Development of high performance computing tools for estimation of high-resolution surface energy balance products using sUAS information

Nassar, Ayman, Torres-Rua, Alfonso, Merwade, Venkatesh, Dey, Sayan, Zhao, Lan, et al.

Ayman Nassar, Alfonso Torres-Rua, Venkatesh Merwade, Sayan Dey, Lan Zhao, I. Luk Kim, William P. Kustas, Hector Nieto, Lawrence Hipps, Rui Gao, Joseph Alfieri, John Prueger, Maria Mar Alsina, Lynn McKee, Calvin Coopmans, Luis Sanchez, Nick Dokoozlian, Nicolas Bambach Ortiz, Andrew J. Mcelrone, "Development of high performance computing tools for estimation of high-resolution surface energy balance products using sUAS information," Proc. SPIE 11747, Autonomous Air and Ground Sensing Systems for Agricultural Optimization and Phenotyping VI, 117470K (12 April 2021); doi: 10.1117/12.2587763

Event: SPIE Defense + Commercial Sensing, 2021, Online Only

Development of High Performance Computing Tools for Estimation of High-Resolution Surface Energy Balance Products Using sUAS Information

Ayman Nassar[1,2], Alfonso Torres[1,2], Venkatesh Merwade[3], Sayan Dey[3], Lan Zhao[3], I Luk Kim[3], William P Kustas[4], Hector Nieto[5], Lawrence Hipps[6], Rui Gao[1,2], Joseph Alfieri[4], John Prueger[7], Maria Mar Alsina[8], Lynn McKee[4], Calvin Coopmans[9], Luis Sanchez[8], Nick Dokoozlian[8], Nicolas Bambach Ortiz[10], Andrew J Mcelrone[10]

(1)Utah State University, Department of Civil and Environmental Engineering, Logan, UT, United States, (2)Utah Water Research Lab, Utah State University, (3) Purdue University, Lyles School of Civil Engineering, Indiana, United States, (4) U. S. Department of Agriculture, Agricultural Research Service, Hydrology and Remote Sensing Laboratory, Beltsville, MD, United States, (5)Complutum Tecnologías de la Información Geográfica, Madrid, Spain, (6) Plants, Soils and Climate Department, Logan, Utah State University, UT 84322, USA, (7) U. S. Department of Agriculture, Agricultural Research Service, National Laboratory for Agriculture and the Environment, Ames, IA, United States, (8) E & J Gallo Winery Viticulture Research, Modesto, CA, United States, (9) Utah State University, Electrical Engineering, Logan, UT, United States, (10) University of California, Davis

1. ABSTRACT

sUAS (small-Unmanned Aircraft System) and advanced surface energy balance models allow detailed assessment and monitoring (at plant scale) of different (agricultural, urban, and natural) environments. Significant progress has been made in the understanding and modeling of atmosphere-plant-soil interactions and numerical quantification of the internal processes at plant scale. Similarly, progress has been made in ground truth information comparison and validation models. An example of this progress is the application of *sUAS* information using the Two-Source Surface Energy Balance (*TSEB*) model in commercial vineyards by the Grape Remote sensing Atmospheric Profile and Evapotranspiration eXperiment - *GRAPEX* Project in California. With advances in frequent *sUAS* data collection for larger areas, *sUAS* information processing becomes computationally expensive on local computers. Additionally, fragmentation of different models and tools necessary to process the data and validate the results is a limiting factor. For example, in the referred *GRAPEX* project, commercial software (*ArcGIS* and *MS* Excel) and Python and Matlab code are needed to complete the analysis. There is a need to assess and integrate research conducted with *sUAS* and surface energy balance models in a sharing platform to be easily migrated to high performance computing (*HPC*) resources. This research, sponsored by the National Science Foundation *FAIR* Cyber Training Fellowships, is integrating disparate software and code under a unified language (Python). The Python code for estimating the surface energy fluxes using TSEB2T model as well as the EC footprint analysis code for ground truth information comparison were hosted in myGeoHub site https://mygeohub.org/ to be reproducible and replicable.

Keywords: surface energy balance, cyberinfrastructure, remote sensing, sUAS, myGeoHub, HPC, Python, TSEB2T, FAIR

2. INTRODUCTION

Evapotranspiration (ET) is a key component for hydrology, agricultural water management[8], and better water resources allocation for ecosystem. For best water resources management practices, accurate estimation of ET is essential to understand the interactions between water and energy cycles[15], droughts [1], climate change[13] and plant growth. ET is

Autonomous Air and Ground Sensing Systems for Agricultural Optimization and Phenotyping VI,
edited by J. Alex Thomasson, Alfonso F. Torres-Rua, Proc. of SPIE Vol. 11747, 117470K

considered the highest component in the hydrologic cycle, constitutes 70% of precipitation on land to the atmosphere[10]. For agricultural water management, ET is the key indicator for quantifying crop water demand and vegetation stress. There are several approaches have been used for quantifying the actual ET such as scintillometers [5], lysimeters[16], and eddy covariance flux towers[14]; however, these methods are limited to small sampling area under assumption of surface homogeneity[7], which is less likely to be met in reality.

Recently, the advent of remote sensing technique with wide range of platforms allow us to produce spatial ET information at different resolutions spanning from sub-meters to kilometers. Spatial information can be acquired by satellites, manned aircraft, and small unnamed aerial systems (sUAS)[6] which then grounded in the theory behind the surface energy balance (SEB) models[9] such as the Two-Source Energy Balance. The inputs for these models include the micrometeorological information (wind speed, air temperature, water vapor pressure, and incoming solar radiation) and other information related to the vegetation cover and land surface temperature which derived from remotely-sensed data. The development of sUAS platforms and various sensors associated with advanced SEB models (e.g., TSEB[12]) nowadays are an example for the progress made in remote sensing being used for estimating ET in complex agricultural environments such as vineyards. Despite the fact of progress made in the understanding and modeling of atmosphere-plant-soil interactions and numerical quantification of the internal processes at plant scale, as well as ground truth information comparison and validation, more work on sUAS data processing is still necessary.

From a general perspective, sUAS information processing becomes computationally expensive on local computers. This requires a more powerful computational platform in order to efficiently derive the value of those data. The advent of high performance computing recent years become very helpful to integrate resources from different locations and analyze real-time big data. In this research effort, sUAS and surface energy balance models are being integrated with a sharing platform namely myGeoHub https://mygeohub.org/ to be easily migrated to the HPC resources to parallelize, streamline and enable seamless integration of modeling components.

2. METHODOLOGY

1.1 Model Overview

In this study, the Two Source energy Balance (TSEB) model was used to calculate surface energy fluxes. TSEB was originally developed by Norman et al, 1995[12] and undergone several revisions to partition the radiative and turbulent energy fluxes between soil and canopy. In this case, net radiation (R_n) and sensible heat flux (H) are partitioned between soil/substrate and canopy. There are several versions of TSEB model including TSEB-PT (Priestley Taylor), TSEB-DTD (Dual Time difference), TSEB-2T (Dual Temperature), and TSEB-2T-DMS (Data-mining sharpening of temperature). TSEB-PT model assumes a composite radiometric temperature (T_{rad}) from the soil/substrate and canopy. The decomposition of radiometric temperature (T_{rad}) between soil/substrate and canopy is based on the fractional cover (f_c). TSEB-DTD is similar to TSEB-PT model with further development by decomposing the T_{rad} into soil/substrate temperature (T_s) and canopy temperature (T_c) using two observations of T_{rad}. The first observation acquired 1.5 h after the sunrise ($T_{rad,0}$), while the second obtained during the daytime ($T_{rad,1}$). TSEB-2T-DMS model is another version of TSEB that partitions T_s and T_c using a data-mining fusion algorithm to sharpen the original LST to be similar to the optical data. This would allow a better discrimination between T_s and T_c. The TSEB2T model is a contextual TSEB that estimates T_s and T_c from LST imagery based on the relationship between vegetation index (VI) and LST, particularly LST-NDVI, to calculate T_s and T_c within a specific spatial domain. As shown in Figure (1), the TSEB2T [11] model separates R_n, H, and LE between vegetation and soil. The equations below describe the mathematical expressions behind the TSEB approach

$$R_n = LE + H + G, \quad (1)$$

$$R_{nc} = H_c + LE_c, \quad (2)$$

$$R_{ns} = H_s + LE_s + G, \quad (3)$$

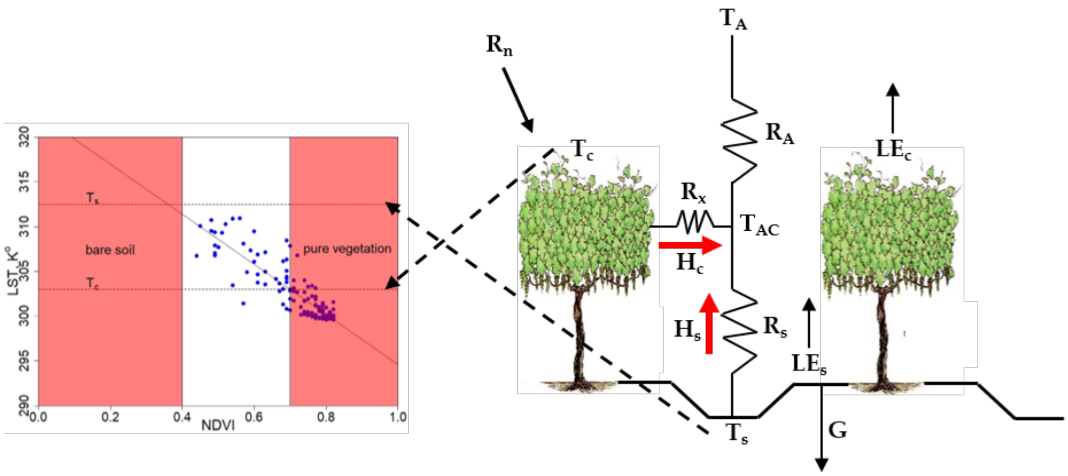


Figure 1. Schematic representation of TSEB2T model

where R_n is the net radiation, H is the sensible heat flux, LE is the latent heat flux, and G is the soil heat flux. T_s and T_c are the soil and canopy temperature, respectively, derived from the LST and high enough resolution of optical data. Subscripts of c and s represent the canopy and soil components, respectively. LE_c and LE_s are solved as residuals when (T_c and T_s) observations are available.

2.2 EC footprint model

Eddy Covariance (EC) footprint models are used for describing the position and size of surface source areas, as well as the relative contribution of passive scalar sources to measured fluxes[3]. The EC flux footprint is defined as a mathematical expression used to transfer between sources and sinks of passive scalars at the surface, Q_c, and the turbulent flux, F_c. There are several parameters influencing the EC footprint estimation which include atmospheric stability, receptor height, and surface roughness, all of which strongly affect the size of the footprint. The mathematical function used to describe the flux footprint is shown in Equation (4)

$$F_c(0,0,z_m) = \int_{\Re} Q_c(x,y) f(x,y) dxdy \qquad (4)$$

where F_c is a flux density (per unit area), $Q_c(x,y)$ is the as source or sink at the surface. Because the footprint function is always estimated at a specific measurement height (receptor height), the vertical reference in f is neglected. From a single unit point source or sink,Q_u, Equation (4) can be simplified as follows

$$f(x,y) = \frac{F_c(0,0,z_m)}{Q_u(x,y)} \qquad (5)$$

In this research, Kljun et al, 2002 [4] model has been considered for footprint analysis. This model uses the three-dimensional Lagrangian stochastic footprint model LPDM-B and found to satisfy the well-mixed condition continuously for convective to stable stratifications and for measurement heights (receptors) within or above the surface layer. Assuming that crosswind turbulent dispersion can be treated independently from vertical/streamwise transport, the mathematical expression for EC can be expressed in terms of a crosswind-integrated footprint, $\overline{f^y}$ and a cross-dispersion function, D_y

$$f(x,y) = \overline{f^y}(x) D_y \qquad (6)$$

$$f(x,y) = \overline{f^y} \frac{1}{\sqrt{2\pi}\sigma_y} \exp\left(-\frac{y^2}{2{\sigma_y}^2}\right) \qquad (7)$$

More details about the derivation of the footprint model and model parameters can be found in Kljun et al. 2015[3]

3. RESULTS

3.1Taxonomy for reusability in environmental modelling

TSEB2T was developed by Norman et al, 1995[12] and further developed by Nieto *et al.*, 2019[11] and hosted in myGeoHub to use the model in conjunction with remotely sensed information and other micrometeorological data to estimate ET and compare against the actual ground data obtained from EC towers installed in the field. The myGeoHub was chosen for several reasons. First, in general, web application can provide users easy access to the data, tools and simulation resources across different locations. Second, myGeoHub was designed to support geospatial modeling[2] such as remote sensing models that deal with big data and require a powerful computational platform for analyzing data in a record time. Third, this hub now is hosting several datasets, groups, and training materials that help users to have a wide range of educational resources. This workflow enables researchers to focus more on the scientific research rather than dealing with software replication and data migration.

The workflow resulting from this research work organizes the reusability spectrum into four levels: findable, accessible, interoperable, and reusable. These levels represent a progression started with the base level, findability, first step to achieve, followed by accessibility, interoperability, and finally reusability.

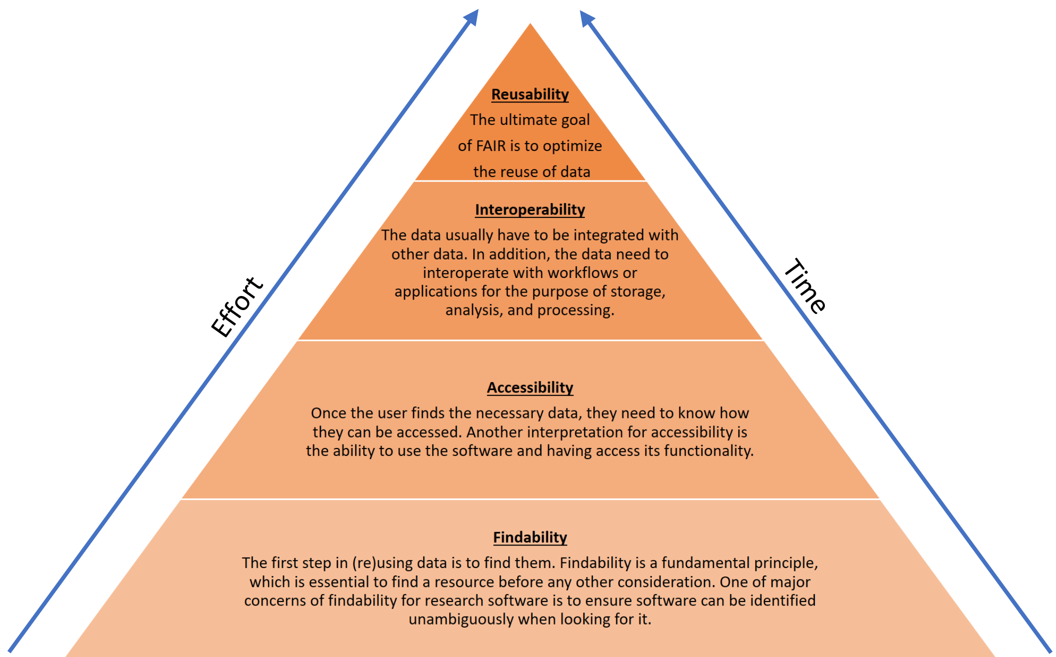


Figure 2. The reusability taxonomy for complex computational studies comprising a progression that requires increased effort and time from findability, through accessibility, interoperability and reusability.

FAIR principles were used in this study to improve the reuse of sUAS data by making it more findable, accessible, interoperable, and reusable by users and machines. This study effort also helps researchers to demonstrate the impact of their work by enabling the reuse of the data, and can foster future and broader collaboration. Using FAIR principles in TSEB2T allow users also to run the model for multiple sUAS images at the same time as well as validate the results by using the ground truth measurements from EC towers.

3.2 Separation of soil and canopy temperatures (T_s and T_c)

Figure (3) shows the Python code deployed in myGeoHub for separating the soil/substrate and canopy temperatures (T_s and T_c) using the relationship between the normalized difference vegetation index (NDVI) and the composite land surface temperature (LST). First, T_s and T_c are calculated by taking the average value of pixels that are pure soil/substrate and pure canopy within a model grid size. For detecting NDVI threshold values of soil, a relationship between Leaf area index (LAI) and NDVI is constructed where the threshold value is identified when LAI is nearly zero. In case of very dense vegetation (pure soil pixels not exist) or sparse vegetation lacking pure vegetation pixels inside the contextual spatial domain, a linear fit between LST and NDVI can by generated to estimate T_s and T_c.

```
MY GEO HUB   jupyter 2T_model_   Last Checkpoint: 10/22/2020 (unsaved changes)

In [5]: ## import the required packages and libaries for the code
import os
import gdal
import numpy as np
import scipy
from scipy.stats import linregress
import pandas as pd

In [4]: # data input include the LST, NDVI, input resolution, output resolution
ndvi_file="/home/mygeohub/aymnassar/Ayman_python_code/NDVI_20140809_0_15.tif"
lst_file="/home/mygeohub/aymnassar/Ayman_python_code/NDVI_20140809_0_15.tif"
lai_file="/home/mygeohub/aymnassar/Ayman_python_code/NDVI_20140809_0_15.tif"

in_res=0.6 # the resolution of input imagery data
out_res=3.6 # the resolution of output imagery data
ndvi_veg=0.746908  # identify the threshold value of NDVI for vegetation
ndvi_soil=0.366844941   # identify the threshold value of NDVI for soil
```

Figure 3. Python code for separating T_c and T_s.

3.3 EC footprint estimation

Figure (4) shows a screenshot of Python code in myGeoHub to estimate the EC footprint using the 2D flux model developed by Kljun et al. 2015[3]. The fetch shape and orientation of the footprint depend on the micro-meteorological conditions at the site measured by the EC towers. Those measurements involve friction velocity, turbulence fluxes, and wind speed, which affect atmospheric stability, and canopy and EC measurement height. Figure (5) represents the EC footprints at different times.

```
MY GEO HUB   jupyter Footprint_Kjlun_model Last Checkpoint: Last Tuesday at 4:18 PM (unsaved changes)

In [13]: zm_s = 5 #Measurement height [m]
h_c = temp_df['Hc'] #Height of canopy [m]
h_s = 3000. #Height of boundary layer [m]
dx =round(3.60,1) ## the spatial resolution of the footprint
origin_d = 500 #Model distance from origin [m]

# air temperature
Ta=temp_df['Ta']

# Calculate MO lenght
Cp=1005.0
rho=1.3079-0.0045*Ta
#L=CalcL (temp_df['ustar'], Ta+273.15, rho, Cp, temp_df['H'], temp_df['LE'])
L=temp_df['Obukov']
## calculate sigmav
sigmav=CalcSigma_v(temp_df['ustar'],L)


#Convert local lat/lon to local UTM (Lodi: UTM 10).
x_coor=round(temp_df['X'],6)
y_coor=round(temp_df['Y'],6)
station_coord = (x_coor,y_coor) #Lodi lat/lon in WGS-84
in_proj = proj.Proj(init='EPSG:4326')
out_proj = proj.Proj(init='EPSG:32610')
(station_x,station_y) = proj.transform(in_proj,out_proj,*station_coord)

new_dat = None
```

Figure 4. Python code for the EC footprint analysis.

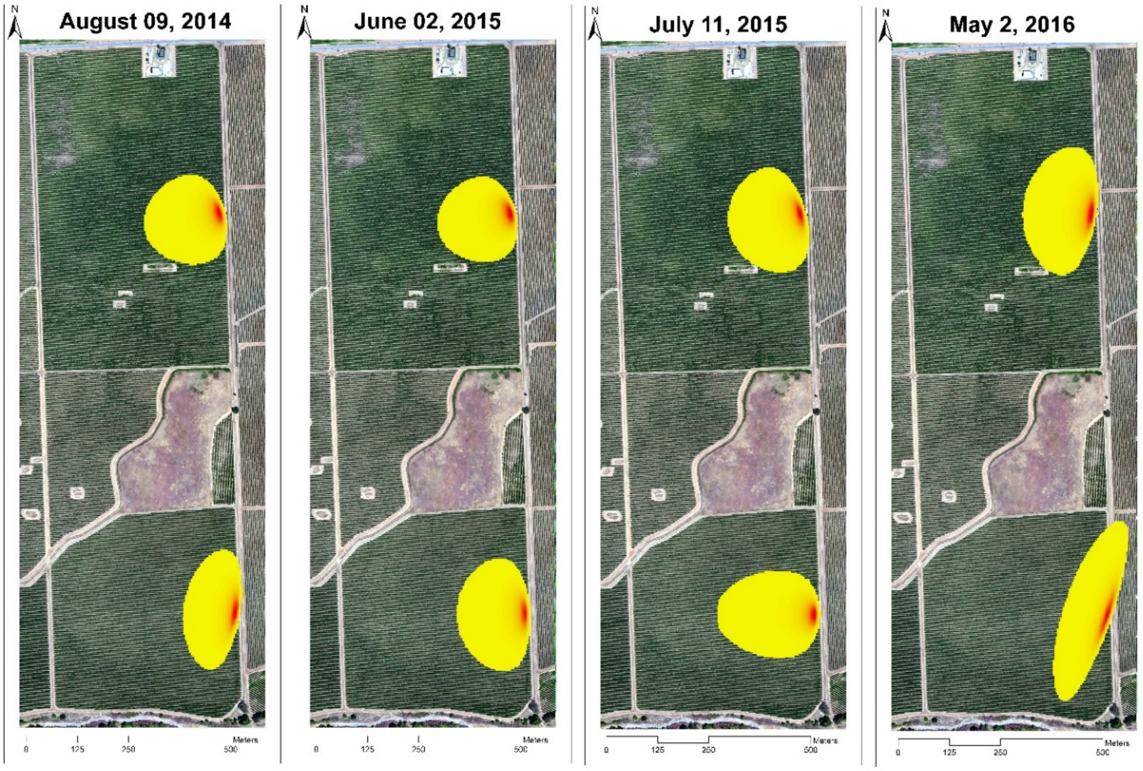
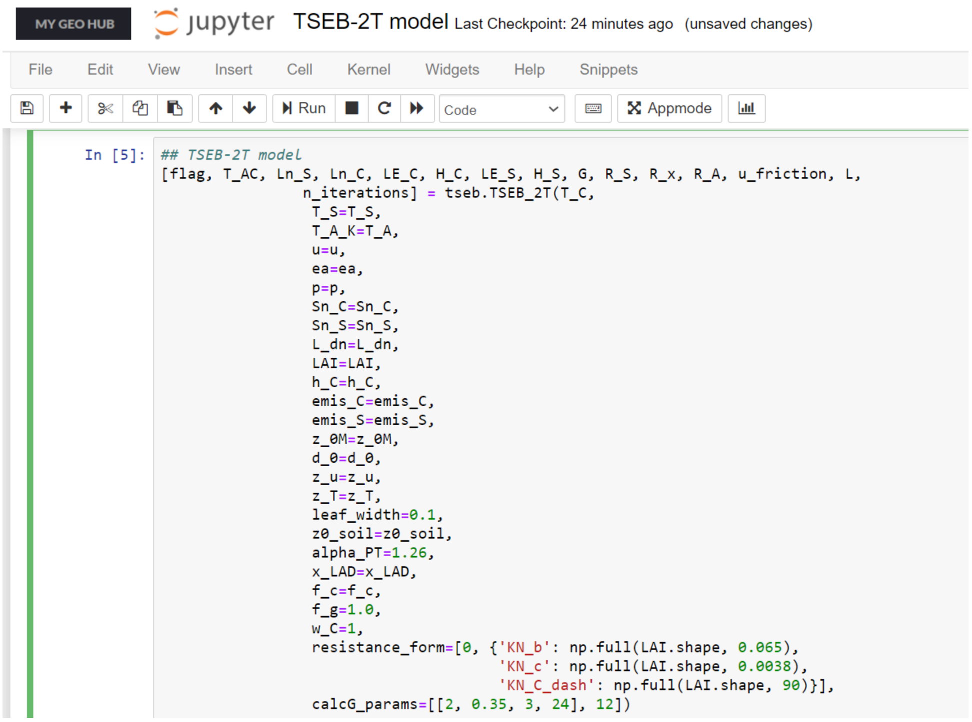


Figure 5. Layout of EC footprints for two towers at different times at Lodi vineyard.

3.4 Surface energy fluxes calculations

Figure (6) shows the Python code for calculating the surface energy fluxes using TSEB2T model. The key inputs for the model include canopy and soil temperatures (T_c and T_s), fractional cover (F_c), Leaf Area Index (LAI), canopy width-to-canopy height ratio (w_C/h_C), Canopy height (h_C), as well as other inputs related to the micrometeorological data. The main model outputs are the energy fluxes (R_n, H, LE, and G) beside other ancillary data.




```
## TSEB-2T model
[flag, T_AC, Ln_S, Ln_C, LE_C, H_C, LE_S, H_S, G, R_S, R_x, R_A, u_friction, L,
                n_iterations] = tseb.TSEB_2T(T_C,
                 T_S=T_S,
                 T_A_K=T_A,
                 u=u,
                 ea=ea,
                 p=p,
                 Sn_C=Sn_C,
                 Sn_S=Sn_S,
                 L_dn=L_dn,
                 LAI=LAI,
                 h_C=h_C,
                 emis_C=emis_C,
                 emis_S=emis_S,
                 z_0M=z_0M,
                 d_0=d_0,
                 z_u=z_u,
                 z_T=z_T,
                 leaf_width=0.1,
                 z0_soil=z0_soil,
                 alpha_PT=1.26,
                 x_LAD=x_LAD,
                 f_c=f_c,
                 f_g=1.0,
                 w_C=1,
                 resistance_form=[0, {'KN_b': np.full(LAI.shape, 0.065),
                                      'KN_c': np.full(LAI.shape, 0.0038),
                                      'KN_C_dash': np.full(LAI.shape, 90)}],
                 calcG_params=[[2, 0.35, 3, 24], 12])
```

Figure 6. Python code for calculating surfaces energy fluxes using TSEB-2T model.

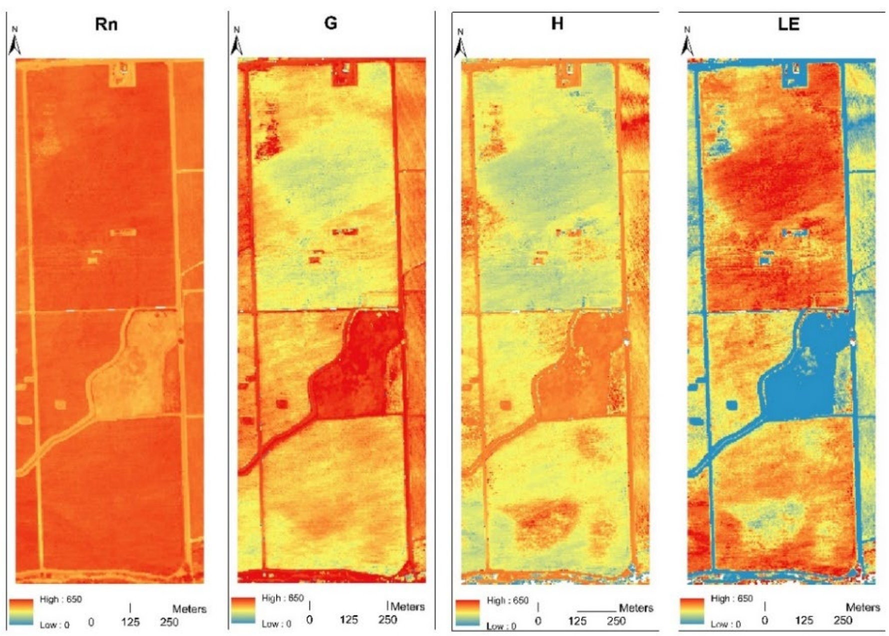


Figure 7. Energy fluxes (R_n, G, H, and LE) using TSEB2T.

4. CONCLUSION

In this study, myGeoHub as a sharing platform has been used to integrate remotely sensed data and other micrometeorological data to estimate ET using TSEB2T model and compare against the ground truth measurements obtained from the EC towers. myGeoHub supports the geospatial modeling, data analysis and visualization needs for research and education communities through hosting datasets, tools, training materials and educational contents. This implies that data can be easily findable, accessible and interoperable with workflows and applications for the purpose of reusability. Python code of the TSEB2T has been deployed in myGeoHub to facilitate the reproducibility and replicability of the model by users. Moreover, the EC footprint analysis code was hosted in the hub which can replicated to generate several footprints/source area using the micrometeorological data obtained from the EC towers.

5. ACKNOWLEDGEMENTS

This study was conducted thanks to the support of the NASA Grant NNX17AF51G, and the Utah Water Research Laboratory Student Fellowship. The authors are also grateful for the extraordinary support from the Utah State University AggieAir UAV and E&J Gallo Scientific Teams, for the data collection support and analysis. We would like also to thank the NSF grant (Award Abstract # 1829764)

6. REFERENCES

[1] Anderson, M. C. *et al.* (2011) 'Evaluation of drought indices based on Thermal remote sensing of evapotranspiration over the continental United States', *Journal of Climate*, 24(8), pp. 2025–2044. doi: 10.1175/2010JCLI3812.1.

[2] Kalyanam, R. *et al.* (2019) 'MyGeoHub—A sustainable and evolving geospatial science gateway', *Future Generation Computer Systems*, 94, pp. 820–832. doi: 10.1016/j.future.2018.02.005.

[3] Kljun, N. *et al.* (2015) 'A simple two-dimensional parameterisation for Flux Footprint Prediction (FFP)', *Geoscientific Model Development*, 8(11), pp. 3695–3713. doi: 10.5194/gmd-8-3695-2015.

[4] Kljun, N., Rotach, M. W. and Schmid, H. P. (2002) 'A three-dimensional backward Lagrangian footprint model for a wide range of boundary-layer stratifications', *Boundary-Layer Meteorology*, 103(2), pp. 205–226. doi: 10.1023/A:1014556300021.

[5] Meijninger, W. M. L. *et al.* (2002) 'Determination of area-averaged water vapour fluxes with large aperture and radio wave scintillometers over a heterogeneous surface - Flevoland field experiment', *Boundary-Layer Meteorology*, 105(1), pp. 63–83. doi: 10.1023/A:1019683616097.

[6] Nassar, A. *et al.* (2019) 'Assessment of High-Resolution Daily Evapotranspiration Models Using Instantaneous sUAS ET in Grapevine Vineyards', *AGUFM*, 2019, pp. B31N-2396. Available at: https://ui.adsabs.harvard.edu/abs/2019AGUFM.B31N2396N/abstract (Accessed: 4 April 2021).

[7] Nassar, A., Torres-Rua, Alfonso F., *et al.* (2020) 'Implications of soil and canopy temperature uncertainty in the estimation of surface energy fluxes using TSEB2T and high-resolution imagery in commercial vineyards', in Thomasson, J. A. and Torres-Rua, A. F. (eds) *Autonomous Air and Ground Sensing Systems for Agricultural Optimization and Phenotyping V*. SPIE-Intl Soc Optical Eng, p. 14. doi: 10.1117/12.2558715.

[8] Nassar, A., Torres-Rua, A., *et al.* (2020) 'Influence of Model Grid Size on the Estimation of Surface Fluxes Using the Two Source Energy Balance Model and sUAS Imagery in Vineyards', *Remote Sensing*, 12(3), p. 342. doi: 10.3390/rs12030342.

[9] Nassar, A., Torres-Rua, Alfonso F, *et al.* (2020) 'Influence of Spatial Heterogeneity in Evapotranspiration Modeling at Natural Areas Using sUAS High Resolution Data'. AGU. Available at: https://agu.confex.com/agu/fm20/meetingapp.cgi/Paper/716416 (Accessed: 4 April 2021).

[10] *NATIONAL WATER SUMMARY 1987 Hydrologic Events and Water Supply and Use* (no date).

[11] Nieto, H. *et al.* (2019) 'Evaluation of TSEB turbulent fluxes using different methods for the retrieval of soil and canopy component temperatures from UAV thermal and multispectral imagery', *Irrigation Science*, 37(3), pp. 389–406. doi: 10.1007/s00271-018-0585-9.

[12] Norman, J. M., Kustas, W. P. and Humes, K. S. (1995) 'Source approach for estimating soil and vegetation energy fluxes in observations of directional radiometric surface temperature', *Agricultural and Forest Meteorology*, 77(3–4), pp. 263–293. doi: 10.1016/0168-1923(95)02265-Y.

[13] Sun, Z. *et al.* (2016) 'Relationship between Evapotranspiration and land surface temperature under energy- and water-limited conditions in dry and cold climates', *Advances in Meteorology*, 2016. doi: 10.1155/2016/1835487.

[14] Swinbank, W. C. (1951) 'THE MEASUREMENT OF VERTICAL TRANSFER OF HEAT AND WATER VAPOR BY EDDIES IN THE LOWER ATMOSPHERE', *Journal of Meteorology*, 8(3), pp. 135–145. doi: 10.1175/1520-0469(1951)008<0135:tmovto>2.0.co;2.

[15] Talsma, C. *et al.* (2018) 'Sensitivity of Evapotranspiration Components in Remote Sensing-Based Models', *Remote Sensing*, 10(10), p. 1601. doi: 10.3390/rs10101601.

[16] Yang, J., Li, B. and Shiping, L. (2000) 'A large weighing lysimeter for evapotranspiration and soil-water-groundwater exchange studies', *Hydrological Processes*, 14(10), pp. 1887–1897. doi: 10.1002/1099-1085(200007)14:10<1887::AID-HYP69>3.0.CO;2-B.