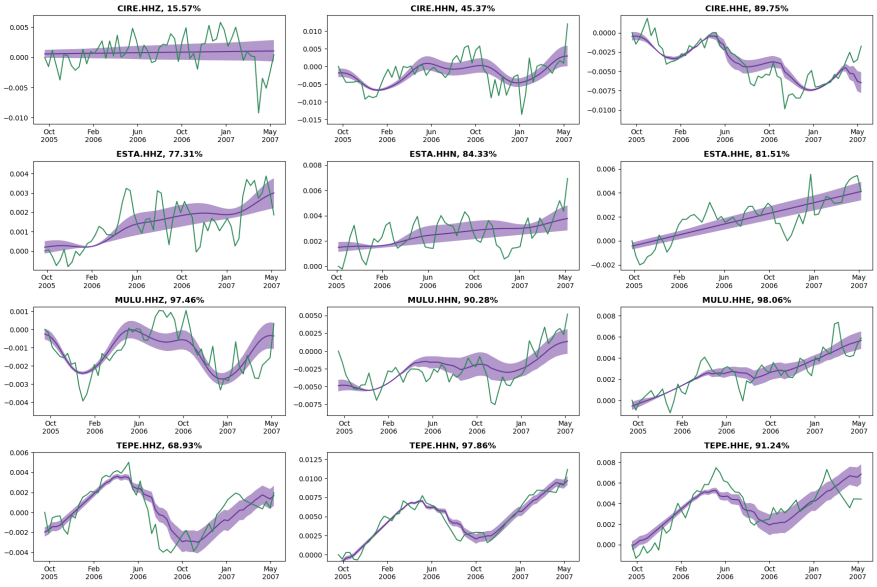

Figure S 6. (continued) TO data and models for 2 – 4 Hz, early coda; see above for description.

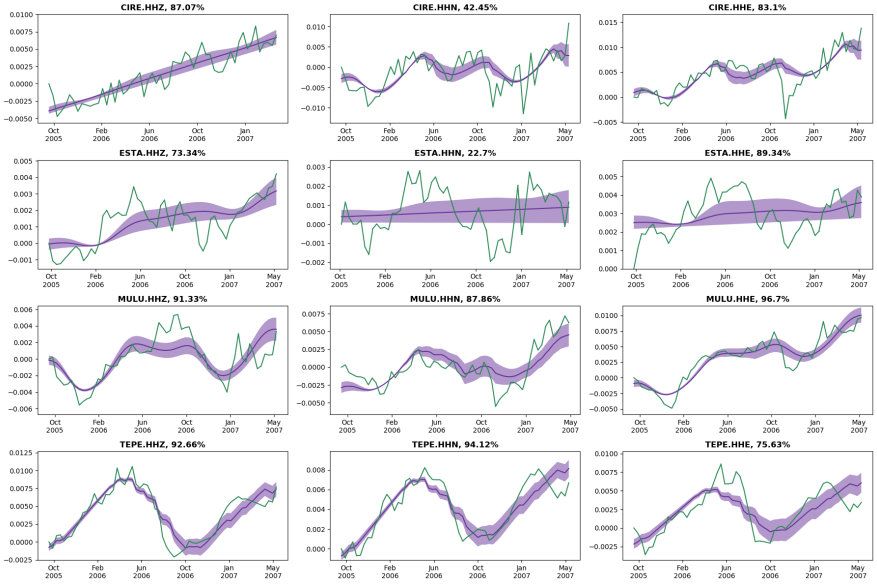

Figure S 6. (continued) TO data and models for 2 – 4 Hz, late coda; see above for description.

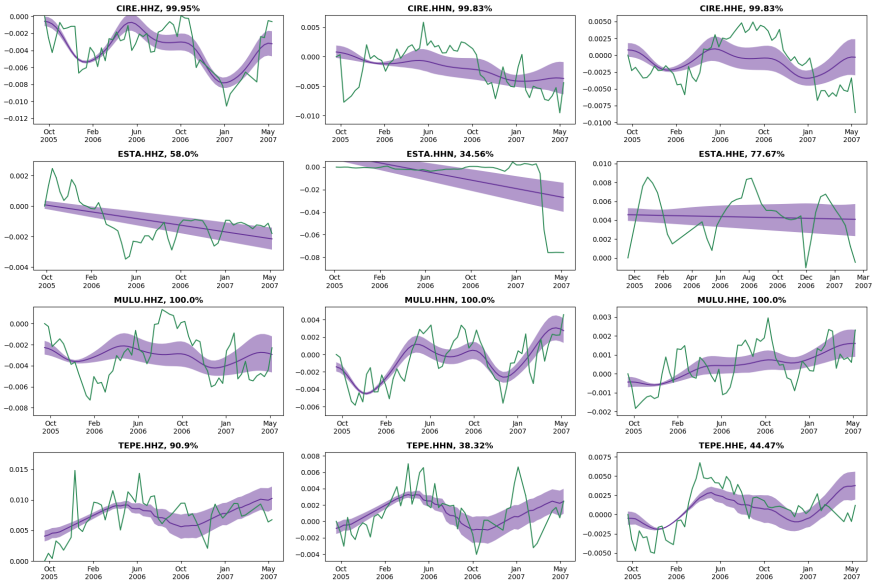

Figure S 6. (continued) TO data and models for 4 – 8 Hz, early coda; see above for description.

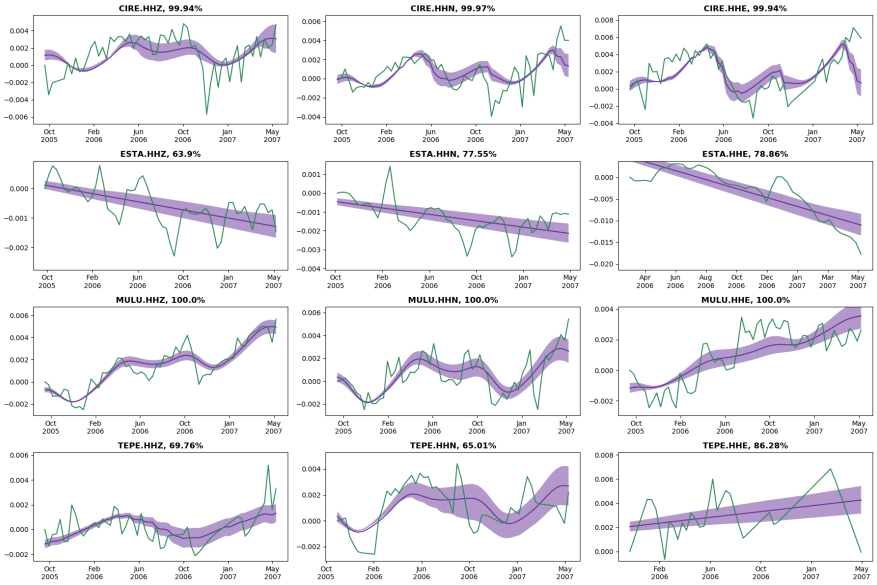

Figure S 6. (continued) TO data and models for 4 – 8 Hz, late coda; see above for description.

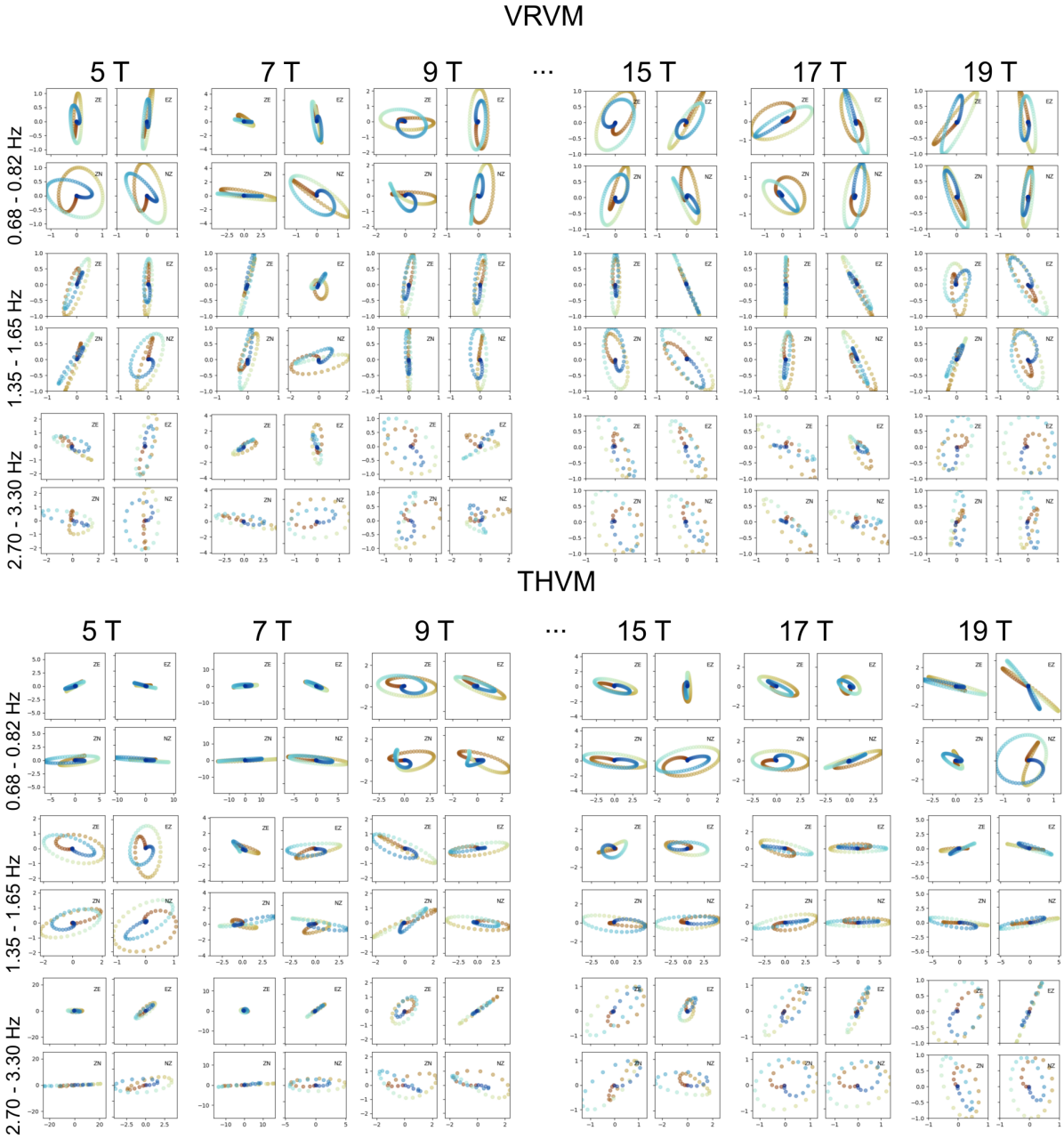

Figure S 7. Particle motion at station VRVM (lake zone, top) and THVM (transition zone, bottom). Tiles show particle motion, color-coded by time, considering the following parameters: **Plane of motion** from top to bottom: NE vs. NZ, NN vs. NZ **Orientation** left negative, right positive lag. **Lag time** from left to right: 4 times to 20 times the longest period. **Central frequency** from top to bottom 0.75 Hz, 1.5 Hz, 3 Hz. This figure illustrates that particle motion is complex, but suggests that Rayleigh surface waves appear as a reasonable, if crude, description.