ACCOUNTING FOR TEMPORAL INFORMATION FROM DENSE TIME-SERIES COARSE-SCALE SATELLITE DATA FOR SPATIO-TEMPORAL DOWNSCALING

Yeseul Kim, No-Wook Park

Inha University, 100 Inha-ro, Michuhol-gu, Incheon 22212, Korea
Email: kim-6674@inha.edu; nwpark@inha.ac.kr

KEY WORDS: Spatio-temporal downscaling; data fusion; component decomposition

ABSTRACT   : A spatio-temporal downscaling model is proposed in this paper to fuse complementary information in spatial and temporal resolution of remote sensing data. The proposed model is based on decomposition of a target attribute of interest into trend and residual components, which is named as a decomposition-based spatio-temporal downscaling (DSPAT). Based on the component decomposition, DSPAT is composed of (1) estimation of trend components and (2) correction of residual components. Unlike conventional spatio-temporal downscaling models, DSPAT fully considers temporal information of dense time-series coarse-scale (DC) data during trends estimation. The residuals are predicted by a geostatistical interpolation at a dense time-series coarse-scale and then incorporated with the trends. To evaluate the applicability of DSPAT, an experiment was conducted using simulated reflectance data. In addition, its performance was compared with conventional models including a spatial and temporal adaptive reflectance fusion model (STARFM) and an enhanced STARFM (ESTARFM). From the experimental results, DSPAT showed better performance than STARFM and ESTARFM even when the difference between the prediction date and the acquisition time of both DC and SF increased. These results demonstrate the potential of DSPAT that can fully account for the temporal information of DC data.

1. INTRODUCTION

Remote sensing data have inevitable a trade-off in temporal and spatial resolutions. That is, remote sensing data with a fine-scale are usually acquired at a sparse temporal resolution (i.e. sparse time-series fine-scale data, SF), while those with a dense temporal resolution are usually acquired at a coarse-scale (i.e. dense time-series coarse-scale data, DC). To fully use above complementary information between DC and SF, data fusion that generates synthetic images with a dense time-series fine-scale can be applied for monitoring spatio-temporal dynamics in environmental applications (Xue et al., 2017). In the sense of providing detailed information in space and time, data fusion can be also called spatio-temporal downscaling.

Many studies have been conducted to produce synthetic data with a temporally dense fine-scale. Among various spatio-temporal downscaling models, a spatial and temporal adaptive reflectance fusion model (STARFM) has been widely applied (Gao et al., 2006). STARFM predicts using a weighted function for neighboring coarse-scale pixels; particularly, it assigns a higher weight for a coarse-scale pixel that only includes one land-cover type. However, since most coarse-scale pixels include various land-cover types (i.e. heterogeneous spatial patterns), the prediction accuracy may not be satisfactory in areas with heterogeneous land-cover patterns. In this regard, an enhanced spatial and temporal adaptive reflectance fusion model (ESTARFM) has been proposed to improve prediction accuracy in heterogeneous patterns (Zhu et al., 2010). In addition, various spatio-temporal downscaling models have been also proposed including dictionary learning, Bayesian model, and unmixing analysis approaches (Huang and Song, 2012; Zhu et al., 2016; Xue et al., 2017; Liu et al., 2019). These models have mainly focused on improvement of the prediction accuracy in heterogeneous patterns. Although the prediction accuracy in heterogeneous patterns could be improved, the conventional models could not fully consider temporal information of DC data because they are only applicable when the DC and SF data are simultaneously acquired. In turn, the prediction performance may degrade when the acquisition time of both DC and SF and the prediction time are significantly different.

To fully consider temporal information of DC data, we propose a decomposition-based spatio-temporal downscaling model (DSPAT) which is based on decomposition of a target attribute into deterministic trend and stochastic residual components. DSPAT can consider both temporal information of DC data and local variations of SF data in trend estimation and incorporate residuals that are not included in the trend estimation by a geostatistical interpolation. The applicability of DSPAT is evaluated via an experiment with simulated reflectance data and its performance is compared with two widely applied models including STARFM and ESTARFM.

2. METHODOLOGY

In DSPAT, a target attribute is regarded as a realization of a random function and decomposed into the trend and residual components (Kyriakidis and Journel, 2001). The spatio-temporal downscaling result is obtained by summing the trend and residual components estimated at a dense time-series fine-scale. More precisely, DSPAT consists of three analytical steps: (1) quantification of temporal information of DC data, (2) application of a deconvolution matrix to estimate dense time-series fine-scale trend components, and (3) correction of residuals estimated at a dense time-series fine-scale via geostatistical interpolation (Figure 1).

First, DSPAT uses a land-cover map to quantify temporal information from DC data. The temporal variations of each land-cover type are extracted, which is the temporal mean values of all coarse-scale pixels that are the same land-cover type. The temporal information is then modelled by quantifying the relationship between the temporal value of each coarse-scale pixel and the temporal mean value of the land-cover type to which the pixel corresponds. The relationship can be estimated by applying a regression model where the temporal values of a coarse-scale pixel and the temporal mean values of the corresponding land-cover type are considered as dependent and independent variables, respectively. Then, these relationships are applied to each coarse-scale pixel with temporal mean values of the corresponding land-cover type to estimate the temporal information.

To obtain the temporal information at a fine-scale, a deconvolution matrix is used since the temporal information is quantified at a coarse-scale. The deconvolution matrix is a linear relationship between the attribute values of DC and SF which are simultaneously acquired. A convolution matrix, on the other hand, is used to estimate the attribute values of DC data from those of SF data. Here, the convolution matrix can be constructed from a point spread function (PSF) for upscaling, which is based on an assumption that the PSF for upscaling is a Gaussian kernel in DSPAT. After the convolution matrix has been constructed, the deconvolution matrix can be initialized as a transposed of the convolution matrix. The initial deconvolution matrix is optimized by applying this initial matrix to the DC and SF data acquired simultaneously. Using the optimized deconvolution matrix, the trend components at a dense time-series fine-scale are then estimated.

The next step is to estimate the residuals at a dense time-series fine scale. The residuals are computed at a coarse-scale but temporally dense by subtracting DC data and upscaled trend components (Figure 1). In DSPAT, area-to-point kriging (ATPK) is applied to estimate temporally dense residuals at a fine-scale. ATPK is a geostatistical interpolation that can account for the difference of spatial resolution between coarse-scale input data and fine-scale prediction locations (Kyriakidis, 2004). Variogram deconvolution is applied to compute block kriging weighting values (Goovaerts, 2008). The final spatio-temporal downscaling result is obtained by summing the trends and residuals estimated at a dense time-series fine-scale.
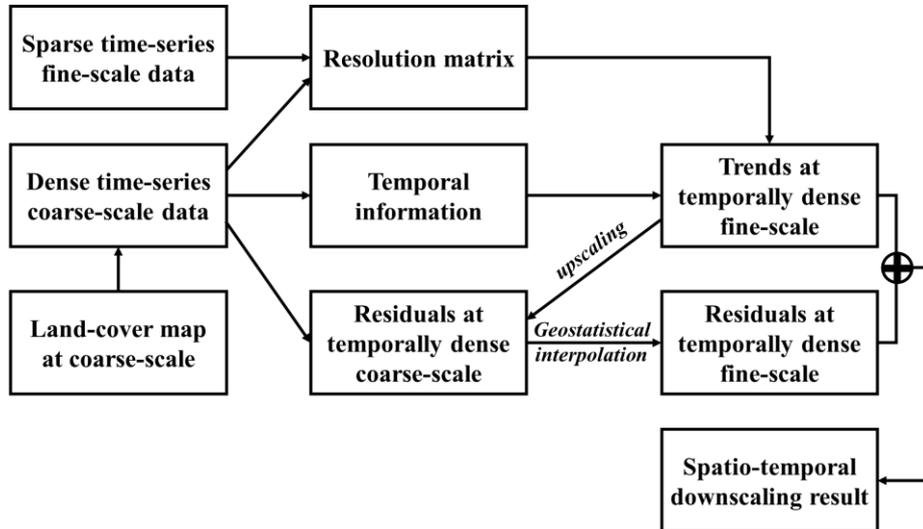

Figure 1. Workflow of DSPAT.

Table 1. Acquisition dates of DC and SF data. Filled circles denote DC-SF pairs.
Open circles denote the prediction date.

Date	Feb. 01	Feb. 04,	Feb. 05	Feb. 06	Feb. 07	Feb. 08	Feb. 13
Prediction date	●		○		○	○	
Date	Feb. 15	Feb. 17	Mar. 03	Mar. 12	Mar. 14	Mar. 23	Mar. 25
Prediction date	○	●	○	○	○	○	○
Date	Mar. 26	Mar. 27	Mar. 28	Apr. 10	Apr. 18	Apr. 19	Apr. 21
Prediction date		○	○	○		○	○
Date	Apr. 28	May 14	May 24	May 26	June 01	June 07	June 22
Prediction date	○	○	●	○		○	○
Date	Aug. 01	Aug. 02	Oct. 12	Oct. 20	Oct. 21	Oct. 30	Oct. 31
Prediction date	○	○	○		○	○	○

3. EXPERIMENT FOR ALGORITHM EVALUATION

3.1 Methods

The proposed DSPAT was evaluated via an experiment with simulated reflectance data. As input data for spatio-temporal downscaling, we used MODIS daily reflectance products (MOD09GQ) acquired from February to October 2018 in cropland of South Korea. Considering the spectral characteristics of crops, red and near-infrared (NIR) channels were used as inputs. MODIS reflectance data at 250 m were first upscaled to 1 km scale, which were considered as DC data. The original reflectance data at 250 m were used as unknown true SF data. At the acquisition date of MODIS products, 35 DC data were used as input data, except for images including clouds and snow (Table 1). Random forest was applied as a regression model to estimate temporal information from these DC data.

Three DC-SF pairs (1 February, 17 February, and 24 May in 2018) were used as input data for estimation of the deconvolution matrix. In the experiment, spatio-temporal downscaling was performed for 25 dates (Table 1) and the performance was evaluated by comparing with the true SF data. In addition, the performance of DSPAT was compared with STARFM and ESTARFM which have been widely used for spatio-temporal downscaling. For performance evaluation, two indices were computed and used as quantitative assessment: (1) root mean square error (RMSE) and (2) structure similarity (SSIM). The SSIM index can evaluate the similarity of overall structures between the true and predicted SF data (Zhu et al., 2016). The closer the value of SSIM is to 1, the more similar the structure is.

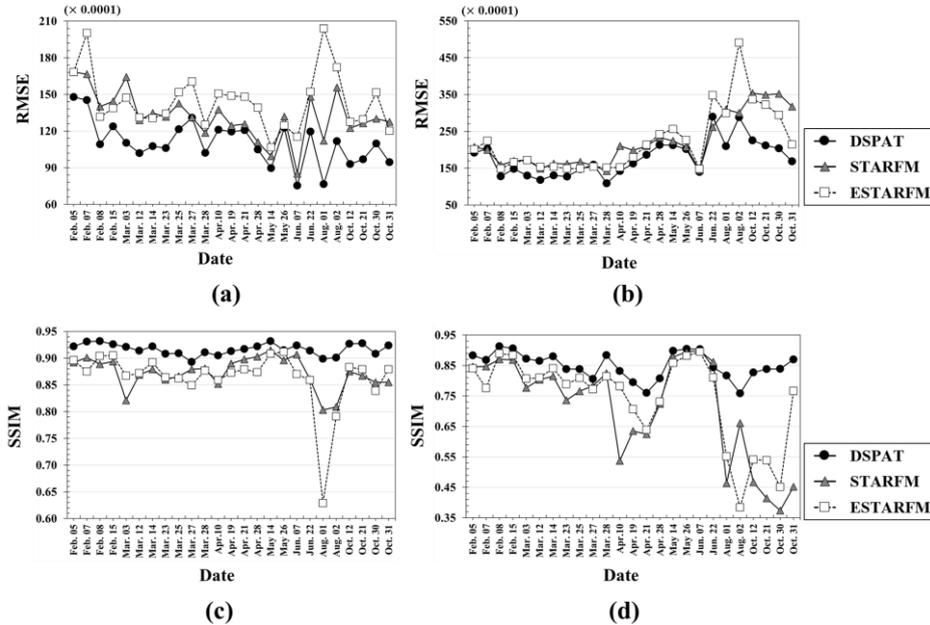### 3.2 Results and Discussion

Figure 2. Comparison of RMSE and SSIM values of different models.
RMSE values of (a) red and (b) NIR channels. SSIM values of (c) red and (d) NIR channels.

First, as the quantitative assessment for the spatiotemporal downscaling results, two indices are presented in Figure 2. When comparing the RMSE and SSIM for each channel, DSPAT showed the relatively lower RMSE values and higher SSIM values than those of STARFM and ESTARFM. In particular, when the SSIM values of downscaling results for NIR channel were compared, the SSIM values of STARFM and ESATRFM became smaller as the difference between the DC-SF pair dates and the prediction date increased. In contrast, the SSIM values of DSPAT are very similar to the true SF data when compared with those of STARFM and ESTARFM (Figure 2(c) and (d)), even when the difference between the DC-SF pair date and the prediction date increased. Figure 3 is also demonstrated these results. These results could be obtained because STARFM and ESTARFM failed to fully account for temporal information of DC data. On the other hand, DSPAT could fully account for temporal information of DC data during estimation of trend components, yielding the downscaling results that are relatively similar to the true SF data.

However, DSPAT showed lower correlation and SSIM values than those of STARFM in the spatio-temporal downscaling results of NIR channel on June 22, 2018. As shown in Figure 4, this may be due to over-estimation for high values in the spatio-temporal downscaling result of DSPAT. This over-estimation for high values may be occurred in the estimation of deconvolution matrix to account for the local variations of SF data because the matrix is estimated by relationship between a fine-scale pixel and a corresponding coarse-scale pixel. For this issue, the possible solution is to estimate the relationship between a fine-scale pixel and not only the corresponding coarse-scale pixel but also the adjacent coarse-scale pixel. This issue should be included in future work to improve the estimation step of deconvolution matrix.
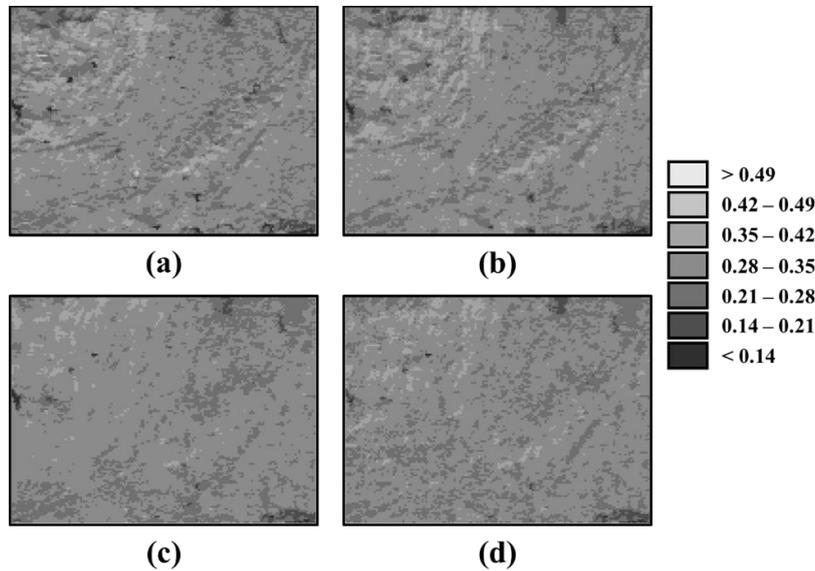

Figure 3. Visual comparison of downscaling results for NIR channel acquired in October 12, 2018: (a) true SF data, the predictions by (b) DSPAT, (c) STARFM, and (d) ESTARFM.
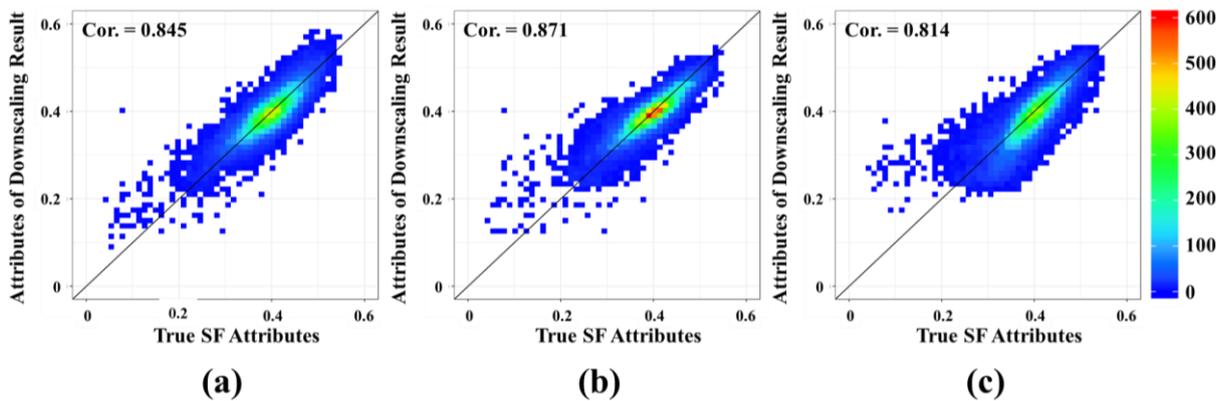

Figure 4. Scatter plots of three spatio-temporal downscaling results with true SF data of NIR channel acquired in June 22, 2018: (a) DSPAT, (b) STARFM, and (c) ESTARFM.

4

4. CONCLUSION

In this study, we proposed a spatio-temporal downscaling model based on component decomposition, named as DSPAT, that can fuse DC and SF data which have complementary information in spatial and temporal resolution. The proposed DSPAT, particularly, can consider both temporal information of DC data and local variations of SF data. We have evaluated DSPAT with an experiment using simulated reflectance data for cropland in South Korea. When comparing conventional spatio-temporal downscaling models including STARFM and ESTARFM, DSPAT could generate the spatio-temporal downscaling results which were spatially and quantitatively similar to true SF data. The prediction accuracy of DSPAT was particularly higher than that of STARFM and ESTARFM even when the difference between prediction date and DC-SF pair date increases, since DSPAT takes into account the temporal information of DC data. Based on these results, it is expected that DSPAT can be effectively applied to spatio-temporal downscaling when the DC-SF pair date and prediction date are significantly different. To enhance the applicability of DSPAT, extensive experiments will be carried out for spatio-temporal downscaling of satellite-based products including soil moisture and land surface temperature, as well as reflectance.

REFERENCES

Gao, F., Masek, J., Schwaller, M., and Hall, F., 2006. On the blending of the Landsat and MODIS surface reflectance: predicting daily Landsat surface reflectance. IEEE Transactions on Geoscience and Remote Sensing, 44 (8), pp. 2207-2218.

Goovaerts, P., 2008. Kriging and semivariogram deconvolution in the presence of irregular geographical units. Mathematical Geosciences, 40 (1), pp. 101-128.

Huang, B. and Song, H., 2012. Spatiotemporal reflectance fusion via sparse representation. IEEE Transactions on Geoscience and Remote Sensing, 50 (10), pp. 3707-3716.

Kyriakidis, P.C., 2004. A geostatistical framework for area-to-point spatial interpolation. Geographical Analysis, 36 (3), pp. 259-289.

Kyriakidis, P.C. and Journal, A.G., 2001. Stochastic modeling of atmospheric pollution: a spatial time-series framework. Part I: methodology. Atmospheric Environment, 35 (13), pp. 2331-2337.

Liu, M., Yang, W., Zhu, X., Chen, J., Chen, X., Yang, L., and Helmer, E.H., 2019. An improved flexible spatiotemporal data fusion (IFSDAF) method for producing high spatiotemporal resolution normalized difference vegetation index time series. Remote Sensing of Environment, 227, pp. 74-89.

Xue, J., Leung, Y., and Fung, T., 2017. A Bayesian data fusion approach to spatio-temporal fusion of remotely sensed images. Remote Sensing, 9 (12), pp. 1310.

Zhu, X., Chen, J., Gao, F., Chen, X., and Masek, J.G., 2010. An enhanced spatial and temporal adaptive reflectance fusion model for complex heterogeneous regions. Remote Sensing of Environment, 114, pp. 2610-2623.

Zhu, X., Helmer, E.H., Gao, F., Liu, D., Chen, J., and Lefsky, M.A., 2016. A flexible spatiotemporal method for fusing satellite images with different resolutions. Remote Sensing of Environment, 172, pp. 165-177.