



Responsible Land Governance: Towards an Evidence Based Approach

ANNUAL WORLD BANK CONFERENCE ON LAND AND POVERTY
WASHINGTON DC, MARCH 20-24, 2017



FINE-SCALE LAND ALLOCATION TOOL FOR GLOBAL LAND USE ANALYSIS

Jingyu Song ^a, Michael S. Delgado ^a, Paul V. Preckel ^a, Shandian Zhe ^b,
Ian J. Campbell ^b, Lan Zhao ^c, and Carol X. Song ^c

^a Department of Agricultural Economics, Purdue University, West Lafayette, IN 47907, USA

^b Department of Computer Science, Purdue University, West Lafayette, IN 47907, USA

^c Rosen Center for Advanced Computing, Purdue University, West Lafayette, IN 47907, USA

Corresponding author: Jingyu Song (song173@purdue.edu)

**Paper prepared for presentation at the
“2017 WORLD BANK CONFERENCE ON LAND AND POVERTY”
The World Bank - Washington DC, March 20-24, 2017**



Responsible Land Governance: Towards an Evidence Based Approach

ANNUAL WORLD BANK CONFERENCE ON LAND AND POVERTY
WASHINGTON DC, MARCH 20-24, 2017



Copyright 2017 by author(s). All rights reserved. Readers may make verbatim copies of this document for non-commercial purposes by any means, provided that this copyright notice appears on all such copies.

Abstract

We describe an open access online tool that provides statistical analysis, land use projections, and model comparison at variable grid cell resolutions (grid cell size 1-60 arc-minutes) at a global or sub-global level. The tool predicts patterns in land use at a grid cell level with aggregate land use data and disaggregate land attribute data. The marginal impact of the disaggregated land attribute data on land use can be computed and used for research and policy analysis. Visualization features facilitate comparison across models, and the user has the option of uploading his/her own data and/or estimation procedure.

Key Words: global land use, grid cell resolution, interactive map, open access

1. INTRODUCTION

Recent studies have advocated using high spatial resolution data measured at a grid cell level (hereafter fine-scale data) to analyze patterns of land use, rather than using state/provincial or nationally aggregated data (Auffhammer et al., 2013; Hendricks et al., 2014; Kyle et al., 2015). The advantage of using fine-scale data is that such data captures heterogeneity that is germane to research and policy interests related to land use patterns; for example, understanding the relationship between temperature and cropland allocation, without restricting the unit of analysis to areas based on political boundaries. As a result, analyses based on fine-scale data tend to provide better support and guidance to policymakers. Yet, despite these empirical lessons, there are few sources of fine-scale cropland allocation data that cover a global or broad sub-global areas, and the existing data are typically based on satellite imagery or data processing techniques that lack transparency (Song et al. 2016). In an effort to provide an alternative, transparent source of fine-scale cropland allocation data that covers a sub-global area, we introduce an online tool that generates land allocation projections at various fine-scale resolutions for global or sub-global regions. These fine-scale resolutions are referred to as “pixels”, and range from 1 arc-minute to 60 arc-minutes latitude/longitude in size depending on user specification.

FLAT is a statistical-based tool that combines dependent variable data measured at an aggregate level (e.g., state/provincial or national) with independent variable data measured at the pixel level to produce estimates of the dependent variable at the pixel level (sometimes called downscaling). The dependent variable can be scalar (for predicting intensity of a single land use) or vector-valued (for predicting shares of alternative land uses). The structured parametric regression model also provides a means of projecting the dependent variable to spatially out-of-sample pixels if observations of the independent variables exist for those out-of-sample pixels but observations of the dependent variable at the aggregate level do not. In addition to the pixel level land shares, FLAT also provides instructions on calculating marginal effects of each independent variable on the pixel level land shares. FLAT also provides visualization of the estimated land shares, and allows users to compare projections from alternative model specifications. All details are provided in Song et al. (2016).

In our examples, the dependent variable is cropland allocation for corn, soybeans, and wheat measured at a state/provincial level and predicted at the 5-arc minute pixel level, and the independent variables are biophysical factors such as temperature, precipitation, and soil conditions, measured at the pixel level. The marginal effects can be used, for example, to evaluate the impact of a one degree Celsius increase in temperature on corn cropland allocation across North America, holding constant other variables such as precipitation or soil conditions.

The estimated pixel level land shares and associated marginal effects can be directly downloaded from FLAT and input into other models for any number of simulations or evaluations. These pixel level

land shares are useful in many research and policy areas. They are essential for examining patterns of human activity and development over time and the resulting impact on the ecological-environmental systems (Newbold, 2015; Pongratz et al., 2011). They can also serve as inputs/components to crop models, general circulation models, and integrated assessment models (Engström et al., 2016; Hertel and Lobell, 2014; Nelson et al., 2014). By providing transparent estimates of pixel level land shares, FLAT enables and supports global level studies for which land use/allocation is an integral part, promotes fine-scale global analysis, and assists in further advancement in multidisciplinary research and policy development. One potential use of FLAT is that it can be combined with the AgMIP tool (Villoria et al., 2016), which was based on the Agricultural Model Intercomparison and Improvement Project, to assess changes in global crop production under different climatic change scenarios.

Although our examples focus on fine-scale cropland allocation, FLAT is a general tool and is not limited to modeling cropland allocation. FLAT is designed to estimate pixel-level shares for data of any type such that the outcome variable(s) are measured at an aggregate level and the input variables are measured at the pixel level. That is, the FLAT interface allows the user to downscale data of any type to any desired pixel scale (from 1 to 60 arc-minutes).

2. MODEL DESCRIPTION

The FLAT user interface consists of three sections, corresponding to model setup and econometric estimation, statistical/numerical results, and visualization (<https://mygeohub.org/tools/flat>). In addition, the tool distinguishes between two types of datasets: the dataset for estimation and the dataset for land share projections. It is possible for one dataset to be used for both estimation and projection, but this is not a requirement. The tool is hosted by Purdue University's MyGeoHUB cyberinfrastructure, which supports geospatial modeling, data analysis, and visualization for researchers and educational communities across disciplines. FLAT is free, open access, and can be accessed using any standard Internet browser. The user can download all the default datasets and programming scripts in R (R Core Team, 2015) and GAMS (GAMS Development Corporation, 2013) from the user interface at no charge.

The first section, shown in Figure 1, is for model setup and execution of the statistical model. The user inputs a model name and chooses a dataset for estimation. The tool provides two default datasets: one for maize (as a single crop), and the other one for maize, soybeans, and wheat (i.e., multiple crops simultaneously). Both datasets cover aggregate areas (states and provinces) for North, Central and South America. Alternatively, a user can choose to upload his/her own dataset, and a sample dataset specifying the formatting requirements for uploaded data is provided in the "Instruction" window.

The user also defines the dataset for which to compute the land share projections. These projections are computed using the estimated parameters from the statistical model that are obtained via the estimation

dataset – that is, the estimated parameters are applied to the projection dataset to compute the projections. A default projection dataset is provided for maize for the Americas; alternatively, the user can upload his/her own projection dataset. Both the default estimation and projection datasets are downloadable by clicking on the “Instruction” window. For the default datasets, the pixel size is set to 5 by 5 arc-minutes, but the user can select any pixel size between 1 and 60 arc-minutes (at a 1 arc-minute interval).

Following dataset selection, the user specifies the number of dependent and independent variables; for example, the number of crops and the number of biophysical factors. (Note that the number of dependent and independent variables are fixed for the default datasets. If the user wishes to explore alternative models based on this data, the datasets should be downloaded, modified, and uploaded as user datasets.) As a final step, the user specifies whether to run the default statistical model or upload his/her own script. The tool uses the General Algebraic Modeling System (GAMS Development Corporation, 2013), so any user selecting to upload his/her own script must upload a GAMS script.

After all the model setup steps are completed, the user can run FLAT. FLAT first performs a check for errors and discrepancies in the dataset(s) and input specifications to make sure that input data meet data format requirements and there is no missing values. The status of each error check will be displayed in the log window, with a warning or error message being reported if encountered. Once the model has passed all necessary checks the GAMS script will run; the log window reports the estimation progress, and announces completion when the GAMS script finishes.

Results from the model are displayed and available for download in the second section of FLAT (see Figure 2). The tool displays the coefficient estimates and standard errors, and provides an option to download the projected pixel level land shares and the covariance matrix for the parameter estimates. Instructions and formulas for calculating marginal effects and odds ratios are available, which allows the user to calculate these metrics. In addition, FLAT can store the results from multiple models, and the user can conveniently choose among models to display and download the results.

The third section of FLAT is for visualization (see Figure 3). FLAT allows the user to display (map) the results from saved estimation and projection results for multiple models. In addition, FLAT allows the user to compare between two different models to directly see the differences in estimates between the two models. The user selects one dependent variable (e.g., one particular crop), for which the land shares are shown on the mapping area with a color scheme that reflects differences in share values. The default map shows country, state/province, and county/district boundaries, and the user can further tailor the map to any desired focus using a zoom-in function. Depending on the level to which the user enlarges the scale, roads, airports, buildings, town, etc. will be shown on the map. When the “Compare” box is checked, comparisons across selected models are displayed. All maps are downloadable in three different formats: .GeoTIFF, .PNG, and .JPG.

3. CASE STUDY

To provide an example, we run the default maize model in FLAT. The default maize dataset includes 196 states from 18 countries across North, Central, and South America. (For details on the selection of variables and regions as well as the theoretical derivation of the estimation procedure, see Song et al., 2016.) The default grid cell size is 5 by 5 arc-minutes, there is one dependent variable, and there are 27 independent variables. The results from the model are shown in Figure 4. The display window shows the crop name, the names of the independent variables, the coefficient estimates, and the standard errors. The estimated and projected cropland allocation fractions are available for download, and Figure 5 shows these allocations on the map. As shown by the color variation, Iowa, Illinois and Indiana have the largest land shares in maize. The northern parts of Iowa and Illinois have as high as 50 percent of the pixel land area in maize.

ACKNOWLEDGEMENTS

The authors are grateful for helpful comments and suggestions from Thomas Hertel. FLAT development is supported in part by USDA grant Agreement #58300010058 and Purdue grant #105651.

REFERENCES

- Auffhammer, M., Hsiang, S.M., Schlenker, W., and Sobel, A. (2013). Using weather data and climate model output in economic analyses of climate change. *Review of Environmental Economics and Policy*, 7(2), 181-198.
- Engström, K., Rounsevell, M.D.A., Murray-Rust, D., Hardacre, C., Alexander, P., Cui, X., Palmer, P., and Arneth, A. (2016). Applying Occam's razor to global agricultural land use change. *Environmental Modelling & Software*, 75 (2016), 212-229.
- GAMS Development Corporation. General Algebraic Modeling System (GAMS) Release 24.2.1. Washington, DC, USA, 2013.
- Hendricks, N.P., Smith, A., and Sumner, D.A. (2014). Crop supply dynamics and the illusion of partial adjustment. *American Journal of Agricultural Economics*, 96(5), 1469-1491.
- Hertel, T.W., and Lobell, D.B. (2014). Agricultural adaptation to climate change in rich and poor countries: Current modeling practice and potential for empirical contributions. *Energy Economics*, 46(2014), 562-575.

- Nelson, E., Sander, H., Hawthorne, P., Conte, M., Ennaanay, D., Wolny, S., Manson, S., and Polasky, S. (2010). Projecting global land-use change and its effect on ecosystem service provision and biodiversity with simple models. *PLOS ONE*, 5(12), e14327.
- Kyle, P., Thomson, A., Wise, M., and Zhang, X. (2015). Assessment of the importance of spatial scale in long-term land use modeling of the Midwestern United States. *Environmental Modelling & Software*, 72(2015), 261-271.
- Nelson, G.C., Valin, H., Sands, R.D., Havlík, P., Ahammad, H., Deryng, D., Elliott, J., Fujimori, S., Hasegawa, T., Heyhoe, E., Kyle, P., Von Lampe, M., Lotze-Campen, H., Mason d’Croz, D., van Meijl, H., van der Mensbrugge, D., Müller, C., Popp, A., Robertson, R., Robinson, S., Schmid, E., Schmitz, C., Tabeau, A., and Willenbockel, D. (2014). Climate change effects on agriculture: Economic responses to biophysical shocks. *Proceedings of the National Academy of Sciences*, 111(9), 3274-3279.
- Newbold, T., Hudson, L.N., Hill, S.L.L., Contu, S., Lysenko, I., Senior, R.A., Börger, L., Bennett, D.J., Choimes, A., Collen, B., Day, J., De Palma, A., Díaz, S., Echeverria-Londoño, S., Edgar, M.J., Feldman, A., Garon, M., Harrison, M.L.K., Alhousseini, T., Ingram, D.J., Itescu, Y., Kattge, J., Kemp, V., Kirkpatrick, L., Kleyer, M., Laginha Pinto Correia, D., Matin, C.D., Meiri, S., Novosolov, M., Pan, Y., Phillips, H.R.P., Purves, D.W., Robinson, A., Simpson, J., Tuck, S.L., Weiher, E., White, H.J., Ewers, R.M., Mace, G.M., Schaelemann, J.P.W., and Purvis, A. (2015). Global effects of land use on local terrestrial biodiversity. *Nature*, 520(7545), 45-50.
- Pongratz, J., Reick, C.H., Raddatz, T., Caldeira, K., and Claussen, M. (2011). Past land use decisions have increased mitigation potential of reforestation. *Geophysical Research Letters*, 38, L15701.
- R Core Team (2015). R: A language and environment for statistical computing. R Foundation for Statistical Computing, Vienna, Austria. URL <https://www.R-project.org/>.
- Song, J., Delgado, M.S., Preckel, P.V., and Villoria, N.B. (2016). Pixel Level Cropland Allocation and Marginal Impacts of Biophysical Factors, *working paper*.
- Villoria, N.B., Elliott, J., Müller, Christoph, Shin, J., Zhao, L., and Song, C.X. (2016). Rapid aggregation of global gridded crop model outputs to facilitate cross-disciplinary analysis of climate change impacts in agriculture. *Environmental Modelling & Software*, 75 (2016), 193-201.

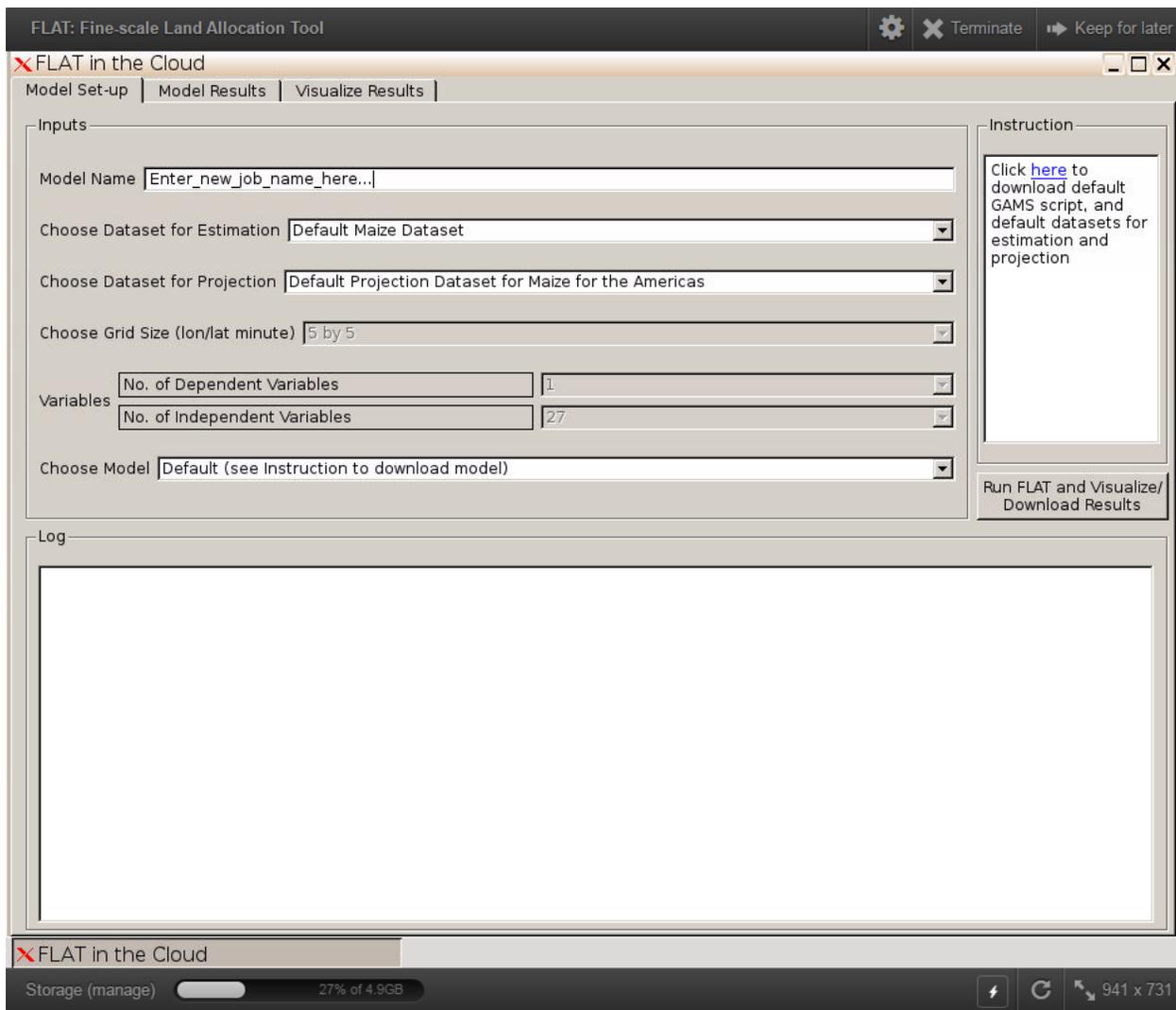


Figure 1. FLAT interface for model specification and data selection

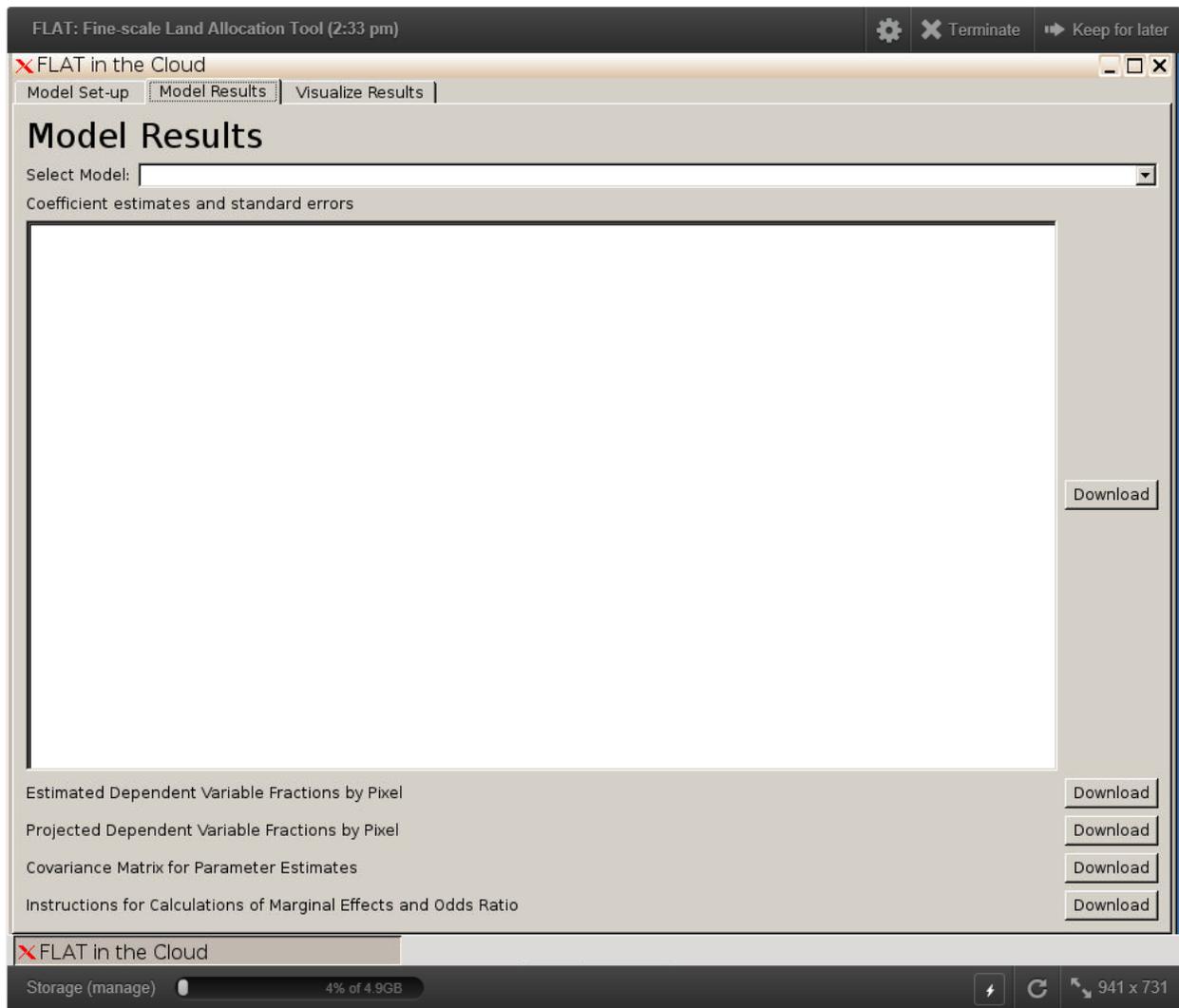


Figure 2. FLAT interface for model results display and downloading

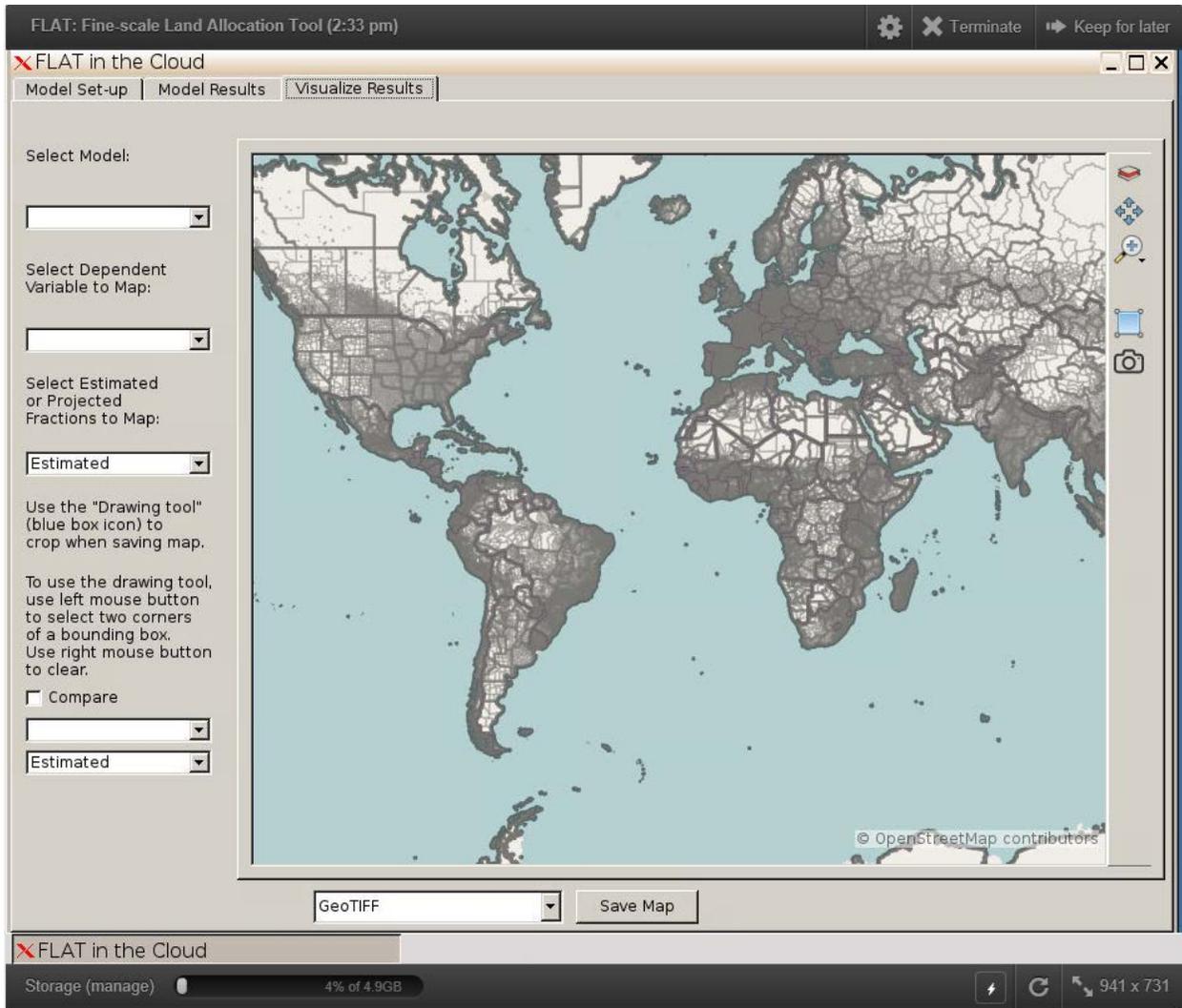


Figure 3. FLAT interface for results visualization

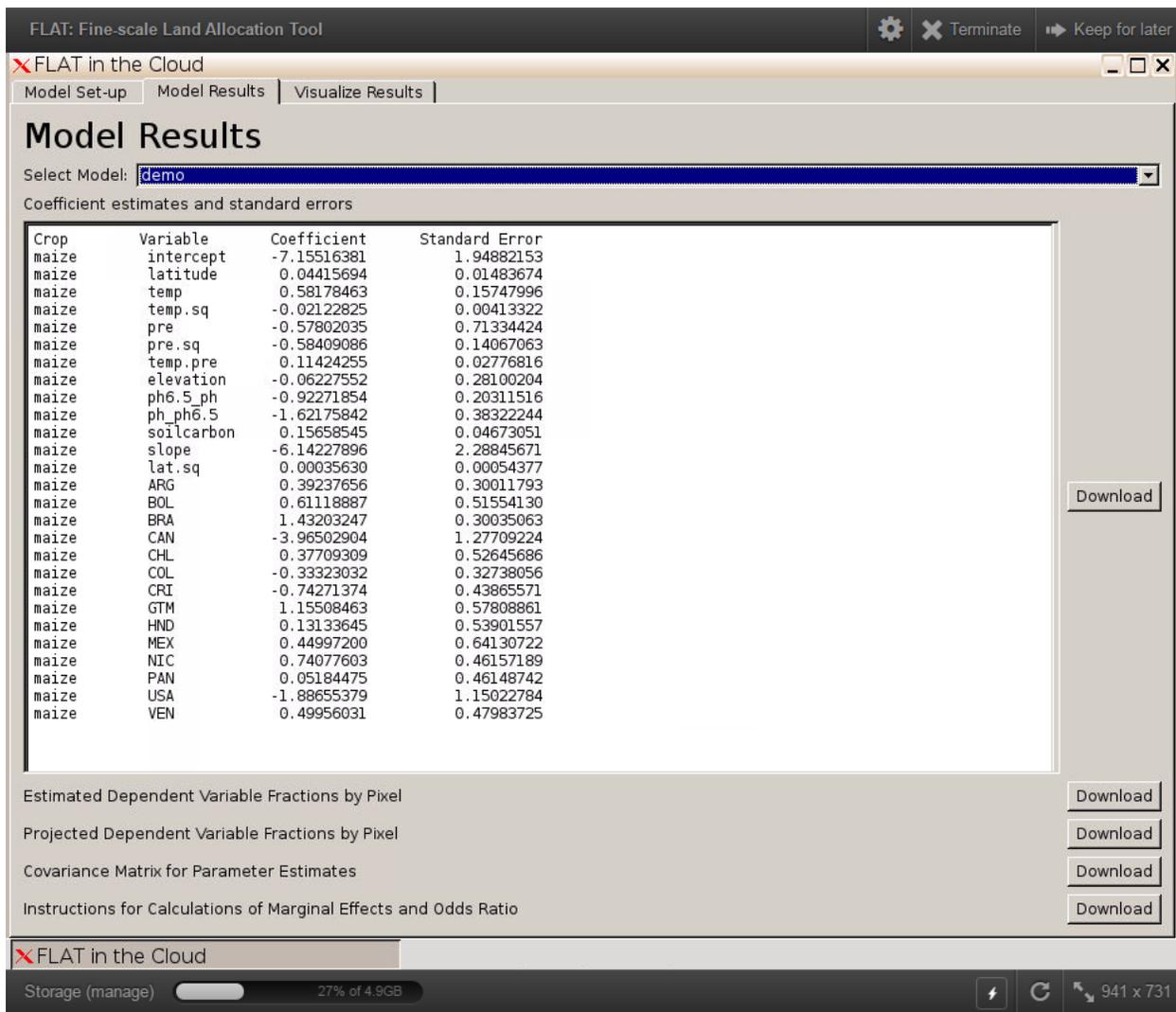


Figure 4. Model results for the default maize model

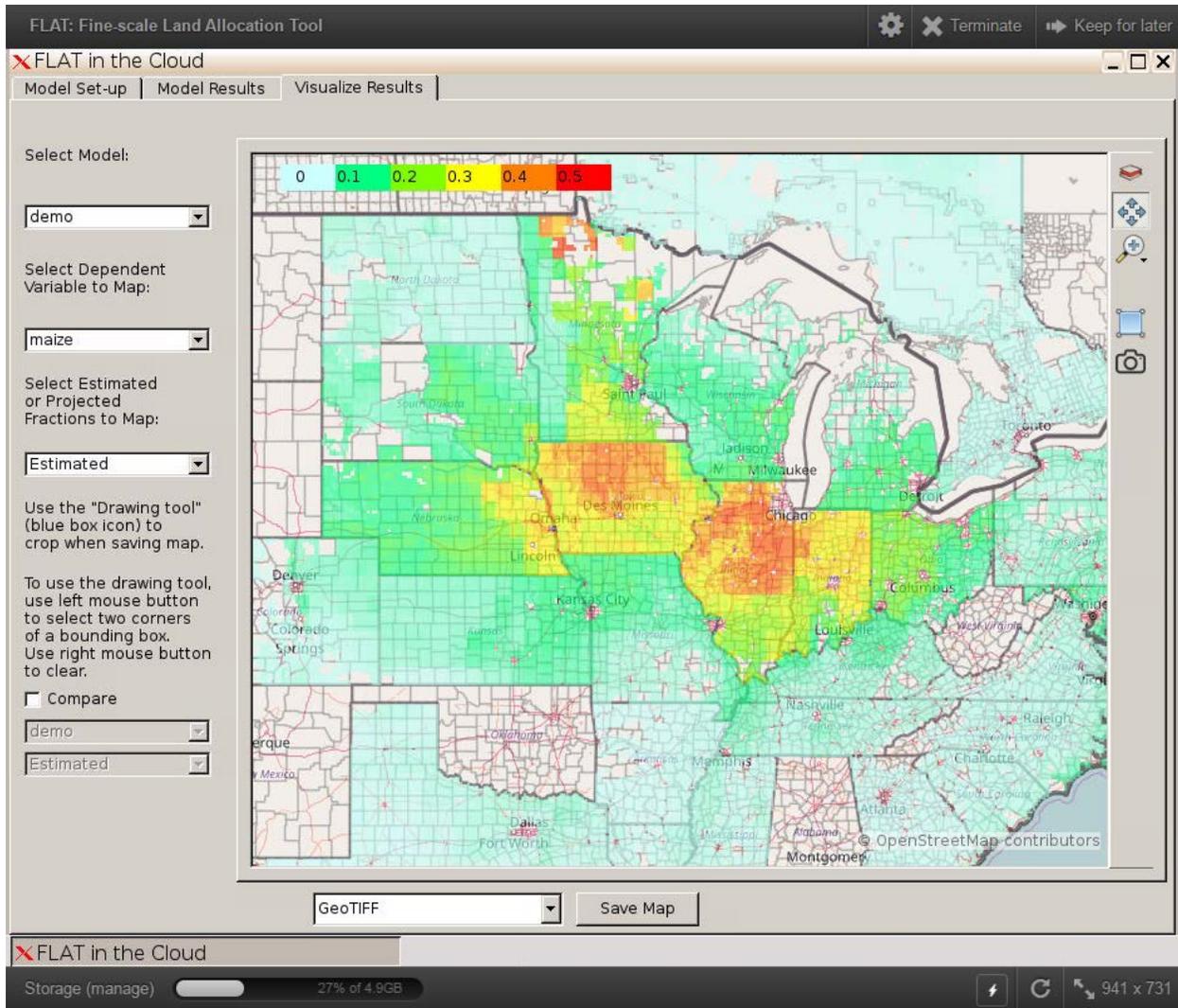


Figure 5. Visualization results for the default maize model