An experimental design to test for the propagation of land cover uncertainty in climate modeling

Xue Li[1,2], *Joseph Messina*[1,2], *Nathan Moore*[1,2]

[1] Department of Geography, Michigan State University

[2] Center for Global Change and Earth Observations, Michigan State University

Abstract

MODIS Land Cover Type (MLCT) is an annual global dataset that is intensively used as input in climate modeling across spatial scales. Despite its wide usage, MLCT's uncertainty and propagation impacts on climate model outputs have not received enough attention. The purpose of this study is to design an experimental framework to test for the propagation impacts of land cover uncertainty on regional climate modeling. Using the Regional Atmospheric Modeling System (RAMS) as example, we developed a three-step workflow, including: (1) characterizing and modeling land cover uncertainty, (2) setting up RAMS to test for uncertainty propagation, and (3) analyzing complex spatio-temporal outputs from RAMS. Each step poses a particular set of challenges that deserve deeper investigation. A case experiment was conducted for Urumqi, China to present the effectiveness of this framework. We argue that such uncertainty analysis should be included routinely in climate modeling work.

Keywords: Uncertainty propagation, land cover, MODIS, climate simulation, RAMS.

1. Introduction

The MODIS Land Cover Type (MLCT) dataset (Friedl *et al.*, 2002) is produced annually by the research group in Boston University beginning 2001. The overall classification accuracy of the MLCT product is claimed at 75%, but it can be much lower for certain classes or in certain regions (MODIS land cover team, 2003; Frey and Smith, 2007; Friedl *et al.*, 2010). MLCT products are used widely as input for climate models, but associated data uncertainties are rarely adequately considered.

Instead of comparing to other datasets produced using different methods and time periods (Ge, 2007; Sertel *et al*., 2010), uncertainty can be revealed from the time series of MLCT products themselves. As pointed out by Liang and Gong (2010), highly unstable locations usually suggest classification inaccuracy across years, rather than actual land cover changes. Thus, we are able to extract uncertainty of land cover data from those unstable locations.

Using Urumqi, China as the case study area, we explored an experimental framework to model and propagate MLCT uncertainties through the Regional Atmospheric Modeling System (RAMS; Cotton *et al.*, 2003), and characterize uncertainty impacts across perspectives.

2. MLCT uncertainty model

Frequent inter-annual changes in MLCT more often result from classification inaccuracies than from actual land cover changes (Liang and Gong, 2011). Furthermore, frequent changes between particular land cover types usually suggest failure to distinguish between those types, in another words, categorical uncertainty. We name pixels with such behavior as “flipping pixels,” and they can be used to identify specific categorical uncertainties.

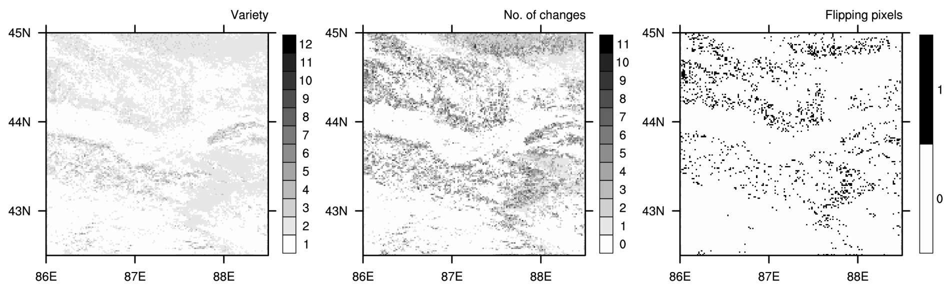

Figure 1: Pixel-based trajectory analysis

The latest version of MLCT (MCD12Q1-V051) products were acquired from 2001 to 2012 for our study area (42.5ºN~45ºN, 86ºE~88.5ºE). The 12-year change trajectory was calculated for each pixel (figure 1). “Flipping pixels” are defined as pixels changed more than 4 times among no more than 3 land cover types. We subdivided the trajectories of “flipping pixels” into their inter-annual change directions and ranked them by their frequencies. The most frequent directions are considered suspicious instead of actual land cover changes. Data for our study area suggest high uncertainties for grasslands vs. croplands and grasslands vs. barren (table 1).

Table 1: Suspicious change directions and their frequencies.

Rank	From	To	Overall frequency
1	Grasslands	Croplands	21.74%
2	Barren	Grasslands	21.66%
3	Grasslands	Barren	21.54%
4	Croplands	Grasslands	20.87%
5	Grasslands	Evergreen needleleaf forest	1.16%
6	Open shrublands	Grasslands	0.98%

To propagate categorical uncertainties into applications, a binary control model was developed to generate equally probable land cover realizations. For any pixels that changed in suspicious directions from one year to the next, we assume 1) equal probabilities for such changes actually happened or not, 2) systematic behavior of all pixels fell into same directions, and 3) independency of different directions. Land cover realizations for the latter year would come from combinations of turning on/off different suspicious directions. A case with N suspicious directions would yield 2^N realizations, which is 16 in our case.

3. RAMS setup

Land cover inputs are passed into RAMS through the LEAF-3 vegetation and surface submodel (Walko *et al*., 2000). LEAF-3 aggregates data into coarser grids by storing up to five land cover types and their respective areas within each grid. Land cover types were translated into a set of biophysical parameters and then used in calculating atmospheric dynamics.

In our case study, we simulated a 25-day period starting March 21, 2011. Sixteen land cover realizations were produced for the 2011 MLCT product to replace RAMS' default Olson Global Ecosystem land cover data. Separate simulations were run using each land cover realizations on two nested domains, with the coarser one at 32 km resolution and finer, which covers the study area, at 8 km. Original 2011 MLCT data were padded outside the study area for the coarser and larger domain simulations. Other model options and initial parameterizations are not trivial issues for any climate simulation, but details are beyond the scope of this paper.

4. Output diagnostic

For applications with specific variables of interest, it is straightforward to assess the impact of MLCT uncertainties by examining the distribution of the target variable from the RAMS output ensemble. RAMS models complex spatio-temporal dynamics of atmospheric conditions, so the behavior of uncertainty propagation can also be complex spatial-temporally, therefore need to be characterized across different dimensions.

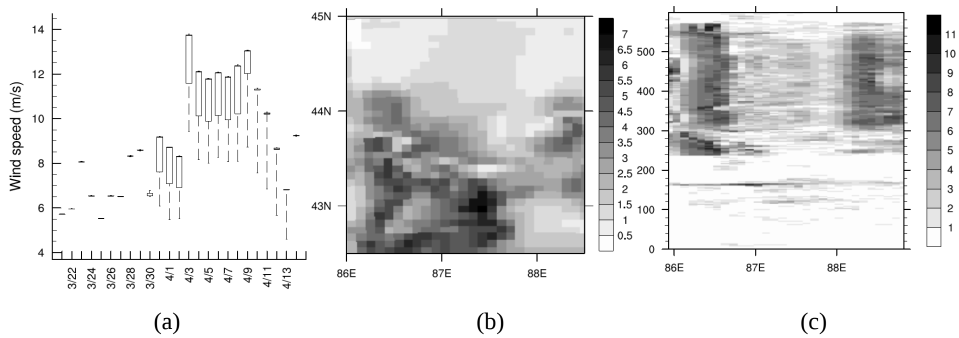

Figure 2: Uncertainty propagation illustrations for wind speed at 10 m above land surface: (a) aggregated temporal profile, (b) spatial profile of output ranges, and (c) spatio-temporal profile of output ranges along 44ºN at hourly resolution.

The overall evolution of uncertainty impacts can be illustrated from a temporal profile using box-plots (figure 2a). Each box depicts the distribution of daily domain-averaged results from different simulations, for a selected variable at a given time. Time series of boxes depict the trends of actual values as well as trends of their spread over time. Spatial heterogeneity of the uncertainty impacts can be revealed by a spatial profile of the variable of interest. A measure of dispersion (e.g. Standard deviation, range) can be calculated by pixel at each time-step and then average over time (figure 2b). Spatial-temporal profile is one way to show patterns across spatial and temporal dimension (figure 2c) for a cross-section of simulation domain. When presentation is not limited to printed media, animations can be used to better illustrate uncertainty patterns and support further explanations on mechanisms that produced such pattern. Climate is the long-term pattern of atmospheric conditions. Short-term differences in atmospheric conditions do not necessarily integrate into different climate patterns. Formal statistical tests, such as T-test and F-test can be used to identify significant deviations among simulations in terms of long-term mean and variance, respectively. Such test can be applied to both domain averages and individual pixels.

5. Discussion

One big challenge that prevents climate modelers from including such land cover uncertainty analysis into their work is the significant computation demand of climate models. For our study case, simulating 25 days in one simulation requires 2 to 4 single-core days on the Michigan State University HPCC facility.

The computational demand for our proposed method increases exponentially with the number of suspicious directions. For areas with more complex MLCT uncertainties, the number of land cover realizations (2^N) could demand tremendous computational time. One solution is to drop the third assumption in our uncertainty model regarding independence of different directions. For each pair of uncertain land cover types (2 directions), we can assume bias towards either type by turning on one direction and off the other. This will reduce the number of realizations to $2^{N/2}$ and maintain most of the dispersion among simulation results.

6. Conclusion

This study explored a three-step experimental framework to test for the propagation impacts of land cover uncertainty on climate modeling. 1) Categorical uncertainties in MLCT can be characterized locally from “flipping pixels” and modeled using a binary control method. 2) Regional climate models, like RAMS, can dynamically downscale global-scale atmospheric conditions to regional/local-scale responding to different land surface characterizations. 3) Outputs from RAMS can be diagnosed graphically and statistically to reveal important uncertainty propagation patterns. Our case study shows the effectiveness of our proposed framework as well as substantial uncertainty impacts in a short-time simulation. Therefore, we recommend the climate modeling community to routinely include land cover uncertainty analysis in their work.

Acknowledgments

This research is supported by National Aeronautics and Space Administration (NNX09AI32G) and National Geospatial-Intelligence Agency (HM02101310002).

References

Cotton, W. R., et al. (2003), “RAMS 2001: Current status and future directions”. *Meteorol. Atmos. Phys.*, Vol. 82(1-4): 5-29.

Frey, K. E., Smith, L. C. (2007), “How well do we know northern land cover? Comparison of four global vegetation and wetland products with a new ground-truth database for West Siberia”. *Global Biogeochem. Cycles,* Vol. 21 GB1016.

Friedl, M. A. *et al*. (2002), “Global land cover mapping from MODIS: Algorithms and early results”. *Remote Sensing of Environment B*, Vol. 83(1-2): 287-302.

Friedl, M. A., *et al.* (2010), “MODIS Collection 5 global land cover: algorithm refinements and characterization of new datasets”. *Remote Sensing of Environment,* Vol. 114(1):168-182.

Ge, J., *et al*. (2007). “Impacts of land use/cover classification accuracy on regional climate simulations”. *Journal of Geophysical Research*, Vol. *112*(5), D05107.

Liang, L. and Gong, P. (2010), “An assessment of MODIS Collection 5 global land cover product for biological conservation studies”. *Proc. 18th Int. Conf. on Geoinformatics*, Beijing, China, GEOINFORMATICS.2010.5567991

MODIS land cover team. (2003), “Validation of the consistent year 2003. V003 MODIS land cover product”. Available at: http://geography.bu.edu/landcover/userguidelc/consistent.htm

Sertel, E., Robock, A., & Ormeci, C. (2010). “Impacts of land cover data quality on regional climate simulations”. *International Journal of of Climatology, Vol.* 30(13): 1942-1953.

Walko, R. L., *et al*. (2000). “Coupled atmosphere-biophysics-hydrology models for environmental modeling”, *J. Appl. Meteorol.* Vol. 39(6): 931-944.