Correction to VS-Lite Model Validation Results

S. E. Tolwinski-Ward

July 5, 2011

Abstract

Here we provide documentation for a correction made to the originally published version of the VS-Lite model of tree-ring width. The originally published code contained a typographical error which affected estimates of potential evapotranspiration, which were in turn used as inputs to the "Leaky Bucket" soil moisture model used by VS-Lite. The main conclusions of the original paper validating the VS-Lite model of tree ring width variability remain unchanged.

1 Description of Error

The VS-Lite code originally published on the National Oceanic and Atmospheric Administration's Paleoclimatology World Data Center Software Library (http://www.ncdc.noaa.gov/paleo/softlib/) contained an typographical error in the evapotranspiration scheme according to *Thornthwaite* (1948). At line 114 in the originally distributed code, the second term in the empirical exponent "a" should be raised to the power -5, rather than -7. The affected estimates of evapotranspiration are inputs to the Climate Prediciton Center's "Leaky Bucket" soil moisture model (*Huang et al.* (1996)), whose outputs are then used to compute synthetic ring width by VS-Lite. In the following sections of this document, we provide corrected results for the analysis detailed in *Tolwinski-Ward et al.* (2010). The new results should be compared with those in *Tolwinski-Ward et al.* (2010) and *Tolwinski-Ward et al.* (2011). The figues and tables here are numbered to correspond with the original versions in *Tolwinski-Ward et al.* (2010).

2 Corrected Results

2.1 Summary of Changes to Great Basin Bristlecone Pine Simulations

- Figure 2: The only chronology whose simulation changes noticably by eye is Methusela Walk, but the correlation of the MWK simulation with observation remains unchanged.
- Figure 3: The growth functions change slightly, but the relative importance of g_T and g_M in the various months does not change for the grouped sites at treeline or below. The interpretation of the controls on modeled growth does not change.

2.2 Summary of Changes to M08 Network Simulations

- Table 3: Changes in percentages are all smaller than the standard deviation across ensemble members.
- Table 4: Some skill is lost in the second-order pattern.
- Figure 4: The main difference is some additional skill gained in the corrected VS-Lite simulations in the coastal south east.

1

2.3 Summary of Changes to Fiveneedle Network Simulations

- Table 5: Changes in results are all smaller than the standard deviation across ensemble members.
- Table 6: Significance of the correlation between observed and simulated first-order patterns is substantially improved in the high-frequency band. The second order pattern also gains skill in both frequency bands.

3 Acknowledgements

References

Huang, J., H. M. van den Dool, and K. P. Georgankakos (1996), Analysis of model-calculated soil moisture over the United States (1931-1993) and applications to long-range temperature forecasts, *J. Clim.*, *9*, 1350–1362.

Thornthwaite, C. (1948), An approach toward a rational classification of climate, *Geogr. Rev.*, *38*, 55–94.

Tolwinski-Ward, S., M. Evans, M. Hughes, and K. Anchukaitis (2010), An efficient forward model of the climate controls on interannual variation in tree-ring width, *Clim. Dyn.*, *36*(11-12), 2419–2439, doi:10.1007/s00382-010-0945-5.

Tolwinski-Ward, S., M. Evans, M. Hughes, and K. Anchukaitis (2011), Erratum to: An efficient forward model of the climate controls on interannual variation in tree-ring width, *Clim. Dyn.*, *36*(11-12), 2441–2445, doi:10.1007/s00382-011-1062-9.

Table 2: Fraction of observed signal variances at frequencies of $1/5$ yr^{-1} and lower at each site; correlation and significance of bristlecone pine simulations with observed chronologies. Low-frequency signals are given by 5-year filtering of the signals; high-frequency signals are the residuals. Low-freq. p-values are corrected for effective number of degrees of freedom. Sites marked "UFB" are at the upper forest border. **Highlighted rows are the updated results.**

Site	Abbrv.	Low Freq. Var. Frac.	Low-freq.	High-freq.
Pearl Peak (UFB)	PRL	0.64	0.55 ($p < 0.01$)	0.12 ($p \approx 0.2$)
Pearl Peak (UFB)	PRL	0.64	0.55 ($p < 0.01$)	0.12 ($p \approx 0.3$)
Sheep Mountain (UFB)	SHP	0.48	0.69 ($p < 0.001$)	0.28 ($p < 0.005$)
Sheep Mountain (UFB)	SHP	0.48	0.69 ($p < 0.001$)	0.31 ($p < 0.002$)
Mount Washington (UFB)	MWA	0.51	0.51 ($p < 0.02$)	0.13 ($p \approx 0.2$)
Mount Washington (UFB)	MWA	0.51	0.51 ($p < 0.02$)	0.12 ($p \approx 0.2$)
Cottonwood Lower	CWL	0.60	0.22 ($p \approx 0.3$)	0.35 ($p < 0.001$)
Cottonwood Lower	CWL	0.60	0.21 ($p \approx 0.3$)	0.35 ($p < 0.001$)
Methuselah Walk	MWK	0.45	0.43 ($p \approx 0.05$)	0.29 ($p < 0.005$)
Methuselah Walk	MWK	0.45	0.42 ($p \approx 0.06$)	0.32 ($p < 0.002$)
Patriarch Lower	PAL	0.55	0.10 ($p \approx 0.7$)	0.25 ($p < 0.01$)
Patriarch Lower	PAL	0.55	0.08 ($p \approx 0.7$)	0.28 ($p < 0.01$)

Table 3: Mean percentage of sites, across an ensemble of 100 simulations, whose simulations correlate significantly with observed tree-ring width chronologies at two significance levels in the M08 network. Results shown for simulations by principal components regression calibrated at each site, simulations by VS-Lite with parameters calibrated at each site, and simulations by VS-lite with a single, "global" parameter set calibrated on the network as a whole. Errors represent 1 standard deviation in the percentages simulated significantly across ensemble members. **Highlighted rows are the updated results.**

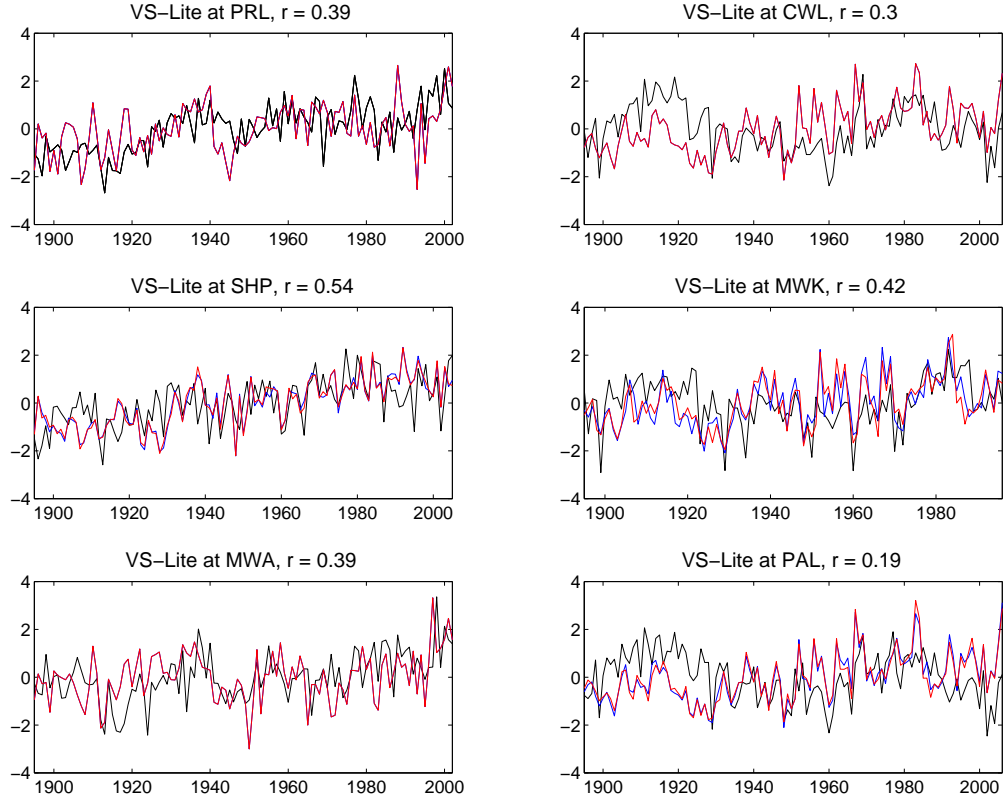|	M08 Network (N = 282)					
	PC Regr., site-by-site		VS-Lite, site-by-site		VS-Lite, global	
	Calibration	Validation	Calibration	Validation	Calibration	Validation
$p < 0.01$	73% $\pm$ 3%	40% $\pm$ 3%	70% $\pm$ 2%	59% $\pm$ 3%	47% $\pm$ 3%	49% $\pm$ 3%
$p < 0.01$	73% $\pm$ 3%	40% $\pm$ 3%	69% $\pm$ 2%	59% $\pm$ 3%	46% $\pm$ 3%	48% $\pm$ 3%
$p < 0.05$	83% $\pm$ 3%	56% $\pm$ 3%	81% $\pm$ 2%	71% $\pm$ 3%	60% $\pm$ 3%	62% $\pm$ 3%
$p < 0.05$	83% $\pm$ 3%	56% $\pm$ 3%	80% $\pm$ 2%	71% $\pm$ 3%	59% $\pm$ 3%	61% $\pm$ 3%

Figure 2: Great Basin bristlecone pine chronologies, observed (solid line) and simulated (dashed). Chronologies from upper forest border sites are displayed in panels at left; chronologies from below treeline are at right. **black is observed chronology, blue is previously simulated chronology with E_p error in the code, and red is simulated chronology with corrected code.**
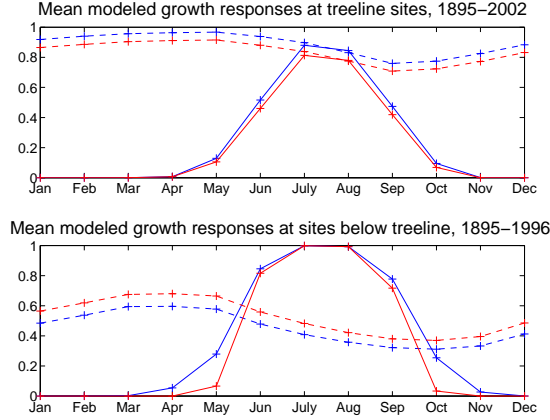
Figure 3: Modeled temperature (solid line) and moisture (dashed) response curves of Great Basin bristlecones at upper forest border (top panel) and below treeline (lower panel). The modeled growth response over the growing season is controlled by the pointwise minimum of these two quantities. VS-Lite therefore models mainly temperature-limited growth in the upper forest border sites, and more strongly moisture-limited growth below treeline. **Previously simulated growth functions with E_p error in the code are plotted in blue; red plots are simulated growth functions with corrected code.**

Table 4: Correlation and significance of temporal loadings of significant patterns of mean M08 network calibration and validation fields, as simulated by VS-Lite and PC regression, with the corresponding principal components of the observed field. Low and high frequency components are given by a 5-year running filter of the temporal loadings and their residuals, and significance of low-frequency correlations are computed using a 2-sided T-test with the effective number of degrees of freedom estimated by the signal length divided by the length of the low-pass filter. **Highlighted rows are the updated results.**

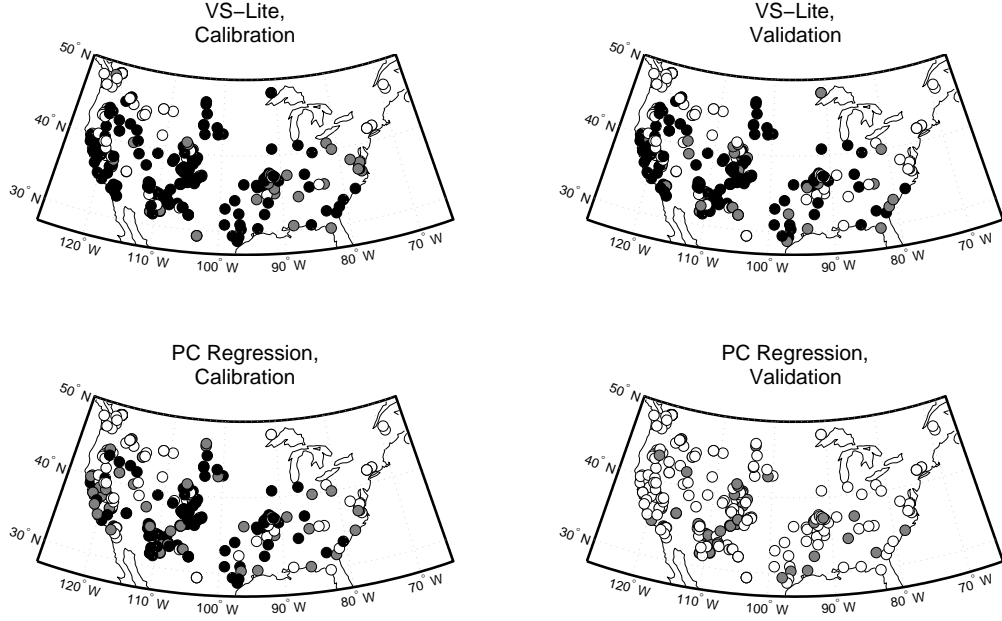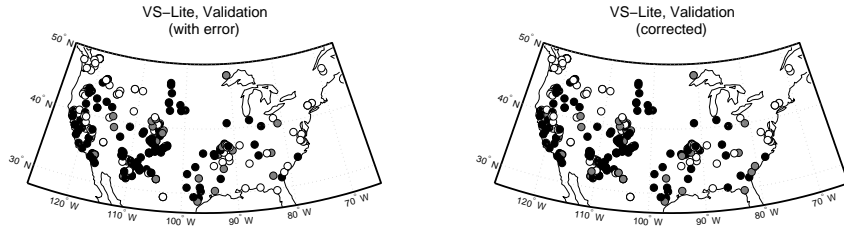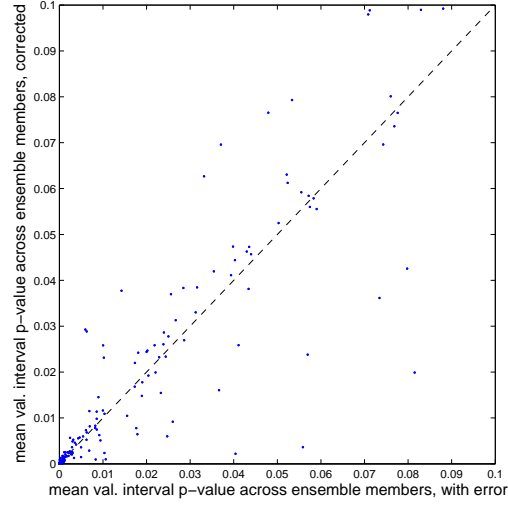| Model	Pattern Order	Low frequency		High frequency	
		Calibration	Validation	Calibration	Validation
VS-Lite	1	$0.68, p < 0.01$	$0.63, p < 0.01$	$0.73, p < 0.001$	$0.69, p < 0.001$
VS-Lite	1	$0.64, p < 0.01$	$0.59, p = 0.01$	$0.71, p < 0.001$	$0.67, p < 0.001$
VS-Lite	2	$0.54, p < 0.05$	$0.52, p < 0.05$	$0.30, p < 0.01$	$0.29, p < 0.01$
VS-Lite	2	$0.42, p = 0.09$	$0.41, p < 0.09$	$0.20, p = 0.07$	$0.20, p = 0.06$
PC Reg	1	$0.82, p < 0.001$	$0.71, p < 0.01$	$0.83, p < 0.001$	$0.74, p < 0.001$
PC Reg	2	$0.51, p < 0.05$	$0.39, p \approx 0.12$	$0.54, p < 0.001$	$0.37, p < 0.001$

Figure 4: *(Corrected)* Mean validation-interval significance of correlations of ring width simulations with observations over a 100-member ensemble of simulations of the M08 network. Ensemble members differ in their randomized calibration intervals. Black circles: $p < .01$, gray circles: $p < .05$, white circles: $p > .05$.

New figure: Previous (left panel) and corrected (right panel) mean validation interval results computed using VS-Lite, generated as in the upper-right panel of Figure 4. *(Note that previous figure is from Erratum)*

New figure: Corrected mean p-values across ensemble members at each point, versus those previously computed with the error in the code, for $0 < p < 0.10$. The correction does not seem to have a systematic effect on skill across sites.

Table 5: Mean percentage of sites, across an ensemble of 100 simulations, whose simulations correlate significantly with observed tree-ring width chronologies at two significance levels in the 5N network. Results shown for simulations by principal components regression calibrated at each site, simulations by VS-Lite with parameters calibrated at each site, and simulations by VS-lite with a single, "global" parameter set calibrated on the network as a whole. Errors represent 1 standard deviation in the percentages simulated significantly across ensemble members. **Highlighted rows are the updated results.**

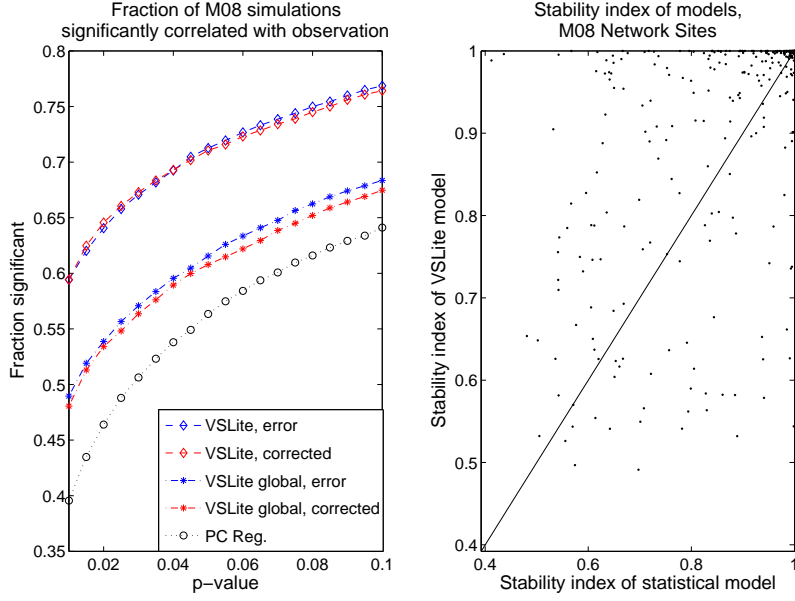| | 5N Network (N = 66) | | | | | |
| | PC Regr., site-by-site | | VS-Lite, site-by-site | | VS-Lite, global | |
	Calibration	Validation	Calibration	Validation	Calibration	Validation
$p < 0.01$	69% ± 5%	17% ± 3%	42% ± 5%	25% ± 5%	22% ± 5%	20% ± 5%
$p < 0.01$	69% ± 5%	17% ± 3%	41% ± 5%	25% ± 5%	21% ± 5%	20% ± 5%
$p < 0.05$	85% ± 6%	31% ± 4%	59% ± 4%	43% ± 5%	36% ± 5%	35% ± 5%
$p < 0.05$	85% ± 6%	31% ± 4%	60% ± 4%	42% ± 5%	37% ± 5%	36% ± 5%

Figure 5: Performance indices of modeling by VS-Lite and principal components regression on the M08 Network. Left panel plots the fraction of network sites whose simulations correlate significantly with observations at a range of p-values for three different simulation approaches. **Previous results are plotted in blue for comparison; corrected results plotted in red.** Right panel plots the stability index (eqn. 3) of simulations by corrected VS-Lite code versus PC regression, with one indicating perfect stability of simulations from the calibration to validation periods, and zero representing complete instability. 202 out of 282 points fall above $y = x$.

Table 6: Correlation and significance of temporal loadings of significant patterns of mean 5N network calibration and validation fields, as simulated by VS-Lite and PC regression, with the corresponding principal components of the observed field. Low and high frequency components are given by a 5-year running filter of the temporal loadings and their residuals, and significance of low-frequency correlations are computed using a 2-sided T-test with the effective number of degrees of freedom estimated by the signal length divided by the length of the low-pass filter. **Highlighted rows are the updated results.**

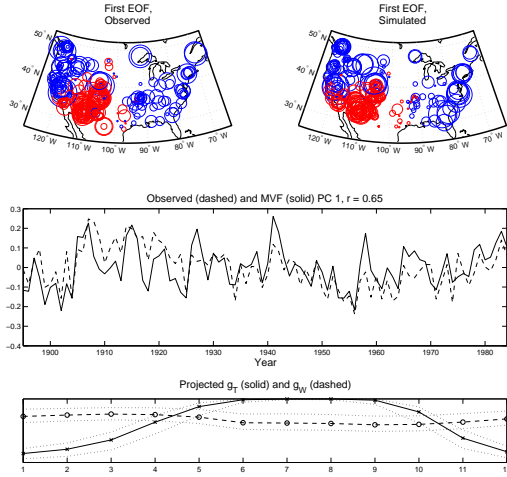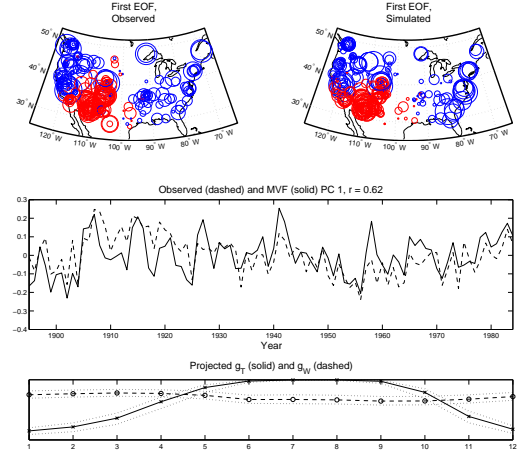|		Low frequency		High frequency	
Model	Pattern Order	Calibration	Validation	Calibration	Validation
VS-Lite	1	$0.76, p < 0.001$	$0.72, p < 0.001$	$0.31, p < 0.005$	$0.20, p = 0.07$
VS-Lite	1	$0.76, p < 0.001$	$0.74, p < 0.001$	$0.46, p < 0.001$	$0.37, p < 0.001$
VS-Lite	2	$0.74, p < 0.001$	$0.71, p = 0.001$	$0.58, p < 0.001$	$0.54, p < 0.001$
VS-Lite	2	$0.81, p < 0.001$	$0.73, p < 0.001$	$0.66, p < 0.001$	$0.63, p < 0.001$
PC Reg	1	$0.60, p < 0.01$	$0.42, p = 0.09$	$0.37, p < 0.001$	$0.09, p = 0.4$

Figure 6: *(From Erratum)* Top: first pattern in observed (left) and simulated (right) data, M08 network, 1895-1984. Center: time series associated with first observed (dashed) and simulated (solid) EOF patterns. Bottom: mean over simulated years of the mean validation field temperature and moisture response functions, projected onto the first simulated MVF EOF. Dashed lines give the 95% confidence bands derived from percentiles of the repeated experiments under randomized calibration intervals.
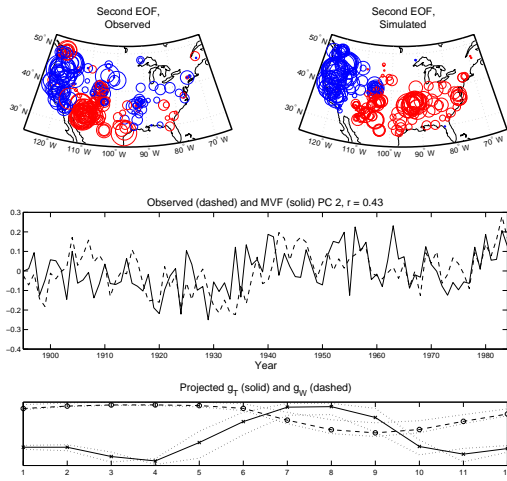
Figure 6: *(Corrected)* Top: first pattern in observed (left) and simulated (right) data, M08 network, 1895-1984. Center: time series associated with first observed (dashed) and simulated (solid) EOF patterns. Bottom: mean over simulated years of the mean validation field temperature and moisture response functions, projected onto the first simulated MVF EOF. Dashed lines give the 95% confidence bands derived from percentiles of the repeated experiments under randomized calibration intervals.

Figure 7: *(From Erratum.)* As in previous figure from Erratum, except displaying results for the second pattern.
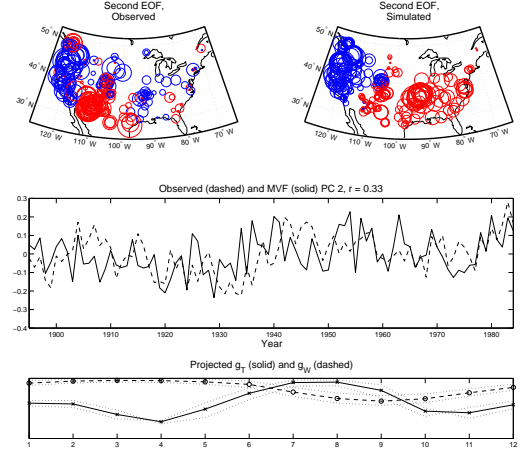

Figure 7: *(Corrected)* As in previous corrected figure, except displaying results for the second pattern.
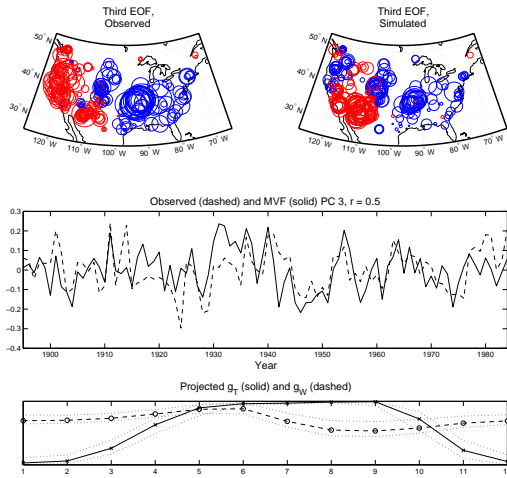

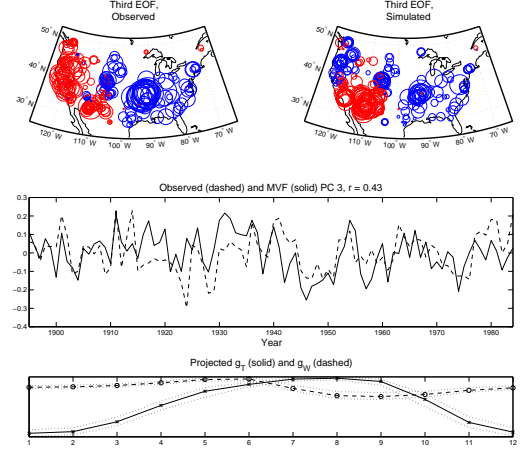Figure 8: *(From Erratum)* As in previous figure from Erratum, except displaying results for the third pattern.

Figure 8: *(Corrected)* As in previous corrected figure, except displaying results for the third pattern.
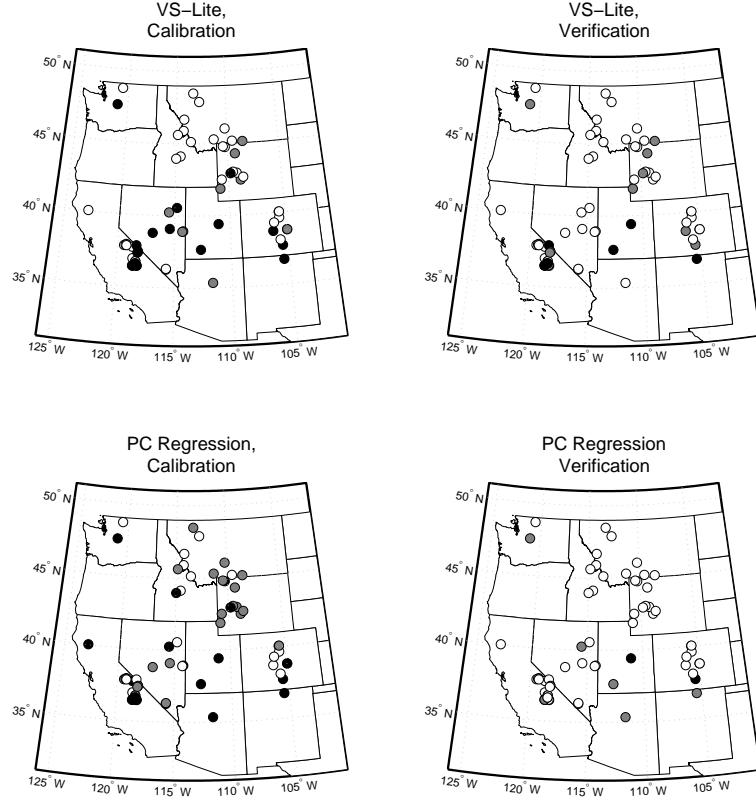
Figure 9: *(Corrected)* Mean validation-interval significance of correlations of ring width simulations with observations over a 100-member ensemble of simulations of the 5N network. Ensemble members differ in their randomized calibration intervals. Black circles: $p < .01$, gray circles: $p < .05$, white circles: $p > .05$.
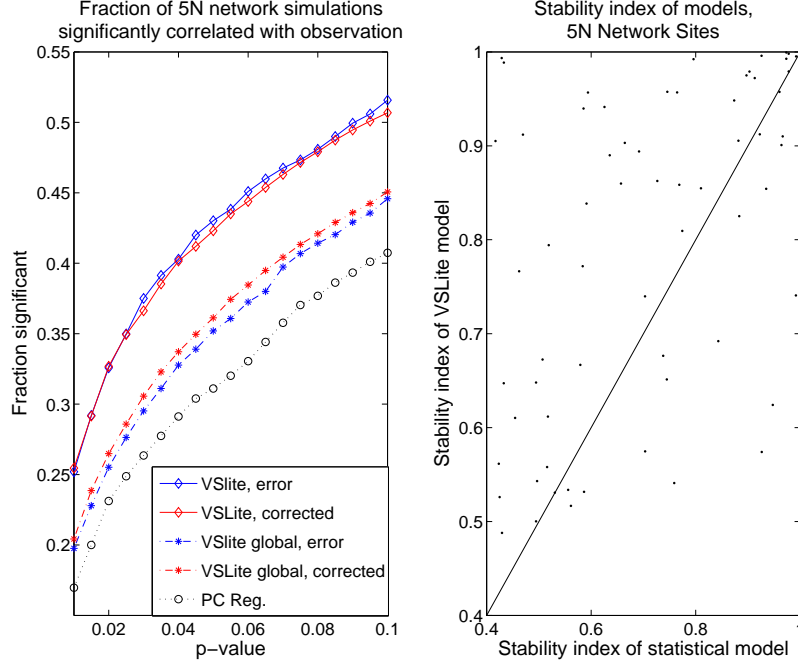
Figure 10: Performance indices of modeling by VS-Lite and principal components regression on the 5N network. Left panel plots the fraction of network sites whose simulations correlate significantly with observations at a range of p-values for three different simulation approaches. **Previous results are plotted in blue for comparison; corrected results plotted in red.** Right panel plots the stability index (eqn. 3) of simulations by corrected VS-Lite code versus PC regression, with one indicating perfect stability of simulations from the calibration to validation periods, and zero representing complete instability. 42 out of 66 points fall above $x = y$.
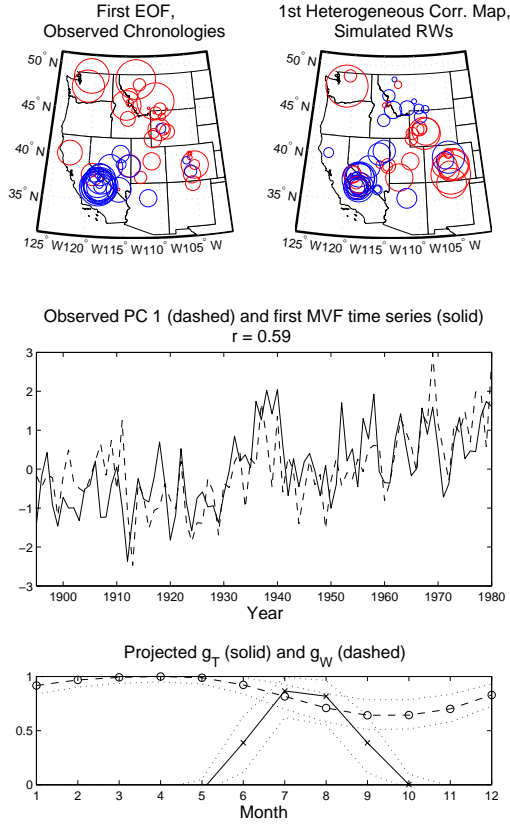
Figure 11: *(Original)* Top left: first pattern in observed data, 5N network, 1895-1980. Top right: Simulated data projected on first pattern of covariance in the observed network. Center: time series associated with first observed pattern and first pattern of covariance in simulated network. Bottom: mean over simulated years of the mean validation field temperature and moisture response functions projected onto the first pattern of observed covariance. Dashed lines give the 95% confidence bands derived from percentiles of the repeated experiments under randomized calibration intervals.
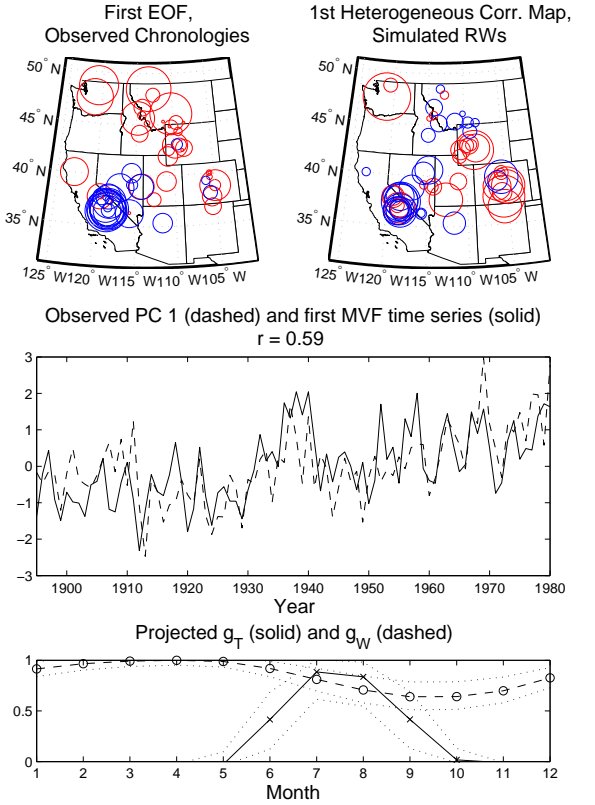
Figure 11: *(Corrected)* Top left: first pattern in observed data, 5N network, 1895-1980. Top right: Simulated data projected on first pattern of covariance in the observed network. Center: time series associated with first observed pattern and first pattern of covariance in simulated network. Bottom: mean over simulated years of the mean validation field temperature and moisture response functions projected onto the first pattern of observed covariance. Dashed lines give the 95% confidence bands derived from percentiles of the repeated experiments under randomized calibration intervals.
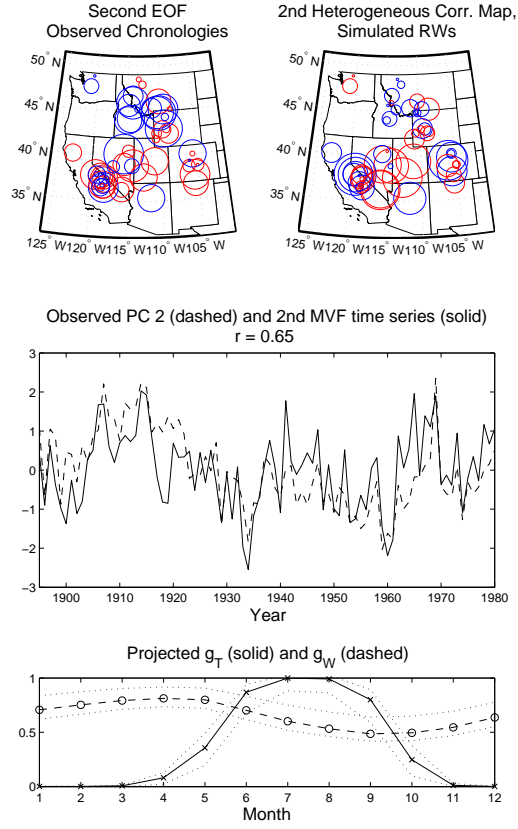
Figure 12: *(Original)* As in previous original figure, except displaying results for the second pattern.
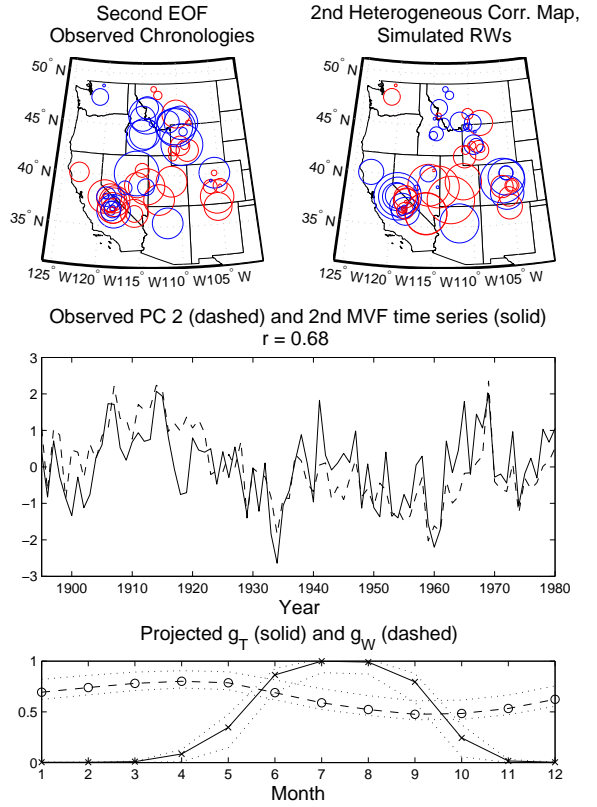

Figure 12: *(Corrected)* As in previous corrected figure, except displaying results for the second pattern.