Uncertainty in Satellite Remote Sensing Derived Evapotranspiration Estimation
Current Status and Assessment Methods

Tran, Bich; Van Der Kwast, Johannes; Mul, Marloes; Seyoum, Solomon; Uijlenhoet, Remko; Jewitt, Graham

DOI
10.3850/IAHR-39WC2521711920221782

Publication date
2022

Document Version
Final published version

Published in
Proceedings of the 39th IAHR World Congress (Granada, 2022)

39th IAHR WORLD CONGRESS GRANADA, SPAIN 2022 From Snow to Sea 19-24 June 2022

Proceedings of the 39th IAHR World Congress
19–24 June 2022, Granada, Spain
doi://10.3850/IAHR-39WC2521716X20221782

Uncertainty in Satellite Remote Sensing Derived Evapotranspiration Estimation: Current Status and Assessment Methods

Bich Tran[(1)], Johannes Van Der Kwast[(2)], Marloes Mul[(3)], Solomon Seyoum[(4)], Remko Uijlenhoet[(5)], Graham Jewitt[(6)]

[(1,2,3,4,6)] IHE Delft Institute for Water Education, Delft, The Netherlands,
e-mail [(1)] b.tran@un-ihe.org [(2)] h.vanderkwast@un-ihe.org [(3)] m.mul@un-ihe.org [(4)] s.seyoum@un-ihe.org
[(1,5,6)] Delft University of Technology, Delft, The Netherlands,
email [(6)] r.uijlenhoet@tudelft.nl.

Abstract

Evapotranspiration (ET), a key variable in both water and energy cycles. It is very challenging to measure or estimate in large regions. Among many approaches to estimate ET indirectly (e.g. through hydrological modelling), models that are based on satellite remote sensing data (RS) are increasingly being used. However, the RS-based models inherit uncertainty from many sources, such as the model's algorithm and parameters, input satellite data, and processing techniques. It is challenging to assess this uncertainty due to limitations of validation data, high volume and high dimensionality of RS data. Many studies have evaluated uncertainty in RS-based estimation of ET using different methods and reference data. The suitability of methods and reference data subsequently affect the validity of these evaluations. Therefore, it is necessary to have an overview of different evaluation methods and their uses. This study aimed to systematically review original research papers that assessed uncertainty or accuracy of RS-ET model or data products. We categorized these papers and quantified based on (i) spatial and temporal scale of ET estimation, (ii) types of uncertainty, and (iii) methods used to assess uncertainty. Studies have been geographically concentrated in North Asia, North America, and Europe. Most studies used the validation method, which quantifies the discrepancy between pixel-based ET estimation with an in-situ estimation. Although a standardized validation approach for satellite-based ET estimates is not yet ready, most validation studies employed Eddy Covariance (EC) flux towers for reference estimation at field-scale. In regions where in-situ measurements are limited, many studies use the residual of the water balance as reference. However, few studies considered uncertainty in the reference estimation and mismatch of spatial and temporal scales. For monitoring agricultural fields, most RS-ET methods have been reported with high accuracy. When applying these methods to larger extent, additional assessments are required to better inform data users of the quality of RS-ET estimation. These include cross-validation, sensitivity, and uncertainty analyses. Overall, this review showed the progress in evapotranspiration estimation using satellite data in terms of uncertainty assessment.

Keywords: Evapotranspiration; Remote sensing; Uncertainty; Surface energy balance models; Satellite data

1. INTRODUCTION

Evapotranspiration (ET) is the key variable linking the water, energy and carbon cycles of our planet. In the water cycle, it is the second-largest flux after precipitation (Korzoun et al., 1978). In agriculture, ET is the amount of water removed by crops from the soil into the atmosphere and, thus, criti-cal for water conservation. ET is a complex phenomenon that depends on many factors: the atmospheric and vegetation conditions (demand), the availability of water in soil, water bodies, and canopy (supply), surface roughness, etc. The complexity in measuring ET directly makes it difficult and expensive to routinely measure and capture the spatial variation of ET. This requires a dense network of ground-based gauging stations. So far, the most successful network of ground-based stations is the Eddy Covariance flux towers network (FluxNet) (Baldocchi et al., 2001). However, most FluxNet stations are located in North America, Europe, and North Asia. Meanwhile, ground-based data is particularly limited in many regions that need to monitor ET, such as Africa and Middle East.

Since ET is very challenging to measure in a large area using ground-based instruments, there has been an increasing interest to estimate ET using satellite remote sensing (RS) observations. As ET cannot be directly measured from space, retrieval algorithms or models have been developed to estimate ET. These models ET from optical and/or thermal remote sensing (RS) data, to name a few SEBS (Su, 2002), TSEB (Kustas and Norman, 1999), SEBAL (Bastiaanssen et al., 1998), METRIC (Allen et al., 2007), and ALEXI (Anderson et al., 2011). They also require ancillary data (such as soil characteristics, land-cover, etc.) and are sometimes combined with in situ data. Different algorithms, parameterization, input RS data sources and processing levels result in a wide range of ET estimates (Long et al., 2014, Chen et al., 2014). While there have been many studies evaluating the different ET models and products, there has not been a single model that performs the best in all situations. Moreover, the methods used to aggregate instantaneous RS observations also cause uncertainty in daily or longer-period estimates, which is needed for water management practices (Liu, 2021).

Driven by community needs, some projects have provided platforms to increase access to various remote sensing-based ET (RS-ET) data products, such as the FAO's WaPOR (FAO, 2018) and OpenET (Melton et al., 2021). With these widely accessible RS-ET data, it is important that data users are informed about their uncertainty or accuracy. Uncertainty has been evaluated and reported in many studies using different assessment methods and reference data. The extent and reference data of these studies are limited and their methods are inconsistent. Therefore, it is still difficult to draw a general conclusion about the uncertainty in RS-based ET models and data products. It is important to have an overview of different uncertainty evaluation methods and how they were used to evaluate RS-based ET data.

This study aimed to systematically review original research papers that assessed uncertainty or accuracy of RS-ET model or data products. We used the systematic quantitative approach, which selects research papers with a set of criteria. We categorized these papers and quantified based on (i) spatial and temporal scale of ET estimation, (ii) types of uncertainty, and (iii) methods used to assess uncertainty. Add 1 or 2 sentences about 'added values' of this review. Based on the quantity of papers in each category, we will address the following research questions:

- What are the conventional and emerging methods used to assess the uncertainty in RS-based ET data in the literature?
- In which contexts can the uncertainty be assessed with these methods?

2. METHODOLOGY

A systematic quantitative literature review (SQLR) was performed following the approach described by Pickering and Byrne (2014). This approach was selected to have more insight in the effort to estimate ET using RS and assess the uncertainty in estimation. Instead of focusing on only papers that use a specific metric or a specific method, this systematic review approach selects all relevant research papers, labels them into groups, and analyses the pattern in the body of literature. Potentially, many groups of research papers that use models and metrics compatible for meta-analysis can be identified.

2.1 Identification and Selection of the Literature

Queries were searched on the academic electronic databases Web of Science (WoS) and Scopus (last access: 21.9.2021). The search query was a combination of the three search terms: "evapotranspiration", "remote sensing", and "uncertainty", with their common variants used in literature (Table 1).

Table 1. Search terms and variants. Search terms were combined using AND operator and synonyms were combined using OR operator. The asterisk * was used to included similar terms.

	Search terms < AND >		
Variants < OR >	**Evapotranspiration**	**Remote sensing**	**Uncertainty**
	Evaporation	Remotely-sensed	Accuracy
	Latent heat	Remotely sensed	Data quality
		Earth observation	Variability
		Satellite*	Reliability
		Global ** product	Evaluat*
		Global ** data*	Validat*

The search result was then refined using the available filters of Scopus and WoS. English articles (>99% of result) which are original research published in scientific peer-reviewed journals were included. Review papers and other reports such as conference proceedings or theses are not included because they follow different formats and quality was not assured by a peer-reviewing system. Duplicated records were removed.

2.2 Selection and Quantification of the Literature

The title and abstract of each paper are screened to identify the relevant papers based on the objective and methodology of the research. The screening process was performed in ASReview software (van de Schoot et al., 2021), which incorporates a classifier that re-orders the list of records based on the previous selection of the reviewer. This allowed us to select the most relevant articles from the search result and record our selection for a transparent reviewing process. Moreover, the screening does not show authors and affiliation to avoid personal bias. The screening process stops when finding 50 irrelevant records in a row.

The inclusion criteria are that studies attempted to quantify accuracy or uncertainty of one or more remote sensing-based diagnostic estimation, models or datasets of terrestrial evapotranspiration. Even when screened with both abstract and title, fault inclusion still occurs in the abstract-title screening stage. For example, some studies use a handheld radiometer or field remote sensors or airborne sensors, but only mentioned as ‘remote sensing techniques’ in abstract and title. Without reading full-text, these papers could not be excluded. Therefore, final inclusion in the database is decided when reading full-text. Articles for which full-text is not accessible to the authors were also excluded.

From each original research paper, the items of information about the paper are recorded in a spreadsheet dataset and it will be classified into categories based on methods, subject of research, and results. In addition, bibliometrics analysis such as co-author network was conducted using VOSviewer software (Eck and Waltman, 2009).

3. RESULTS

3.1 Studies that Assess Uncertainties in RS-ET

There is an increasing trend in number of studies that have ‘Evapotranspiration’ and ‘Remote sensing’ aspects in title, abstract, and keywords (Figure 1), from 243 to 787 records per year from 2010 to 2021 from Scopus search results. However, only 50-60% of these studies mentioned keywords of ‘Uncertainty’ aspect. This percentage did not change considerably with the years. After screening titles and abstracts, we found that about 11% of these studies in Scopus search results are relevant to the objectives and scope of this review (635 in total). Research articles were published in 149 journals from diverse disciplines including hydrology, ecology, meteorology, remote sensing, and agronomy.

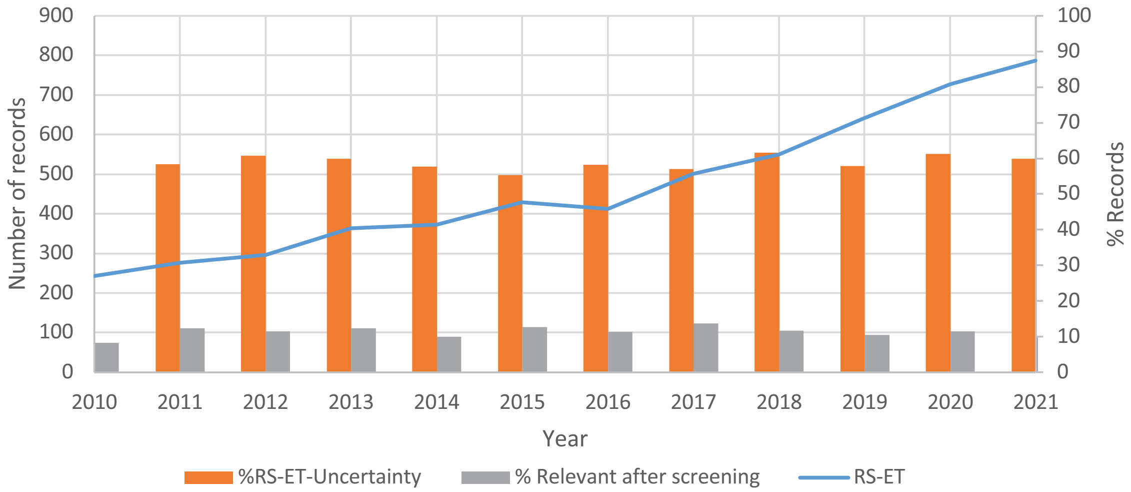

Figure 1. Number of studies on ‘Remote sensing’ and ‘Evapotranspiration’ aspects and the percentage of studies including ‘Uncertainty’ aspect from Scopus search results.

Figure 2 shows the co-author network of the relevant articles (after abstract-title screening). There are 10number of co-author clusters, defined by the number of studies they collaborated. Each cluster tends to publish more studies using a certain RS-ET model. For example, Anderson M.C and Kustas W.P. studied mainly about METRIC model, Senay G.B. studied mainly SSEB/SSEBop, Fisher J. B. studied mainly PT-JPL, Miralles D.G. studied mainly GLEAM. The cluster of Chinese co-authors have more publications in the recent years (with average publication year >2015).

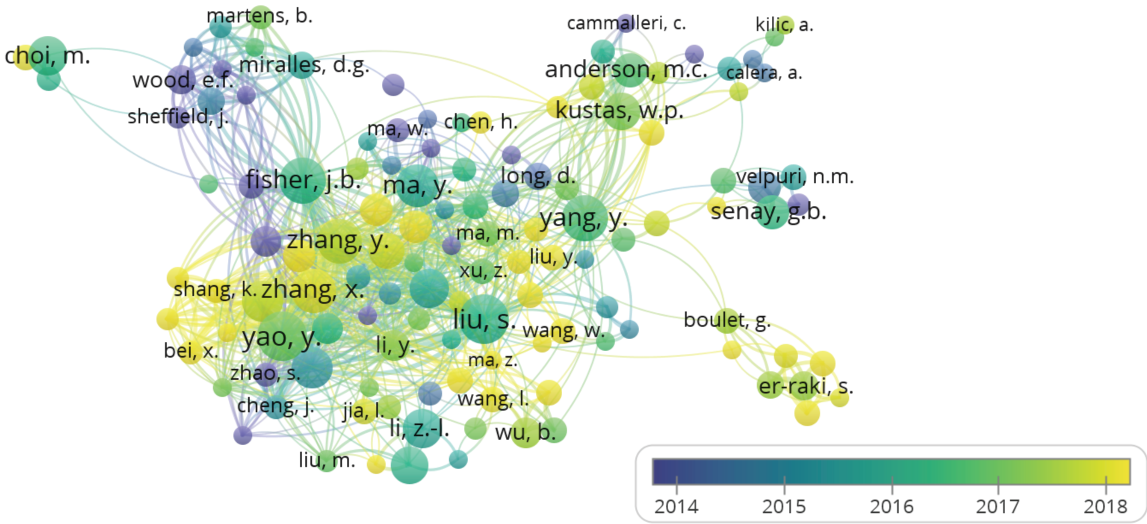

Figure 2. Co-authors network of relevant studies and average year of publication

3.2 Types of Evapotranspiration Estimates

Figure 3 shows the number of studies grouped by application of ET, temporal support, spatial support, and spatial extent. Most of the reviewed studies estimate ET at daily support and at spatial support (often same as resolution in RS) of less 500m-5km, which is the spatial resolution of MODIS. The second most used spatial support is <100m, which is the spatial resolution of Landsat series (30m). This shows the spatial support of ET estimate depends on the resolution of RS datasets available.

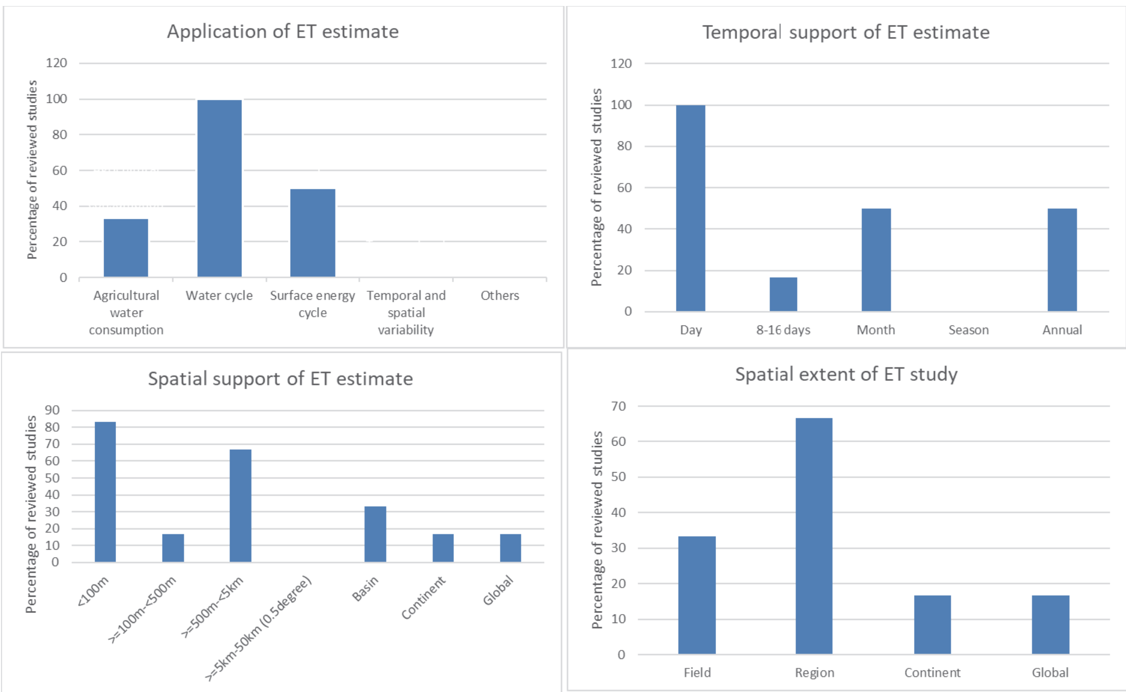

Figure 3. Distribution of reviewed studies by application, temporal support, spatial support and extent of ET estimate.

3.3 Sources of Uncertainties and Assessment Methods

Figure 4 shows the number of studies grouped by objective, uncertainty assessment approach, types of uncertainties and metrics used. Mostly, uncertainty was assessed during model implementation using validation approach. The most common validation approach was eddy covariance flux and catchment water balance. The focus of uncertainty assessment was still mainly focused on compound uncertainty, not the sources such as input data, spatial support, etc. The most used metric to report uncertainty is Root Mean Square Error (RMSE)

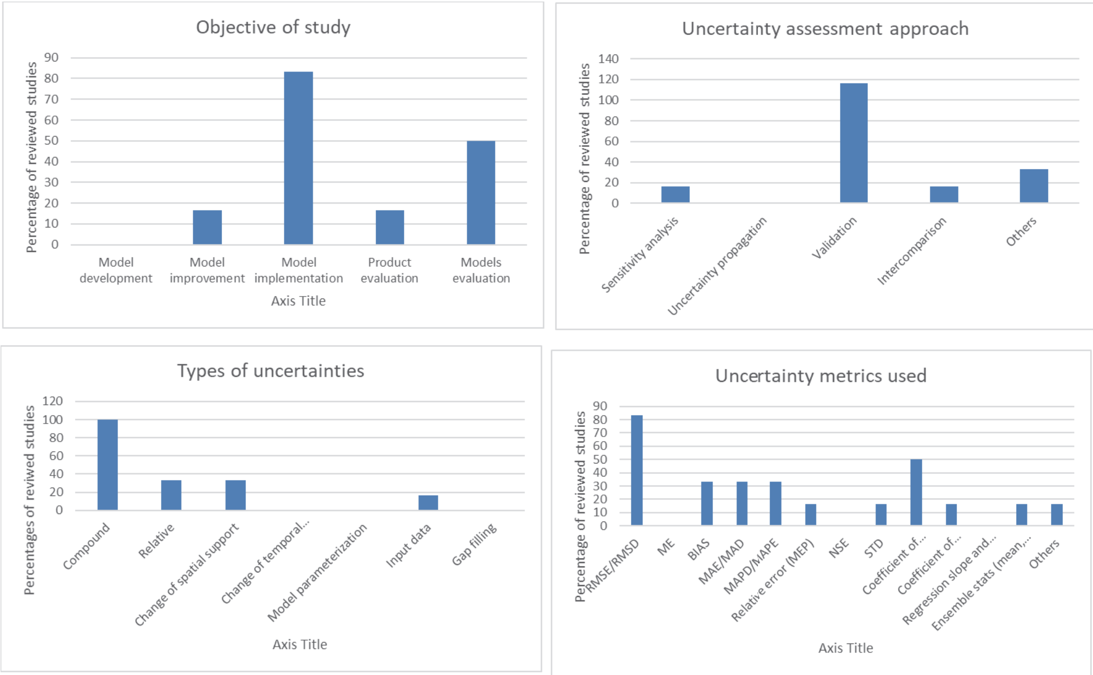

Figure 4. Distribution of reviewed studies by objective, uncertainty assessment approach, types of uncertainties and metrics used.

4. DISCUSSION

4.1 Applicability of Uncertainty Assessment Methods

The most common method is validation, which focusses on compound uncertainty and relies on reference ground-based measurements (analogous to ground-truthing). The most common ground-based reference is Eddy-Covariance flux measurements. Dealing with different support and uncertainty of reference ground-based measurements is challenging and there is no consistent method in the reviewed literature. When support of RS-based estimates is smaller than reference measurements, some studies average values from pixels approximately within station footprint to account for the different in support (Singh and Irmak, 2011; Mu et al., 2011). When support of RS-based estimate is larger than the reference, some studies report the model simulation using ground-based input net radiation data (Miralles et al., 2011). Some studies only mentioned that the scale mismatch between reference measurement and RS-base estimates (here ‘scale’ is mainly about ‘support’) and the uncertainty of the measurement when discussing the validation result. It is practically understandable that this information is not always available to researchers if they collect reference data from other sources. However, without reporting the support and uncertainty of measurements, we might easily have biased conclusions: when the comparison is good, we conclude that model has good result without questioning the quality of our reference, but when the comparison is not so good, we conclude that it is because of the imperfect reference measurements and the model still has good result. Hence, the recommendation for future research is to always report support and uncertainty of the measurements used for validation.

4.2 Reporting Uncertainty in RS-ET Estimates

The uncertainty of ET estimates is often reported as Root-Mean-Squared-Error (RMSE), where the error is defined as the difference between estimated values and reference values. The second most used evaluation metric is the coefficient of determination (R^2). This should not be confused with the squared Pearson’ correlation coefficient (r^2) as one shows the determination strength of the model (R^2) while the other only shows non-negative correlation strength between model estimates and reference values (r^2). In hydrological modelling, R^2 was proposed to evaluate model prediction efficiency (Nash-Sutcliffe, 1970), and it is later called Nash-Sutcliffe Efficiency (NSE). Some modelers then use the symbol R^2 to refer to squared Pearson’s correlation coefficient

(r^2), and even call this 'coefficient of determination' . This can be misleading about model performance since r^2 (correlation) is often closer to the optimal value than R^2 (determination).

5. CONCLUSIONS

This paper presents a systematic quantitative review of the current methods used to assess uncertainty in RS-ET estimates. RS-ET assessment has been geographically concentrated in North Asia, North America, and Europe. Most studies used the validation method, which quantifies the discrepancy between pixel-based ET estimation with an in-situ estimation. Most validation studies employed Eddy Covariance (EC) flux towers for reference estimation. In regions where in-situ measurements are limited, many studies use residual of the water balance as reference. However, few studies considered uncertainty in the reference estimation and mismatch of spatial and temporal scales. For crop water consumption monitoring, most RS-ET methods have been reported with high accuracy. However, when upscaling to larger regions including non-crop areas, additional assessments are required to better inform data users of the quality of RS-ET estimation, including cross-validation, sensitivity, and uncertainty analyses

6. ACKNOWLEDGEMENTS

This work has been conducted as a part of the second phase of the Water Productivity through Open-access of Remotely sensed derived data (WaPOR) project, which has been supported by the Ministry of Foreign Affairs of the Netherlands

7. REFERENCES

Allen, R.G., Tasumi, M., Trezza, R., 2007. Satellite-Based Energy Balance for Mapping Evapotranspiration with Internalized Calibration (METRIC)—Model. J. Irrig. Drain Eng. 133, 380–394. https://doi.org/10.1061/(ASCE)0733-9437(2007)133:4(380)

Anderson, M.C., Kustas, W.P., Norman, J.M., Hain, C.R., Mecikalski, J.R., Schultz, L., González-Dugo, M.P., Cammalleri, C., d'Urso, G., Pimstein, A., Gao, F., 2011. Mapping daily evapotranspiration at field to continental scales using geostationary and polar orbiting satellite imagery. Hydrology and Earth System Sciences 15, 223–239. https://doi.org/10.5194/hess-15-223-2011

Baldocchi, D., Falge, E., Gu, L., Olson, R., Hollinger, D., Running, S., Anthoni, P., Bernhofer, C., Davis, K., Evans, R., Fuentes, J., Goldstein, A., Katul, G., Law, B., Lee, X., Malhi, Y., Meyers, T., Munger, W., Oechel, W., U, K.T.P., Pilegaard, K., Schmid, H.P., Valentini, R., Verma, S., Vesala, T., Wilson, K., Wofsy, S., 2001. FLUXNET: A New Tool to Study the Temporal and Spatial Variability of Ecosystem-Scale Carbon Dioxide, Water Vapor, and Energy Flux Densities. Bulletin of the American Meteorological Society 82, 2415–2434. https://doi.org/10.1175/1520-0477(2001)082<2415:FANTTS>2.3.CO;2

Bastiaanssen, W.G.M., Menenti, M., Feddes, R.A., Holtslag, A.A.M., 1998. A remote sensing surface energy balance algorithm for land (SEBAL). 1. Formulation. Journal of Hydrology 212–213, 198–212. https://doi.org/10.1016/S0022-1694(98)00253-4

Bayat, B., Camacho, F., Nickeson, J., Cosh, M., Bolten, J., Vereecken, H., Montzka, C., 2021. Toward operational validation systems for global satellite-based terrestrial essential climate variables. International Journal of Applied Earth Observation and Geoinformation 95, 102240. https://doi.org/10.1016/j.jag.2020.102240

Chen, Y., Xia, J., Liang, S., Feng, J., Fisher, J.B., Li, Xin, Li, Xianglan, Liu, S., Ma, Z., Miyata, A., Mu, Q., Sun, L., Tang, J., Wang, K., Wen, J., Xue, Y., Yu, G., Zha, T., Zhang, L., Zhang, Q., Zhao, T., Zhao, L., Yuan, W., 2014. Comparison of satellite-based evapotranspiration models over terrestrial ecosystems in China. Remote Sensing of Environment 140, 279–293. https://doi.org/10.1016/j.rse.2013.08.045

Eck, N. van, Waltman, L., 2009. Software survey: VOSviewer, a computer program for bibliometric mapping. Scientometrics 84, 523–538. https://doi.org/10.1007/s11192-009-0146-3

FAO, 2018. WaPOR Database Methodology: Level 1. Remote Sensing for Water Productivity (Technical Report). Food and Agriculture Organization of the United Nations, Rome.

Korzoun, V.I., Sokolov, A.A., Budyko, M.I., Voskresensky, K.P., Kalinin, G.P., Konoplyantsev, A.A., Korotkevich, E.S., Kuzin, P.S., Lvovich, M.I., 1978. World water balance and water resources of the earth. Studies and Reports in Hydrology (UNESCO).

Kustas, W.P., Norman, J.M., 1999. Evaluation of soil and vegetation heat flux predictions using a simple two-source model with radiometric temperatures for partial canopy cover. Agricultural and Forest Meteorology 94, 13–29. https://doi.org/10.1016/S0168-1923(99)00005-2

Liu, Z., 2021. The accuracy of temporal upscaling of instantaneous evapotranspiration to daily values with seven upscaling methods. Hydrology and Earth System Sciences 25, 4417–4433. https://doi.org/10.5194/hess-25-4417-2021

Long, D., Longuevergne, L., Scanlon, B.R., 2014. Uncertainty in evapotranspiration from land surface modeling, remote sensing, and GRACE satellites. Water Resources Research 50, 1131–1151. https://doi.org/10.1002/2013WR014581

Melton, F.S., Huntington, J., Grimm, R., Herring, J., Hall, M., Rollison, D., Erickson, T., Allen, R., Anderson, M., Fisher, J.B., Kilic, A., Senay, G.B., Volk, J., Hain, C., Johnson, L., Ruhoff, A., Blankenau, P., Bromley, M., Carrara, W., Daudert, B., Doherty, C., Dunkerly, C., Friedrichs, M., Guzman, A., Halverson, G., Hansen, J., Harding, J., Kang, Y., Ketchum, D., Minor, B., Morton, C., Ortega-Salazar, S., Ott, T., Ozdogan, M., ReVelle, P.M., Schull, M., Wang, C., Yang, Y., Anderson, R.G., 2021. OpenET: Filling a Critical Data Gap in Water Management for the Western United States. JAWRA Journal of the American Water Resources Association n/a. https://doi.org/10.1111/1752-1688.12956

Miralles, D.G., Holmes, T.R.H., De Jeu, R. a. M., Gash, J.H., Meesters, A.G.C.A., Dolman, A.J., 2011. Global land-surface evaporation estimated from satellite-based observations. Hydrology and Earth System Sciences 15, 453–469. https://doi.org/10.5194/hess-15-453-2011

Moher, D., Liberati, A., Tetzlaff, J., Altman, D.G., PRISMA Group, 2009. Preferred reporting items for systematic reviews and meta-analyses: the PRISMA statement. PLoS Med 6, e1000097. https://doi.org/10.1371/journal.pmed.1000097

Mu, Q., Zhao, M., Running, S.W., 2011. Improvements to a MODIS global terrestrial evapotranspiration algorithm. Remote Sensing of Environment 115, 1781–1800. https://doi.org/10.1016/j.rse.2011.02.019

Pickering, C., Byrne, J., 2014. The benefits of publishing systematic quantitative literature reviews for PhD candidates and other early-career researchers. Higher Education Research & Development 33, 534–548. https://doi.org/10.1080/07294360.2013.841651

Singh, R.K., Irmak, A., 2011. Treatment of anchor pixels in the METRIC model for improved estimation of sensible and latent heat fluxes. Hydrological Sciences Journal 56, 895–906. https://doi.org/10.1080/02626667.2011.587424

Su, Z., 2002. The Surface Energy Balance System (SEBS) for estimation of turbulent heat fluxes. Hydrology and Earth System Sciences 6, 85–100. https://doi.org/10.5194/hess-6-85-2002

van de Schoot, R., de Bruin, J., Schram, R., Zahedi, P., de Boer, J., Weijdema, F., Kramer, B., Huijts, M., Hoogerwerf, M., Ferdinands, G., Harkema, A., Willemsen, J., Ma, Y., Fang, Q., Hindriks, S., Tummers, L., Oberski, D.L., 2021. An open source machine learning framework for efficient and transparent systematic reviews. Nat Mach Intell 3, 125–133. https://doi.org/10.1038/s42256-020-00287-7

