Evaluation of the event driven phenology model coupled with the VegET evapotranspiration model through comparisons with reference datasets in a spatially explicit manner.

V. Kovalskyy, [1] G.M. Henebry, [1] B. Adusei, [1] M. Hansen, [1] D.P. Roy, [1] G. Senay,[2] D.M. Mocko [3].

[1]South Dakota State University Geographic Information Science Center of Excellence, Brookings, South Dakota, USA

[2]United States Geological Survey's National Center for Earth Resources Observation and Sciences , Sioux Falls, South Dakota, USA

[3]GMAOSAIC at the Hydrological Sciences Branch and the Global Modeling and Assimilation Office, NASA Goddard Space Flight Center, Greenbelt, Maryland, USA

Abstract

A new model coupling scheme with remote sensing data assimilation was developed for estimation of daily actual evapotranspiration (ET). The scheme represents a mix of the VegET, a physically based model to estimate ET from a water balance, and an event driven phenology model (EDPM), where the EDPM is an empirically derived crop specific model capable of producing seasonal trajectories of canopy attributes. In this experiment, the scheme was deployed in a spatially explicit manner within the croplands of the Northern Great Plains. The evaluation was carried out using 2007-2009 land surface forcing data from the North American Land Data Assimilation System (NLDAS) and crop maps derived from remotely sensed data of NASA's Moderate Resolution Imaging Spectroradiometer (MODIS). We compared the canopy parameters produced by the phenology model with normalized difference vegetation index

(NDVI) data derived from the MODIS nadir bi-directional reflectance distribution function (BRDF) adjusted reflectance (NBAR) product. The expectations of the EDPM performance in prognostic mode were met, producing determination coefficient (r^2) of 0.8 ±0.15. Model estimates of NDVI yielded root mean square error (RMSE) of 0.1 ±0.035 for the entire study area. Retrospective correction of canopy dynamics with MODIS NDVI brought the errors down to just below 10% of observed data range. The ET estimates produced by the coupled scheme were compared with ones from the MODIS land product suite. The expected r^2=0.7 ±0.15 and RMSE = 11.2 ±4 mm per 8 days were met and even exceeded by the coupling scheme functioning in both prognostic and retrospective modes. Minor setbacks of the EDPM and VegET performance (r^2 about 0.5 and additional 30 % of RMSR) were found on the peripheries of the study area and attributed to the insufficient EDPM training and to spatially varying accuracy of crop maps. Overall the experiment provided sufficient evidence of soundness and robustness of the EDPM and VegET coupling scheme, assuring its potential for spatially explicit applications.

1. Introduction.

There is growing consensus in the climate science community that the ability to precisely partition energy and matter fluxes on the land surface is key to improving our understanding of mesoscale atmospheric dynamics, ecosystem, responses to climate change, and interactions with human activities and institutions [*Pitman*, 2003; *Ibanez et al.*, 2010]. Since *Manabe* [1969], land surface modules (LSM) have become increasingly complex modules within most general circulation models (GCMs). The complexities of LSMs have grown substantially as scientists ask questions about the pace and consequences of climate change that require more precise answers. In pursuit of these answers, researchers have been coupling global and regional climate models with a spectrum of modules detailing interactions between the land surface and the lowest level of the atmospheric boundary layer. Modules range from a set of simplified surface energy and water balance procedures to more detailed interactive systems like dynamic soil and vegetation modules, complete with light transfer, photosynthesis, and hydrological schemes. Computational resources often limit the level of detail in LSMs especially in regional studies that require finer spatial resolution. Also, there is a trade-off between the number of land surface characteristics that can be tracked and the greater spatial detail often needed for regional to local projects [*Stensrud,* 2007].

Applications of land surface models in regional studies were often focused on just a few variables of interest. In many instances, this narrower focus has led to the use of simplified schemes of land surface processes. Numerous local impact studies are turning to empirical methods based on relationships of modeled land surface characteristics to net radiation, precipitation, air temperature and other variables [*Nagler et al.,* 2005; *Godfrey et al.,* 2007; *Senay et al.,* 2007; *Abramowitz et al.,* 2008; *Jang et al.,* 2009; *Gao et al.,* 2010]. However, being developed on microclimatological data, empirical models were often unable to predict well when transferred to a different location, even under similar conditions [*Li et al.,* 2009]. Yet, deployment of process-based LSMs to address local questions are often hindered by computational expense and a lack of appropriate ground level data to calibrate and validate at the level of spatial detail required. Also, several studies have expressed concerns about model assumptions, process parameterizations, and a limited range of parameters available for tuning [*Sabater et al.,* 2007; *Kiniry et al.,* 2008; *Kang et al.,* 2009; *Stancalie et al.,* 2010], all of which increase doubts about the likelihood of successful deployment of LSMs in regional to local studies. Alternatives solutions are needed to provide robust schemes capable of replacing complex LSMs in finer spatial resolution studies. This paper presents a recent development in land surface modeling combining both physics-based and empirical approaches to take advantage of the strengths of each approach while yielding results on an appropriately fine scale.

Our research focuses on how potential futures for rainfed agricultural production in the Northern Great Plains may affect regional hydrometeorology. Actual evapotranspiration (ET_a) was the key flux of interest. We chose to use a simplified simulator of ET_a called VegET [*Senay* 2008]. Similar to *Godfrey et al.* [2007], *Kang et al.* [2009] and *Yuan et al.* [2010], *Senay's* scheme relies on the Penman-Monteith equation [*Monteith*, 1964] to calculate reference ET (ET_0) and handles the influences of soil water status and canopy phenology through the two coefficients: K_s for soil water status and K_{cp}, for canopy phenology. The Penman-Monteith method is a physics-based one source model of evapotranspiration in cereal crops with fully developed canopies, used extensively by FAO [*Allen et al.*, 1998]. A key innovation of VegET is the modulation of ET_a by a canopy phenology coefficient using a climatology of the normalized difference vegetation index (NDVI) as observed from spaceborne sensors [*Senay*, 2008].

The original implementation of VegET, however, could not serve our purpose because we were seeking how ET_a would change in response to both interannual variability and changes in the crop area. Since they were derived from averages of past observations, a static retrospective climatology for K_{cp} would not reflect changes in growing conditions [*Godfrey et al.*, 2007; *Wegehenkel*, 2009] or in the extent of cultivation [*Kovalskyy and Henebry*, 2011b]. Therefore, we replaced a static phenological parameterization with an interactive vegetation growth module. The use of fully functional crop growing modules with energy balance models in point based studies has been common practice [*Maruyama and Kuwagata*, 2010; *Sancalie et al.*, 2010]. However, our study case required spatially explicit ET_a estimates that would entail additional parameterization, tuning and running time for the models like ALMANAC [*Kiniry et al.,* 2008], CERES [*Mearns et al.,* 1999], CROPWAT [*Sancalie et al.,* 2010] or MODWht [*Kang et al.,* 2009]. Moreover, these specific crop models did not have freely available versions capable of working with raster inputs and producing spatially explicit estimates. Conversely, the vegetation growth modules in global LSMs were developed to deliver spatially explicit results [*Dickinson et al.,* 1998; *Foley et al.,* 2000; *Bondeau et al.,* 2007; *Campo et al.,* 2009]. Even the most advanced modules do not provide crop specific canopy behavior; instead, they were designed to mimic seasonal patterns of very broad classes of vegetation functional types [*Bonan et al.,* 2003; *Lawrence and Chase*, 2007].

Here we have used the Event Driven Phenology Model (EDPM), which was recently developed as a phenology model that can simulate seasonal dynamics of canopy properties (e.g., in terms of a vegetation index) [*Kovalskyy and Henebry,* 2011a, 2011b]. The model was shown to capture fine temporal details of canopy behavior [*Kovalskyy and Henebry*, 2011a, 2011b] which has been called "crucial" for ET and other surface fluxes [*Dickinson et al.* 1998; *Foley et al.* 2000; *Pitman*, 2003; *Gorfrey et al.* 2007; *Prihodko et al.* 2008; *Rosero et al.* 2009; *Rötzer et al.* 2010; *Zha et al.* 2010]. The EDPM uses virtually the same set of forcings as the Penman-Monteith equation to build seasonal trajectories of canopy properties. Essentially, the model provides a computationally inexpensive replacement for a dynamic vegetation model with a phenology sub-module. The model also has an option of a simple, fast 1D data assimilation scheme for satellite observations which is a great advantage for spatially explicit simulation studies. The EDPM has been coupled with VegET and evaluated against flux tower observations of ET_a [*Kovalskyy and Henebry*, 2011b], where it performed better or comparable to the results obtained by *Nagler et al.* [2005] and *Abramowitz et al.* [2008].

This paper presents an assessment of the performance of the EDPM on its own and also in conjunction with VegET within a spatially explicit application. Our task was to select appropriate sources of scientifically sound data products that would enable pixelwise comparisons of daily canopy states and ET_a estimates. We were looking to assess two aspects of the coupled model performance. First and foremost, we focused on temporal and spatial behavior of differences between our estimates and reference data. Analyzing the results from the three years (2007-2009) within the study region delimited by croplands of Northern Great Plains, we tried to capture both inter-annual and intra-annual variability of residuals as well as correlation between reference data and estimates produced by our model. Second, we looked at the ability of the EDPM to capture the key dates of the three growing seasons. We contrasted phenological dates reported to National Agricultural Statistics Service (NASS) by farmers with the dates produced by the EDPMs phenophase control module. Specifically, the Start of Season (SoS), End of Season (EoS), and Length of Season (LoS) became the main criteria for the evaluation. We also tried to incorporate spatial and temporal variability of phenological metrics into the evaluation process. This assessment helped us to identify the strong points of the EDPM and to prioritize directions for model improvement.

2. Methods and materials.

2.1 Study area.

The study area includes Nebraska, Iowa, Minnesota, North and South Dakota entirely and parts of Illinois and Indiana. Together these states have more than half of the nation's maize and soybean crops and comprise the major part of the US corn and soybean belts. There strong gradients of ET across the region. The northern tier has only 600 mm ET annually; whereas at the southern end, the annual ET can reach 1000 mm, but only 400 mm at the western extreme.

Maize and soybean are the most prevalent crops across the region. Farmers use different genetic varieties of these crops to match the growing conditions of their farms [*Ransom et al.*, 2004]. The green-up of the area starts in early May on the southeast end but for the northwestern part of the region it can happen as late as mid-June if spring comes late. The length of the growing cycle also varies greatly; it can last almost five months in the south and barely more than three months in the north.

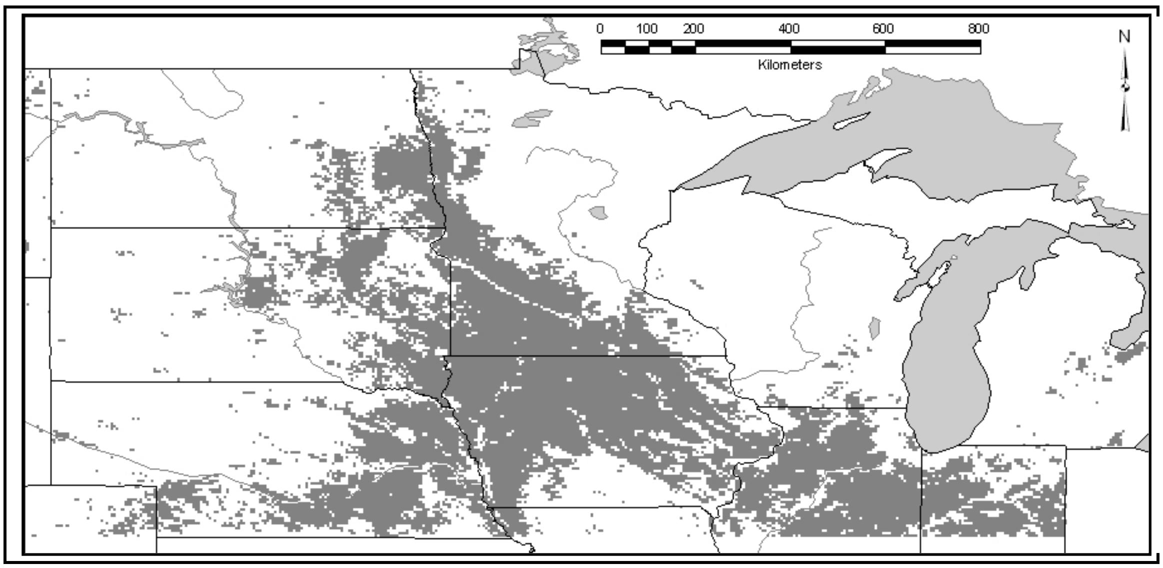


Figure 1. The study area (dark gray) depicted as at least 50% corn/soybean crop cover during 2007-2009.

2.2 Coupling the VegET and the EDPM in a spatially explicit manner.

The idea motivating the development of VegET was the use the time series of remotely sensed vegetation indices to drive the canopy factor that modifies ET_0 as calculated by the Penman-Monteith model. The original design of VegET [*Senay*, 2008] used very simple empirical transformations from the normalized difference vegetation index (NDVI) to phenology driven coefficients based on thresholds and the observed variability in NDVI climatologies derived from long-term AVHRR observations. However, *Kovalskyy and Henebry* [2011b] demonstrated the use of an interactive event driven phenology model [*Kovalskyy and Henebry,* 2011a] to replace the climatologies in the VegET for point-based estimation of daily ET_a. The coupling scheme was shown to account for contemporaneous fluctuations in the canopy component of evapotranspiration

The experiment described here evaluates the performance of the EDPM and the VegET after the coupled models were extended for deployment in a spatially explicit manner. In addition to simulating the temporal dynamics of maize and soybean over the three year period (2007-2009), the EDPM was also providing seasonal canopy trajectories for a third vegetation type: grassland. Using weather forcing from the North American Land Data Assimilation System (NLDAS), the EDPM transformed the data in to events (rain, heat stress, frost, insufficient insolation, adequate insolation, and growing degree-days) and further produced daily values of Tower NDVI [*Huemmrich et al.*, 1999], The model simultaneously estimated phenological transition dates in the three growing seasons for each vegetation type (Fig. 2). The TNDVI trajectories were mixed linearly based on the proportion of their cover within areal units (0.05 degree pixels) and later transformed into phenology coefficients as described in *Kovalskyy and Henebry* [2011b]. The percentages of cover for each crop and grassland were taken from MODIS based crop maps products [*Chen et al.*, 2007] and aggregated into standard 0.05 degree (~5km) pixels that form the spatial unit of analysis for this investigation.

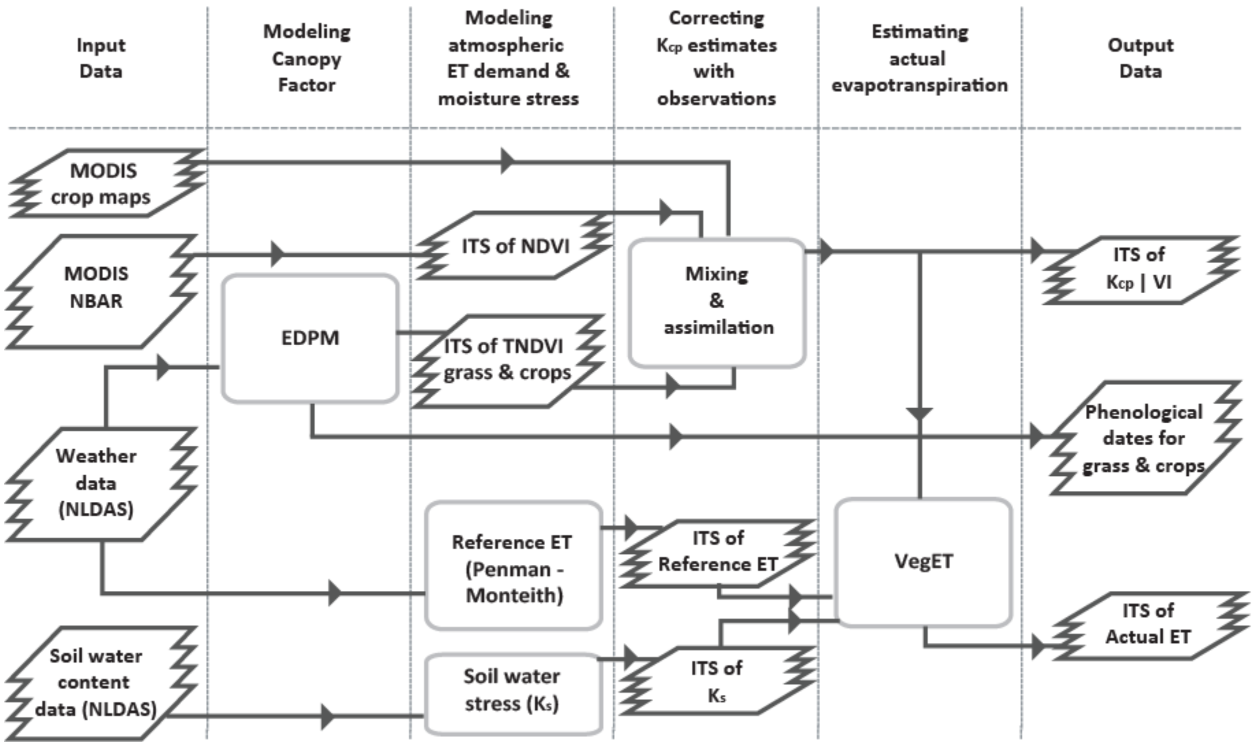


Figure 2. Data processing scheme for the experiment. Rounded boxes are modeling and data preparation procedures; stacks are image time series (ITS) of data; squared boxes are maps of results.

Using the workflow shown in Figure 2 the coupled scheme was tested in the prognostic mode (running the forcing only) and the diagnostic mode (involving data assimilation scheme with MODIS NDVI observations). The use of data assimilation techniques is becoming increasingly popular in evapotranspiration studies [*Meng et al.*, 2009; *Anderson et al.*, 2010; *Miralles et al.*, 2010; *Godfrey and Stensrud*, 2010]. Most of these projects relied on remotely sensed data to improve their estimates of ET addressing the spatial variability of land surface. While bringing improvements in performance, these techniques have been criticized as being temporally constrained and scene dependent [*Li et al.*, 2010]. Our study took a more general approach to data assimilation using an unambiguous relationship between Tower NDVI and MODIS NDVI established previously [*Kovalskyy et al.*, 2011]. Relying on this relationship *Kovalskyy and Henebry* [2011a] presented a one-dimensional Kalman filter (1DKF) data assimilation scheme in which the EDPM used MODIS NBAR NDVI to adjust its estimates of canopy states.

2.2.1 Differences from point based deployment of the EDPM.

Several features in the EDPM model were added and modified so that the model could represent spatial variability of canopy development during growing season. First of all, the model received the ability to represent pixels with mixed vegetation cover. The Figure 2 shows the linear mixing procedure used to derive values of vegetation index in a pixel with partial covers of grassland and the two crops. Direct linear mixing of NDVI values based on their share of pixel area has been criticized in the literature due to its impact on outcomes [*Settle and Campbell,* 1998; *Roy,* 2000; *Busetto et al.*, 2008]. Since the NDVI is not a linear function of red and near infrared reflectances, linear mixing should be performed on reflectances first so that "unbiased" NDVI values can be obtained later. However, a relatively small impact from direct linear mixing (DLM) of NDVI values may not be entirely prohibitive since the reflectances coming from MODIS products have their own errors of the estimate [*Roy et al.*, 2005]. If the differences between the DLM NDVI and the NDVI derived from reflectances can be kept within the margin of error propagated into the unbiased NDVI, then one can successfully perform linear mixing directly using NDVI values. The magnitude of differences with true values depends on the number of endmembers used in linear mixing and varies greatly across space and time. We evaluated a real (empirical) example to demonstrate how the direct linear mixing procedure impacts the resulting values of NDVI.

First we took a 1000 by 1000 pixel subset from the MODIS Nadir Bi-directional Reflectance Distribution Function (BRDF) Adjusted Reflectance (NBAR) MOD43A4 [*Schaaf et al.,* 2002] version product covering the central part of the study area. We screened for snow and clouds and based on averages aggregated the 500m reflectance values from MODIS bands 1 and 2 to produce image time series with 1000 m by 1000 m, 2500 m by 2500 m, 2500 m by 5000 m (rectangular shape pixel) and 5000 m by 5000 m pixel sizes. For each time series less than 25 km^2 in size, we calculated NDVI [1] that was later mixed directly into 5000 m pixels. Correspondingly, the four sets of results represented 100, 25, 4, and 2 endmember mixing. Out of 5000 m reflectance data we calculated "true" NDVI [1] and expected error [2] propagated from reflectances: 0.004 for band 1 and 0.015 for band 2 [*Roy et al.*, 2005].

$$NDVI = \frac{\rho_N - \rho_R}{(\rho_N + \rho_R)^2} \quad [1]$$

$$\sigma^2_{NDVI} = \left(\frac{-2\rho_N}{(\rho_N + \rho_R)^2} \right)^2 \sigma^2_R + \left(\frac{2\rho_R}{(\rho_N + \rho_R)^2} \right)^2 \sigma^2_N \quad [2]$$

where ρ_N and ρ_R are the reflectance values of near infrared and red bands respectively, and σ^2_N and σ^2_R are the associated variances. In this setup where all resolutions of NDVI data were nested within 5000 m pixels, we expected to see the difference coming just from linear mixing without other effects such as re-projection or resampling that may otherwise contribute to the difference [*Roy*, 2000]. Figure 3 demonstrates the temporal dynamics of average differences between true NDVI and the four DLM NDVI sets coming from linear mixing with different number of endmembers.

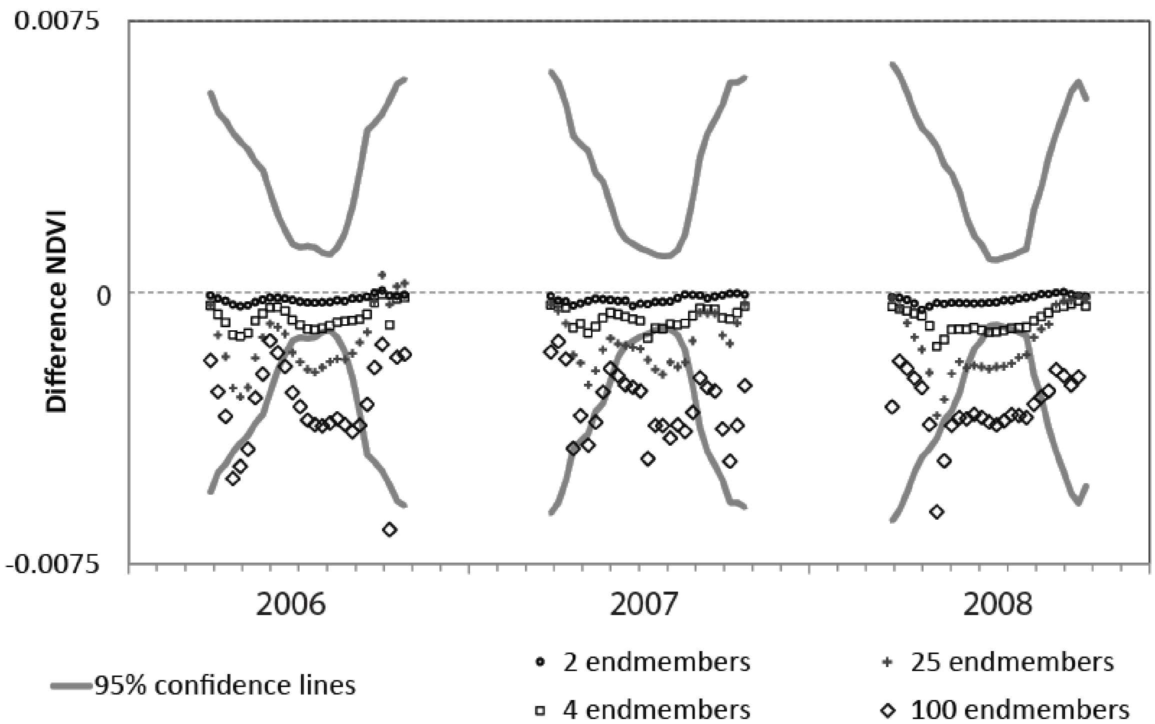


Figure 3. Consequences of linear mixing – an observation based example. Difference NDVI = Mixed NDVI minus "unbiased" NDVI.

It is seen clearly from the figure above that the impacts (differences) from linear mixing increase as the number of endmembers grows. The differences become significant when the number of mixing endmembers reaches 4. In our study we used only 3 endmembers (maize, soybeans and grassland) to be mixed into 0.05 by 0.05 degree pixel representing the trajectory of Tower NDVI values. Considering the magnitude of errors from demonstrated direct linear mixing examples it was safe for us to assume that 3 endmember linear mixing did not make a significant impact on the seasonal trajectories of TNDVI produced by the EDPM. We also have to point out here that the minor impacts from linear mixing appeared to be negligible (20 times smaller) compare to the estimate errors of the EDPM reported in *Kovalskyy and Henebry* [2011a]. In this context, the impacts from direct linear mixing could hardly make a difference for comparisons undertaken in this experiment.

Next, the transformation of mixed TNDVI into phenology driven coefficient K_{cp} had to be generalized (unified). In the prior point based studies we found TNDVI and K_{cp} relationships to be different for crops and grassland [*Kovalskyy and Henebry* 2011b]. For the later land cover type, the linear model carried substantial noise that we tried to compensate with modeling of residuals through their relationship with vapor pressure deficit. We did not find the same relationship in residuals for crops assuming the bias in grassland was due to differences in equipment calibration. Therefore, we used a single linear model with the slope of 1.22 and offset of 0.01 to transform modeled TNDVI into K_{cp} in this spatially explicit experiment. This relationship was derived on observations of TNDVI and K_{cp} on crops and proved its efficacy in *Kovalskyy and Henebry* [2011b].

Finally, the spatially explicit application of the event driven phenology model required some amendments in the functioning of the phenological phase control module described *in Kovalskyy and Henebry* [2011a]. Trained on specific locations and tested on locations with similar climatic conditions, the EDPM required a supplementary mechanism to match the variability of the growing season dates within a much wider range of conditions than in the initial testing studies. Latitudinal gradients for the emergence of vegetated cover which marks the start of the season were applied to the two controlling variables: thermal time and elapsed days since January 1. For the elapsed days we applied a 4 days per degree northward gradient suggested by *Hopkins* [1918]. The thermal time triggers for the onset of greening was also modified with the latitude using slopes and intercepts tuned for three vegetation types following the phenological transfer functions described in *Henebry* [2010]. To deal with the southward increase in the duration of the growing seasons, we adjusted the dynamic triggering for transitions between phenological phases. The adjustment made the transition probability threshold vary inversely with the latitude.

This helped to postpone transitions between phenological phases for locations to the South of training sites, while accelerating the transitions to the North.

2.3 Data sources and preparations for the experiment

The experiments conducted within this investigation had to use various data sources to reach their goals: (1) running the EDPM plus VegET coupling scheme required weather forcing data; (2) percent crop cover data were necessary for the EDPM to produce seasonal canopy trajectories of pixels with mixed vegetation cover; (3) NDVI observations were needed to verify the EDPM's prognosis of seasonal canopy trajectories and later to produce retrospective outcomes; (4) observations of actual ET were needed to evaluate the quality of estimates produced by the EDPM plus VegET coupling scheme; and (5) crop progress reports were crucial for assessment of the EDPM module responsible for estimating dates of phenological transitions.

The meteorological forcings for the EDPM and the VegET were supplied by the North American Land Data Assimilation System (NLDAS) in native GRIB1 format (1 hour temporal and 0.125 degree spatial resolutions). The choice of NLDAS [*Mitchell et al.*, 2004] was based in part on the fact that these forcings were validated on the southern Great Plains adjacent to our study region [*Luo et al.*, 2003]. The original time series of weather data were aggregated into daily image time series and resampled into 0.05 degree grid using nearest neighbor procedure to match the MODIS Climate Modeling Grid (CMG) projection. The last transformation preserved most of the original data and allowed for the fusion of MODIS NDVI data with the EDPM produced seasonal canopy trajectories and later calculation of the ET_a at 0.05 degree resolution. The list of forcing variables included: 2 meter air temperature [K]; 2 meter specific humidity [kg/kg]; surface pressure [Pa]; U wind component [m/s]; V wind component [m/s]; downward shortwave radiation [W/m^2]; downward longwave radiation [W/m^2]; total precipitation [kg/m^2]. The forcing dataset and LSM simulations of NASA's Mosaic model from NLDAS Phase 2 were obtained from the NASA Goddard Earth Sciences Data and Information Services Center (GES DISC) at: *http://disc.sci.gsfc.nasa.gov/hydrology/data-holdings.*

Land cover data came in the form of MODIS based crop maps for 2007 through 2009. The 0.5 km resolution maps were provided directly by members of the product development team [*Chang et al*, 2007]. The procedures for deriving the percent covers of maize and soybeans were based on decision tree techniques applied on level 2 MODIS reflectance in seven bands covering visible and infrared portions of the electromagnetic spectrum. Additional metrics capturing the temporal development of land surface properties relevant to the vegetation were also fused into the procedure. The 2007 paper reported drawbacks in using the universal sampling approach in the development of their decision tree model which were related mostly to the differences in cropping timing and radiometric properties of underlying soils in different areas of CONUS. As an alternative, they proposed a single state based modeling of percent crop type cover which significantly improved the performance of the procedure in independent tests.

Here we were able to use the latest versions of crop cover maps derived from state based decision tree models of Nebraska, Iowa, Minnesota, North and South Dakota, where the most variability was captured in the least amount of training. Some adjacent areas of other states were combined in the final areal compositing procedure. Within delineated study area where each pixel had at least 50% crop cover, the proportion of grassland was assumed to be the remainder of a pixel cover. This assumption was based on the NLDAS land cover scheme that considered grassland as the second most abundant land cover within our region [*Luo et al.,* 2003]. We also masked out all non-grassland or non-cropland land covers from our study area based on the

MODIS land cover product (MCD12C1, IGBP classification type available at ftp://e4ftl01.cr.usgs.gov/MOTA/MCD12C1.005).

Verifying the estimated canopy states and actual evapotranspiration on a spatially explicit basis posed some difficulties only because there are so few observational datasets available for such analyses. One such comparison was the TNDVI time series generated by the EDPM in prognostic mode with the MODIS NDVI time series. The reference NDVI image time series were produced from MODIS NBAR data (MCD43C4 version) available at ftp://e4ftl01.cr.usgs.gov/MOTA/MCD43C4.005. First, bands 1 and 2 of NBAR data in CMG projection were extracted and screened for insufficient quality records using QA bits. Then, the NDVI [1] was computed out of screened red and near infrared reflectance data and organized into image time series.

Potentially, daily ET_a estimates from the EDPM plus VegET scheme had several sources of reference data since at the time of our study two ET monitoring products were on their way to public release [*Mu et al.,* 2007; *Anderson et al.,* 2010]. However, only the MODIS evapotranspiration product (MOD16) data had become publically accessible [*Mu et al.*, 2009]. Therefore the MOD16 product had become our first choice for reference when assessing the quality of results from the coupled EDPM+VegET scheme. This product presents estimates of 8 day sums of actual and reference ET modeled from weather forcings and remotely sensed properties of the land surface [*Mu et al.,* 2007]. Standard HDF files were obtained from *ftp.ntsg.umt.edu/pub/MODIS/Mirror/MOD16/MOD16A2.105_MERRAGMAO/.* The original actual ET layers of MOD16A2 version of the product with 1 km resolution were spatially aggregated and then resampled into 0.05 degree grid again to match the MODIS CMG projection adopted as the basis for this experiment.

The MOD16 product has been closely approaching the *in situ* measured ET_a with each improvement to its procedures [*Mu et al.*, 2009, 2011], yet it is still a product with varying degree of spatial and temporal uncertainty. Therefore, we retained an alternative set of ET_a estimates with which to compare our results. We selected the outcomes of NASA's Mosaic LSM [*Koster and Suarez*, 1994, 1996] from NLDAS as an alternative reference point for comparison based on the validation studies of Mosaic LSM [*Koster and Suarez*, 2003; *Koster et al.*, 2004]. To match the formatting of the first reference product (MOD16), the daily ET_a estimates from the coupled EDPM + VegET scheme and from Mosaic LSM were each temporally aggregated into 8 day ET_a totals.

The accuracy in estimating phenological dates has always been a subject of point based validation studies [*Menzel et al.*, 2006; *Schwartz et al.*, 2006; *Richardson et al.*, 2009; *Zhang et al.*, 2009; *White et al.*, 2009; *Dufour and Morin*, 2010]. In the search of reference data we examined the National Agricultural Statistics Service (NASS) weekly Crop Progress reports on the percentages of crops achieving crop specific phenophases. From this source we could only obtain information about growing season progress on the state level since the county level reports were inconsistent. Therefore, we reorganized the pixel based EDPM reports into the daily state level growing season progression time series to see the parallels between reported and observed dates. These dates were compared with two available years (2008, 2009) of state level crop progress reports obtained for the five states (Nebraska, Iowa, Minnesota, North and South Dakota) from the NASS archives: http://www.nass.usda.gov/Data_and_Statistics/Quick_Stats_1.0/index.asp

Considering the spatial mismatch, temporal precision differences, and the differences in biogeophysical meaning between reported events and dates estimated by the EDPM, we have chosen to rely mostly on the midpoints of distributions in phenological metrics for our comparison. Therefore in the analysis we used midpoint dates (when 50% of crops went through start of season [SoS] or end of season [EoS]) and their inter-quartile range (IRQ) as a measure of data variability. Based on SoS and EoS dates we also calculated the lengths of seasons (LOS) together with their inter-quartile ranges. The LoS values from the NASS reports were calculated by subtracting the 50% EoS date from 50% EoS date and the IQR 75% EoS date minus 25% EoS date and 25% EoS date minus 75% EoS date. The IQR in the LoS data generated during our experiment were collected directly from the EDPM reported pixel phenology dates.

2.4 Road map for analysis.

Resulting test runs of the EDPM and the coupled scheme with VegET produced several sets of results for the evaluation. First, the image time series of TNDVI estimated by the EDPM in prognostic mode were compared with MODIS NBAR NDVI data. Despite the discrepancy in temporal resolution (8 day for MODIS products and daily for our estimates), the comparison could give a good idea of how close our predictions were to the observations. In preparation for such comparison, the EDPM outcomes went through the transformation into MODIS NDVI using the relationship developed in *Kovalskyy and Henebry* [2011a] and confirmed in *Kovalskyy et al.* [2011]. Avoiding the comparison of data beyond the growing season where the EDPM cannot produce TNDVI, we allocated only the results and reference data representing the period from early March (97th day of the year) to late October (305th day of the year). In addition to that, only the dates matching the beginnings of 8 day compositing periods of MODIS products (not the averages over compositing period) were selected for comparison.

In diagnostic mode the EDPM used the former reference--MODIS NBAR NDVI data—to correct its outcomes via the built-in data assimilation scheme [*Kovalskyy and Henebry*, 2011a].

Therefore, to assess the model performance in diagnostic mode, we had to rely on error propagation to infer the accuracy of the assimilation-enhanced EDPM estimates of TNDVI.

Prognostic and diagnostic versions of the EDPM outcomes were used to parameterize VegET to produce corresponding ET_a outcomes. Aggregated into 8 day totals to match the format of first reference data, the ET_a estimates from both prognostic and diagnostic runs of the scheme were compared with the temporally matching image time series of actual evapotranspiration from MOD16 product validated by *Mu et al.* [2009] and Mosaic LSM validated by *Luo et al.* [2003]. Only the time series of ET_a from early March to late October were used in the comparisons.

In our assessment we relied generally on the two most common measures of performance: coefficient of determination (r^2) and root mean square error (RMSE). The first measure showed the ability of produced estimates to follow the observed developments of the modeled variable. RMSE showed the overall level of departure of modeled TNDVI and ET_a from what we assumed to be the reality (reference datasets). Based on the results received in *Kovalskyy and Henebry* [2011a] and *Kovalskyy and Henebry* [2011b], the expected performance levels for the canopy state estimates (viz., NDVI) simulated by the EDPM were r^2=0.8 ±0.1 and RMSE = 0.1 ±0.025. For ET_a, the expected performance levels were r^2=0.7 ±0.15 and RMSE = 1.4 ±0.5 mm per day, but transformed into 8 day values by simple multiplication yields RMSE = 11.2 ±4 mm per 8 days. Additionally, the results were examined for the presence of biases in the residuals. Analyzing differences with reference data, we aimed to assess both temporal and spatial aspects of their distributions to receive clear contrasts between sets of our modeling results and reference data.

In its collection, the NASS archive offered emergence and maturity dates for maize as well as emergence and leaf drop dates for soybeans. We assumed these phenological turn points to be closely related to the SoS and EoS dates produced by the EDPM. Comparing phenological data we plotted our estimates against references expecting to see connections between plant physiological events and their manifestation in the temporal dynamics of optical properties of the vegetated surface.

3. Results.

3.1 Contrasting the EDPM derived NDVI against MODIS product.

The maps representing performance measures for each year were produced to show how the ability of the EDPM to represent the canopy conditions varies in space. We also included the maps of average seasonal propagated errors into Figure 4 from results received after the data assimilation (retrospective mode) to contrast those with RMSE obtained during uncorrected (prognostic) estimation.

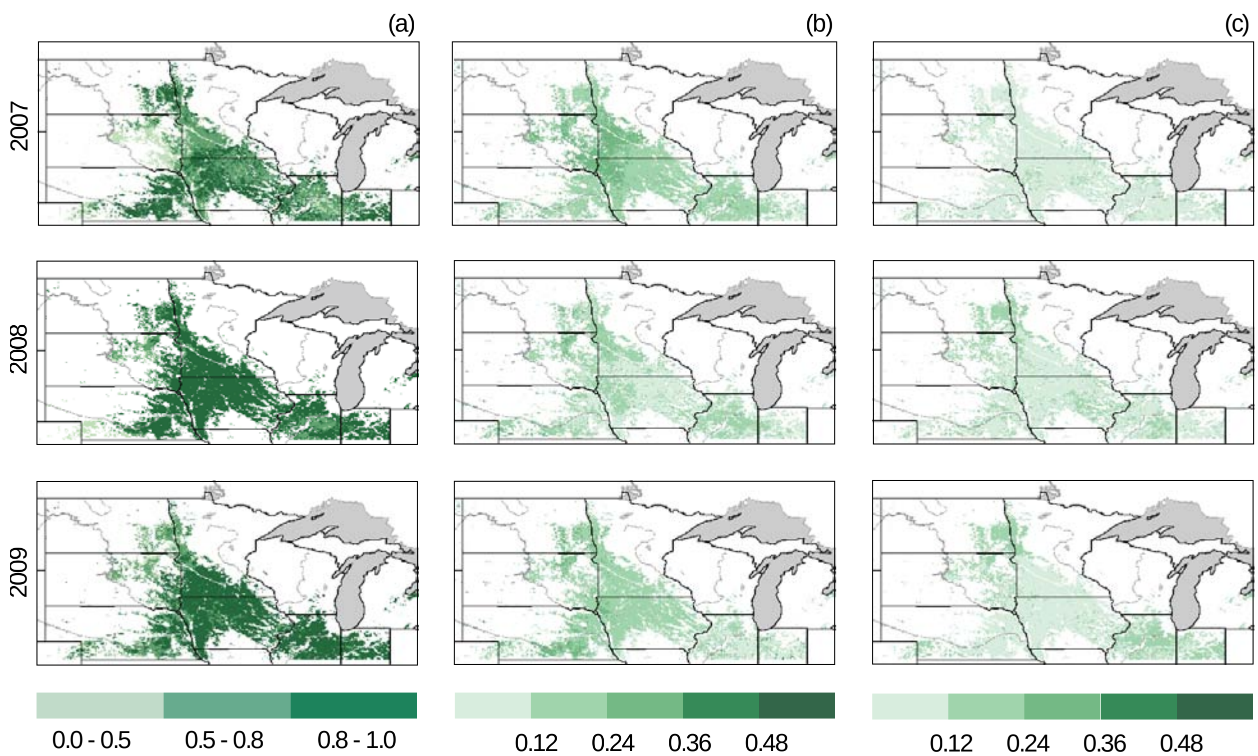


Figure 4. Comparison of the EDPM produced vegetation index against MODIS NDVI within the study area. (*a*) Coefficient of determination (r^2); (*b*) Root mean square error; (*c*) Seasonally averaged propagated daily NDVI error after assimilation of MODIS NDVI observations.

The figure above clearly demonstrates that the EDPM was well fit for to the task of following the dynamics of observed MODIS NDVI. Maps in the left column are dominated by dark color representing r^2 of 0.8 and higher. The r^2 values had a tendency to decrease toward the borders of the study area and whereas 2007 was the year with the worst performance, 2008 the best. The same conclusion was supported by the RMSE maps in Figure 4. The overall level of error reached 0.18 for 2007, but dropped to just above 0.11 for 2008. The right column of Figure 4 shows the uniform distribution of average seasonal propagated errors throughout the study area after EDPM predictions were updated with MODIS NDVI observations. The general level of propagated errors was very close for all three years and constituted slightly less than 0.1.

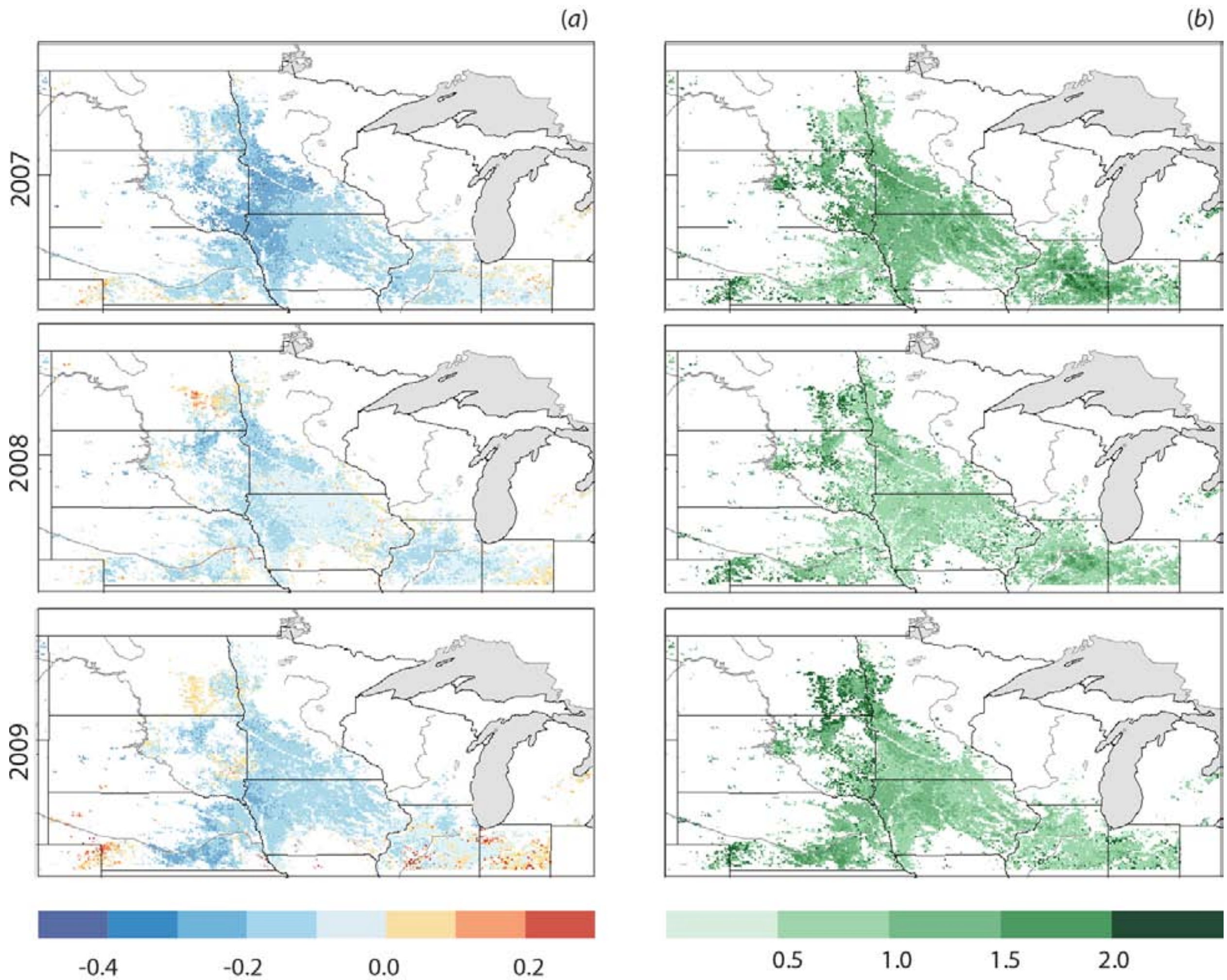


Figure 5. Spatial distributions of residuals ($NDVI_{EDPM} – NDVI_{MODIS}$): (*a*) seasonal means of residuals; (*b*) standard deviations of residuals.

Figure 5 above reveals that the EDPM was mostly underestimating the value of NDVI. Again the picture changed for different years and the character of bias reversed towards the peripheral areas of the study region. The year of 2007 came out as the most biased having the mean of residuals -0.2 to -0.3 spread along the western Iowa and Minnesota borders. For 2008 and 2009, most of the seasonally averaged differences between observed and modeled NDVI varied between -0.2 and 0.1. The variability of the residuals grew from the center towards the borders for each year.

However, similar absolute values of RMSE (Fig. 4b) and mean residuals (Fig.5a) point that the bias was rather uniform in time for most of the study area.

A closer look into intra-annual dynamics of residuals (Fig.6) reveals similarities in developments seen in both the mean difference with observations and the standard deviation of residuals within the three growing seasons.

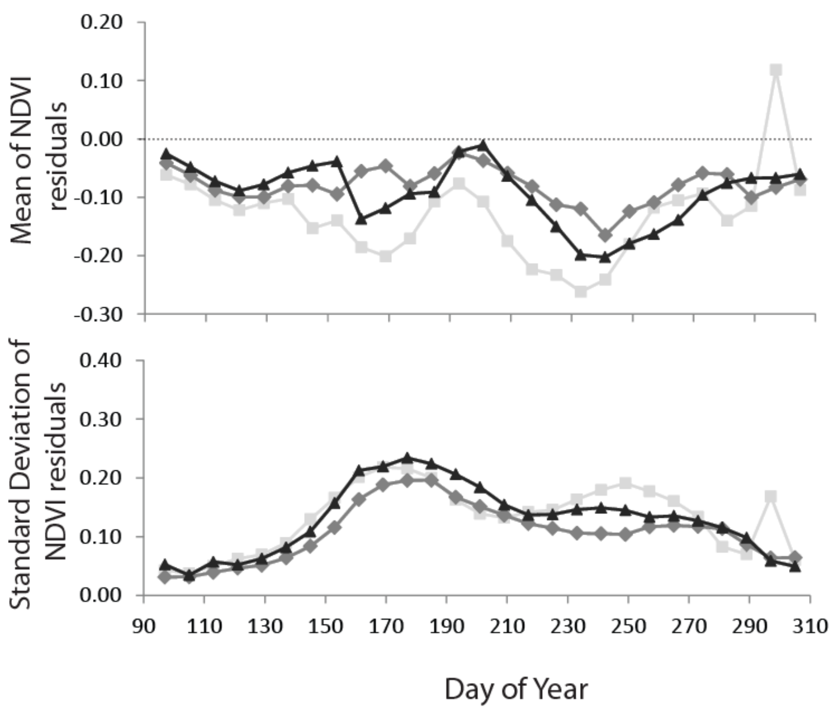


Figure 6. Temporal dynamics of residuals ($NDVI_{EDPM}$ - $NDVI_{MODIS}$) during the 2007-2009 growing seasons. Light grey squares represent season of 2007; darker grey diamonds are 2008; and black triangles are 2009.

The trajectories in Figure 6 represent temporal dynamics of residuals averaged over the entire study region (18.7k pixels). It is seen clearly that the biases from different years went through similar seasonal patterns. The graphs show that the EDPM in prognostic mode was starting up seasons with minor underestimation and kept it at this level till the growing season started for maize and soybeans. The mean of residuals was dropping at every phenological transition point which constitutes a source of performance problems in the EDPM [*Kovalskyy and Henebry*, 2011a, 2011b]. After the change of the phenological phase, the differences with observations came back to the initial level. This pattern means that corrections of the model outcomes during phase change were needed to decrease the bias and make the bias more stable. Overall, the analysis of the EDPM performance suggests that although the errors from EDPM were higher, they were still within the expected range based on prior performance. However, the results from the EDPM can be found reasonably accurate for prognosis or retrospective temporal gap filling in observations, considering the fact that the NDVI from EDPM carries the uncertainty from transformation to MODIS NDVI (standard error of the slope of 0.11[*Kovalskyy and Henebry*, 2011a]), and the uncertainties of the crop maps proliferated through the mixing process.

3.2 Contrasting ET_a estimates from the EDPM+VegET scheme against MOD16 product.

Before comparing the results from the EDPM+VegET with references, it is important to note that the gaps between the two reference datasets were substantial. Plot (*a*) in Figure 7 clearly shows that compared with MOD16 product, Mosaic ET_a first overestimated and then brought bias close to 0 in the middle of the growing season, but later it returned to overestimation. The two versions of the EDPM+VegET estimates representing ET_a derived with and without assimilation via 1DKF scheme also had their differences shown in plot (*b*) of the figure 7. Following the previously noted pattern of underestimation of canopy properties by the EDPM working in prognostic mode, the prognosis of ET_a values was lower than ET_a produced in diagnostic mode (with 1DKF). The variability of residuals in Figure 7b exhibited similar temporal behavior to the one found in the bottom plot of Figure 6.

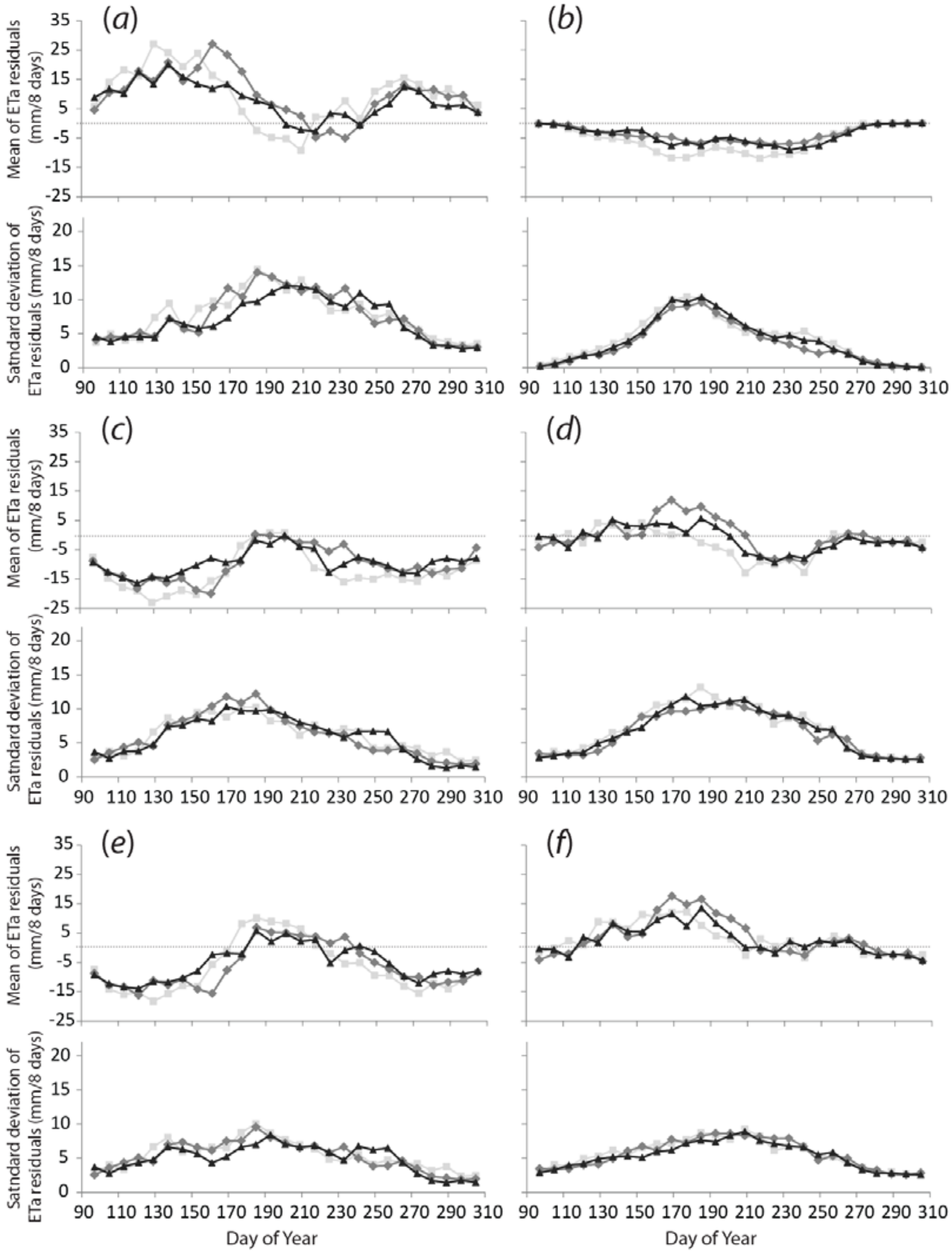


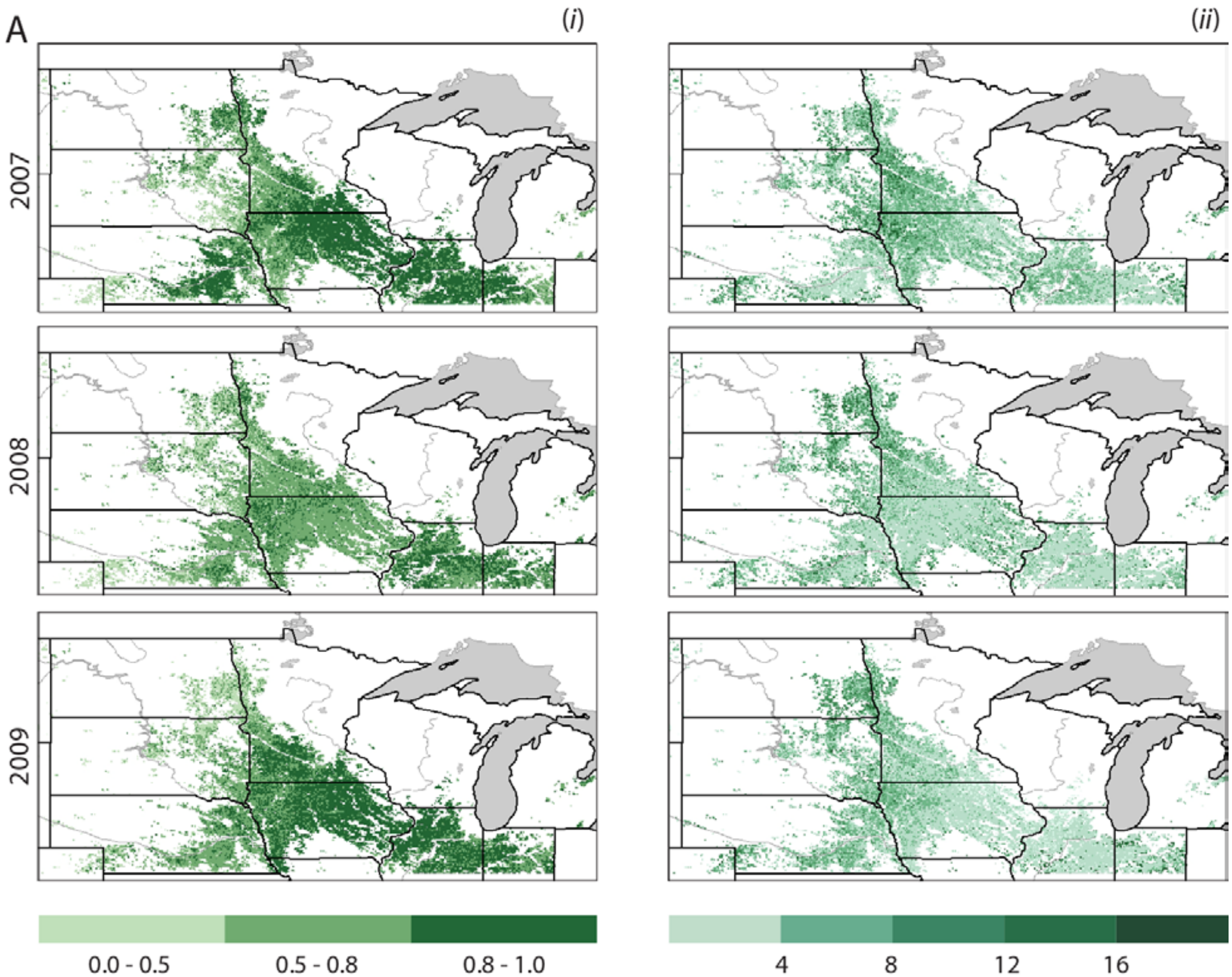
Figure 7. Temporal dynamics of ET_a residuals during the 2007-2009 growing seasons. (*a*) $ET_{a\ Mosaic}$ – $ET_{a\ MOD16}$; (*b*) $ET_{a\ EDPM\ +VegET}$ – $ET_{a\ EDPM\ with\ 1DKF\ +VegET}$; (*c*) $ET_{a\ EDPM\ +VegET}$ - $ET_{a\ Mosaic}$; (*d*) $ET_{a\ EDPM\ +VegET}$ - $ET_{a\ MOD16}$; (*e*) $ET_{a\ EDPM\ with\ 1DKF\ +VegET}$ - $ET_{a\ Mosaic}$; (*f*) ET_a

$_{\text{EDPM with 1DKF +VegET}}$ - $ET_{a\ MOD16}$. Light grey squares represent season of 2007; darker grey diamonds are2008; and black triangles are 2009.

Retaining the main features from plots a and b, the remaining graphics of Figure 7 show the temporal dynamics of differences between two reference datasets and the two sets of 8 day ET_a estimates from the EDPM+VegET scheme. In prognostic mode the EDPM+VegET results were starting growing seasons with underestimation of 15 mm per 8 days compare to ET_a produced by Mosaic. In the midseason the difference came close to zero, but later a smaller (~10 mm per 8 days) underestimation prevailed again (Fig. 7c). Meanwhile compared to the ET_a from MOD16, the prognosis from the EDPM+VegET showed close to 0 difference for most of the season with slight overestimation in early June (up to 7 mm per 8 days) and underestimation of the same magnitude in late August (Fig. 7d). The variability of residuals for prognostic ET_a estimates remained high and had a clear temporal pattern driven by phenology.

The estimates of ET_a obtained with the EDPM+VegET working in diagnostic mode (with 1DKF) exhibited similar behavior of residuals when compared to reference datasets. Differences with Mosaic were negative at the beginnings of growing seasons (Fig. 7e), but in the mid-season the curves drifted toward slight (up 5 mm per 8 days) overestimation which later changed back to the underestimation of 15 mm per 8 days again (Fig. 7e). Compared with MOD16 the EDPM+VegET diagnostic estimates produced residuals that signal slight overestimation early in the growing season. Later, however, the residuals came close to 0 and remained there till the end of growing season indicating good match (Fig. 7f). The variability of residuals for retrospective/ diagnostic ET_a estimates from EDPM+VegET dropped quite dramatically in both comparisons (Fig. 7e,f) showing the relative efficacy of data assimilation for this method of ET_a estimation.

Overall, the EDPM+VegET scheme showed closer temporal resemblance with MOD16 product and therefore further we present figures representing the spatial particularities of the coupled model performance compared to the MODIS product. (Analogous figures showing the comparison with Mosaic can be found in Appendix A.)



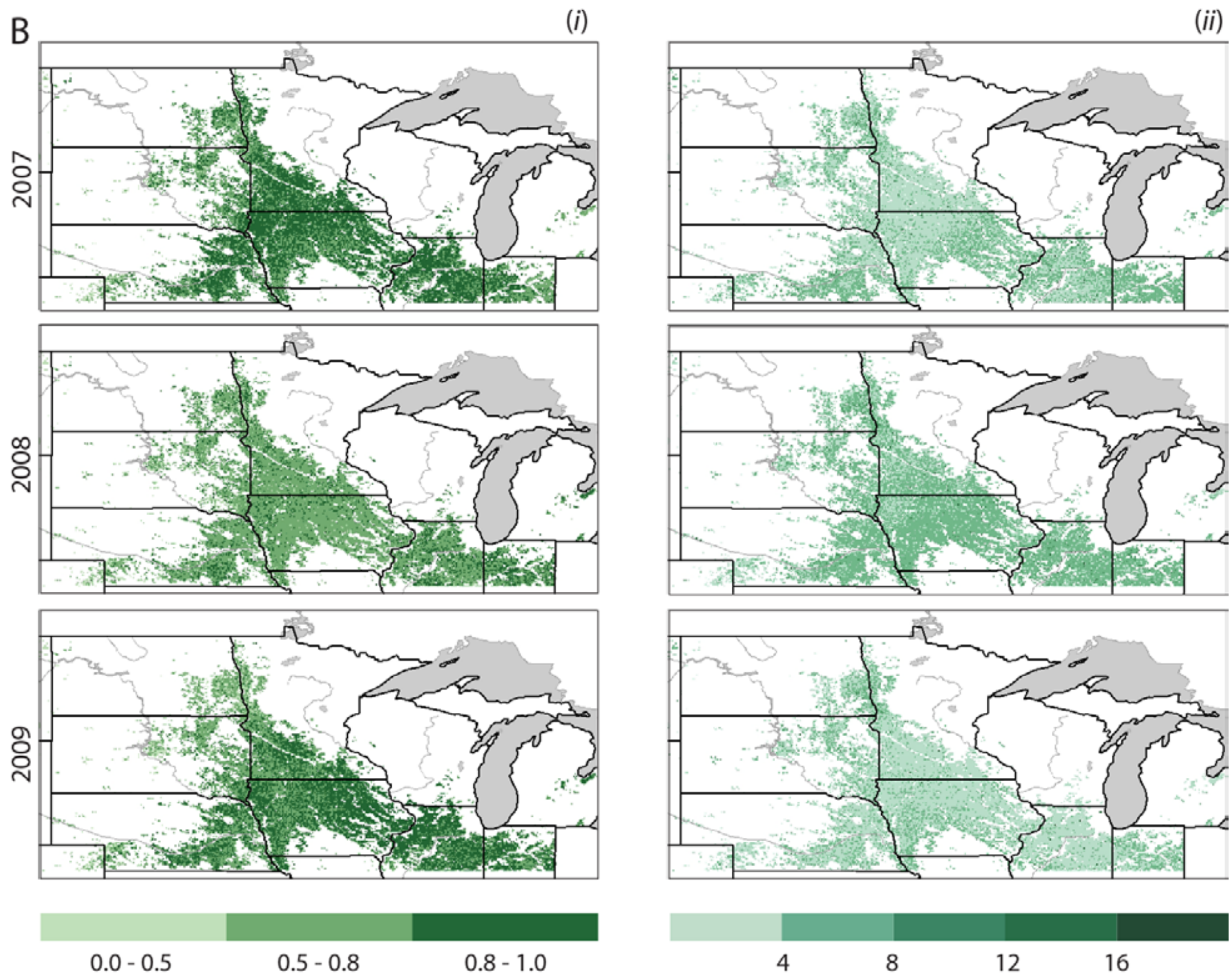
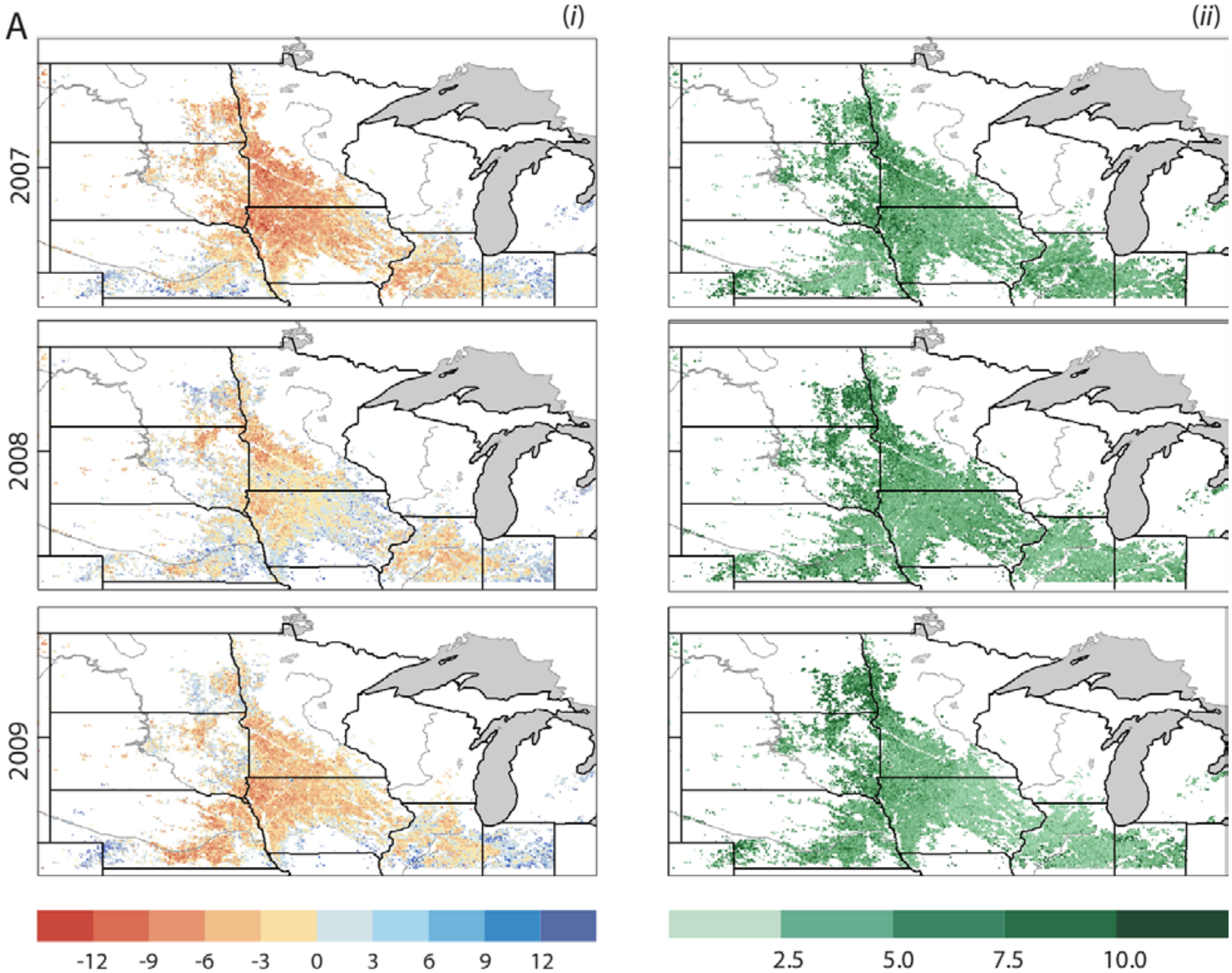


Figure 8. Comparison of MOD16 ET$_a$ with the ET$_a$ produced by EDPM+VegET working in (A) prognostic mode and (B) diagnostic mode involving 1DKF assimilation. (*i*) Coefficient of determination (r^2); (*ii*) Root mean square error (mm per 8 days).

Both parts of figure 8 show that EDPM+VegET scheme was able to follow the dynamics of ET$_a$ in the reference dataset and produced high values of determination coefficient exceeding the expectations set in previous section. Average coefficient of determination was above 0.8 level for the scheme working in both prognostic and diagnostic modes. In 2008, however, the average value of r^2 dropped to the expected 0.7 level for both versions of derived ET$_a$ (Fig. 8A and B). The distribution of r^2 values within the study area was more even in the results from the coupled scheme working in diagnostic mode involving 1DKF assimilation with MODIS NDVI data (Fig. 8A). In both modes the EDPM+VegET scheme showed lower r^2 in the western peripheral regions where the accuracy of crop cover maps was lower. Correspondingly, the RMSE values in those regions were higher especially in the results of the scheme working in prognostic mode. In the results from diagnostic mode RMSE had more uniform distribution and constituted around 6 mm per 8 days on average which is half of what was expected. The average RMSE for EDPM+VegET outcomes derived in prognostic mode was about 8 mm per 8 days. Transformed into corresponding units, this performance would be comparable to *Nagler et al.* [2005] or *Abramowitz et al.* [2008], if the ET_a data from MOD16 product approximated the reality with the accuracy of flux tower instruments [*Mu et al.*, 2009]. A point based flux tower validation study has shown that the scheme can approximate daily ET_a in crops with similar accuracy [*Kovalskyy and Henebry*, 2011b].


A
(i)
(ii)
2007
2008
2009
-12 -9 -6 -3 0 3 6 9 12
2.5 5.0 7.5 10.0

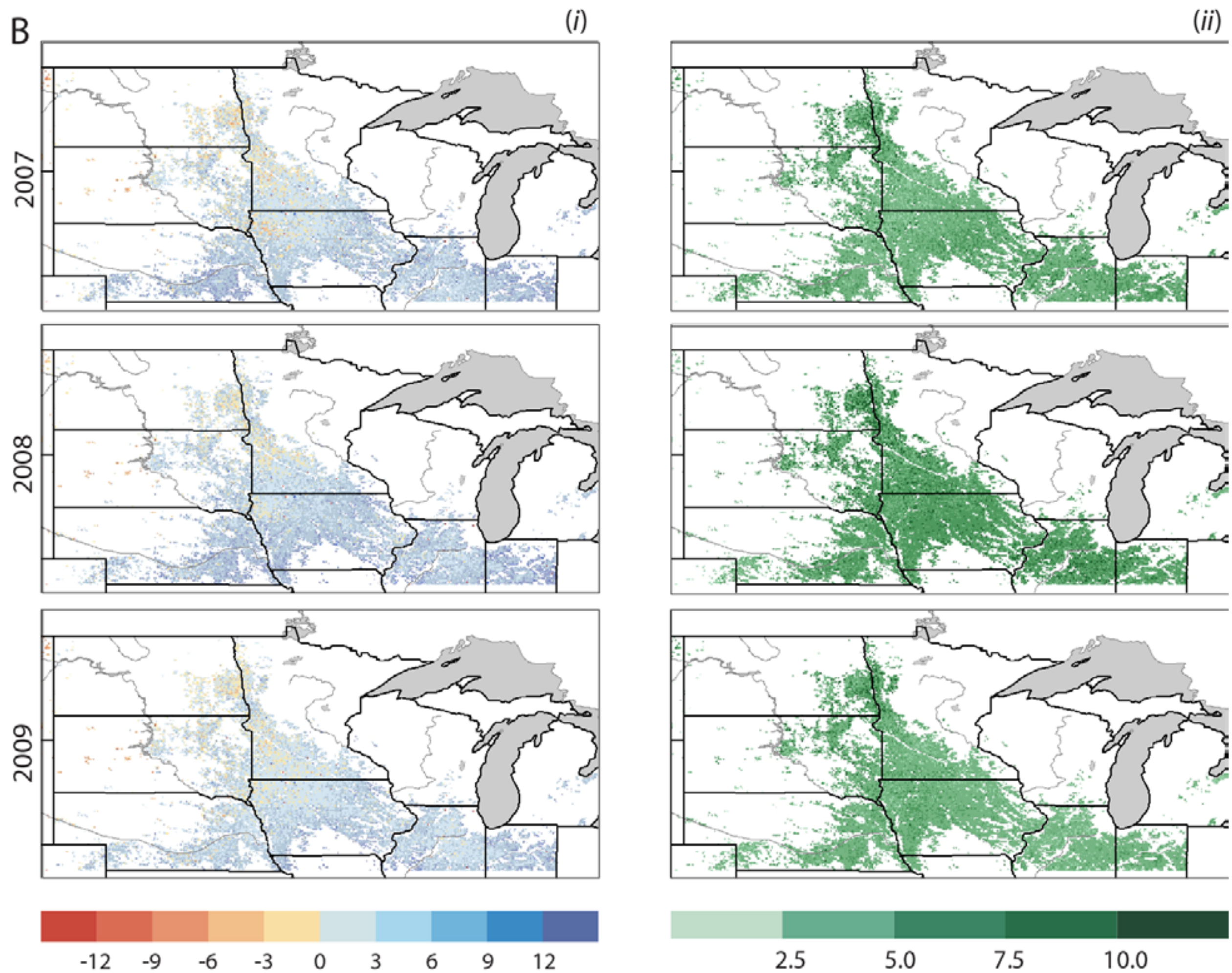


Figure 9. Spatial distributions of residuals (A) $ET_{a\ EDPM+VegET} - ET_{a\ MOD16}$ (B) $ET_{a\ EDPM\ with\ 1DKF\ +VegET} - ET_{a\ MOD16}$. (*i*) annual mean of residuals (mm per 8 days) ; (*ii*) standard deviation of residuals (mm per 8 days).

The contrast between the two sets of ET_a estimates from the EDPM+VegET scheme can be easily depicted from the Figure 9 (A and B). In the left column (*i*) of panel A of Figure 9, the prognoses of ET_a had mostly negative bias changing to overestimation in the peripheral areas of the study region (both east and west). The magnitudes of the mean residuals deviated not too far from 0 giving a peak of up to 12 mm per 8 days in 2007 in the central part of the study region. Left column (*ii*) of Figure 9A shows uneven distribution of variability in residuals revealing clusters of instability in performance coming from EDPM+VegET scheme working in prognostic mode. Panel B of the Figure 9 shows that performance of the EDPM+VegET scheme was more stable during the work in diagnostic mode. The bias in the left column of the Figure 9A was mostly positive fluctuating no more than 9 mm per 8 days. There was less contrast between years and also less difference between various parts of the study region. Smaller and more homogenously distributed standard deviations of residuals (Fig. 9B column *ii*) also indicated a greater stability in performance compare to prognostic mode (Fig. 9A column *ii*).

Contrasted with the ET_a estimates from Mosaic (Appendix A) the results from the EDPM+VegET scheme were less correlated and had greater spatial variability in RMSE and residuals. Figures in Appendix A clearly demonstrate the problem in the central part of the study area (especially during 2007 growing season) that came from numerous differences in approaches to the ET_a modeling and the associated assumptions made about the parameter datasets e.g. land cover types, soil types, LAI, etc [*Koster and Suarez*, 1996; *Mitchell et al.*, 2004]. Nevertheless, the expected performance of r^2=0.7 ±0.15 and RMSE = 11.2 ±4 mm per 8 days were achieved by the coupled models working only in retrospective/diagnostic mode using MODIS observations for correction of simulated TNDVI trajectories.

3.3 Comparison of growing season parameters.

The need to evaluate the performance of the phenological control module in the EDPM was well motivated by the patterns in residuals seen in Figures 6 and 7. Therefore, we highlight the contrasts between the EDPM estimated and in situ start of season [SoS] and end of season [EoS] dates reported to NASS.

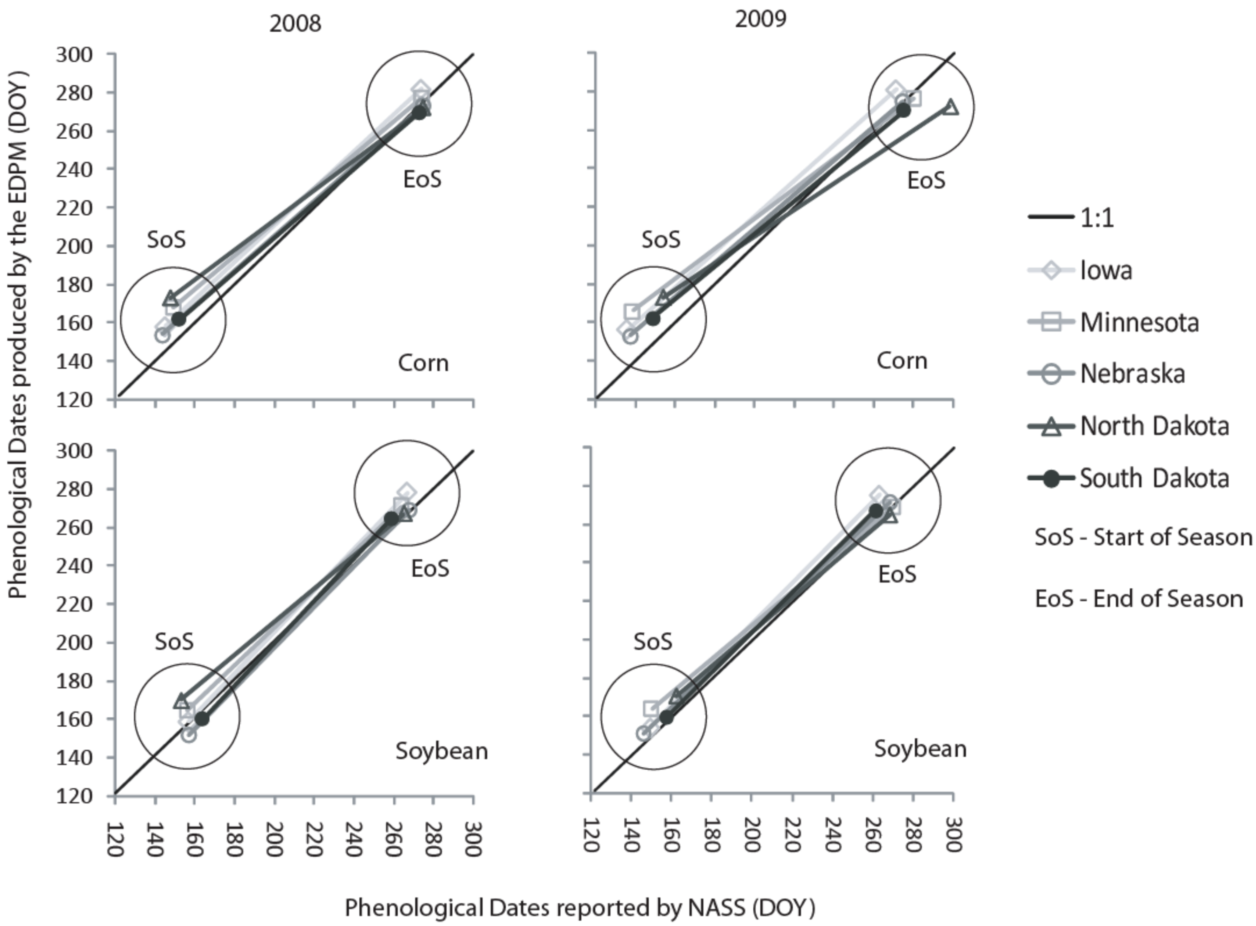


Figure 10. Contrasting start and end dates of the growing season for the two crops and two years.

Figure 10 shows fairly good agreement between observed and estimated parameters of the two growing seasons. It also reveals the persisting delays in SoS for maize crops within all five states. Nevertheless, the 2 weeks delays in SoS prognoses were comparable with errors encountered in retrospective analyses by *Fisher et al.* [2006] *Zhang et al.* [2009] and *Kovalskyy et al.* [2011]. Meanwhile, the estimates of both SoS and EoS for soybeans were even more precise and consistent. Figure 10, however, does not show the variability of the start and end of season dates where dramatic differences arise between NASS reports and the EDPM. To condense the graphical information, we brought the variability measure, the interquartile range

(IQR) into Figure 11, which also shows the scatterplots in length of season (LoS). Similar patterns occur in the variability of SoS and EoS (data not shown).

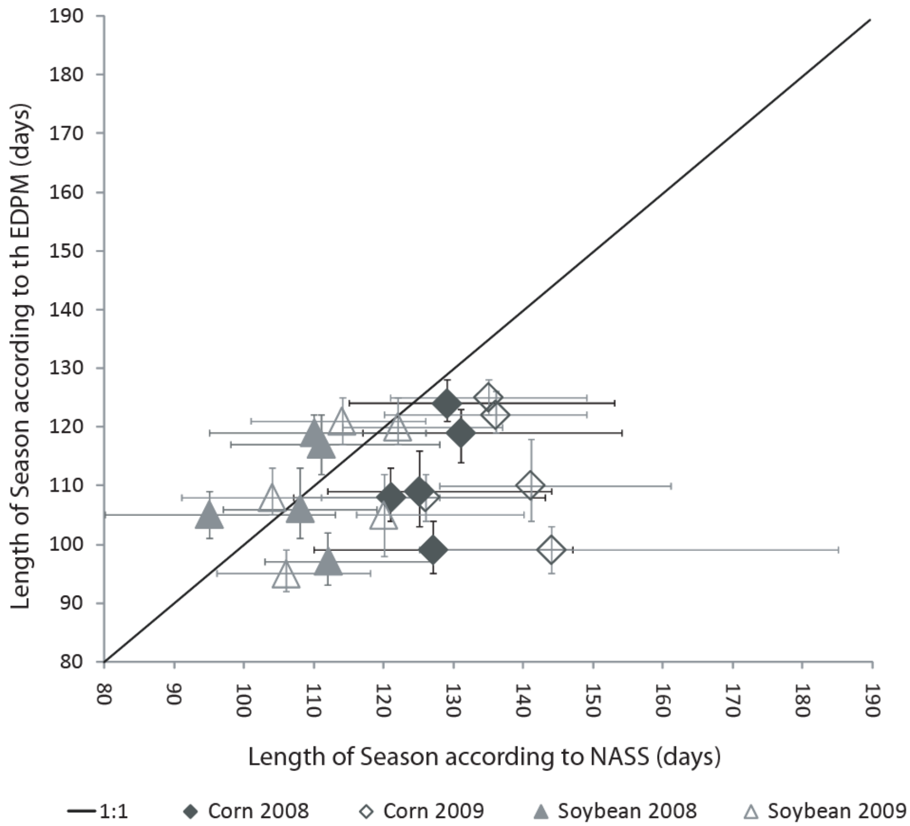


Figure 11. Contrasts between the EDPM estimates and NASS reports in the length of season and its variability for the two crops and two years.

The most apparent feature of Figure 11 is the error bars showing the inter-quartile range of the length of the season. The contrast between NASS and the EDPM dates went to the edge of the anticipated differences due to disparities between compared datasets. We expected the variability in LoS to be driven by gradients in some climatic factors such as rain, temperatures, duration of daylight etc. What we found in NASS reports was that the states with more variability in seasonal precipitation (viz., Nebraska and the Dakotas) had more variability in phenological timing. The EDPM did not have the precipitation in the list of phenological controls [*Kovalskyy and Henebry,* 2011a] and, therefore, the vast difference between observed and estimated IQRs in LoS came as a result of limitations in number of factors considered as drivers of phenological timing. Moreover, the EDPM could not take into account the progress of agricultural work in spring as well as other anthropogenic factors affecting the development of crops. Nevertheless, the central tendencies were captured quite well for soybeans. The SoS delays in maize became the reason for underestimation of LoS for this crop. Yet, with all the shortcomings, the EDPM estimates of phenological dates for all crops and years managed to stay within the range of state reports from NASS.

4. Discussion.

Planned as a validation study this experiment took the form of a comparison between products while still providing insight on the performance of the EDPM +VegET scheme. In this context the discrepancies between estimates found in this study have to be considered just as relative indicators of better or worse performance. Lacking the actual spatially explicit observations, we managed to obtain the reference points for the future application studies where the results will receive interpretation. It is clear now that the outcomes of this experiment helped reaching the goal of this investigation, yet they raised a number of other issues that need to be clarified. In each of the three sets of comparisons we presented spatial and temporal dynamics of error measures but we did not talk in details about the structure of uncertainties or about the reasons behind the observed patterns. Many of these issues are interconnected, and therefore we kept them for this section where the linkages can be explained. Every issue here is discussed in terms of its impact on the abilities of the EDPM alone and the EDPM plus VegET scheme to meet nominal performance expectations. We also present ideas about how these impacts can be mediated at this point and draw perspectives on possible corrections of the problems in future versions of the event driven phenology model.

Comparison between the MODIS NDVI and the vegetation index produced by the EDPM had both temporal and spatial issues in performance. High r^2 was definitely a plus to the EDPM, but the RMSE and bias of 2007 in prognostic mode pushed the performance to the edge of what was expected of the model. Introduction of noise from the NDVI-TNDVI relationship could not be the reason for this error jump since such noise should have been present constantly and not just during late season drought on just about one-fifth of the study area. Apparently, the reaction of the EDPM to this development was too strong (residuals dropped to -0.25), most likely due to inability to account for irrigation. An appropriate solution for the 2007 error spike problem would be extra training of the EDPM on irrigated flux tower sites during the drought years. During other years, the bias appeared to be quite consistent throughout the area and could be arithmetically removed from the results. Possibly, the bias can be corrected by obtaining better estimates of background vegetation-free TNDVI values for growing season initiation as suggested by *Zhang at al.*, [2003]. This correction would, most probably, draw the overall RMSE close to 0.1 level. This performance mark was also achieved through the data assimilation.

Patterns in temporal dynamics of residuals constitute a problem that cannot be corrected with a simple transformation. It requires collecting new data for parameterization of phenophase control module in the EDPM. Inclusion of precipitation as a control variable for phase transitions should help to address the issue of temporal variability in PTPs within states in addition to increasing the overall accuracy of the phenophase control procedures. With the current level of accuracy, we should refrain from interpreting the results based on uncorrected (prognostic) daily NDVI data in places where the variability of residuals goes beyond the level of two seasonal standard deviations. This warning, however, would be less applicable for time averaged (weekly or monthly) or composited prognoses. Meanwhile, the NDVI outcomes received from the EDPM's data assimilation scheme carried significantly smaller traces of phase control errors. Therefore further analysis can be conducted on the retrospective 1DKF corrected daily VI records and interpretations would be valid throughout the study region. The issues with temporal stability in performance also came out in the ET_a estimates produced by the EDPM+VegET scheme. Exceeding the expectation in r^2 and RMSE in comparison with MOD16 product, the results from prognostic mode exhibited a small bump and a dip of similar magnitude in temporal dynamics of the residuals. These fluctuations appeared exactly in the times of phenological transitions from green-up to reproductive phase and then from reproductive phase to senescence respectively. Present in results from all three years, the features indicated a systematic problem in phenological control module of the EDPM that, if removed, could further increase the performance of the coupling scheme. In retrospective mode the results still had the issue of the early season overestimation. This indicates that while decreasing the level of variability in residuals the assimilation of MODIS NDVI could not completely suppress all the setbacks for the EDPM+VegET scheme. Improvements in the functioning and parameterization of phenological phase control module requires further training on long term flux tower records that will be undertaken in the future. However, all observed magnitudes of the deviations in temporal pattern would not pose a significant obstacle for the use of these results in further analyses.

From the comparison of the EDPM+VegET scheme outcomes with ET_a estimates from Mosaic LSM, we received the diverse spatial dynamics in r^2 and RMSE complemented with clear seasonal patterns in temporal dynamics of residuals. These discrepancies persisted even after the assimilation of MODIS data into the EDPM and VegET results. In fact, the pattern became even more pronounced since the variability in residuals dropped. It is most likely that a better sensitivity of the EDPM to ongoing weather conditions contributed to the temporal dynamics of differences between two ET_a estimates as the energy balance scheme in NASA's Mosaic LSM [*Koster and Suarez,* 1996] uses static climatological trajectories of leaf area index as a phenology driven factor of canopy resistance. However, other patterns could not be explained entirely by the lack of sensitivity to contemporaneous vegetation development in the Mosaic model. It is also possible that numerous discrepancies came out as consequences of different assumptions about land cover on the 0.125-degree NLDAS grid and/or the ET flux partitioning between canopy and underlying soil.

A unique feature of this study was the comparison of growing season metrics estimated by the EDPM with ones reported to NASS. In our analysis, we were missing proper geographic and temporal precision in the NASS reports for each of the five states. Nevertheless, we tried to preserve the temporal and spatial variability of growing season dates by organizing our SoS and EoS estimates to match the structure of reference data. We also kept in mind the fact that the transition points in NDVI dynamics and the actual phenological event for crops had different physical meanings. A good matching was achieved between reported and estimated state averaged SoS and EoS. Their variability, however, became problematic for the EDPM giving the ground to include more controlling variables into the automatic estimation of phenophase transition dates.

Despite all the issues listed in this section the overall impression from the comparisons is quite positive for the VegET+EDPM coupling scheme. The scheme managed to keep the departures from references within nominal boundaries. The results matched and even exceeded most of the expected measures of model performance obtained on point based validations [*Kovalskyy and Henebry*, 2011a, 2011b]. The biggest problem for the TNDVI trajectories estimated by the EDPM was the model's overreaction to late season drought in 2007 that accentuated the usually small underestimation. Meanwhile, the ET_a estimates followed closely the reference records from MOD16 products. Even in the worst cases, the error measures in ET_a were also comparable with those of *Senay* [2008], *Mu et al.* [2007], and *Abramowitz et al.* [2008]. Remarkably, this level of performance was achieved during the spatially explicit deployment of the coupled models. Plus, the results from the scheme were complemented with estimates of phenological metrics for grassland and crops that matched well the central tendencies of NASS reports. Combined with the ability of the scheme to produce daily estimates of vegetation index and actual evapotranspiration the performance characteristics of the VegET+EDPM coupling scheme justified its use in a real life application study.

The lessons learned from this experiment will help to analyze and interpret the results of the greater investigation of recent shifts in the phenology and ET regime in the Northern Great Plains. After the undertaken comparisons we can confidently say that consistency of received errors still allows for the trend analysis especially after correcting with MODIS observations. The delays of season starts in maize will be accounted for in the assessment of inter-annual variability of growing season parameters. Also, we intend to scale the variability in phenological dates from the EDPM to match the variability in NASS reports through inclusion of precipitation in the phenophase control mechanism. Special attention will be paid to the peripherals of the study region as those are most likely to carry land cover mapping errors. Finally, we will use appropriate testing methods and critical values when relating the shifts in ET_a regime to crop cover change insuring a more conservative interpretation of their correlation.

5. Conclusion

The purpose of the experiment described in this paper was to provide the rationale for the use of the EDPM+VegET coupling scheme in a spatially explicit application. Such rationale was attained via assessing the performance of the scheme through comparison of modeled variables with reference data. First, we compared the image time series of vegetation index produced by the phenology model with MODIS NDVI derived from MCD43C4 product. The expectations of model performance in producing seasonal NDVI trajectories were met yielding r^2 of 0.8 ±0.15 and RMSE of 0.1 ±0.035 for the entire study area. Retrospective correction of canopy dynamics with MODIS NDVI brought the variability in errors closer to the 0.1 level. Estimation of growing season metrics by the EDPM matched the NASS reports with reasonable accuracy – up to 2 weeks of difference in key dates. The estimates of actual evapotranspiration produced by the coupled scheme were compared with ET_a from NASA's Mosaic model from NLDAS and with MOD16 data from MODIS land product suite. In both comparisons, the expected r^2=0.7 ±0.15 and RMSE = 1.4 ±0.5 mm per day were met by the coupling scheme working in retrospective mode using MODIS observations for correcting seasonal trajectories of canopy development.

Minor issues of model performance were encountered during this experiment as well. The EDPM produced trajectories of vegetation index biased towards underestimation but the bias was relatively uniform in space and time and therefore removable. Actual ET estimates from the VegET+EDPM were closer to MOD16 product while producing greater differences with Mosaic LSM that also had persisting spatial and temporal patterns in them. While spatial patterns in differences could be attributed to distinct assumptions about land cover in Mosaic LSM [*Mitchel et al.*, 2004], the seasonal profiles of differences between our estimates and reference data exhibited clear patterns driven by phenology. The impacts of these issues on performance of the

EDPM and the VegET models, however, were relatively small and therefore they could not pose an obstacle for the analysis and interpretation of the outcomes. In general, this study provided sufficient assurance that the interpretations of future results derived the the planned spatially explicit application study will be valid and sound, provided that the detected issues are properly addressed in the analysis.

Acknowledgements.[gmh1]

Research was supported in part by NASA grant NNX07AT61A to GMH.
Funding from the National Aeronautics and Space Administration supported this research under grant NNH07ZDA001N.
The NLDAS data used in this effort were acquired as part of the activities of NASA's Science Mission Directorate, and are archived and distributed by the Goddard Earth Sciences (GES) Data and Information Services Center (DISC).

References.

Abramowitz, G., R. Leuning, M. Clark, and A. Pitman (2008), Evaluating the Performance of Land Surface Models, Journal of Climate, 21(21), 5468-5481.

Allen R. G., Pereira L, Raes D., and S. M. (1998), Crop Evapotranspiration - Guidelines for Computing Crop Water Requirement - FAO Irrigation and Drainages Paper 56. , UNFAO, Rome, Italy.

Anderson, M. C., et al. (2010), Mapping daily evapotranspiration at field to global scales using geostationary and polar orbiting satellite imagery, Hydrol. Earth Syst. Sci. Discuss., 7(4), 5957-5990.

Bondeau, A., et al. (2007), Modelling the role of agriculture for the 20th century global terrestrial carbon balance, Global Change Biology, 13(3), 679-706.

Busetto, L., M. Meroni, and R. Colombo (2008), Combining medium and coarse spatial resolution satellite data to improve the estimation of sub-pixel NDVI time series, Remote Sensing of Environment, 112(1), 118-131.

Campo, L., F. Castelli, D. Entekhabi, and F. Caparrini (2009), Land-atmosphere interactions in an high resolution atmospheric simulation coupled with a surface data assimilation scheme, Nat. Hazards Earth Syst. Sci., 9(5), 1613-1624.

Chang, J., M. C. Hansen, K. Pittman, M. Carroll, and C. DiMiceli (2007), Corn and Soybean Mapping in the United States Using MODIS Time-Series Data Sets, Agron. J., 99(6), 1654-1664.

Dickinson, R. E., M. Shaikh, R. Bryant, and L. Graumlich (1998), Interactive Canopies for a Climate Model, Journal of Climate, 11(11), 2823-2836.

Dufour, B., and H. Morin (2010), Tracheid production phenology of Picea mariana and its relationship with climatic fluctuations and bud development using multivariate analysis, Tree Physiology, 30(7), 853-865.

Fisher, J. I., J. F. Mustard, and M. A. Vadeboncoeur (2006), Green leaf phenology at Landsat resolution: Scaling from the field to the satellite, Remote Sensing of Environment, 100(2), 265-279.

Foley, J. A., S. Levis, M. H. Costa, W. Cramer, and D. Pollard (2000), Incorporating Dynamic Vegetation Cover Within Global Climate Models, Ecological Applications, 10(6), 1620-1632.

Gao, Z. Q., C. S. Liu, W. Gao, and N. B. Chang (2010), A coupled remote sensing and the Surface Energy Balance with Topography Algorithm (SEBTA) to estimate actual evapotranspiration under complex terrain, Hydrol. Earth Syst. Sci. Discuss., 7(4), 4875-4924.

Godfrey, C. M., and D. J. Stensrud (2010), An Empirical Latent Heat Flux Parameterization for the Noah Land Surface Model, Journal of Applied Meteorology and Climatology, 49(8), 1696-1713.

Godfrey, C., D. Stensrud, and L. Leslie (2007), A new latent heat flux parameterization for land surface models, Preprints of 21st Conf. on Hydrology, San Antonio, TX, Amer. Meteor. Soc.,, 6A.3.

Henebry, G. M. (2010), Land surface phenology as an integrative diagnostic for landscape modelling. Proceedings of LANDMOD2010, Montpellier, France, 3-5 February. Available at: http://www.symposcience.org/exl-doc/colloque/ART-00002394.pdf

Hopkins, A. D. (1918), Periodical events and natural law as guides to agricultural research and practice, *Monthly Weather Review, Suplement 9*, 42.

Huemmrich, K. F., T. A. Black, P. G. Jarvis, J. H. McCaughey, and F. G. Hall (1999), High temporal resolution NDVI phenology from micrometeorological radiation sensors, Journal of Geophysical Research, 104(D22), 27935-27944.

Ibanez, I., R. B. Primack, A. J. Miller-Rushing, E. Ellwood, H. Higuchi, S. D. Lee, H. Kobori, and J. A. Silander (2010), Forecasting phenology under global warming, Philosophical Transactions of the Royal Society B: Biological Sciences, 365(1555), 3247-3260.

Jang, K., S. Kang, J. Kim, C. B. Lee, T. Kim, J. Kim, R. Hirata, and N. Saigusa (2009), Mapping evapotranspiration using MODIS and MM5 Four-Dimensional Data Assimilation, Remote Sensing of Environment, 114(3), 657-673.

Kang, S., W. A. Payne, S. R. Evett, C. A. Robinson, and B. A. Stewart (2009), Simulation of winter wheat evapotranspiration in Texas and Henan using three models of differing complexity, Agricultural Water Management, 96(1), 167-178.

Kiniry, J., M. Schmer, K. Vogel, and R. Mitchell (2008), Switchgrass Biomass Simulation at Diverse Sites in the Northern Great Plains of the U.S, BioEnergy Research, 1(3), 259-264.

Koster, R. D., and M. J. Suarez (1994), The components of a 'SVAT' scheme and their effects on a GCM's hydrological cycle, Advances in Water Resources, 17(1-2), 61-78.

Koster, R. D., and M. J. Suarez (1996), Energy and water balance calculations in the Mosaic LSM, edited, p. 59, NASA.

Koster, R. D., and M. J. Suarez (2003), Impact of Land Surface Initialization on Seasonal Precipitation and Temperature Prediction, Journal of Hydrometeorology, 4(2), 408-423.

Koster, R. D., M. J. Suarez, P. Liu, U. Jambor, A. Berg, M. Kistler, R. Reichle, M. Rodell, and J. Famiglietti (2004), Realistic Initialization of Land Surface States: Impacts on Subseasonal Forecast Skill, Journal of Hydrometeorology, 5(6), 1049-1063.

Kovalskyy, V., and G. M. Henebry (2011a), A new concept for simulation of vegetated land surface dynamics – Part 1: The event driven phenology model, Biogeosciences Discuss., 8(3), 5281-5333.

Kovalskyy, V., and G. M. Henebry (2011b), Alternative methods to predict actual evapotranspiration illustrate the importance of accounting for phenology – Part 2: The event driven phenology model, Biogeosciences Discuss., 8(3), 5335-5378.

Kovalskyy, V., D. P. Roy, X. Y. Zhang, and J. Ju (2011. In Press), The suitability of multi-temporal Web-Enabled Landsat Data (WELD) NDVI for phenological monitoring – a comparison with flux tower and MODIS NDVI, Remote Sensing Letters.

Lawrence, P. J., and T. N. Chase (2007), Representing a new MODIS consistent land surface in the Community Land Model (CLM 3.0), J. Geophys. Res., 112(G1), G01023.

Li, Z.-L., R. Tang, Z. Wan, Y. Bi, C. Zhou, B. Tang, G. Yan, and X. Zhang (2009), A Review of Current Methodologies for Regional Evapotranspiration Estimation from Remotely Sensed Data, Sensors, 9(5), 3801-3853.

Luo, L., et al. (2003), Validation of the North American Land Data Assimilation System (NLDAS) retrospective forcing over the southern Great Plains, J. Geophys. Res., 108(D22), 8843.

Manabe, S. (1969), Climate and the Ocean Circulation, Monthly Weather Review, 97(11), 739-774.

Maruyama, A., and T. Kuwagata (2010), Coupling land surface and crop growth models to estimate the effects of changes in the growing season on energy balance and water use of rice paddies, Agricultural and Forest Meteorology, 150(7-8), 919-930.

Mearns, L. O., T. Mavromatis, E. Tsvetsinskaya, C. Hays, and W. Easterling (1999), Comparative responses of EPIC and CERES crop models to high and low spatial resolution climate change scenarios, J. Geophys. Res., 104(D6), 6623-6646.

Meng, C. L., Z. L. Li, X. Zhan, J. C. Shi, and C. Y. Liu (2009), Land surface temperature data assimilation and its impact on evapotranspiration estimates from the Common Land Model, Water Resour. Res., 45(2), W02421.

Menzel, A., T. H. Sparks, N. Estrella, and D. B. Roy (2006), Altered geographic and temporal variability in phenology in response to climate change, Global Ecology and Biogeography, 15(5), 498-504.

Miralles, D. G., T. R. H. Holmes, R. A. M. De Jeu, J. H. Gash, A. G. C. A. Meesters, and A. J. Dolman (2010), Global land-surface evaporation estimated from satellite-based observations, Hydrol. Earth Syst. Sci. Discuss., 7(5), 8479-8519.

Mitchell, K. E., et al. (2004), The multi-institution North American Land Data Assimilation System (NLDAS): Utilizing multiple GCIP products and partners in a continental distributed hydrological modeling system, Journal of Geophysical Research, 109(D7), D07S90.

Mu, Q., M. Zhao, and S. W. Running (2011), Improvements to a MODIS global terrestrial evapotranspiration algorithm, Remote Sensing of Environment, In Press, Corrected Proof.

Mu, Q., F. A. Heinsch, M. Zhao, and S. W. Running (2007), Development of a global evapotranspiration algorithm based on MODIS and global meteorology data, Remote Sensing of Environment, 111(4), 519-536.

Mu, Q., L. A. Jones, J. S. Kimball, K. C. McDonald, and S. W. Running (2009), Satellite assessment of land surface evapotranspiration for the pan-Arctic domain, Water Resources Research, 45(9), W09420.

Nagler, P. L., J. Cleverly, E. Glenn, D. Lampkin, A. Huete, and Z. Wan (2005), Predicting riparian evapotranspiration from MODIS vegetation indices and meteorological data, Remote Sensing of Environment, 94(1), 17-30.

Pitman, A. J. (2003), The evolution of, and revolution in, land surface schemes designed for climate models, International Journal of Climatology, 23(5), 479-510.

Prihodko, L., A. S. Denning, N. P. Hanan, I. Baker, and K. Davis (2008), Sensitivity, uncertainty and time dependence of parameters in a complex land surface model, Agricultural and Forest Meteorology, 148(2), 268-287.

Ransom, J., D. Franzen, P. Glogoza, K. Hellevang, V. Hofman, M. McMullen, and R. Zollinger (2004), Basics of corn production in North Dakota, North Dakota Extension Service, 20, accsessed 05/02/2011, at http://www.ag.ndsu.edu/pubs/plantsci/rowcrops/a834.pdf.

Richardson, A. D., D. Y. Hollinger, D. B. Dail, J. T. Lee, J. W. Munger, and J. Oâ€™keefe (2009), Influence of spring phenology on seasonal and annual carbon balance in two contrasting New England forests, Tree Physiology, 29(3), 321-331.

Rosero, E., Z.-L. Yang, L. E. Gulden, G.-Y. Niu, and D. J. Gochis (2009), Evaluating Enhanced Hydrological Representations in Noah LSM over Transition Zones: Implications for Model Development, Journal of Hydrometeorology, 10(3), 600-622.

Rötzer, T., M. Leuchner, and A. Nunn (2010), Simulating stand climate, phenology, and photosynthesis of a forest stand with a process-based growth model, International Journal of Biometeorology, 54(4), 449-464.

Roy, D. P. (2000), The impact of misregistration upon composited wide field of view satellite data and implications for change detection, Geoscience and Remote Sensing, IEEE Transactions on, 38(4), 2017-2032.

Roy, D. P., Y. Jin, P. E. Lewis, and C. O. Justice (2005), Prototyping a global algorithm for systematic fire-affected area mapping using MODIS time series data, Remote Sensing of Environment, 97(2), 137-162.

Sabater, J. M., C. Rüdiger, J.-C. Calvet, N. Fritz, L. Jarlan, and Y. Kerr (2008), Joint assimilation of surface soil moisture and LAI observations into a land surface model, Agricultural and Forest Meteorology, 148(8-9), 1362-1373.

Schaaf, C. B., et al. (2002), First operational BRDF, albedo nadir reflectance products from MODIS, Remote Sensing of Environment, 83(1-2), 135-148.

Schwartz, M. D., R. Ahas, and A. Aasa (2006), Onset of spring starting earlier across the Northern Hemisphere, Global Change Biology, 12(2), 343-351.

Senay, G. (2008), Modeling Landscape Evapotranspiration by Integrating Land Surface Phenology and a Water Balance Algorithm, Algorithms, 1(2), 52-68.

Senay, G., M. Budde, J. Verdin, and A. Melesse (2007), A Coupled Remote Sensing and Simplified Surface Energy Balance Approach to Estimate Actual Evapotranspiration from Irrigated Fields, Sensors, 7(6), 979-1000.

Settle, J., and N. Campbell (1998), On the errors of two estimators of sub-pixel fractional cover when mixing is linear, Geoscience and Remote Sensing, IEEE Transactions on, 36(1), 163-170.

Stancalie, G., A. Marica, and L. Toulios (2010), Using earth observation data and CROPWAT model to estimate the actual crop evapotranspiration, Physics and Chemistry of the Earth, Parts A/B/C, 35(1-2), 25-30.

Stensrud, D. J. (2007), *Parameterization schemes: keys to understanding numerical weather prediction models*, 459 pp., Cambridge University Press.

Wegehenkel, M. (2009), Modeling of vegetation dynamics in hydrological models for the assessment of the effects of climate change on evapotranspiration and groundwater recharge, Adv. Geosci., 21, 109-115.

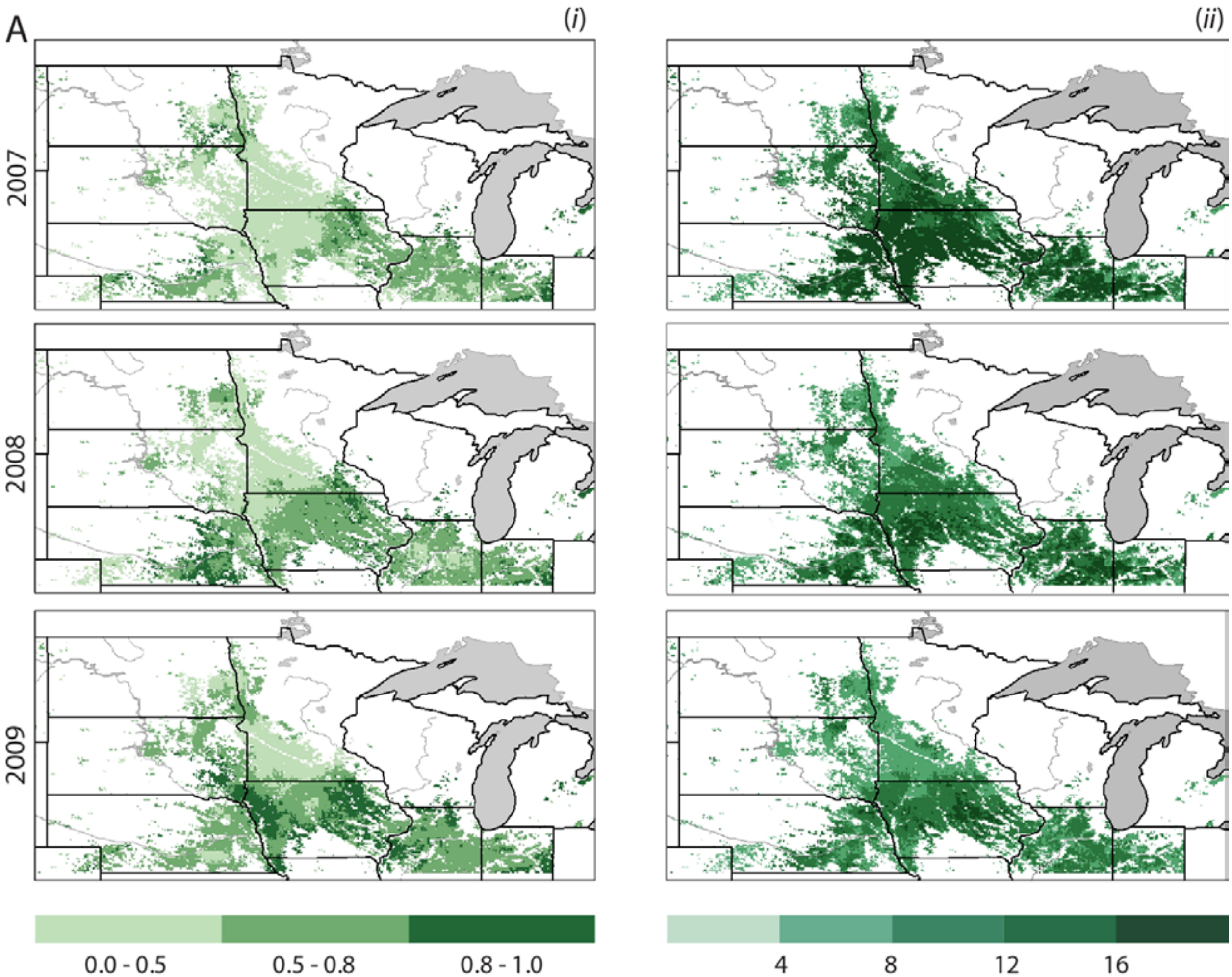
Yuan, W., et al. (2010), Global estimates of evapotranspiration and gross primary production based on MODIS and global meteorology data, Remote Sensing of Environment, 114(7), 1416-1431.

Zha, T., A. G. Barr, G. van der Kamp, T. A. Black, J. H. McCaughey, and L. B. Flanagan (2010), Interannual variation of evapotranspiration from forest and grassland ecosystems in western canada in relation to drought, Agricultural and Forest Meteorology, 150(11), 1476-1484.

Zhang, X., M. A. Friedl, and C. B. Schaaf (2009), Sensitivity of vegetation phenology detection to the temporal resolution of satellite data, International Journal of Remote Sensing, 30(8), 2061 - 2074.

Zhang, X., M. A. Friedl, C. B. Schaaf, A. H. Strahler, J. C. F. Hodges, F. Gao, B. C. Reed, and A. Huete (2003), Monitoring vegetation phenology using MODIS, Remote Sensing of Environment, 84(3), 471-475.

Appendix A.



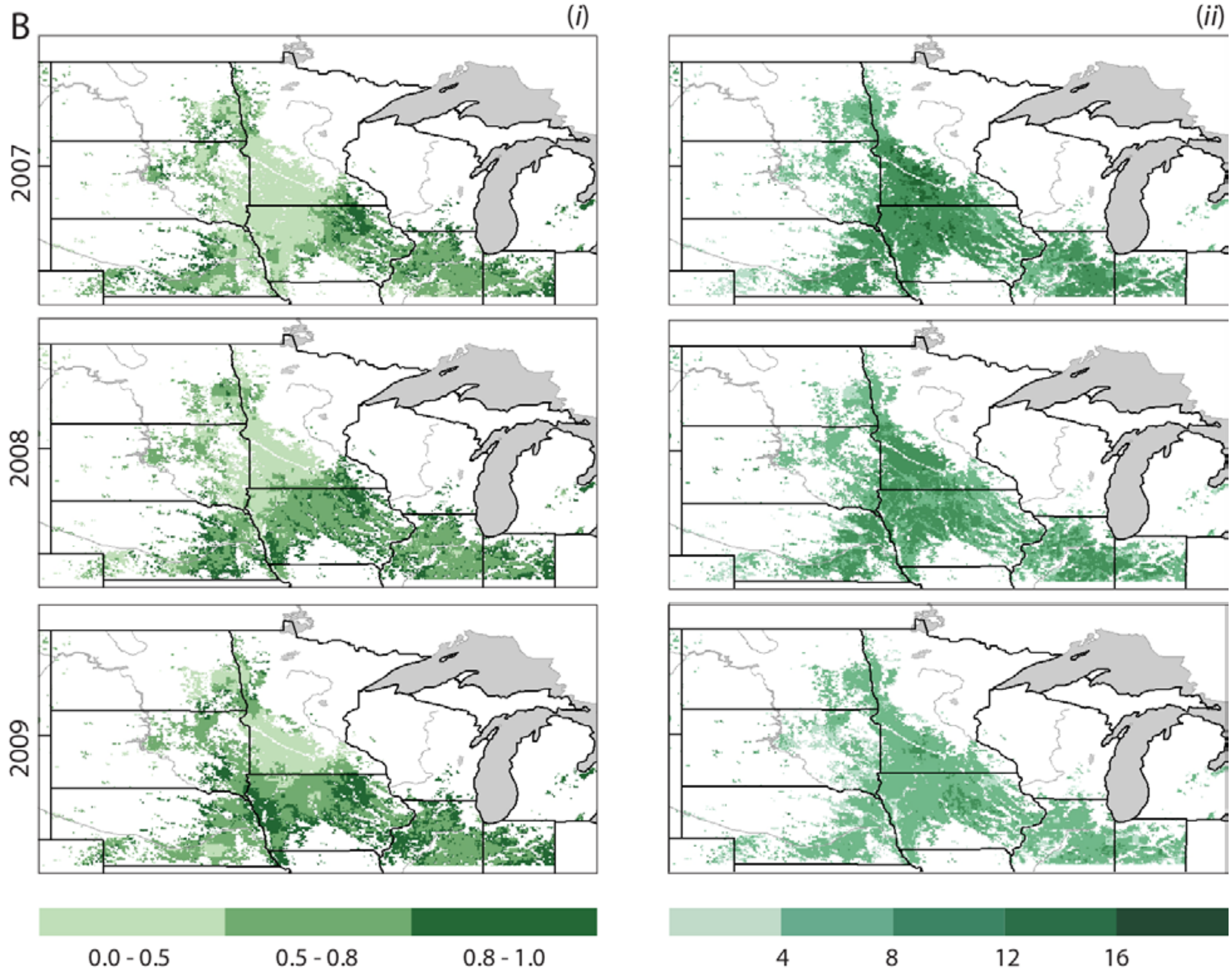
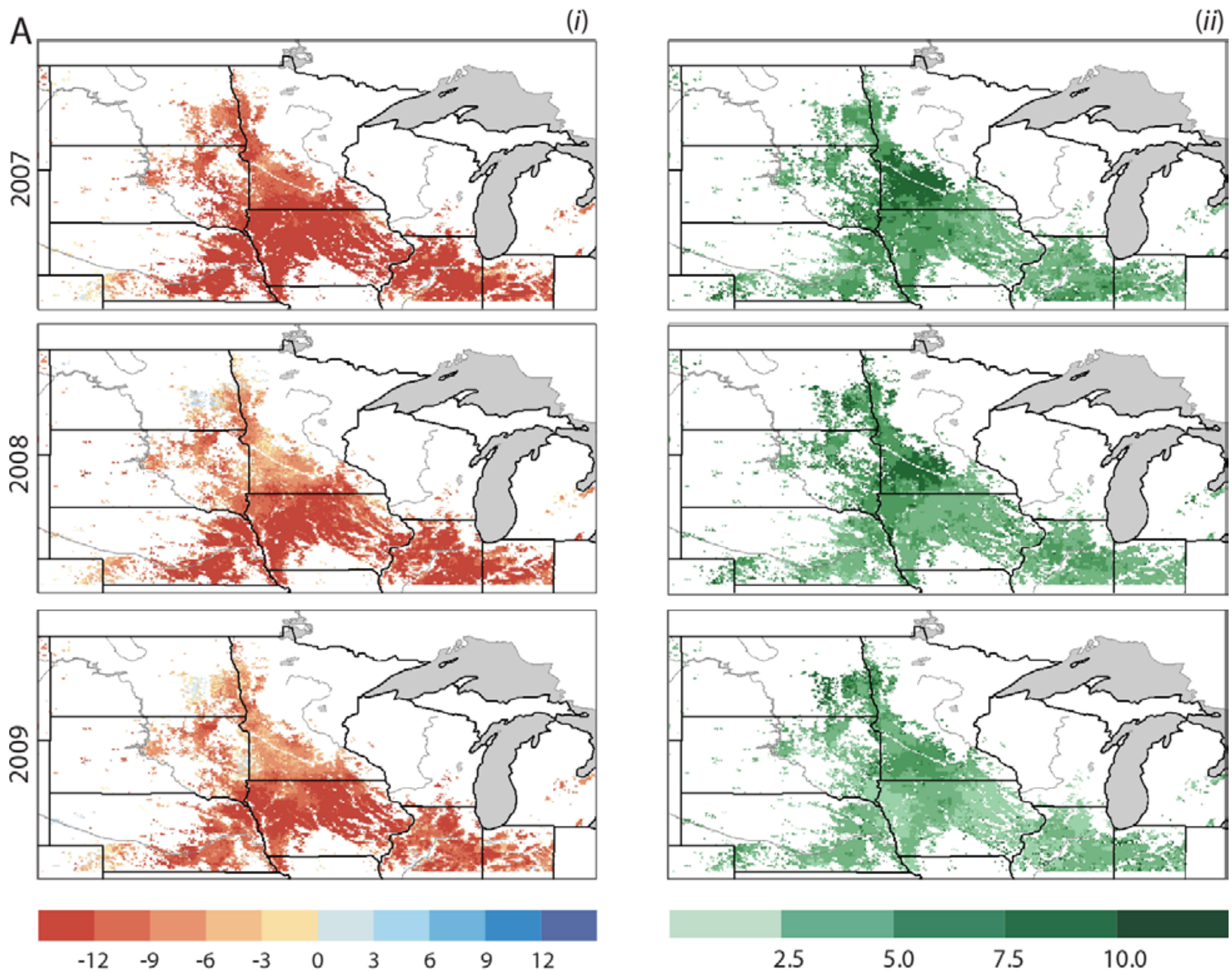


Figure A1. Comparison of ET_a from Mosaic LSM with the ET_a produced by EDPM plus VegET coupling scheme deployed in (A) prognostic mode and (B) diagnostic mode involving 1DKF assimilation. (*i*) Coefficient of determination (r^2); (*ii*) Root mean square error (mm per 8 days) .


A
(i)
(ii)
2007
2008
2009
-12
-9
-6
-3
0
3
6
9
12
2.5
5.0
7.5
10.0

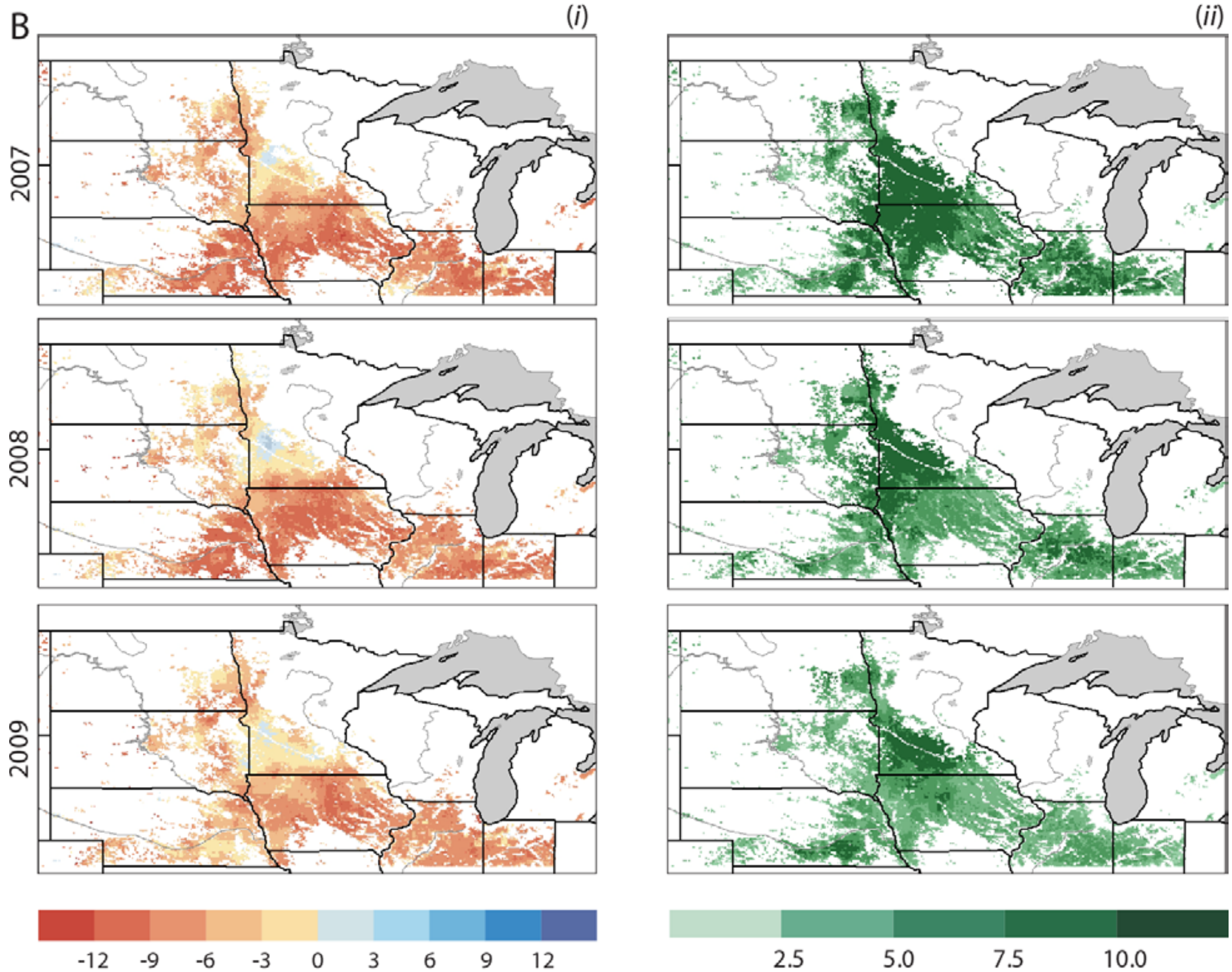


Figure A2. Spatial distributions of residuals (A) $ET_{a\ EDPM+VegET} - ET_{a\ Mosaic}$ (B)$ET_{a\ EDPM\ with\ 1DKF\ +VegET} - ET_{a\ Mosaic}$. (*i*) seasonal means of residuals (mm per 8 days) ; (*ii*) standard deviations of residuals (mm per 8 days) .