Alternative Methods to Predict Actual Evapotranspiration Illustrate the Importance of Accounting for Phenology: The Event Driven Phenology Model Part II.

V. Kovalskyy and G. M. Henebry

Anonymous Referee #1

Kovalskyy and Henebry present an application of their recently developed EDPM phenology model. The EPDM model is coupled to a simple model of evapotranspiration (VegET) and simulations of ET using this version of the model is contrasted with phenology based on satellite NDVI data (climatologies are retrospective time series). The authors aim at illustrating the advantages of the EPDM model version over the more traditional approaches. Overall, the manuscript is well written and the topic is interesting. However, there are many points which I criticize below.

Major points:

+ I don't understand why the authors do not directly compare the phenology from EPDM and satellite data but use an additional modeling step and compare it with flux tower data. Looking at the aims and objectives, it is clear that the authors are interested in how the phenology differs among the different approaches. Using the VegET model on top causes additional problems such as the bias of eddy co-variance ET data, which are used as benchmark, but also the modification of the model to take VPD into account is problematic in this context (see below).

Response. Comparison of phenology metrics from satellites and from the EDPM can be an interesting topic for a different paper. In this paper, however, we tried compare EDPM's predictions of seasonal trajectories with most common techniques used to drive seasonal changes. VegET provided a simple and straight forward framework for such comparison that, in addition, presented the results in terms of actual land surface flux and not just in terms of a spectral index that gives only an indication of vegetation state. Nevertheless, we tried to address referees comments about eddy co-variance ET data and VPD correction (see below).

+ The authors do not make the purpose of the model clear, i.e. if it is meant to run in diagnostic or prognostic mode. Depending on diagnostic vs prognostic, different validations/model comparisons should be made. I assume that the model is meant to be used in prognostic mode such as forecasts since for the past phenology data from satellites are available. Therefore, I expected to see a validation where some recent years of potentially available training data are left out in the training and used for validation. In such a comparison the authors could contrast their results with those based on MODIS phenology data of the specific years (not using climatologies). In this context, a comparison with another state-of-the-art interactive phenology model implemented in biosphere models would have been interesting to see to judge on the actual value of the EPDM model.

Response. We shared the curiosity of the referee #1 and in the next paper, that presents results from spatially explicit assessment of EDPM performance, we compare NDVI produced by EDPM and NDVI observations from MODIS for 3 years and more than 18k pixels in Northern Great Plains. In this BG manuscript, however, we tried to get an idea of how the newly developed EDPM can work in a coupled scheme and how those results compare to the best available references such as flux tower ET records and to the typical solutions coming from climatologies. Such information provides us with a starting point and with a ground to put the EDPM in competition with another state-of-the-art interactive phenology model. To clarify this, we inserted the following statement into the last paragraph of Introduction: "*We compare each modeled ET_a outcome with the best available references—ET measured at flux towers—to characterize performance of the interactive EDPM in the coupled scheme relative to typical solutions coming from climatologies.*"

As for phenology products from MODIS (MOD09PHN), they are still on their way to become operational. Unlike the EDPM which predicts seasonal TNDVI trajectories and phenological transition times, the PHN products include only metrics (e.g. start of season, end of season, peak of season etc.) derived from curves retrospectively fitted into EVI or NDVI observations (Tan et al., 2007). The PHN metrics has not been fully assessed and validated to become a reference point (Zhang et al., 2006, 2009).

+ In my opinion using the tower NDVI values for training is not a good test for the actual application of the model because the TNDVI values are only available at measurement stations. Training instead with satellite data would make more sense to me and be a more realistic test of the model. The comparison of results based on satellite NDVI and TNDVI is quite unfair because the towers of the used sites are short so that they have a small footprint of only a few 10 meters (I guess) while the satellite footprint is in the order of 1 km and coarser. These footprint differences may cause differences in the performance of the models caused by the resolution and may not be due to 'errors' in phenology.

Response. We agree with the referee #1 that footprint sizes have an impact on the data. This issue was discussed in a previous paper (Kovalskyy et al., 2011) where we showed comparability of TNDVI and NDVI from satellites for phenological monitoring. There we also reported footprint sizes (24-30m) of pyranometers and albedo meters (relevant to TNDVI) mounted on flux towers at Fermi and Brookings. As for the energy fluxes estimated with equipment used for eddy covariance method we could not be at the position to model temporal dynamics of footprint since such a correction (Kim et al. 2006) requires finer spatial resolution image time series that have not been available to us. Instead we followed the lead of many studies (Nagler et al., 2005; Yang et al., 2006; Cleugh et al.,2007; Mu et al.,2007; Senay 2007, 2008;Mu et al., 2009; Zhang et al., 2009) and assumed that the towers were located in homogeneous areas where the fluctuation of footprint due to wind and heterogeneity of land cover have minimal effect on energy fluxes quantified per unit area.

Finally, we clarified the footprint sizes for the data in the section 2.4 that now reads:

"Remotely sensed observations from NASA's Moderate Resolution Imaging Spectroradiometer (MODIS) and from NOAA's Advanced Very High Resolution Radiometer (AVHRR) were obtained from the following two sources: (1) MODIS NBAR 0.5 km resolution product (2000-2009) at ftp://e4ftl01u.ecs.nasa.gov/MOTA/MCD43A4.005/; and (2) AVHRR 1.1 km resolution NDVI composites by USGS (1989-2007) at http://edcsns17.cr.usgs.gov/EarthExplorer/."

+ The authors motivate their approach based on the poor representation of phenology in existing models, in particular regarding the representation of interannual variability (which is e.g. not captured when using a climatology). However, the authors also do not explicitly analyze the performance regarding anomalies (i.e. deviations from the mean seasonal cycle) although this is

expected given the motivation presented in the introduction (interannual variability etc). This could have been done if the analysis was based on a validation in forward model where recent years are left out from training.

Response. We considered comparing anomaly sizes with EDPM errors during these anomalies to be a part of this manuscript. In our internal work we saw that the advantage coming from the interactive model was especially high during anomalous years. However, we did not make a distinction of any anomalous year because there was incomplete temporal overlap in flux tower data from different sites. Plus, a major drought affected Brookings and Mead sites in 2007 but not Bondville and Fermi; late spring was in our data for Brookings 2008 and Mead 2002 only. Here is the statement we included into the discussion section to acknowledge this gap in our analysis: *"The analysis conducted for this study is incomplete without year by year comparison of performance between the five arrangements of VegET during different phenophases. It would help reveal reasons for poor performance by climatologies during anomalous years with shifts in the timing of spring or late season droughts. Lack of complete temporal overlap and vast distances between the flux tower locations prevented us from including such analysis into this study. "*

+ ET measurements from the eddy covariance method are usually low biased by 10-30%. The authors do not account for that problem nor discuss it properly. Given the often slight differences in performance (RMSEs) of the different ET models and the problems with the tower ET data I find it very difficult to judge on the adequacy of some of the models.

Response. According to literature listed further in the response, the problem mentioned by the referee #1 is not directly related to errors in latent heat flux (LE or ET) measurements. LE is only one of several factors that may cause **lack of energy balance closure** in flux tower records together with Sensible Heat Flux (H), Storage heat (S), Ground Heat Flux (G) (Twine et al., 2000;Tanaka et al., 2001 ; Wilson et al., 2002; Ham and Heilman 2003; Ma et al., 2003; Tanaka et al., 2003; Meyers and Hollinger, 2004; Oliphants et al., 2004; Mauder et al., 2006; Cava et al., 2008; Foken 2008; Wolf et al., 2008). Some studies suggest that the energy balance closure is not possible without footprint modeling for eddy covariance measurements since the sensors for

LE estimation and Net Radiation (Rn) estimation are based on different principles and have variable (for LE) and static (for Rn) footprints that should be brought to some common ground (Kim et al 2006) . To force energy balance closure as suggested by Twine et al. (2000), we would need H, S records that were not kept consistently at all sites. Nevertheless, our task in this manuscript was not to solve the currently debated problem in eddy covariance data but to compare results of coupled modeling with best available reference data.

+ Page 5346, lines 2-8: The authors use additional information of VPD to correct some biases of predicted ET by EPDM; which means they change the model and it is no longer only a phenology model. This modification seems not to be made for other approaches that are used in the comparison, which in my opinion is unfair.

Response. VPD correction was used **only on grassland sites** and we pointed out in the manuscript that the EDPM did not bring a great advantage there. The problem with equipment calibration on complicated sensor arrays persists even in the level II Ameriflux data products. In this case it was not a measurement or modeling problem, it was an inversion problem where the discrepancy might have come from any of the 7 variables in Penman-Monteith model, as well as from gaps in the soil water content records. Our attempt to compensate for obvious correlation between VPD and differences between inverted and modeled Kcp on grassland sites appeared reasonable to us, but it was not particularly successful. We inserted the following text into the discussion section: *"… the correction for drift in inverted K_{cp} on grassland sites that correlated with VPD appeared to bring no advantage to the EDPM and should not be used in future research or application."*

+ The performance of the EPDM (using automatic PTPs) based ET models measured by RMSEs is also not really outstanding (only in 2 out of 4 sites is (slightly) better than using MODIS NDVI) despite the fact that the EPDM model uses VPD in addition and is based on much better training data (tower NDVI vs satellite NDVI). I do not find that very encouraging.

Response. We do admit small differences between those two sets of results in our discussion. However the difference between MODIS NDVI based results and EDPM based results is that the

EDPM was predicting daily TNDVI values while MODIS NDVI data were just a time series of interpolated observations. Having predictions close to observations is actually quite desirable for modeling; our guess is that MODIS NDVI based results are as good as the VegET can deliver with contemporary satellite data. Even though we do not share the referee's interpretation, we are happy to see the same minor difference between the two sets of results. And as for VPD, please refer to the previous response.

Some specific points:

+ Page 5337, lines 14-28. The authors may also mention data driven approaches to modeling ET based on eddy covariance, remote sensing, and machine learning approaches (e.g. Yang et al 2006 IEEE; Jung et al 2010 Nature).

Done. The following statement was inserted into the second paragraph of the introduction: "*Several attempts were made to use empirical machine learning techniques for ET modeling (Yang et al., 2006; Kaheil et al., 2008; Jung et al., 2010).*"

+ Page 5338, lines 16-28. The authors refer to models with very simplistic representations of phenology (e.g. climatologies). Many models, in particular those designed for the biosphere model dynamic LAI most of them on daily time-step based on daily carbon allocation.

Response. We are well aware that such models exist, but those are rarely used in narrowly focused ET estimation studies, most likely because verifying their performance on daily bases is hardly possible, since LAI is not collected on a fine temporal resolution at flux tower sites, but is instead interpolated between periodic measurements.

+ Page 5345, lines 18-20: not clear if it's only due to limited data or also due to limitations of EPDM.

Done. (Rephrased). Now the sentence reads as: "*Differences in performance between these two regimes should reveal discrepancies created by lack of data to properly define distributions of phenological transition dates*"

+ Page 5347, line 25: absolute residuals?

Done. Corrected to absolute residual.

+ The usage of prescribed PTPs seems to be useless to me and unfair comparison. If we knew the PTPs then we wouldn't need a model. What is the point of including them in the analysis?

Response. We respectfully disagree: this arrangement was made not for giving the EDPM an unfair advantage but for showing its imperfection at current stage of development. For this reason we show both automatic PTP and prescribed PTP results. Following explanation was inserted into 4th paragraph of section 2.5(Data preparation). "*Contrasting the performance of the EDPM with prescribed PTPs and automatically estimated PTPs is thought to show how far the fully automatic EDPM results are from its ideal performance and to emphasize that error in phenological timing affects the performance of the EDPM.*"

+ I don't understand why PET is included in the comparison. I suggest to remove PET.

Response. It is not a potential evapotranspiration, it is reference ET (1st is from water surface and 2nd is from well watered grass covered surface LAI=2 Allen et al., 1998). PM model was brought in because during the preparation to this research we found lots of criticism of not accounting for phenology but no actual quantification of accuracy losses. These results were thought to be used as quantitative references in further research by us and by other members of scientific community.

List of References.
Cava, D., Contini, D., Donateo, A., and Martano, P.: Analysis of short-term closure of the surface energy balance above short vegetation, Agricultural and Forest Meteorology, 148, 82-93, 10.1016/j.agrformet.2007.09.003, 2008.

Cleugh, H. A., Leuning, R., Mu, Q., and Running, S. W.: Regional evaporation estimates from flux tower and MODIS satellite data, Remote Sensing of Environment, 106, 285-304, 2007.

Foken, T.: THE ENERGY BALANCE CLOSURE PROBLEM: AN OVERVIEW, Ecological Applications, 18, 1351-1367, 10.1890/06-0922.1, 2008.

Ham, J. M., and Heilman, J. L.: Experimental Test of Density and Energy-Balance Corrections on Carbon Dioxide Flux as Measured Using Open-Path Eddy Covariance, Agron. J., 95, 1393–1403, 2003.

Kim, J., Guo, Q., Baldocchi, D. D., Leclerc, M. Y., Xu, L., and Schmid, H. P.: Upscaling fluxes from tower to landscape: Overlaying flux footprints on high-resolution (IKONOS) images of vegetation cover, Agricultural and Forest Meteorology, 136, 132-146, 10.1016/j.agrformet.2004.11.015, 2006.

Kovalskyy, V., Roy, D. P., Zhang, X., and Ju, J.: 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 3:4, 325-334, 2011.

Ma, Y., Su, Z., Koike, T., Yao, T., Ishikawa, H., Ueno, K. i., and Menenti, M.: On measuring and remote sensing surface energy partitioning over the Tibetan Plateau—from GAME/Tibet to CAMP/Tibet, Physics and Chemistry of the Earth, Parts A/B/C, 28, 63-74, 10.1016/s1474-7065(03)00008-1, 2003.

Mauder, M., Liebethal, C., Göckede, M., Leps, J.-P., Beyrich, F., and Foken, T.: Processing and quality control of flux data during LITFASS-2003, Boundary-Layer Meteorology, 121, 67-88, 10.1007/s10546-006-9094-0, 2006.

Monteith, J. L.: Evaporation and environment, Symposium Society Experiment Biology London, 1965, 205-234, 1965.

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

Mu, Q., Jones, L. A., Kimball, J. S., McDonald, K. C., and Running, S. W.: Satellite assessment of land surface evapotranspiration for the pan-Arctic domain, Water Resources Research, 45, W09420, 10.1029/2008wr007189, 2009.

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

Oliphant, A. J., Grimmond, C. S. B., Zutter, H. N., Schmid, H. P., Su, H. B., Scott, S. L., Offerle, B., Randolph, J. C., and Ehman, J.: Heat storage and energy balance fluxes for a temperate deciduous forest, Agricultural and Forest Meteorology, 126, 185-201, 10.1016/j.agrformet.2004.07.003, 2004.

Senay, G., Budde, M., Verdin, J., and Melesse, A.: A coupled remote sensing and simplified surface energy balance approach to estimate actual evapotranspiration from irrigated fields, Sensors, 7, 979-1000, 2007.

Senay, G.: Modeling landscape evapotranspiration by integrating land surface phenology and a water balance algorithm, Algorithms, 1, 52-68, 2008.

Tan, B., Morisette, J., Wolfe, R., Gao, F., Nightingale, J. M., Pedelty, J., and Ederer, G.: User guide for MOD09PHN and MOD15PHN products: http://accweb.nascom.nasa.gov/project/docs/User_guide_PHN.pdf, access: 05/15/2011, 2007.

Tanaka, K., Ishikawa, H., Hayashi, T., Tamagawa, I., and Ma, Y.: Surface Energy Budget at Amdo on the Tibetan Plateau using GAME/Tibet IOP98 Data, Journal of the Meteorological Society of Japan, 79, 505-517, 2001.

Tanaka, K., Tamagawa, I., Ishikawa, H., Ma, Y., and Hu, Z.: Surface energy budget and closure of the eastern Tibetan Plateau during the GAME-Tibet IOP 1998, Journal of Hydrology, 283, 169-183, 10.1016/s0022-1694(03)00243-9, 2003.

Twine, T. E., Kustas, W. P., Norman, J. M., Cook, D. R., Houser, P. R., Meyers, T. P., Prueger, J. H., Starks, P. J., and Wesely, M. L.: Correcting eddy-covariance flux underestimates over a

grassland, Agricultural and Forest Meteorology, 103, 279-300, 10.1016/s0168-1923(00)00123-4, 2000.

Wilson, K., Goldstein, A., Falge, E., Aubinet, M., Baldocchi, D., Berbigier, P., Bernhofer, C., Ceulemans, R., Dolman, H., Field, C., Grelle, A., Ibrom, A., Law, B. E., Kowalski, A., Meyers, T., Moncrieff, J., Monson, R., Oechel, W., Tenhunen, J., Valentini, R., and Verma, S.: Energy balance closure at FLUXNET sites, Agricultural and Forest Meteorology, 113, 223-243, 10.1016/s0168-1923(02)00109-0, 2002.

Wolf, A., Saliendra, N., Akshalov, K., Johnson, D. A., and Laca, E.: Effects of different eddy covariance correction schemes on energy balance closure and comparisons with the modified Bowen ratio system, Agricultural and Forest Meteorology, 148, 942-952, 10.1016/j.agrformet.2008.01.005, 2008.

Yang, F., White, M. A., Michaelis, A. R., Ichii, K., Hashimoto, H., Votava, P., Zhu, A. X., and Nemani, R. R.: Prediction of Continental-Scale Evapotranspiration by Combining MODIS and AmeriFlux Data Through Support Vector Machine, Geoscience and Remote Sensing, IEEE Transactions on, 44, 3452-3461, 2006.

Zhang, X., Friedl, M. A., and Schaaf, C. B.: Global vegetation phenology from Moderate Resolution Imaging Spectroradiometer (MODIS): Evaluation of global patterns and comparison with in situ measurements, Journal of Geophysical Research, 111, G04017, 10.1029/2006jg000217, 2006.

Zhang, J., Hu, Y., Xiao, X., Chen, P., Han, S., Song, G., and Yu, G.: Satellite-based estimation of evapotranspiration of an old-growth temperate mixed forest, Agricultural and Forest Meteorology, 149, 976-984, 2009.

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