Supplementary Material

Some Refinements on the Comparison of Areal Sampling Methods via Simulation

J. H. Gove[a]

[a] *USDA Forest Service, Northern Research Station, 271 Mast Road, Durham, NH 03824 USA*

Contents

S.1   Introduction

The material in the following sections presents more details on certain results given in the main text for those that are interested. All results both here and in the main manuscript were derived using the **R** (R Core Team, 2017) software system. The two main **R** packages used in the analysis are `sampSurf` (Gove, 2017a) and `ssWavelets` (Gove, 2017b) as noted in the main manuscript. Other related information is available from the author upon request.

S.2   Sampling Surfaces

S.2.1   A simple example

Figure S.1 presents a simple set of sampling surfaces that are much less complex and busy than those in Figure 1 in the main text where a large number of overlapping inclusion zones obscure the finer details of surfaces with the respective methods. The figure illustrates the differences in the surfaces when only two trees are present. Note that the same two trees are used in each subfigure, the only change is the sampling method. HPS and CPS have flat or constant sampling surfaces within an individual tree's inclusion zone. CHS and CMC have surfaces that vary from sample point to sample point with a given tree's inclusion zone.

Similarly, Figure S.2 presents a corresponding set of figures to those in Figure S.1, but in perspective view. The two figures may be compared for a better idea of sampling surface form under these sampling methods.

(S.1a) HPS

(S.1b) CHS

(S.1c) CMC

(S.1d) CPS
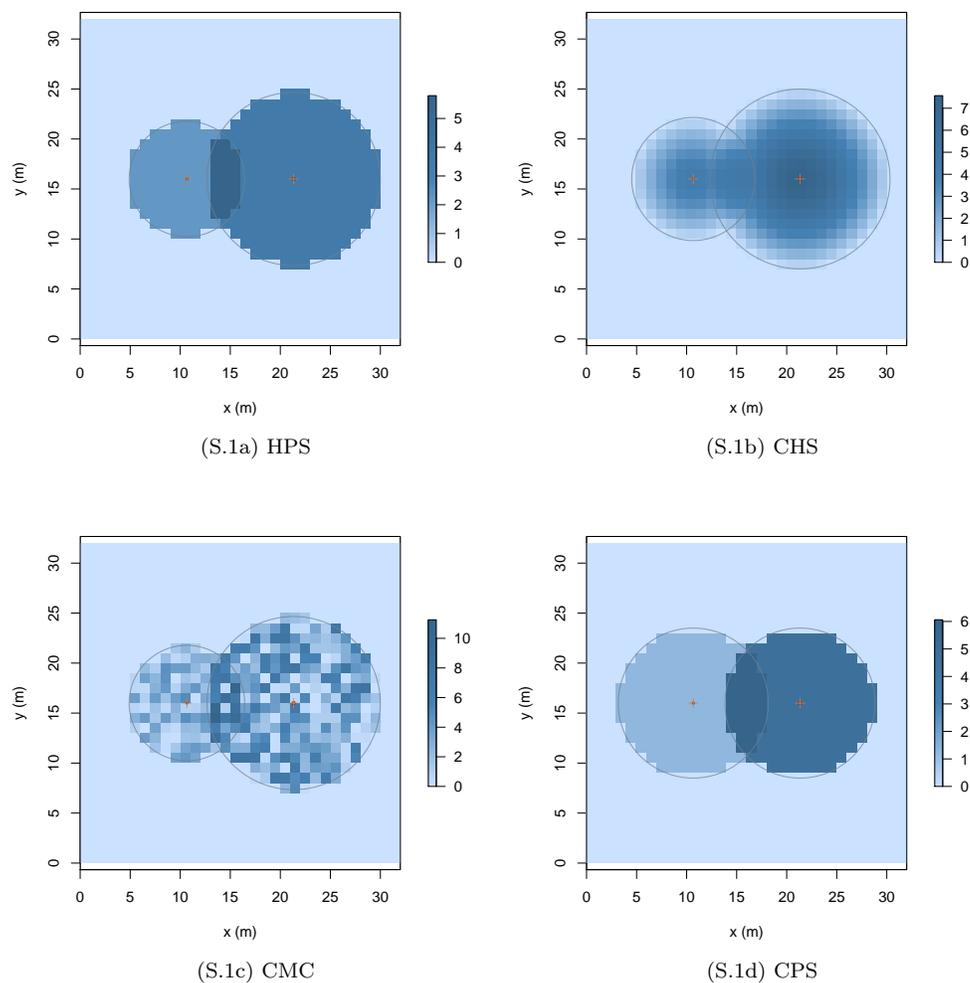
Figure S.1: Four simple sampling surfaces using two trees with DBH= 20 (left) and 30 cm (right) and respective total heights of 15 and 25 m, illustrating the four sampling methods: (a) HPS, (b) CHS, (c) CMC all with $\mathcal{F} = 3$ m^2 ha^{-1}, and (d) CPS with $R = 7.5$ m. Tree positions are denoted by '+', while the circles correspond each individual tree's inclusion zone.

S.2.2   Raster sampling surfaces

The sampling surface simulations shown in Figure S.3 correspond to the perspective view in Figure 1. The benefit of the raster view is that is allows one to see the individual inclusion zones, and it more clearly demonstrates the scale variation between the images. However, it is difficult to get a clear understanding of the shape of the overall surfaces without viewing individual larger versions of the subfigures.

(S.2a) HPS

(S.2b) CHS

(S.2c) CMC

(S.2d) CPS
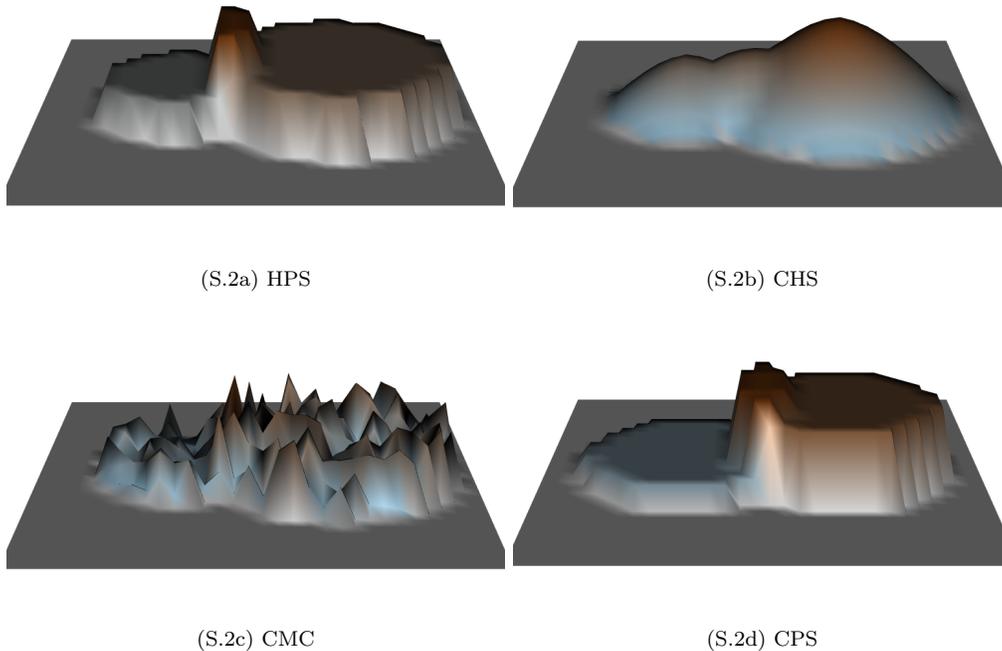
Figure S.2: Perspective view of the sampling surfaces in Figure S.1 using two trees with DBH= 20 (left) and 30 cm (right) and respective total heights of 15 and 25 m, illustrating the four sampling methods: (a) HPS, (b) CHS, (c) CMC all with $\mathcal{F} = 3 \text{ m}^2 \text{ ha}^{-1}$, and (d) CPS with $R = 7.5$ m.

S.3 Taylor Diagrams

The information in Table S.1 was computed directly from the `sampSurf` simulations, whose main results appear in Table 1 and may be cross-referenced for convenience. The table presents the variance differences, covariances and correlations required in (5) and (6) for presentation in the Taylor diagram in Figure 3

S.4 Wavelets

The following sections provide material in support of the wavelet analysis component of the manuscript.

S.4.1 Wavelet levels, scale, etc.

Table S.2 presents the levels, j, physical scale, τ'_j, and the number of Haar wavelet and scaling filter coefficients, L_j, in the MODWT analysis. In addition, the total area covered by the wavelet and

(S.3a) HPS

(S.3b) CHS

(S.3c) CMC

(S.3d) CPS
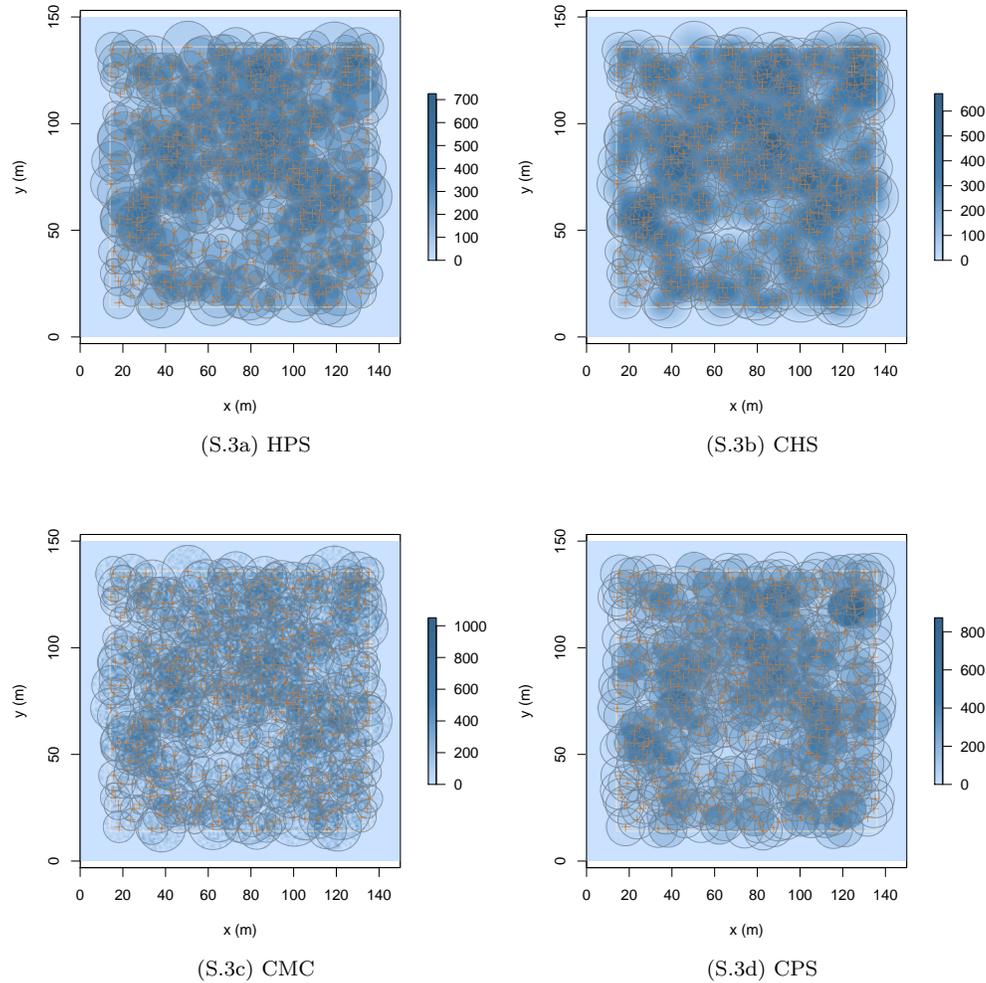
Figure S.3: Raster view of the sampling surfaces from the simulations in Table 1 illustrating the four sampling methods: (a) HPS, (b) CHS, (c) CMC all with $\mathcal{F} = 3$ m^2 ha^{-1}, and (d) CPS with $R = 8.92$ m. Tree positions are denoted by '+', while the circles correspond each individual tree's inclusion zone.

scaling filters, and the area difference—or change in area—associated with the wavelet filters are also presented. Note that since the resolution of the images are $\Delta xy = 1$ m in the simulations, one can think of the various area quantities in either m^2 or pixels.

Table S.1: Standard deviation difference, covariance and correlation results from the `sampSurf` simulations required for the Taylor diagram in Figure 3 with HPSbaf3 as the reference sampling method.

Sampling Method	$\mathcal{F}$ or R	Method ID	St. Dev. (m³)	$\sqrt{\mathrm{Var}(x-y)}$ (m³)	$\mathrm{Cov}(x,y)$ (m³)²	$\mathrm{Corr}(x,y)$
HPS	3.0	HPSbaf3	157.76	0.00	24888.11	1.00
HPS	5.0	HPSbaf5	177.95	88.36	24373.73	0.87
HPS	7.0	HPSbaf7	194.20	120.04	24095.50	0.79
HPS	9.0	HPSbaf9	208.79	143.76	23906.56	0.73
CHS	3.0	CHSbaf3	163.29	56.45	24183.30	0.94
CHS	5.0	CHSbaf5	184.48	106.44	23794.95	0.82
CHS	7.0	CHSbaf7	202.88	137.31	23598.14	0.74
CHS	9.0	CHSbaf9	220.07	162.26	23494.55	0.68
CMC	3.0	CMCbaf3	182.96	92.89	24866.66	0.86
CMC	5.0	CMCbaf5	214.11	149.53	24186.03	0.72
CMC	7.0	CMCbaf7	241.45	185.99	24297.19	0.64
CMC	9.0	CMCbaf9	264.82	217.51	23852.90	0.57
CPS	8.92	CPSrad9	169.97	87.28	23080.03	0.86
CPS	6.89	CPSrad7	196.04	127.44	23539.40	0.76
CPS	5.64	CPSrad6	222.11	164.65	23555.71	0.67
CPS	5.05	CPSrad5	239.80	188.00	23523.68	0.62

Note: the second column lists the BAF (m² ha⁻¹) for methods based on HPS, and plot radius (m), for CPS. The third column lists identifiers used in the main text for each sampling protocol (sampling method plus design parameters).

Table S.2: The levels with associated scale and area dimensions for MODWT analysis.

Level j	Scale τ'_j (m)	L_j	Coverage Area (m²)	Area Change (m²)
1	1	2	4	2
2	2	4	16	8
3	4	8	64	32
4	8	16	256	128
5	16	32	1024	512
6	32	64	4096	2048

Note: The coverage area equals $L_j^2 \Delta xy^2$.

S.4.2 Wavelet Variance Plot

The wavelet variance includes those components of the MODWT decomposition that include the wavelet filter; that is, all variances except the scaling-scaling, or smooth variances. The latter can dominate the former in terms of magnitude (depending on the level, J, to which the filters are applied) if included in a summary. Note that (12) is in the form $\mathrm{E}[X^2] - \mathrm{E}[X]^2$, and implies that

(e.g., Mondal and Percival, 2012, first equation, p. 543)

$$\widehat{\sigma}_X^2 = \sum_{j=1}^{J} \widehat{\sigma}_{DVH_j}^2 + \widehat{\sigma}_{Z_J}^2$$

where, analogous to the above,

$$\widehat{\sigma}_{DVH_j}^2 = \frac{1}{NM} \left(\|\mathbf{D}_j\|^2 + \|\mathbf{V}_j\|^2 + \|\mathbf{H}_j\|^2 \right) \tag{1}$$

$$\widehat{\sigma}_{DVH}^2 = \sum_{j=1}^{J} \widehat{\sigma}_{DVH_j}^2 \tag{2}$$

$$\widehat{\sigma}_{Z_J}^2 = \frac{1}{NM} \|\mathbf{Z}_J\|^2 - \bar{X}^2 \tag{3}$$

where $\widehat{\sigma}_{DVH_j}^2$ is the jth level *empirical wavelet variance* with total $\widehat{\sigma}_{DVH}^2$, and $\widehat{\sigma}_{Z_J}^2$ is the *empirical scaling variance*.

Figure S.4 displays the overall isotropic empirical wavelet variances, $\widehat{\sigma}_{DVH_j}^2$, plotted against distance, τ_j, for the four sets of simulations in Table 1. Similar plots can be generated for anisotropic components as well (e.g., Geilhufe et al., 2013). Critical height sampling and HPS both show a similar overall trend with variance increasing as the BAF increases. Notably, depending on the BAF, variance peaks at approximately 16 m or greater for HPS, while it peaks at a shorter distance ($\sim$ 8 m and larger) under CHS. The variance for CPS more clearly peaks at approximately 8–16 m, and the variances uniformly increase with decreasing plot size. The results for CMC are almost inverted from the others. It is conjectured that this result is due to the completely spatially random nature of the sampling surface: whereas each of the other surfaces have some degree of spatial structure, CMC does not (Figure S.2). An interesting observation is that, even with the unstructured surface, the variances are consistent across each of the BAF's, yielding a minimum variance at approximately 4 m. Each of the sampling methods tends to converge to what appears to be a stable variance (resembling a variogram sill at approximately 6,800 to 8,000) for all BAF's (plot sizes) as the scale (distance) increases.

S.4.3   Wavelet Taylor Diagrams

The full set of Talyor diagrams for the MODWT wavelet decomposition of the simulated sampling surfaces is given in Figure S.5.
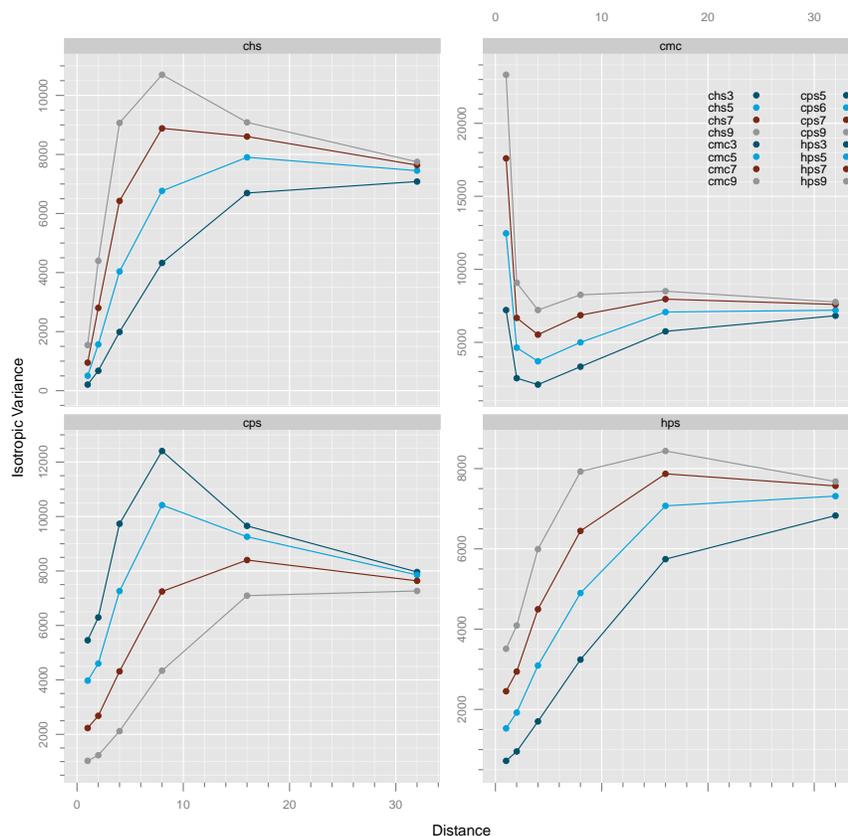
Figure S.4: Empirical wavelet variance plot of the simulations in Table 1 by distance.

References

M. Geilhufe, D. B. Percival, and H. L. Stern. Two-dimensional wavelet variance estimation with application to sea ice SAR images. *Computers and Geosciences*, 54:351–360, 2013. 6

J. H. Gove. The sampSurf package: Sampling surface simulation, 2017a. URL http://sampsurf.r-forge.r-project.org/. 1

J. H. Gove. The ssWavelets package: Wavelet functionality for package sampSurf, 2017b. URL http://sswavelets.r-forge.r-project.org/. 1

D. Mondal and D. B. Percival. Wavelet variance analysis for random fields on a regular lattice. *IEEE Transactions on Image Processing*, 21(2):537–549, 2012. 6

R Core Team. *R: A Language and Environment for Statistical Computing.* R Foundation for Statistical Computing, Vienna, Austria, 2017. URL https://www.R-project.org. 1
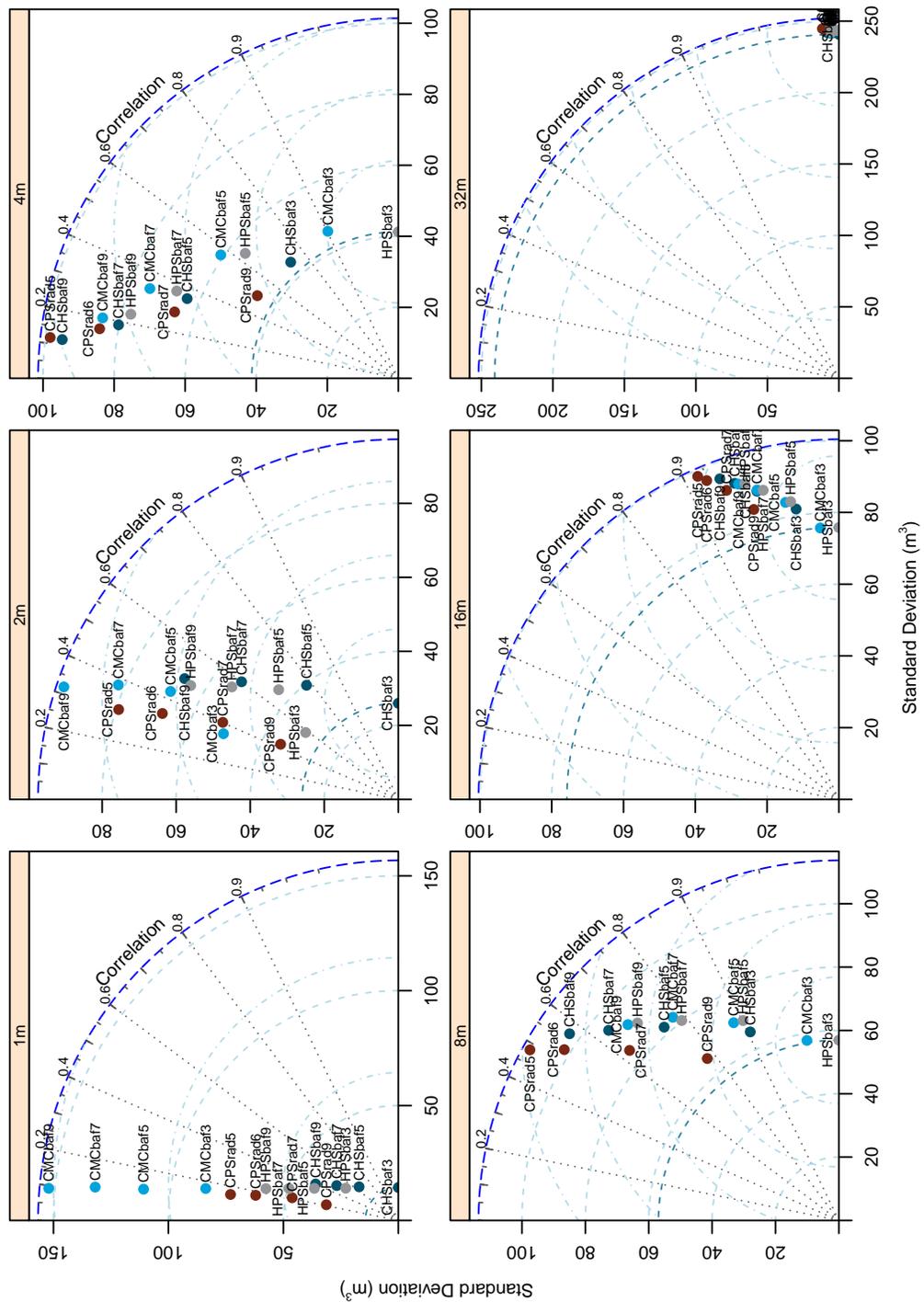
Figure S.5: MODWT Taylor diagrams for the simulations in Table 1 where the individual panels represent relationship (6) by wavelet decomposition distance $\tau'_j, j = 1, \ldots, 6$. All panels include wavelet variances (in terms of their standard deviations), while only the coarsest scale at 32 m contains the scaling (smooth) variance component as well.