Supplement to: "Reliability of Resilience Estimation based on Multi-Instrument Time Series"

Taylor Smith[1], Ruxandra-Maria Zotta[2], Chris A. Boulton[3], Timothy M. Lenton[3], Wouter Dorigo[2], and Niklas Boers[3,4,5,6]

[1]Institute of Geosciences, Universität Potsdam, Germany

[2]Department of Geodesy and Geo-Information, Vienna University of Technology, Vienna, Austria

[3]Global Systems Institute, University of Exeter, Exeter, UK

[4]Earth System Modelling, School of Engineering & Design, Technical University of Munich, Germany

[5]Potsdam Institute for Climate Impact Research, Germany

[6]Department of Mathematics, University of Exeter, UK

Corresponding author:

Taylor Smith

Email: tasmith@uni-potsdam.de

Synthetic Data Parameterization

Sensor	**Time frame**	**Relative Reliability**
SSMI	1987-01-01 to 2002-04-01	0.7
TMI	1997-12-01 to 2015-04-01	0.5
WindSat	2003-02-01 to 2012-07-01	0.44
AMSR-E	2002-06-01 to 2011-10-01	0.88
AMSR-2	2012-07-01 onwards	1

Table S1. Time frame and relative reliability for the satellites making up the VODCA data set. AMSR-2 is considered to be the most reliable (value of 1), with WindSat being the least (value of 0.44). Values calculated roughly from the variability of the underlying data sets, and to match aggregated synthetic patterns with those of the global VODCA patterns in lag-one autocorrelation and variance.

Sensor	Time frame	Relative Reliability
NOAA 9F	1985-02-01 to 1988-09-01	0.87
NOAA 11H	1988-09-01 to 1994-09-30	1
NOAA 9F-d	1994-09-01 to 1995-01-01	0.88
NOAA 14J	1995-01-01 to 2000-11-01	0.79
NOAA 16L	2000-11-01 to 2003-12-01	0.85
NOAA 17M	2003-12-01 to 2009-01-01	0.83
NOAA 18N	2005-08-01 onwards	0.7
NOAA 19N	2009-06-01 onwards	0.65

Table S2. Time frame and relative reliability for the satellites making up the AVHRR GIMMS3g data set. NOAA 11H is considered to be the most reliable (value of 1), with NOAA 19N being the least (value of 0.65). Values calculated roughly to match the aggregated synthetic patterns with those of the global NDVI patterns in lag-one autocorrelation and variance, and are not drawn directly from computations of NDVI variance in GIMMS3g.

Supplemental Figures

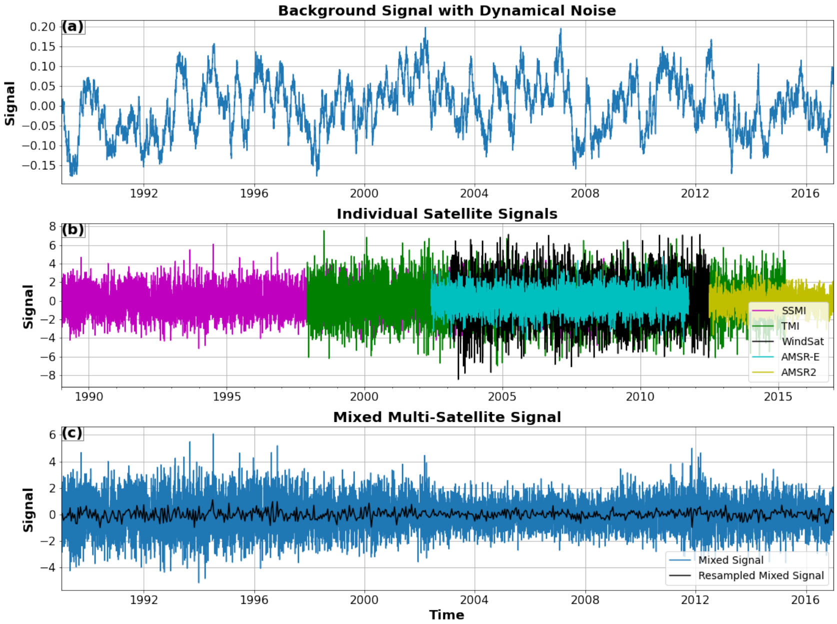

Figure S1. Synthetic Experiment for Vegetation Optical Depth (VOD), signal-to-noise ratio set at 0.1. Relative measurement noise scaling ($R_{satellite}$, see Methods) set to values between 1 for the most reliable sensor and 0.44 for the least reliable. (a) Ornstein-Uhlenbeck process with dynamical noise mimicking an underlying signal to be measured (see Methods). (b) Underlying signal plus additional white Gaussian measurement noise by individual synthetic sensor scaled by reliability $R_{satellite}$, based on the characteristics of the satellites used in the VOD data set (see Supplemental Table S1 and Methods for details). (c) Combined synthetic signal via taking the daily (blue) and bi-weekly (black) means.

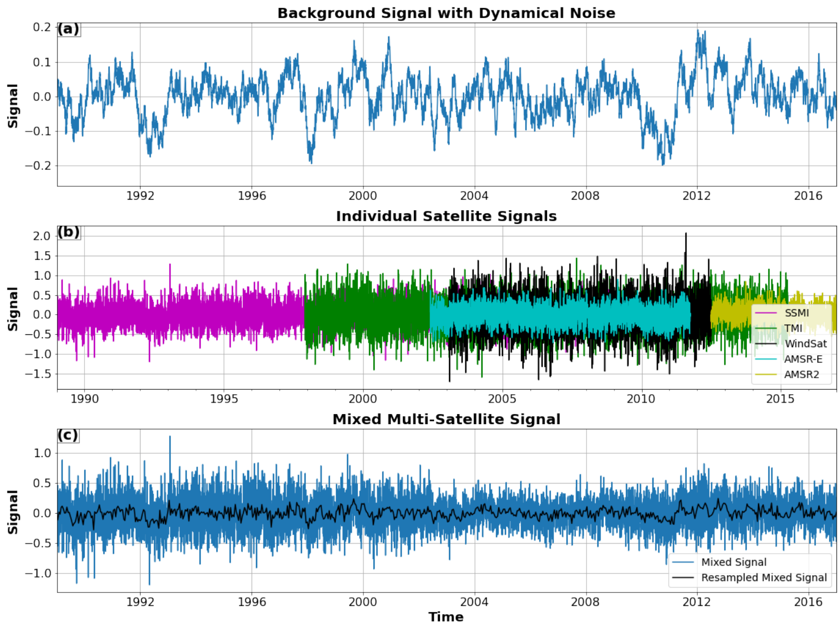

Figure S2. Synthetic Experiment for Vegetation Optical Depth (VOD), signal-to-noise ratio set at 0.5. Relative measurement noise scaling ($R_{satellite}$, see Methods) set to values between 1 for the most reliable sensor and 0.44 for the least reliable. (a) Ornstein-Uhlenbeck process with dynamical noise mimicking an underlying signal to be measured (see Methods). (b) Underlying signal plus additional white Gaussian measurement noise by individual synthetic sensor scaled by reliability $R_{satellite}$, based on the characteristics of the satellites used in the VOD data set (see Supplemental Table S1 and Methods for details). (c) Combined synthetic signal via taking the daily (blue) and bi-weekly (black) means.

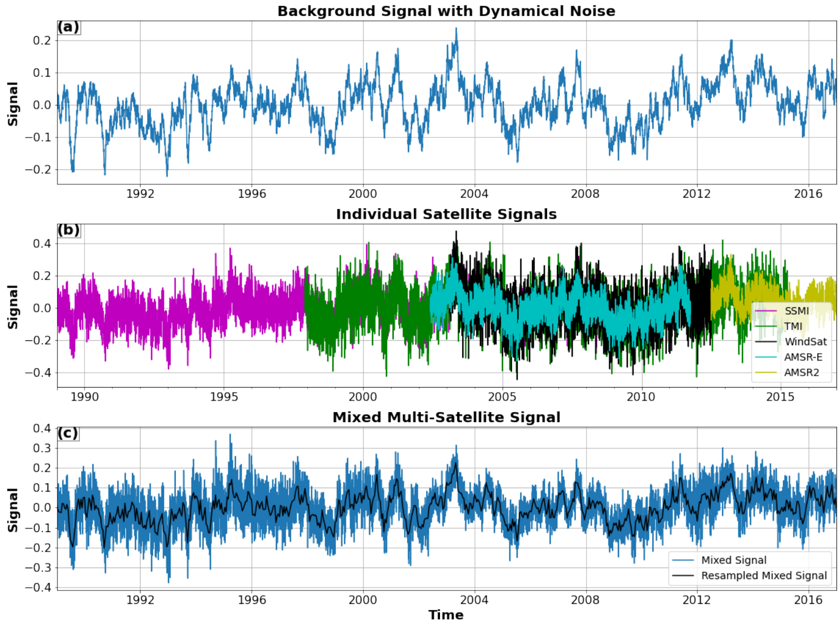

Figure S3. Synthetic Experiment for Vegetation Optical Depth (VOD), signal-to-noise ratio set at 2. Relative measurement noise scaling ($R_{satellite}$, see Methods) set to values between 1 for the most reliable sensor and 0.44 for the least reliable. (a) Ornstein-Uhlenbeck process with dynamical noise mimicking an underlying signal to be measured (see Methods). (b) Underlying signal plus additional white Gaussian measurement noise by individual synthetic sensor scaled by reliability $R_{satellite}$, based on the characteristics of the satellites used in the VOD data set (see Supplemental Table S1 and Methods for details). (c) Combined synthetic signal via taking the daily (blue) and bi-weekly (black) means.

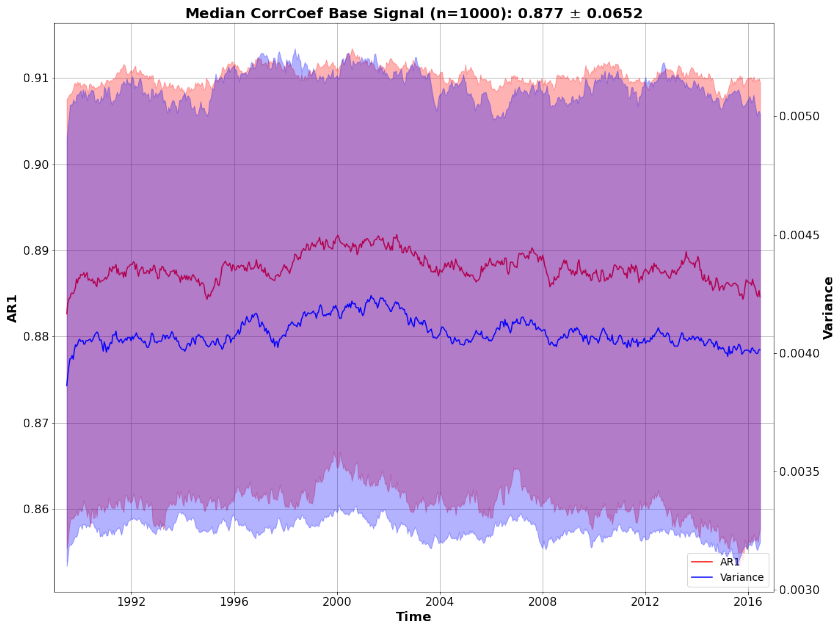

Figure S4. Median correlation coefficient between AR1 and variance (title, ± one standard deviation) for 1000 iterations of the underlying Ornstein-Uhlenbeck process without additional measurement noise. AR1 and variance are highly co-correlated, as is to be expected when the driving process and noise structure do not change through time or between synthetic samples.

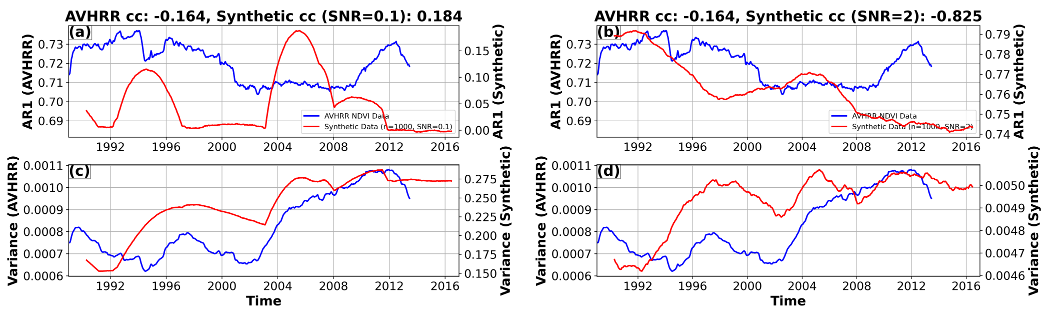

Figure S5. Comparison between real and synthetic data. (a,b) AR1, and (c,d) variance for synthetic data (red) and globally-averaged AVHRR GIMMS3g NDVI data (blue). Left column shows low signal-to-noise ratio (SNR=0.1), right column shows SNR=2. AR1 and variance calculated on a five-year rolling window. Correlation coefficients (cc) between AR1 and variance plotted in titles. The data sets show both negative and positive correlations between AR1 and variance depending on SNR. Note that satellite and synthetic data are not plotted on identical y-scales.