Reconstruction of multivariate satellite images using multiple-point statistics

Gregoire Mariethoz[1], Philippe Renard[2] and Matthew McCabe[3]

Abstract With the widespread use of satellite imaging, a wealth of information is available to help in the understanding and modeling of earth system processes. In particular, these data play a key role in the analysis of climate variability. However, satellite images can present data gaps due to partial coverage of the domain by the different orbital characteristics of the satellites. We present a method to fill these gaps and reconstruct the missing data with realistic property values using Direct Sampling multiple-point geostatistical simulation. The training images used are based on the known parts of similar images taken at previous dates.

Introduction

Satellite-based retrievals present spatial discontinuities due to incomplete coverage of the domain resulting from satellite orbital characteristics, or through occlusion by cloud cover and other atmospheric effects. We propose a method based on multivariate multiple-point geostatistics to fill these gaps with realistic values.

Important auxiliary information is often available at a gap location when other satellites that may inform upon a different (but related) variable cover the missing area. For example, there may be locations where soil moisture is not informed, but where a different satellite provides cloud coverage, temperature, or both. By using multivariate training images, our method allows maximizing the use of such auxiliary information, even if the relationships between the different reconstructed variables are highly non-linear.

The method is applied on synthetic imagery derived from a regional climate model of south-eastern Australia. The variables considered are soil moisture, temperature, latent heat flux (evaporation) and shortwave downward radiation (as

[1] The University of New South Wales, School of Civil and Environmental Engineering, Sydney, Australia, gregoire.mariethoz@unsw.edu.au

[2] University of Neuchatel, Centre of Hydrogeology and Geothermics, Neuchatel, Switzerland

[3] The University of New South Wales, School of Civil and Environmental Engineering, Sydney, Australia

a proxy for cloud coverage). Artificial gaps corresponding to satellites passages are made in the images (representing up to 40% of the domain). The gap locations are then reconstructed with our method and compared against the original data. The reconstructions illustrate that our method is able to accurately reconstruct the missing part of the satellite images, and to accurately reproduce the complex dependencies occurring between the variables considered.

Methodology

The problem of gap filling in spatially discontinuous data sets, including those inherent in retrievals from Earth observing systems, has been the focus of many research investigations [e.g. 1, 2-5]. In general terms, the gap filling problem can be formulated as determining the value of a pixel with spatial constraints (it must be coherent with the surrounding values), temporal constraints (it must be coherent with the preceding values) and also constraints related to any dependence with covariates (which may not necessarily be linear dependencies). For example, topography is a covariate which is known to be influential on the spatial distribution of rainfall [6].

A popular approach to gap-filling is cokriging [7]. Zhang et al [8] applied the technique to multispectral images to impose correlation with the same gap free image taken four months earlier. However, kriging and its variations have two major limitations: they are smoothing interpolators and can only account for linear relations with covariates. Using kriging can result in the interpolated areas presenting unrealistic continuous textures and, if point measurements are available, artifacts near these locations.

In this paper, we investigate the use of multivariate multiple-point geostatistics with continuous variable. The method employed here is the Direct Sampling approach [9, 10]. The training images used are constituted of similar satellite retrievals as the ones we want to reconstruct, but taken at previous dates. These images represent the typical patterns to be expected. Although gaps also exist in those past images, the informed areas are large enough to contain a large diversity of patterns, especially when one considers that such data are available at daily intervals for decades.

Application

The reconstruction method is applied to synthetic data derived from a regional climate model of south-eastern Australia [11]. Artificial gaps corresponding to exaggerated satellites scan tracks are imposed upon the images, masking an area of up to 40% of the domain. These artificial gaps are then reconstructed with

Direct Sampling and compared with the original model output. It is assumed that each variable has different gaps, as in practice they might be informed by a different satellite. Three variables are considered partially informed: the surface latent heat (LH), the surface temperature (TSK) and the soil moisture (SMOIS). Another variable, the shortwave downward radiation (SWDOWN) is considered exhaustively known. The purpose of using synthetic data derived from a model is to have a reference to compare the results against.

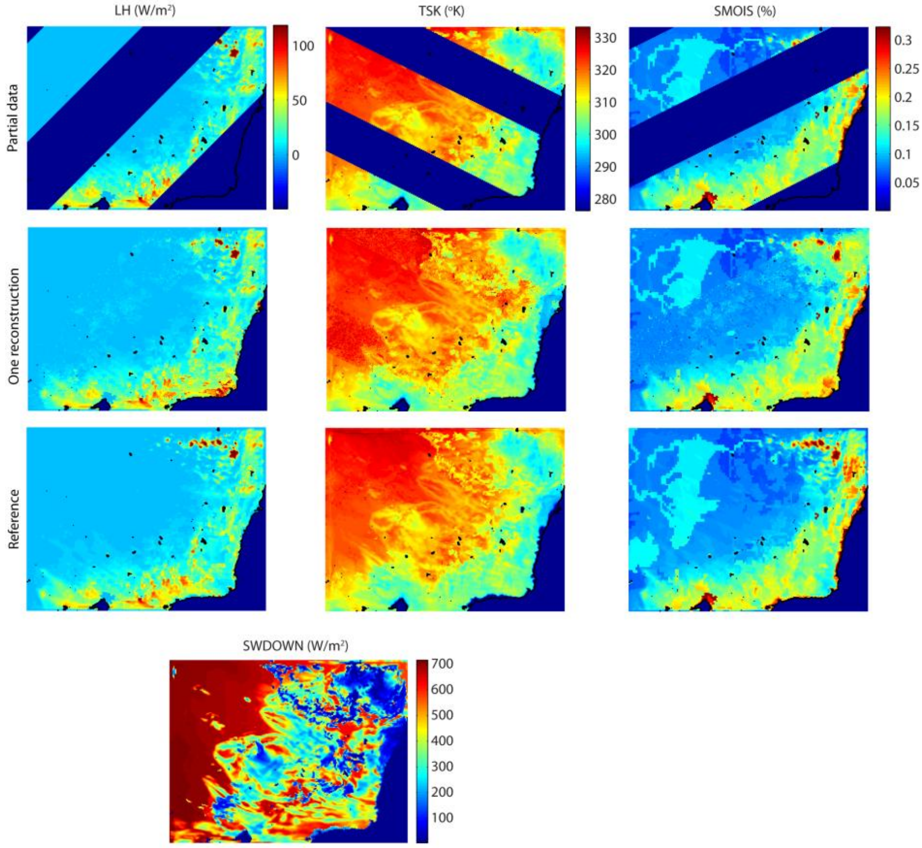

Figure 1 Gaps caused by orbital characteristics in a sample reconstruction for January 15th 2006. The tree columns representing WRF simulated data fields are (from left to right): latent heat flux (LH), surface temperature (TSK) and soil moisture (SMOIS). Each of the rows details (from top to bottom): the artificially gap enforced simulation, the reconstructed image, and the original continuous WRF simulation. Downward shortwave radiation is included at the bottom of the panel, with the influence of cloud evident throughout the lower left portion of the image.

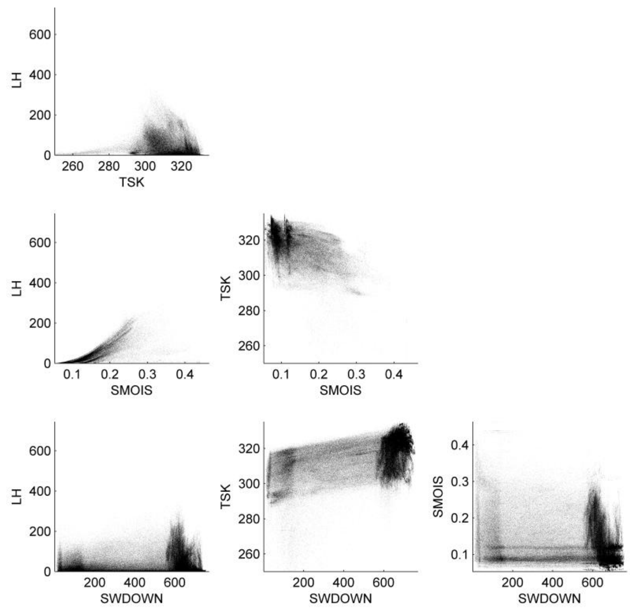

Figure 2 Sample scatter plots of the WRF reference variables for all dates in January, illustrating varying degrees of non-linear relationships between variables.

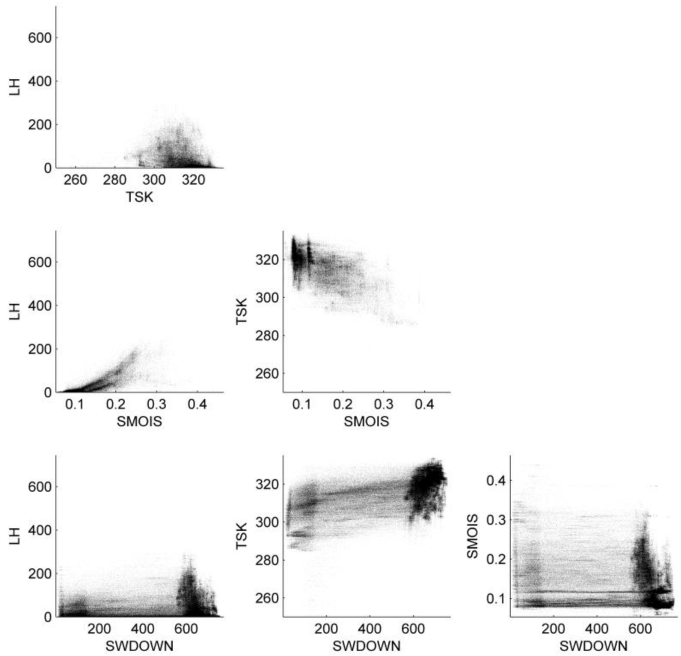

Figure 3 Scatter plots of the reconstructed gaps locations for all dates in January. As can be seen, there is good visual agreement of the non-linear relationships exhibited in the reference data.

The patterns in the reconstructed variables are realistic, especially given the significant amount of gaps present in the images, and the errors to the reference are small (Figure 1). More importantly, the complex relationships between the variables modeled are faithfully reproduced, which is very difficult to accomplish with other geostatistical approaches [12]. Figure 2 shows the complex relationships between variables at the gaps locations for the month of January 2006. Figure 3 shows the same relationships for the reconstructed values, which are strikingly similar to the reference.

Conclusion

Realistic reconstruction of missing areas in remote sensing retrievals is a challenging problem that goes far beyond simple interpolation. Here we present a newly developed geostatistical approach that accounts for the complex spatial, structural and textural properties of the variables considered and the inherent non-linear relationships in Earth surface variables.

Results from this analysis show that our Direct Sampling based gap-filling method is able to accurately reconstruct the missing elements of the synthetic satellite images. Complex spatial patterns can be resolved that also reproduce the often non-linear dependencies observed between the variables considered. Furthermore, the stochastic nature of the methodology makes it possible to ascertain the uncertainty related to the reconstruction. The governing principle is to use past occurrences of observed multivariate multiple-point relationships and then apply these to the present. This approach however assumes that the realm of possible outcomes is contained in the past observations, an assumption that may not always be the case.

Bibliography

1. Boloorani, A.D., S. Erasmi, and M. Kappas, *Multi-source remotely sensed data combination: Projection transformation gap-fill procedure.* Sensors, 2008. **8**(7): p. 4429-4440.

2. Maxwell, S.K., G.L. Schmidt, and J.C. Storey, *A multi-scale segmentation approach to filling gaps in Landsat ETM+ SLC-off images.* International Journal of Remote Sensing, 2007. **28**(23): p. 5339-5356.
3. Wang, G., et al., *A three-dimensional gap filling method for large geophysical datasets: Application to global satellite soil moisture observations.* Environmental Modelling & Software, 2012. **30**: p. 139-142.
4. Yuan, H., et al., *Reprocessing the MODIS Leaf Area Index products for land surface and climate modelling.* Remote Sensing of Environment, 2011. **115**(5): p. 1171-1187.
5. Zhang, C., W. Li, and D. Travis, *Gaps-fill of SLC-off Landsat ETM+ satellite image using a geostatistical approach.* International Journal of Remote Sensing, 2007. **28**(22): p. 5103-5122.
6. Goovaerts, P., *Geostatistical approaches for incorporating elevation into the spatial interpolation of rainfall.* Journal of hydrology, 2000. **228**(1-2): p. 113-129.
7. Zhang, X., et al., *Geostatistical interpolation of missing data and downscaling of spatial resolution for remotely sensed atmospheric methane column concentrations.* International Journal of Remote Sensing, 2012. **33**(1): p. 120-134.
8. Zhang, C., W. Li, and D.J. Travis, *Restoration of clouded pixels in multispectral remotely sensed imagery with cokriging.* International Journal of Remote Sensing, 2009. **30**(9): p. 2173-2195.
9. Mariethoz, G. and P. Renard, *Reconstruction of incomplete data sets or images using Direct Sampling.* Mathematical Geosciences, 2010. **42**(3): p. 245-268.
10. Mariethoz, G., P. Renard, and J. Straubhaar, *The Direct Sampling method to perform multiple-point geostatistical simulations.* Water Resources Research, 2010. **46**(W11536).
11. Evans, J.P. and M.F. McCabe, *Regional climate simulation over Australia's Murray-Darling basin: A multitemporal assessment.* Journal of Geophysical Research D: Atmospheres, 2010. **115**(14).
12. Mariethoz, G., P. Renard, and R. Froidevaux, *Integrating collocated auxiliary parameters in geostatistical simulations using joint probability distributions and probability aggregation.* Water Resources Research, 2009. **45**(W08421).