31 August 2015

Dear Prof. Janne Rinne,
Editor,
Atmospheric Chemistry and Physics,

We have completed the revision of our manuscript (acp-2015-150). We thank the two anonymous Referees for their valuable comments and suggestions. We addressed all comments bellow point-by-point. Some major changes were incorporated to the manuscript:

1) We deleted the word “solely” from the title (as suggested by Referee #2) and added that with this method ET could be estimated at 250 m spatial resolution. The new title: **“Annual evapotranspiration retrieved from satellites’ vegetation indices for the EM at 250 m spatial resolution”**.
2) We moved former Section 3.5 to the data Section (Section 2.3 in the revised MS) adding the relevant information on data sources.
3) We rewrote Section 4.2 following the comments of Referee #2.
4) We added a new Figure 4 (in the revised MS) showing the mean annual ET for the EM for 2000-2014, supplying a web connection to readers for downloading this product.
5) We added Figures S3 and S4 in the SI showing the seasonal and inter-annual time series of ET from eddy covariance at the FLUXNET sites, following the suggestion of Referee #1.

We think the MS has improved because of these changes and the Referee’s suggestions. We thank you for considering it for publication in *ACP* (to keep track changes in the revised MS we used blue text).

Sincerely,

David, Itamar and Amir.

We thank **Anonymous Referee (#1)** for his insightful comments. Following the Editors' comment (at the initial stage), we also addressed here the suggestion posted at the initial evaluation of the manuscript by **referee #1** (comment **1**).

REFEREE #1:

1) An interesting study showing that empirical modeling for ET is as good as more physically based modeling. This is fine to say the empirical way is 'good enough'. However, it would be better to say something about the generality of the PAVIE approach and commentary on the way forward for better ET estimation accuracy. The impression the paper leaves me is that we have reached the practical limits of accuracy for estimating ET from satellites, so it would be helpful for authors to state their opinion on this issue explicitly.

We thank you for this suggestion. Regarding the limits of accuracy for estimating ET from satellites we cannot say much because it was out of the scope and goals of this paper. However, what we can say is that VIs could not be used for estimating seasonal ET at complex Mediterranean vegetation systems comprising both annual (ephemeral) and perennial (mostly evergreen) vegetation. This is primarily because those two vegetation components have distinct phenology in Mediterranean systems (Helman et al., 2015). Annual and perennial vegetation usually constitute two different vertical layers in those systems and satellites only detect their combined signal. This is illustrated in Figure 1 and was explained in the revised MS:

Lines 3-12(P06):

Perennial and annual vegetation in Mediterranean regions have distinct phenology contributing differently to the VIs signal (Helman et al., 2015; Karnieli et al., 2003; Lu et al., 2003). Here we examined VIs - ET relationships in vegetation systems comprising both annual and perennial vegetation (i.e., forests, woodlands, savannah and shrublands, hereafter PA) separately from those comprising only annual vegetation (i.e., croplands and grasslands, hereafter AN).

We found that annual vegetation in the understory of PA systems might contribute significantly to VIs while having very small contribution to the total ecosystem ET. In some cases this results in an apparent phase shift between ET and VIs (Fig. 1) leading to negative or a lack of correlations."

And **Lines 27-31(P13)**:

"Intra-annual relationships were poor probably due to the mixed VI signal contributed by annual and perennial vegetation that constitute different vertical layers (Helman et al., 2015). While the annual vegetation (in the understory) was the main contributor to the intra-annual VI change, it was only a minor contributor to the total ecosystem ET in complex Mediterranean systems."

2) The scientific value of this work is not significant due to its statistical, locally calibrated approach. A more productive research avenue would be to develop greater, not less, modeling sophistication, and to find better ways to combine all lines of evidence into the ET estimation process. The currently available data sets, even confining oneself to remote sensing data alone, are rich; so it is hard to justify approaches that avoid using such data.

We acknowledge that developing a more general model for estimating ET at a global scale would be of a greater value. However, such task is out of the scope of this paper, though it is a possible future direction. Nevertheless, the importance of developing regional scale models in the EM should not be underestimated. Chen et al. 2014 is a good example for the importance of developing models at a regional scale presenting a model for China (published in ACP and cited in this MS). Moreover, ET models were hardly proposed or tested and validated at the EM. Developing remote sensing ET models without using ancillary data (except for the initial calibration using the EC data) is important for this region that lacks the required density of weather stations to assess ET at a high to moderate spatial resolution (250-1000m). Providing an ET product at such spatial resolution (250 m) is essential for the scientific community and forest and water resource managers of this region. We added this to the title:

"Annual evapotranspiration retrieved from satellites' vegetation indices for the Eastern Mediterranean at 250 m spatial resolution"

And the following lines to the revised MS:

Lines 28(P01) – 2(P02):

"In the lack of high-resolution (<1 km) ET models for the EM the proposed model is expected to contribute to the hydrological study of this region assisting in water resource management, which is one of the most valuable resources of this region."

And **Lines 6-7(P14)**:

"PaVI-E is the first ET model with such high-resolution (250 m) for the Eastern Mediterranean region."

While the need for simplicity can be important for common adoption of the method by others, the practitioners of ET science are a sophisticated group and can handle more complexity. Considering the ground work done with simplified approaches using vegetation indices and temperature more than 30(!) years ago by people such as Carlson, Price, Seguin, Gurney…plus many others I don't mention… I am wondering how proposing such a simple approach could be considered tenable. Has the world not progressed with ET research much since then? My opinion is that this work goes in the wrong direction.

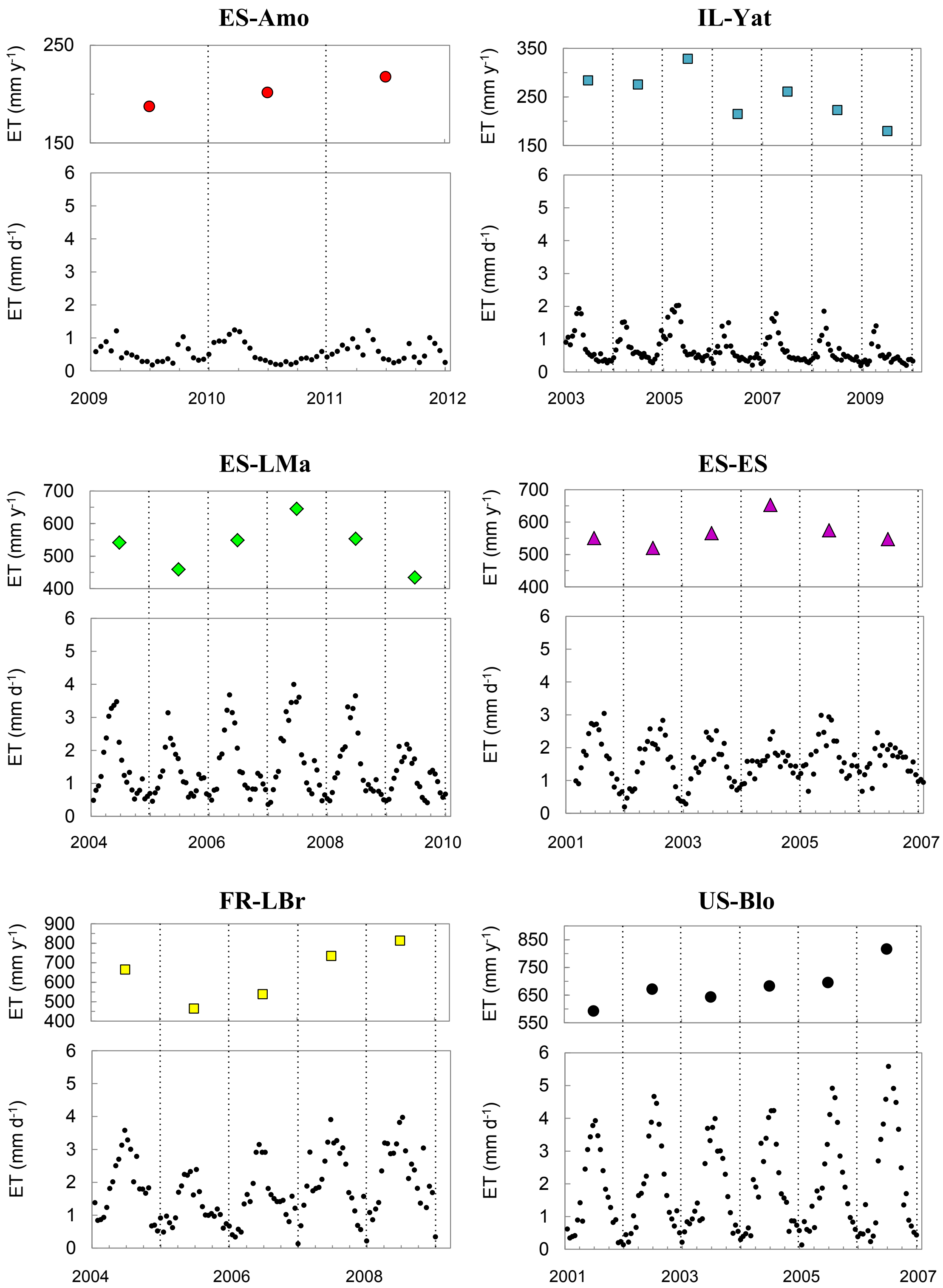
The practitioners of the remote sensing ET science are certainly a sophisticated group that can handle complex models and multiple data sources. However, from our experience, water resource and forest managers do not implement much of the models proposed by the remote sensing community because of their complexity. Moreover, some models report a limited accuracy at the sites they were tested and do not guarantee a better accuracy at the site to be implemented. Those models also have often too coarse spatial resolution for local to regional scale studies. In this paper, we propose a simpler model at relatively high spatial resolution (250m) that could be easily implemented by the scientific community and stakeholders at the EM with an acceptable accuracy, similar to (or better than) the proposed by more complex models. We see no reason to introduce additional data that would complicate the model while its current accuracy is within that of the EC data used for its calibration/validation. Glenn et al. (2010) in their review of existing remote sensing-based ET models already concluded that: *"The methods…range from the very simple (e.g. Groeneveld and Baugh 2007) to the more complex (e.g., Fisher et al. 2008), without a concomitant increase in accuracy."* They also stated that: *"One reason that simple models work as well as multi- factor models is that plant productivity (and transpiration) tend to be limited by a single factor at a given time and place (Liebig's law of the minimum) (Paris 1992)."*

The excellent work presented by Carlson, Price, Gillies, Smith, Kustas and more… founded the principles for using VIs and LST in ET (and soil moisture) assessment from remote sensing. However, they all used air temperature data from weather stations (or models) and a SVAT model to derive ET. The "triangle method" that does not use meteorological data is based on the simple relationship between VIs, LST and ET (or soil moisture), which are not straightforward at complex Mediterranean vegetation systems as illustrated by **Figure 1** (see also results from the 16-day regressions in **Tables 4-5**). The aim of our empirical approach was to test directly VIs/LST-ET relationships and examine the feasibility of using those simple relationships to estimate actual ET. Results imply that satellite-derived VIs could not be used for seasonal estimation of ET in complex Mediterranean systems due to the mixed contribution of annual and perennial vegetation to the VIs signal. However, ET could be estimated on an annual basis in those systems. This is an important contribution of this paper. As far as we know this is the first study that derives annual ET at such spatial resolution (250m) and accuracy (70-90% and MAE of 12%) from satellite data in the EM. Such annual ET product is essential for studying inter-annual changes in ET at the EM.

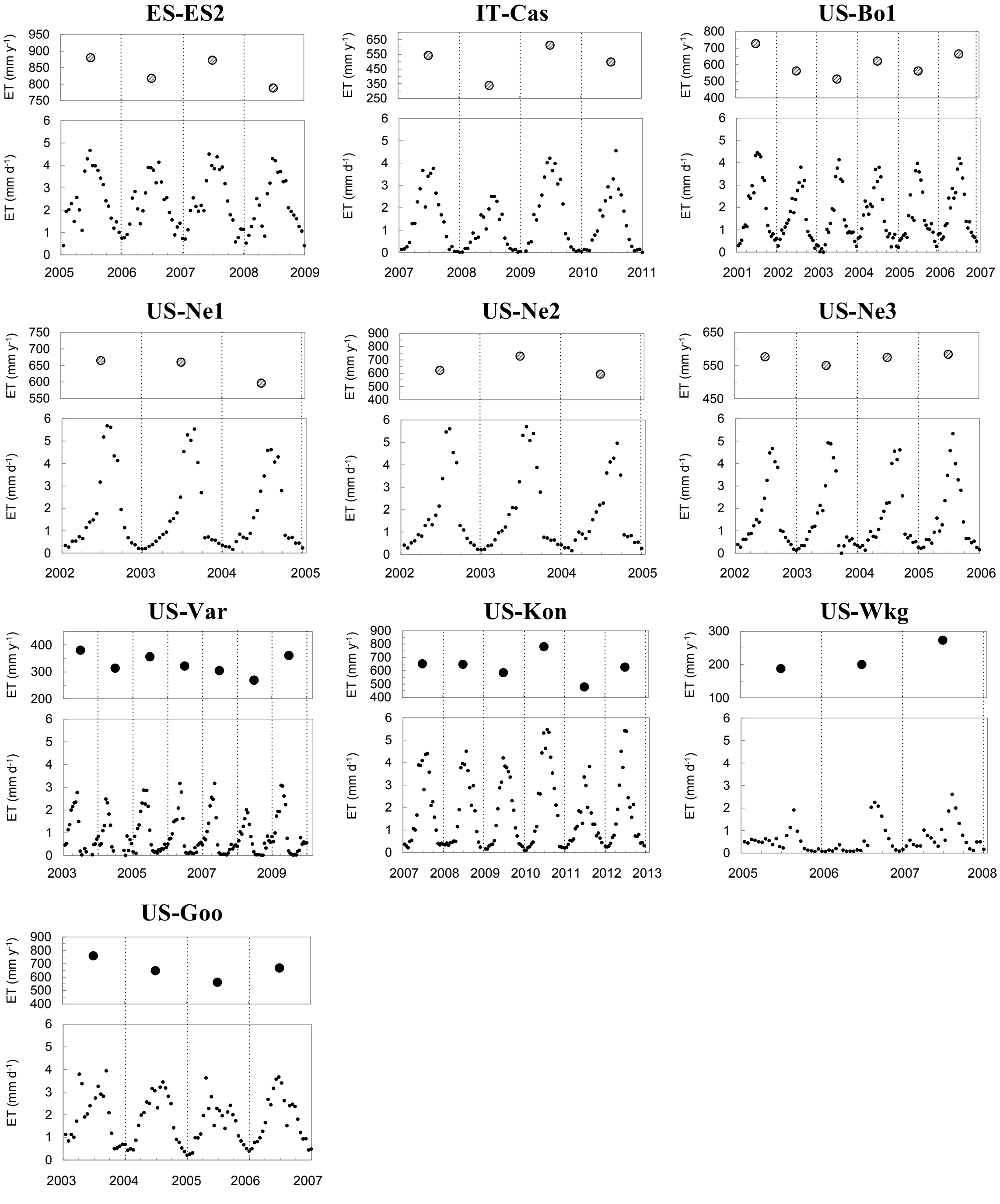
3) More details are needed to show readers statistical aspects of the 15-year data set: seasonal and annual variability of ET at each of the flux sites would help.

We added seasonal and interannual time series of the eddy covariance ET for each one of the 16 FLUXNET sites. However, to keep the flow of this paper we added this in the Supplementary Information (new **Figures S3** and **S4** in the revised SI).

Figure S3:


ES-Amo
ET (mm y-1)
250
150
ET (mm d-1)
0
1
2
3
4
5
6
2009
2010
2011
2012
IL-Yat
ET (mm y-1)
350
250
150
ET (mm d-1)
0
1
2
3
4
5
6
2003
2005
2007
2009
ES-LMa
ET (mm y-1)
700
600
500
400
ET (mm d-1)
0
1
2
3
4
5
6
2004
2006
2008
2010
ES-ES
ET (mm y-1)
700
600
500
400
ET (mm d-1)
0
1
2
3
4
5
6
2001
2003
2005
2007
FR-LBr
ET (mm y-1)
900
800
700
600
500
400
ET (mm d-1)
0
1
2
3
4
5
6
2004
2006
2008
US-Blo
ET (mm y-1)
850
750
650
550
ET (mm d-1)
0
1
2
3
4
5
6
2001
2003
2005
2007

and **Figure S4**:



4) What can you say about the accuracy of ET from the flux sites?

Following Glenn et al. (2010) and Kalma et al. (2008) that reviewed previous works that used the same data we used here, we can say that in general the accuracy of the ET flux data is ~70-90% (**Lines 18-20 P02** and **28-30 P09**).

5) How can you verify your confidence in your regression equations?

We dedicated two sections to validate and compare the results from our regression function models. First, a comparison with ET products from MODIS (based on Penman-Monteith) and MSG (based on H-TESSEL SVAT) for one year was conducted. Second, we validated our results using our mean annual estimates (2000-2013) with those calculated from water balances at the EM. Results showed good agreement in both cases. This is the best we can do with the scarcity of ET data for this region.

6) How representative is 2011 for comparing your empirical ET against MODIS and MSG ET estimates.

The ET product from MSG is only available for 2011-2013. We used one of those years for the comparison. PaVI-E was used for the same year, so it should be representative at least for that specific year.

7) Your statistical analyses, when considered as cover classes, do not seem strong with only ~7 samples per regression.

Although the number of per-site samples is relatively small (one sample per year), most relationships were statistically significant at $p<0.05$. The aim of the per-site comparison was to show possible relationships between VIs/LST and ET at annual time scales. Following this comment we indicated which correlations were statistically significant in **Table 3**. Cross-sites correlations were all statistically significant at $p<0.001$.

8) A lesser concern is your choice of R vs. R^2 for statistics; of course using R^2 values will decrease the apparent strength of your regressions.

The coefficient of determination (R^2) gives the proportion of the variance (fluctuation) of one variable that is predictable from the other variable, while R measures the strength and also the direction of the relationship. We used R to indicate strength and type of relationships (positive or negative) but referred to the percent of the variance in ET explained by VIs (R^2) in the text (**Lines 27-28 P09**).

9) No improvement of results at annual scales using LST data is not especially significant: VI data represent the long term vegetation patterns while LST data excel in identifying shorter term water stress events.

The 16-day LST reflect actual soil moisture conditions while mean annual LST is more representative of the water availability for ET in a specific site at an annual timescale. Good examples are the IL-Yat and ES-LMa sites, which had strong and significant ($p<0.05$) negative ET correlations with annual LST. In those sites, LST did improve correlations at the annual timescale for the multiple regression models (with EVI).

We thank **Anonymous Referee (#2)** for his thorough review and insightful comments; we think they have helped improve the content of the manuscript.

REFEREE #2:

1) The discussion paper "Annual evapotranspiration retrieved solely from satellites' vegetation indices for the Eastern Mediterranean" by D. Helman, I. M. Lensky and A. Givati fits the scope of ACP and addresses a relevant scientific question of estimating the evapotranspiration of spatially extended areas. The work presented in the paper is on a solid scientific base. The methods and assumptions are outlined clearly with an exception of the data intensity question I will discuss below in more detail. The conclusions reached by the authors are therefore well justified and bring new understanding to this field of research. The paper is well structured and reasonably well written. Unfortunately, it still contains language errors. Some of the errors cause ambiguity in text and should be corrected. A careful proofreading is strongly suggested. The language of the paper is not 'fluent and precise' as required by the journal.

We thank you for the positive assessment of the manuscript. We will send the final version for English editing prior to publication to meet the standards of *ACP*.

2) The biggest shortcoming of the paper is the lack of new concepts or ideas. While the research is scientifically sound, it resembles a technical report of successfully applying well-established methods with basic statistics in a specified geographical area. However, such reassuring knowledge is required by the scientific community and other stakeholders.

Using satellite remote sensing data to retrieve ET is not a novel idea and many papers have been published on the subject. However, the methods described in those papers require ancillary (mostly meteorological) data, which are not always available at a reasonable spatial resolution. This is particularly true for dry regions like North Africa and the Eastern Mediterranean where there is a great need for quantifying ET and the density of meteorological stations is low. In this paper we present a simple method to retrieve annual ET at a resolution of 250 m for the EM after calibrating with ET from several flux towers. Using only empirical relationships between VIs and ET to retrieve ET is a novel idea. Till date, VIs were used to replace the crop coefficient in FAO56 formulation in order to retrieve ET mostly in crop sites (with the use of additional meteorological information). Such approach was less successful in natural vegetation systems. Here we show that for inter-annual timescales ET and VIs have strong relationships that might be used to retrieve ET in space and time. Such relationships were not found before and we present them here for the first time. We also show that this method is accurate as (or more than) complex physical-based models. This simple approach introduces a new concept into satellite-based ET modeling and thus it is not a technical report of applying an established method in a specified geographical area.

3) The manuscript suffers from a fundamental or philosophical issue of the data or information richness of the empirical approach used by the authors. The authors stress that their approach requires less information compared to the physically based ones (line 12, page 15399). The same is implied by the word ’solely’ in the title of the paper. In reality, the method described in the paper requires spatially and temporally extensive field measurement data to establish and validate the empirical relationships. Thus, the results are not based solely on satellite data, but rather include a very significant ground-based information component. I suggest deleting the word ’solely’ from title and mentioning this in manuscript text.

We deleted the word ‘solely’ from the title and text as suggested. What we meant was that ET could be retrieved only from MODIS VIs data after first establishing the empirical relationships with ET measured from eddy covariance towers. While physical-based models need meteorological and radiative information to retrieve ET for a given time and place, our method uses a pre-established relationship (established here for the first time) to derive ET for any given time and place from MODIS VIs alone. We added the following clarifying line in the Abstract:

Lines 22-23(P01):

“After establishing empirical relationships, PaVI-E was used to retrieve ET_{Annual} *for the EM from 2000 to 2014.”*

4) A second shortcoming is the use of exponential equations in the model proposed to estimate ET from VIs. This should be discussed in more detail and more thoroughly. Exponential model does indeed allow to produce the low ET values reported by the authors on page 15409. However, it increases considerably ET at high VI values compared with the linear one. This should be quantified, or a linear model should be used with negative ET values discarded for producing ET maps. The continuity of the ET prediction mentioned on page 15410 is indeed rather a disadvantage. The continuity is achieved by using the ad hoc exponential function. No theoretical justification is given for it. Scientifically, it would be more appropriate to use a linear model indicated by the flux site data with negative values omitted or set to zero.

This issue was unclear in the former version of our paper and is better explained in the revised version [Lines **23-27(P09)** and **17-19(P11)**]. The exponential function is preferred here because it better explains ET-VIs relationships and not only because ET gets negative values at low VI (as we previously stated). ET-VIs is expected to have exponential relationships for two reasons:

(1) VIs have an exponential relationship with leaf area index (LAI) [Baret et al., 1989; Duchemin et al., 2006] and LAI is directly related to ET because plants reduce their LAI under water stress conditions [Mailhol et al., 1997].

(2) ET is greater than zero in places with very low vegetation cover (VIs≤0.1) due to soil evaporation.

The implication of (1) could be understood from the following example: The annual ET from the linear function (e.g., for PA systems) and NDVI of 0.85 (a value close to its saturation limit) is ca. 900 mm y^{-1}, while when using the exponential function ET reaches higher values by ca. 300 mm y^{-1}. Assuming that NDVI of 0.85 corresponds to dense forests with LAI ~6 m^2 m^{-2}, the reported annual ET for such LAI is within the range of ~1200-1400 mm y^{-1} [Kergoat et al., 2002], which is more close to the value obtained from the exponential function.

Lines 23-27(P09):

"Although a linear regression function is usually preferred to explain simple relationships between two parameters, the exponential relationship is more realistic in the case of ET-VIs. This is because VIs exhibit exponential relationships with LAI (Baret et al., 1989; Duchemin et al., 2006), which is directly related to water consumption and ET. Also, ET is usually greater than zero in places with low vegetation cover (VIs≤0.1) due to soil evaporation. "

Lines 19-21(P11):

"We used the exponential function because VIs exhibit exponential relationships with LAI, which is directly related to ET and because ET is greater than zero in areas with low vegetation cover due to soil evaporation."

References:

Baret, F., Guyot, G. and Major, D. J.: Crop biomass evaluation using radiometric measurements, Photogrammetria, 43(5), 241–256, doi:10.1016/0031-8663(89)90001-X, 1989.

Duchemin, B., Hadria, R., Erraki, S., Boulet, G., Maisongrande, P., Chehbouni, A., Escadafal, R., Ezzahar, J., Hoedjes, J. C. B., Kharrou, M. H., Khabba, S., Mougenot, B., Olioso, A., Rodriguez, J.-C. and Simonneaux, V.: Monitoring wheat phenology and irrigation

in Central Morocco: On the use of relationships between evapotranspiration, crops coefficients, leaf area index and remotely-sensed vegetation indices, Agric. Water Manag., 79(1), 1–27, doi:10.1016/j.agwat.2005.02.013, 2006.

Mailhol, J.C., Olufayo, A.A., Ruelle, P., 1997. Sorghum and sunflower evapotranspiration and yield from simulated leaf area index. *Agricultural Water Management* 35(1-2), 167–182.

Kergoat, L., S. Lafont, H. Douville, B. Berthelot, G. Dedieu, S. Planton, and J.F. Royer., 2002. Impact of doubled CO_2 on global-scale leaf area index and evapotranspiration: Conflicting stomatal conductance and LAI responses, *J. Geophys. Res.*,107(D24), 4808, doi:10.1029/2001JD001245.

---

5) A technical shortcoming of the paper is the lack of the description of the data source for the six catchments in Table 2. The relevant section in section Data is missing. Thus, it is impossible to objectively validate the performance of the various satellite products on the catchments.

Indeed this part of the Data section was missing. The former **Section 3.5** was moved to the data Section (**Section 2.3** in the revised paper) and given the name: “Evapotranspiration from water catchments balances for validation”. We added the relevant information on data sources in the revised **Section 2.3**.

Lines 8- 19(P05):

“Precipitation data (P) were collected for 2000-2013 from a total of 30 stations of the Israel Meteorological Service: 5 in Kziv, 2 in HaShofet, 21 in the Mountain Aquifer (north, centre and south) and 2 stations in the Mamashit catchment. Data were interpolated for the entire catchments area using ArcGIS and the inverse-distance weighting (IDW) methodology (Lu and Wong, 2008). Discharges (Q) were measured for the same period (2000-2013) for Kziv, Hashofet and Mamashit catchments using runoff gauges of the Hydrological Service of Israel (HSI) in: Gesher Haziv hydrometric station for Kziv, HaShofet-Hazorea for HaShofet and Mamashit station for the Mamashit catchment. Annual runoffs at the upper parts of the Mountain Aquifer (drainage areas without hydrometric stations at the Hedera, Alexander, Yarkon, Ayalon, Soreq and Lachish basins) were calculated using the HEC-HMS (Hydrologic Engineering Centre – Hydrologic Modelling System) model (Feldman, 2000) run by the HSI (http://www.water.gov.il).”

Minor comments

6) page 15401 line 3. Delete "well" which is not a quantitative descriptor. It may be argued that for many applications, the phenological changes are rapid and correlated with e.g. cloudiness thus that the 16-day window may lead to statistically relevant artifacts.

We deleted the word “well” from the text.

7) page 15402 line 1. Why were only the images for NAfr tile averaged remains unclear. It unclear why this is brought out in the manuscript. line 20. Change ’singular’ to ’single’. Also, rephrase the sentence, as it is not grammatically correct.

The NAfr tile is the relevant MSG tile that covers the EM region. We deleted this line and added to the next line the following information:

Lines 24-25(P04):

“The annual MODIS (MOD16A3) and daily MSG (LSA-SAF MSG Eta) ET products were downloaded for 2011 for the EM region.”

We rephrased **Section 3.1**:

Lines 3-12(P06):

“Perennial and annual vegetation in Mediterranean regions have distinct phenology contributing differently to the VIs signal (Helman et al., 2015; Karnieli, 2003; Lu et al., 2003). Here we examined VIs - ET relationships in vegetation systems comprising both annual and perennial vegetation (i.e., forests, woodlands, savannah and shrublands, hereafter PA) separately from those comprising only annual vegetation (i.e., croplands and grasslands, hereafter AN).

We found that annual vegetation in the understory of PA systems might contribute significantly to VIs while having very small contribution to the total ecosystem ET. In some cases this results in an apparent phase shift between ET and VIs (Fig. 1) leading to negative or a lack of correlations”

8) page 15403 line 19. Replace ’alone’ by single’. line 19. The sentence starting with ’In AN we used...’ VI does not have a growth season, the vegetation does. Rephrase the sentence as "In AN we subtracted the annual minimum VI before integrating it over the growing season..."

We meant that TG was designed to work only with data at 16-day intervals. We deleted the word “alone” and rephrased this and the following lines as suggested.

Lines 4-9(P07):

“We used all models with 16-day ET averages and 16-day VIs and/or LST data but only the first two models with total annual ET and mean annual VIs and/or LST because the TG model was designed to work only with 16-day data (Sims et al. 2008). In AN, we subtracted the annual minimum VIs before integrating it over the growing season instead of using the original 16-day VIs data (see in Helman et al., 2014a, 2014b). The integral over the VIs during the growth season was used in the two regression models against total annual ET.”

9) page 15404 line 16. Replace ’rational’ by ’rationale’. line 22. Replace ’considered as those’ by ’chosen as the ones.’ Alternatively, reorder the sentence. It’s no good English. line 23. Replace ’student’ by ’Student’s’

We change to “rationale” in line **26(P07)** and rephrased **Section 3.3**.

Lines 2-8(P08) [Section 3.3]:

“Pearson’s correlation coefficient (R) and mean absolute error (MAE) were chosen as accuracy metrics to evaluate the VIs-based ET models. The best model is considered as the one with the highest |R| and lowest MAE or at least lower than the eddy covariance accuracy (<30%). If two (or more) models fulfil these requirements, the one with the best performance with respect to its complexity i.e., with respect to the number of variables and operations needed, is preferred. A two-tailed Student’s t-test was used to examine statistical differences between the models p-values.”

10) page 15405 line 1. Delete the parentheses around ’p-value.’ line 3. Delete parentheses. line 17. Delete ’automatically.’ line 18. Chose a more appropriate word for ’a bit.’ line 18. Replace ’upon’ by ’to.’

We deleted parentheses (please see also the previous answer). We deleted the word “bit” and changed the word “upon” to “to” in line **21(P08)**.

Lines 21-22(P08):

“Although this classification procedure might be coarse, we preferred it to the MODIS land cover product for two reasons…”

11) page 15406 line 6. The source of P and Q should be specified, preferably already much earlier. line 21. Replace ’In average’ by ’On average.’ line 22. Use a more quantitative word instead of ’better.’ Name the quantity which increased by 40 and 60%, respectively.

We moved **Section 3.5** to the Data Sections (**Section 2.3** in the revised paper) and added a description of the sources for the water catchments components (see previous comment **5**).

Lines 8- 19(P05):

“Precipitation data (P) were collected for 2000-2013 from a total of 30 stations of the Israel Meteorological Service: 5 in Kziv, 2 in HaShofet, 21 in the Mountain Aquifer (north, centre and south) and 2 stations in the Mamashit catchment. Data were interpolated for the entire catchments area using ArcGIS and the inverse-distance weighting (IDW) methodology (Lu and Wong, 2008). Discharges (Q) were measured for the same period (2000-2013) for Kziv, Hashofet and Mamashit catchments using

runoff gauges of the Hydrological Service of Israel (HSI) in: Gesher Haziv hydrometric station for Kziv, HaShofet-Hazorea for HaShofet and Mamashit station for the Mamashit catchment. Annual runoffs at the upper parts of the Mountain Aquifer (drainage areas without hydrometric stations at the Hedera, Alexander, Yarkon, Ayalon, Soreq and Lachish basins) were calculated using the HEC-HMS (Hydrologic Engineering Centre – Hydrologic Modelling System) model (Feldman, 2000) run by the HSI (http://www.water.gov.il).”

We changed “In average” to “On average” in line **4(P09)**. We moved former line 22 to the beginning of this paragraph and rephrased it as follows:

Lines 4-11(P09):

“On average, the |R| for the ET-VIs linear regressions using annual data were higher by 60% (for NDVI) and 40% (for EVI) than the |R| for the 16-day regressions in PA sites. Total annual ET was highly correlated with mean annual NDVI in PA sites, 0.85<R<0.93 (Table 3; Fig. 2). In contrast, 16-day ET averages were only poorly correlated with 16-day NDVI (0.17<R<0.63). The same was for total annual ET and mean annual EVI with 0.66<R<0.94 compared to 0.28<R<0.70 when using 16-day EVI and ET data. The year-to-year changes in mean annual NDVI and EVI were significant enough to detect even small interannual changes in ET of 20 – 40 mm yr^{-1} (e.g., ES-Amo in Fig. 2).”

12) page 15407 line 9. Add ’annual’ before ’data.’ line 10. Replace ’high as in’ by ’as high as for’. line 10. Replace ’using’ by ’for.’ Add ’both’ before ’linear.’ Currently, the text is ambiguous. line 11. Add ’estimating’ before ’functions.’ line 11. The use of non-linear regression should be justified here, not later in the manuscript. line 18. Add ’a’ before ’dominant.’ line 18. Replace ’significant to’ by ’significant for.’

We rephrased former line 9 page 15407 (lines **18-19(P09)** in the revised paper). We changed to “as high as for” and to “high for both linear…” in lines **18-19(P09)**. We added the word “estimating” in line **20(P09)** and “a” in line **2(P10)** and replaced to “insignificant for…” in line **3(P10)**. We justified the use of the exponential function in lines **23-27(P09)**:

Lines 18-20(P09):

“Correlation coefficients for the cross-site comparisons were as high as for the site-specific regressions when using annual data in PA sites (Fig. 3). Correlations were high for both linear and exponential functions (R = 0.94, p<0.05 for both VIs and estimating functions).”

Lines 23-27(P09):

“Although a linear regression function is usually preferred to explain simple relationships between two parameters, the exponential relationship is more realistic

in the case of ET-VIs. This is because VIs exhibit exponential relationships with LAI (Baret et al., 1989; Duchemin et al., 2006), which is directly related to water consumption and ET. Also, ET is usually greater than zero in places with low vegetation cover (VIs≤0.1) due to soil evaporation.”

Lines 1-3(P10):

“The contribution of annual and perennial vegetation to VIs at the sub pixel level is most difficult to distinguish in PA systems. In some cases, one of those components might have a dominant contribution to VIs but insignificant for the ecosystem flux exchange (Fig. 1).”

13) page 15408 line 10. Add ’, respectively’ after the first occurrence of ’EVI.’ line 10. Unclear. What is the simple model? Did include LST? why is "(for LST with NDVI or EVI)" repeated? line 12. Replace ’substantially improved’ by ’were substantially better.’ line 14. Add ’, respectively’ after both occurrences of ’EVI.’ line 15. Delete ’But.’ line 17. Add ’, respectively’ after ’EVI.’ line 19. Delete ’alone.’ Add a ’, respectively after ’EVI.’

We rewrote **Section 4.2** following these comments.

Lines 9(P10)-5(P11):

“In AN, correlation coefficients from the cross-site regressions of ET against VIs (i.e., the integrals over the growing season period) using annual data were comparable to those with the 16-day data (Table 4). The R for 16-day regressions was 0.86 for both indices (p<0.001). The R for the annual ET-NDVI regression was higher (R = 0.88, p<0.001) than that for the ET-EVI regression (R = 0.79, p<0.001). However, the mean relative error (i.e., MAE/mean) was much lower for the annual regressions (12-16%) than for 16-day regressions (32-33%, Table 5). The relatively high R for the 16-day ET-VIs regressions in AN supports the biomass-ET-VIs relationship in those systems described elsewhere (Glenn et al., 2010).

Correlations did not significantly improve (p>0.1) when LST was added and multiple regressions were applied in AN sites (Tables 4 and 5). The R for multiple regressions of LST and VIs against ET using 16-day data was 0.87 (for LST with each one of the VIs compared to 0.86 for ET-NDVI and ET-EVI, p<0.001 for both). R for multiple regressions using annual data were 0.89 and 0.79 (ET against LST with NDVI or EVI, respectively and p<0.001 for both) compared to 0.87 and 0.82 (p<0.001) for ET-NDVI and ET-EVI regressions, respectively.

In PA, correlations from the multiple regressions of ET against 16-day LST and VIs were substantially better (p<0.05 using LST with each one of the VIs) than those from the simple ET-VIs regressions. R from multiple regressions were 0.71 and 0.73 for 16-day ET against LST with NDVI and EVI, respectively compared to 0.51 and 0.61 for ET against NDVI and EVI, respectively. R for single and multiple regressions *were not statistically different (p>0.1) when using annual data in PA. The R was 0.94 and 0.96 for LST with NDVI and EVI, respectively and 0.94 for ET against both VIs.*

The modified TG model resulted in significantly higher R (p<0.05 for both indices) only for PA and for the 16-day data (R = 0.80 and 0.78 using NDVI or EVI in Eq. (5)). However, it was still significantly lower (p<0.05 for both VIs) than the R for ET–VIs regressions when using annual data (Table 4 and Fig. S6B). In AN, R from TG and 16-day ET–VIs regressions were not significantly different (p>0.1, Table 4 and Fig.S6A).”

14) page 15410 line 6. Replace ’relatively high’ by ’higher.’ line 11. Add ’a’ after ’such.’ line 19. Relative biases should be in plural. line 21. The slash is used to denote division. Use parentheses as in ’...the relatively higher (lower) MOD 16...’

We changed to “higher” in line **13(P12)** and added “a” to “such” in line **17(P12)**. We changed to “biases” in line **24(P12)** and replaced the slash by a parenthesis in line **27(P12)**.

15) page 15411 line 2. Replace ’from’ by ’for.’ line 20. Replace ’VIs relationships’ by ’ET-VI relationships.’ line 23. Delete ’Yet.’

We changed to “for” in line **4(P13)** and to “ET-VI relationships” in line **23(P13)**. We deleted the word “Yet” from line **24(P13)**.

16) page 15412 line 3. The text is overly cumbersome and difficult to read. Delete the words ’Following a performance-simplicity criterion we used’ and ’interannual relationships to retrieve total annual ET for the Eastern Mediterranean. This’ line 5. Add comma before ’had.’ line 6. Add comma at the end of the line. line 11. The sentences ’Improvement in the estimation of ET is essential for water budget calculations and water resource management especially in water limited regions. Here we propose the use of a simple model to retrieve annual ET at 250 m spatial resolution suitable for the Eastern Mediterranean region’ are not conclusions and need to be deleted.

We changed these lines and deleted superfluous sentences as suggested.

Lines 3-10(P14):

“The simple ET-VIs model, named here the parameterized vegetation index for ET estimates model (PaVI-E), had comparable estimates to MODIS and MSG ET products in the Eastern Mediterranean. Models’ estimates also agreed well to ET calculated from six water catchments balances along the south-north EM rainfall gradient. PaVI-E is the first ET model with such high-resolution (250 m) for the Eastern Mediterranean region. Its advantage is its simplicity and spatial resolution compared to the coarser resolutions of MODIS and MSG ET products (1 and 3.1 km, respectively). We are confident that using PaVI-E will enhance the hydrological study

in this region where ET plays a major role in the hydrological cycle."

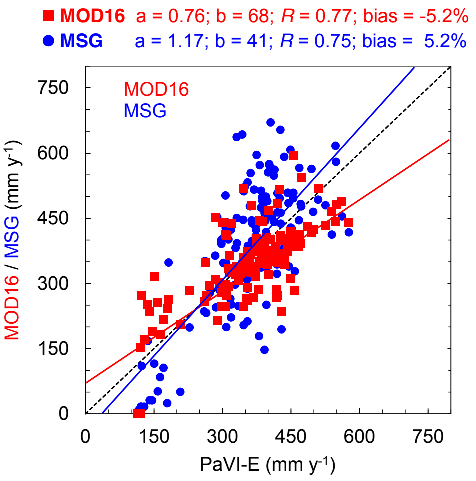
17) Caption to Table 1. Use 'Top' instead of 'Up.' Table 2. The names in the table are technical abbreviation, not names. Give clear and unambiguous names allowing the catchments to be identified in nature. Legend to Figure 3. Give here or elsewhere also the coefficients of the linear models. Figure 5. Use also different symbols, not just different color.

We changed to "Top" in caption of Table 1. In **Table 2**, those are the names of the catchments used by the Israel Hydrological Service (we have no other names for those catchments). Except for the Mountain Aquifers that have no other identifying names all other names are geographically identified. The coefficients for the linear regressions were given in the text [lines **20-22(P09)**]. We changed the symbols in the revised **Figure 5** as suggested.

Lines 20-22(P09):

"The linear functions were ET = 1277 NDVI – 189 and ET = 2844 EVI – 300 (mm y^{-1}*). Exponential functions were ET = 85* $e^{3.12\ NDVI}$ *and ET = 65* $e^{6.31\ EVI}$ *(mm* y^{-1}*)."*

Figure 5 was revised:



Annual evapotranspiration retrieved from satellites' vegetation indices for the Eastern Mediterranean at 250 m spatial resolution

D. Helman[1], **A. Givati**[2] **and I. M. Lensky**[1]

[1]{Department of Geography and Environment, Bar Ilan University, Ramat-Gan, Israel}

[2]{Israeli Hydrological Service, Water Authority, Jerusalem, Israel}

Correspondence to: D. Helman (davidhelman.biu@gmail.com)

Abstract

We present a simple model to retrieve actual evapotranspiration (ET) from satellites' vegetation indices (PaVI-E) for the Eastern Mediterranean (EM) at a spatial resolution of 250 m. The model is based on the empirical relationship between satellites' vegetation indices (NDVI and EVI from MODIS) and total annual ET (ET_{Annual}) estimated at 16 FLUXNET sites representing a wide range of plant functional types and ET_{Annual}. Empirical relationships were first examined separately for (a) annual vegetation systems (i.e., croplands and grasslands) and (b) systems with combined annual and perennial vegetation (i.e., woodlands, forests, savannah and shrublands). Vegetation indices explained most of the variance in ET_{Annual} in those systems (71% for annuals, and 88% for combined annuals and perennials systems) while adding land surface temperature data in multiple regression and modified Temperature and Greenness models did not result in better correlations ($p>0.1$). After establishing empirical relationships, PaVI-E was used to retrieve ET_{Annual} for the EM from 2000 to 2014. Models' estimates were highly correlated ($R = 0.92$, $p<0.01$) with ET_{Annual} calculated from water catchments balances along rainfall gradient of the EM. They were also comparable to the coarser resolution ET products of MSG (LSA-SAF MSG ETa, 3.1 km) and MODIS (MOD16, 1 km) at 148 EM basins with R of 0.75 and 0.77 and relative biases of 5.2 and -5.2%, respectively ($p<0.001$ for both). In the lack of high-resolution (<1 km) ET models for the EM the proposed model is expected to contribute to the hydrological study of this region

assisting in water resource management, which is one of the most valuable resources of this region.

1 Introduction

Actual evapotranspiration (ET) is a primary component of the global water cycle. Its assessment at global and regional scales is essential for forecasting future atmospheric feedback (Jung et al., 2010; Oki and Kanae, 2006; Zemp et al., 2014). Estimating ET at such scales though, is not straightforward and requires the use of models (Chen et al., 2014; Hu et al., 2015; Jung et al., 2009; Trambauer et al., 2014). Data-driven models using satellite information benefit from a continuous spatio-temporal direct observation (Ma et al., 2014; Shi and Liang, 2014).

Satellite-based ET models are classified into two: (1) empirical, using the relationship between in situ ET and satellites-derived vegetation indices (VIs) (Glenn et al., 2011; Nagler et al., 2012; Tillman et al., 2012) and (2) physical, using surface temperature from satellites to solve energy balance equations (Anderson et al., 2008; Colaizzi et al., 2012). While some models combine the two approaches (Tsarouchi et al., 2014).

Although physical-based models are much more common their performance is comparable to that of the empirical-based models (Glenn et al., 2010). The accuracy of both approaches is within that of the eddy covariance measurements (70-90%) used for their calibration or validation (Kalma et al., 2008). Yet, the empirical approach is simpler than the physical-based model and requires less additional information.

The basis for the empirical model is the resource optimisation theory. This theory suggests that plants adjust their foliage density to the environmental capacity to support photosynthetic activity and transpiration (Glenn et al., 2010). Accordingly, changes in vegetation foliage cover (and VIs) will affect ET resulting in high ET-VIs correlations. Then, the empirical equation could be used to retrieve ET in space and time.

This approach is mostly used in vegetation systems with annual cycle of growth and drying where VIs define well the phenological stages (Glenn et al., 2011; Senay et al., 2011). However, in complex systems comprised of annual (i.e., herbaceous) and perennial (i.e., woody) vegetation the model must be adjusted with additional meteorological data (Maselli et al., 2014).

The main drawback of the empirical-based approach is that it is limited to a specific site and vegetation type (Glenn et al., 2010; Maselli et al., 2014; Nagler et al., 2012). No common relationship was found between ET and VIs for different sites and climatic conditions.

Here we used MODIS VIs and land surface temperature (LST) products and eddy covariance ET from 16 FLUXNET sites with different plant functional types to establish the empirical relationships between VIs (and/or LST) and ET in Mediterranean vegetation systems. We first examine relationships in annual vegetation systems and complex systems comprising both annuals and perennials vegetation. Three empirical models were used: (1) simple regression, (2) multiple regression and (3) modified Temperature and Greenness models with 16-day and mean annual data. We used a performance-simplicity criterion to choose the best model to retrieve ET for the EM. Estimates were compared with MODIS and MSG ET operational products and evaluated against ET calculated from water catchments balances in the EM.

2 Data

2.1 Evapotranspiration from eddy covariance towers

In situ ET was derived from eddy covariance towers that constitute the international flux towers net (FLUXNET). Two open FLUXNET sources were used to acquire the datasets: the Oak Ridge National Laboratory Distributed Active Archive Centre (available online [http://fluxnet.ornl.gov] from ORNL DAAC, Oak Ridge, Tennessee, U.S.A) and the European fluxes database [http://gaia.agraria.unitus.it/home]. Half-hourly level 4 ET data were checked for acceptable quality (Reichstein et al., 2005) and gap-filled using methods described in Reichstein et al. (2005) and Moffat et al. (2007). Then, data were aggregated to 16 days means (mm d^{-1}) and total annual ET (mm yr^{-1}). Only ET data since the time MODIS VIs products are available were used (i.e., since 2000).

2.2 Satellite products

We used 16-day NDVI and EVI at a spatial resolution of 250 m (MOD13Q1) and 8-day LST at 1 km spatial resolution (MOD11A2) from MODIS on board Terra satellite. Although Terra provides LST twice a day (around 10:30 a.m./p.m. local time) here we used only daytime LST, which is the relevant for ET processes. The 8-day LST was averaged to match the 16-day temporal resolution of the VIs product.

The MODIS 16-day VIs product is a composite of a single day value selected from 16 days period based on a maximum value criterion (Huete et al., 2002). It represents the vegetation status of the entire 16-day period because of the gradual development of the vegetation. This enables regressing this MODIS VIs product against 16-day averages of ET. NDVI is defined as (Rouse et al., 1974):

$$NDVI = \frac{R_{0.8} - R_{0.6}}{R_{0.8} + R_{0.6}} \tag{1}$$

and EVI as (Huete et al., 2002):

$$EVI = 2.5 \times \frac{R_{0.8} - R_{0.6}}{R_{0.8} + 6R_{0.6} - 7.5R_{0.5} + 1} \tag{2}$$

where $R_{0.8}$, $R_{0.6}$ and $R_{0.5}$ are the reflectance at near infrared (0.8 μm), red (0.6 μm) and blue (0.5 μm) bands, respectively. NDVI suffers from asymptotic problems (saturation) over high density of vegetation biomass while EVI is more sensitive in such cases (Huete et al., 2002).

For the model development, time series of NDVI, EVI and LST at each FLUXNET site were obtained from MODIS Land Product Subsets [http://daac.ornl.gov/MODIS/modis.html] (ORNL DAAC, Oak Ridge, Tennessee, U.S.A., last accessed December 2014) for the years when ET data was available since 2000 (see ‘Period’ column in Table 1). NDVI and EVI time series were smoothed using local weighted scatterplot technique (LOWESS) as in Helman et al. (2014a, 2014b). For model implementation, tiles h20v05, h21v05, h20v06 and h21v06 of the MOD13Q1 product were downloaded for 2000–2014 using the USGS EarthExplorer tool [http://earthexplorer.usgs.gov]. These tiles fully cover the Eastern Mediterranean region.

Model results were compared with two satellite operational ET products from MODIS (MOD16) and MSG (LSA-SAF MSG ETa) in 2011 at 148 main basins in the Eastern Mediterranean. MODIS and MSG ET products are based on different physical models, and have different spatial and temporal resolutions (1km/8day for MODIS, and 3.1km/daily for MSG) (Hu et al., 2015). The annual MODIS (MOD16A3) and daily MSG (LSA-SAF MSG Eta) ET products were downloaded for 2011 for the EM region. The basins layer map was taken from HydroSHEDS, a mapping product based on high-resolution elevation developed by the Conservation Science Program of World Wildlife Fund (http://hydrosheds.cr.usgs.gov). Only main basins with an area greater than 10 km^2 were selected (Fig. S1).

2.3 Evapotranspiration from water catchments balances for validation

We evaluated our model with mean annual ET calculated from six water catchments balances along a north-south rainfall gradient (130 – 800 mm yr^{-1}) in the Eastern Mediterranean (Fig. S2 and Table 2). The calculation follows the classical water balance equation:

$$ET = P - Q - \frac{dS}{dt} \tag{3}$$

where P and Q are the total annual precipitation and discharge measured in the catchment, and dS/dt is the change in water storage.

Precipitation data (P) were collected for 2000-2013 from a total of 30 stations of the Israel Meteorological Service: 5 in Kziv, 2 in HaShofet, 21 in the Mountain Aquifer (north, centre and south) and 2 stations in the Mamashit catchment. Data were interpolated for the entire catchments area using ArcGIS and the inverse-distance weighting (IDW) methodology (Lu and Wong, 2008). Discharges (Q) were measured for the same period (2000-2013) for Kziv, Hashofet and Mamashit catchments using runoff gauges of the Hydrological Service of Israel (HSI) in: Gesher Haziv hydrometric station for Kziv, HaShofet-Hazorea for HaShofet and Mamashit station for the Mamashit catchment. Annual runoffs at the upper parts of the Mountain Aquifer (drainage areas without hydrometric stations at the Hedera, Alexander, Yarkon, Ayalon, Soreq and Lachish basins) were calculated using the HEC-HMS (Hydrologic Engineering Centre – Hydrologic Modelling System) model (Feldman, 2000) run by the HSI (http://www.water.gov.il).

For timescales of several years dS/dt is assumed to be negligible so the mean annual ET could be simply calculated from P minus Q (Conradt et al., 2013). Because our water balances were averaged over 14 years (i.e., 2000-2013), this assumption was valid in our case. The water balances components (P and Q) and the calculated mean annual ET for the six catchments are presented in Table 2.

The water balances approach has some drawbacks like the difficulty to properly estimate precipitation distribution over the catchment and uncertainties about catchment boundaries (Conradt et al., 2013). However, it is the best existing approach to compare in situ ET with satellite-derived ET at a basin scale.

3 Methods

3.1 Sites selection

Perennial and annual vegetation in Mediterranean regions have distinct phenology contributing differently to the VIs signal (Helman et al., 2015; Karnieli, 2003; Lu et al., 2003). Here we examined VIs - ET relationships in vegetation systems comprising both annual and perennial vegetation (i.e., forests, woodlands, savannah and shrublands, hereafter PA) separately from those comprising only annual vegetation (i.e., croplands and grasslands, hereafter AN).

We found that annual vegetation in the understory of PA systems might contribute significantly to VIs while having very small contribution to the total ecosystem ET. In some cases this results in an apparent phase shift between ET and VIs (Fig. 1) leading to negative or a lack of correlations. Moreover, AN sites exhibited one single ET–VI relationship under wide range of rainfall conditions while significantly differ for similar PA systems under different climatic regimes (Unpublished results).

Therefore, the AN sites (FLUXNET sites in AN systems) were selected from wide range of climatic regimes while PA sites (FLUXNET sites in PA systems) were selected only from Mediterranean-climate regions. Selection of the FLUXNET sites had to fulfil the following criteria: (1) at least three years of satellite and eddy covariance data in the FLUXNET site; (2) missing data less than 30 days yr^{-1} for ET and 15% for VIs; and (3) homogeneous vegetation cover near the FLUXNET tower within at least the 250 m spatial resolution of MODIS VIs products. The last criterion was manually assured using Google Earth™. These led us to select 16 FLUXNET sites that represent a wide range of plant functional types and ET rates (Table 1, Figures S3 and S4).

3.2 Empirical ET models using VIs and LST

Three regression models using VIs and/or LST and ET from eddy covariance towers were tested:

(1) Simple regressions between ET against VIs or LST with 16-day or annual data.
(2) Multiple regressions of VIs and LST as dependent variables with 16-day or annual data.

(3) Modified version of the Temperature and Greenness (TG) model proposed by Sims et al. (2008) using LST as a proxy for radiation and reference ET (Maeda et al., 2011) with 16-day data alone.

We used all models with 16-day ET averages and 16-day VIs and/or LST data but only the first two models with total annual ET and mean annual VIs and/or LST because the TG model was designed to work only with 16-day data (Sims et al. 2008). In AN, we subtracted the annual minimum VIs before integrating it over the growing season instead of using the original 16-day VIs data (see in Helman et al., 2014a, 2014b). The integral over the VIs during the growth season was used in the two regression models against total annual ET. Multiple regressions were applied only on NDVI and LST data or EVI and LST data, but not on NDVI with EVI data because NDVI and EVI were highly correlated ($R > 0.95$, $p<0.001$).

The original TG model is based on the observed correlations between MODIS-EVI and FLUXNET GPP, which were further refined by incorporating LST data (Sims et al., 2008):

$$GPP = a \times EVI_{scaled} \times LST_{scaled} \,, \qquad (4)$$

where EVI_{scaled} is the scaled EVI set to zero at EVI = 0.1 (i.e., $EVI_{scaled} = EVI - 0.1$) due to absence of photosynthetic activity at this value (Sims et al., 2006); *a* is the slope of the relationship that enables parameterization of the model; and LST_{scaled} is daytime LST scaled to 1 at an optimum temperature for leaf photosynthetic response around 30 °C, decreasing towards 0 at lower and higher temperatures as follows (Sims et al., 2008):

$$LST_{scaled} = \min\left[\left(\frac{LST}{30}\right); (2.5 - 0.05 \times LST)\right]. \qquad (5)$$

Note that LST_{scaled} in Eq. (4) is negative at LST higher than 50°C. In such case, LST_{scaled} is set to 0 assuming no photosynthetic activity at those high temperatures following a stomata closure (Sims et al., 2008).

Here, we modified the TG model by using ET instead of GPP in Eq. (4):

$$ET = a \times EVI_{scaled} \times LST_{scaled} \qquad (6)$$

The rationale is that GPP and ET are correlated through trade-offs between carbon gains and water loss during photosynthesis processes. We used the modified TG model with EVI and NDVI alternatively in Eq. (6).

3.3 Models evaluation

Pearson's correlation coefficient (*R*) and mean absolute error (MAE) were chosen as accuracy metrics to evaluate the VIs-based ET models. The best model is considered as the one with the highest |*R*| and lowest MAE or at least lower than the eddy covariance accuracy (<30%). If two (or more) models fulfil these requirements, the one with the best performance with respect to its complexity i.e., with respect to the number of variables and operations needed, is preferred. A two-tailed Student's t-test was used to examine statistical differences between the models *p*-values.

3.4 Land cover map for model implementation

ET was assessed for the Eastern Mediterranean using the best models for AN and PA systems separately. To produce the required land cover map, we classified pixels as AN and PA based on their NDVI during the year. Low NDVI during the dry season (<0.25) implies absent or dry vegetation typical for AN systems (Lu et al., 2003). Yet, some PA systems (e.g., open shrublands) also have low NDVI during this period but differ from AN systems by smaller NDVI change (<0.4) during the growth season (Lu et al., 2003; Roderick et al., 1999).

Hence, we classified pixels with minimum NDVI < 0.25 as AN only if their NDVI increased by more than 0.4 during the growth season. To account for the high NDVI in agricultural fields of the Nile delta, pixels with minimum NDVI smaller or equal to 0.35 were also classified as AN only if their NDVI increased by more than 0.35. All remaining pixels were classified as PA (Fig. S5).

Although this classification procedure might be coarse, we preferred it to the MODIS land cover product for two reasons. First, a significant discrepancy was found between MODIS-based land cover product and actual land cover type distribution in the Eastern Mediterranean (Sprintsin et al., 2009a). Second, this procedure produces a mask at the spatial resolution of the model (250 m), while the MODIS-derived land cover product is available at coarser resolution (500 m).

The produced AN/PA land cover map showed the general pattern known for this region (Fig. S5). Moreover, the total AN area estimated for Israel not considering the Golan Heights grasslands (i.e. considering mostly Israel's croplands) was $255 \cdot 10^3$ ha. This agreed well with the total cropland area reported by the Israeli Central Bureau of Statistics for the same years ($220 \cdot 10^3$ ha, CBS 2014).

4 Results and discussion

4.1 ET-VIs in systems with both- annual and perennial vegetation

On average, the $|R|$ for the ET-VIs linear regressions using annual data were higher by 60% (for NDVI) and 40% (for EVI) than the $|R|$ for the 16-day regressions in PA sites. Total annual ET was highly correlated with mean annual NDVI in PA sites, $0.85<R<0.93$ (Table 3; Fig. 2). In contrast, 16-day ET averages were only poorly correlated with 16-day NDVI ($0.17<R<0.63$). The same was for total annual ET and mean annual EVI with $0.66<R<0.94$ compared to $0.28<R<0.70$ when using 16-day EVI and ET data. The year-to-year changes in mean annual NDVI and EVI were significant enough to detect even small interannual changes in ET of 20 – 40 mm yr^{-1} (e.g., ES-Amo in Fig. 2).

LST was negatively correlated with 16-day and total annual ET in all PA FLUXNET sites. This implies the role of transpiration in attenuating thermal load (Rotem-Mindali et al., 2015). Mean annual LST was highly correlated with total annual ET ($|R| > 0.84$, $p< 0.05$) particularly in sites with low canopy cover (IL-Yat – 30-45% and ES-LMa – 20-30%; Casals et al., 2009; Sprintsin et al., 2009b). Those sites had relatively high interannual variability in LST (2 – 3.5 °C; Fig. 2).

Correlation coefficients for the cross-site comparisons were as high as for the site-specific regressions when using annual data in PA sites (Fig. 3). Correlations were high for both linear and exponential functions ($R = 0.94$, $p<0.05$ for both VIs and estimating functions). The linear functions were $ET = 1277\ NDVI - 189$ and $ET = 2844\ EVI - 300$ (mm y^{-1}). Exponential functions were $ET = 85\ e^{3.12\ NDVI}$ and $ET = 65\ e^{6.31\ EVI}$ (mm y^{-1}).

Although a linear regression function is usually preferred to explain simple relationships between two parameters, the exponential relationship is more realistic in the case of ET-VIs. This is because VIs exhibit exponential relationships with LAI (Baret et al., 1989; Duchemin et al., 2006), which is directly related to water consumption and ET. Also, ET is usually greater than zero in places with low vegetation cover ($VIs\leq0.1$) due to soil evaporation. The mean annual NDVI and EVI explained 71 and 88% of the variability (R^2) in the total annual ET using these functions. This is within the accuracy of the eddy covariance technique for estimating ET (Glenn et al., 2010; Kalma et al., 2008). Cross-site correlations of annual ET and LST in PA were also high with a negative relationship ($R = -0.89$, $p<0.05$, Fig. 3).

The contribution of annual and perennial vegetation to VIs at the sub pixel level is most difficult to distinguish in PA systems. In some cases, one of those components might have a dominant contribution to VIs but insignificant for the ecosystem flux exchange (Fig. 1). This is probably one of the reasons that VIs could not be used to assess ET at a seasonal timescale (i.e., using 16-day data) in such systems. However, at interannual timescales (i.e., using the annual mean) relationships between ET and VIs were strong and might be used to retrieve total annual ET in PA systems.

4.2 Comparison between empirical VIs-based ET models

In AN, correlation coefficients from the cross-site regressions of ET against VIs (i.e., the integrals over the growing season period) using annual data were comparable to those with the 16-day data (Table 4). The R for 16-day regressions was 0.86 for both indices ($p<0.001$). The R for the annual ET-NDVI regression was higher ($R = 0.88$, $p<0.001$) than that for the ET-EVI regression ($R = 0.79$, $p<0.001$). However, the mean relative error (i.e., MAE/mean) was much lower for the annual regressions (12-16%) than for 16-day regressions (32-33%, Table 5). The relatively high R for the 16-day ET-VIs regressions in AN supports the biomass-ET-VIs relationship in those systems described elsewhere (Glenn et al., 2010).

Correlations did not significantly improve ($p>0.1$) when LST was added and multiple regressions were applied in AN sites (Tables 4 and 5). The R for multiple regressions of LST and VIs against ET using 16-day data was 0.87 (for LST with each one of the VIs compared to 0.86 for ET-NDVI and ET-EVI, $p<0.001$ for both). R for multiple regressions using annual data were 0.89 and 0.79 (ET against LST with NDVI or EVI, respectively and $p<0.001$ for both) compared to 0.87 and 0.82 ($p<0.001$) for ET-NDVI and ET-EVI regressions, respectively.

In PA, correlations from the multiple regressions of ET against 16-day LST and VIs were substantially better ($p<0.05$ using LST with each one of the VIs) than those from the simple ET-VIs regressions. R from multiple regressions were 0.71 and 0.73 for 16-day ET against LST with NDVI and EVI, respectively compared to 0.51 and 0.61 for ET against NDVI and EVI, respectively. R for single and multiple regressions were not statistically different ($p>0.1$) when using annual data in PA. The R was 0.94 and 0.96 for LST with NDVI and EVI, respectively and 0.94 for ET against both VIs.

The modified TG model resulted in significantly higher *R* ($p<0.05$ for both indices) only for PA and for the 16-day data ($R = 0.80$ and 0.78 using NDVI or EVI in Eq. (6)). However, it was still significantly lower ($p<0.05$ for both VIs) than the *R* for ET–VIs regressions when using annual data (Table 4 and Fig. S6B). In AN, *R* from TG and 16-day ET–VIs regressions were not significantly different ($p>0.1$, Table 4 and Fig.S6A).

4.3 PaVI-E model

NDVI and EVI explained most of the interannual changes in ET in both AN and PA systems (Table 4). This means that a single ET–VIs regression function could be used to estimate total annual ET in those systems. Multiple regression and TG modified models had higher *R* and lower MAE in some cases (Table 5), but differences were not significant ($p>0.05$). Hence, we chose the simple regression functions as the best models following the performance-simplicity criterion. The functions obtained from ET-NDVI and ET-EVI regressions were averaged for PA:

$$ET_{Annual} = \frac{85\exp(3.1 \cdot NDVI) + 65\exp(6.9 \cdot EVI)}{2} \qquad (7)$$

and AN systems:

$$ET_{Annual} = \frac{187\exp(0.23 \cdot NDVI) + 224\exp(0.26 \cdot EVI)}{2} \qquad (8)$$

Where ET_{Annual} is the total annual ET in mm yr^{-1}. NDVI and EVI in Eq. (7) are the mean annual NDVI and EVI. $NDVI_{GSI}$ and EVI_{GSI} in Eq. (8) are the integrals over the NDVI and EVI during the growth season, respectively. We used the exponential function because VIs exhibit exponential relationships with LAI, which is directly related to ET and because ET is greater than zero in areas with low vegetation cover due to soil evaporation.

Finally, we named this model PaVI-E, the Parameterization of Vegetation Indices for ET estimation model. The mean relative error of PaVI-E was 13 and 12% for AN and PA, respectively. This is within the accuracy of the eddy covariance measurements that were used for calibration and much lower than the reported for more complex models (Glenn et al., 2010; Kalma et al., 2008). PaVI-E was used to assess total annual ET at a spatial resolution of 250 m for the Eastern Mediterranean (EM) after using the land cover map created for AN and PA as a mask (Section 3.3 and Fig. S5).

Figure 4 shows the mean annual ET at the EM for the period of 2000-2014. The annual products of PaVI-E are currently downloadable at 1 km spatial resolution for the EM and at 250 m for Israel from the web (http://davidhelman.weebly.com).

4.4 Model evaluation in the Eastern Mediterranean

4.4.1 Comparison with MODIS and MSG ET products

ET estimates from PaVI-E were compared with two operational remote sensing ET products in 148 large basins (>10 km^2). The spatial patterns of annual ET for 2011 from PaVI-E, MOD16 and MSG were generally similar over the EM (Fig. 5). The three models show a general west to east and south to north ET gradients along the eastern coastline, matching the rainfall gradients of this region (Ziv et al., 2014). Also, all three models show higher ET estimates over agricultural fields in the Nile delta compared to the surrounding desert.

However, some discrepancies also exist. MOD16 estimates were lower along the EM coast compared to PaVI-E and MSG. ET estimates from MSG were higher along the eastern coast especially to the east of the Galilee Sea (mean ET of ~800 mm yr^{-1}). Differences between models were particularly noted over the Nile delta. The average annual ET over the Nile delta for 2011 was 160 mm yr^{-1} from MSG, 530 mm yr^{-1} from MOD16, and 680 mm yr^{-1} from PaVI-E. While MSG estimates seem extremely low for a such highly productive area, PaVI-E and MOD16 estimates agreed well with the high ET reported from in situ measurements (Elhag et al., 2011). Besides the advantage of an improved spatial resolution (250 m compared to 1 km and 3.1 km of MOD16 and MSG) PaVI-E also has the ability to produce spatially continuous ET compared to MSG and MODIS products (Fig. 5).

Comparing the three models at a basin scale resulted in good agreement between them (R = 0.77 and 0.75 for PaVI-E vs. MOD16 and MSG, respectively, p<0.001 for both; Fig. 6). MOD16 and MSG products had small biases with respect to PaVI-E with relative biases (i.e., bias/mean) of -5.2% and 5.2% and slopes of 0.76 and 1.17 for MOD16 and MSG ET products, respectively.

The relatively higher (lower) MOD16 estimates in xeric (mesic) Mediterranean areas (Fig. 6) was already pointed out by Trambauer et al. (2014) that compared this product with several independent ET models. Furthermore, comparison of MOD16 and MSG ET products in Europe showed that correlations with in situ ET (from 15-eddy covariance sites) were better for MSG (Hu et al., 2015), and that MOD16 underestimate ET in Mediterranean dry regions similarly to the observed in this study (Fig. 5).

4.4.2 Evaluation against ET from water balances along rainfall gradient

ET estimates for PaVI-E were evaluated against ET calculated from six water catchments along rainfall gradient in the Eastern Mediterranean (EM). PaVI-E estimates were highly correlated with ET from water balances ($R = 0.92$, $p<0.01$) at six catchments along the north – south rainfall gradient in the EM (Fig. 7a). ET from MOD16 and MSG were also significantly correlated with the water balances-derived ET ($p<0.05$, Fig. S7). All three models had very similar ET estimates in the mountain aquifer catchments (MA-N, MA-CS, and MA-S), lower than the calculated from water balances (Fig. 7b). Still, within the accuracy of the models (~12%) and gauging/rainfall distribution uncertainties (~10 – 15%, Conradt et al., 2013).

As shown in Fig. 5, ET estimates derived from PaVI-E are significantly higher than those from MOD16 and MSG in the dry areas of the EM. This is due to the exponential functions used in PaVI-E (Eq. (7) and (8)). It generated a comparable ET to that calculated from water balances at the dry catchment of Mamashit with a slight overestimation (15%) of PaVI-E (Fig. 7b). MSG largely underestimated the calculated ET in Mamashit (by more than 85%) while MOD16 had no data for this area.

5 Conclusions

Three VIs-based ET models using only eddy covariance ET and MODIS vegetation indices and land surface temperature data for Mediterranean vegetation systems were tested. Vegetation systems comprising mostly annual vegetation (i.e., grasslands and croplands) had strong ET-VIs relationships with intra-annual (16-day ET averages) and interannual (total annual ET) ET estimates. The mean relative error was larger for intra-annual relationships compared to interannual relationships (32% compared to 12%). In systems with annual and perennial vegetation (i.e., forests, woodlands, savannah and shrublands) ET-VIs relationships were strong only at interannual timescales (i.e., using annual data). Intra-annual relationships were poor probably due to the mixed VI signal contributed by annual and perennial vegetation that constitute different vertical layers (Helman et al., 2015). While the annual vegetation (in the understory) was the main contributor to the intra-annual VI change, it was only a minor contributor to the total ecosystem ET in complex Mediterranean systems. Multiple regression and modified TG models using VIs and LST were not significantly better than simple ET-VIs regression models for both PA and AN vegetation systems (*p*>0.1).

The simple ET-VIs model, named here the parameterized vegetation index for ET estimates model (PaVI-E), had comparable estimates to MODIS and MSG ET products in the Eastern Mediterranean. Models' estimates also agreed well to ET calculated from six water catchments balances along the south-north EM rainfall gradient. PaVI-E is the first ET model with such high-resolution (250 m) for the Eastern Mediterranean region. Its advantage is its simplicity and spatial resolution compared to the coarser resolutions of MODIS and MSG ET products (1 and 3.1 km, respectively). We are confident that using PaVI-E will enhance the hydrological study in this region where ET plays a major role in the hydrological cycle.

Acknowledgements

This research was supported by the Hydrological Service of Israel (HSI), Water Authority (Grant No. 4500962964). David Helman acknowledges personal grants provided by the JNF-Rieger Foundation, USA and the HSI. Authors thank to Shilo Shiff for his assistance with the LSA-SAF MSG files and two anonymous referees that helped improve the previous version of this manuscript with insightful comments and suggestions. This study used eddy covariance data acquired by the FLUXNET community and in particular by the following networks: AmeriFlux (U.S. Department of Energy, Biological and Environmental Research, Terrestrial Carbon Program), CarboItaly and CarboEuropeIP. MODIS subsets and tiles land products were acquired from the Oak Ridge National Laboratory Distributed Active Archive Centre (ORNL DAAC) and the U.S. Geological Survey (USGS).

References

Anderson, M. C., Norman, J. M., Kustas, W. P., Houborg, R., Starks, P. J. and Agam, N.: A thermal-based remote sensing technique for routine mapping of land-surface carbon, water and energy fluxes from field to regional scales, Remote Sensing of Environment, 112(12), 4227–4241, doi:10.1016/j.rse.2008.07.009, 2008.

Baldocchi, D. D., Xu, L. and Kiang, N.: How plant functional-type, weather, seasonal drought, and soil physical properties alter water and energy fluxes of an oak–grass savanna and an annual grassland, Agricultural and Forest Meteorology, 123(1–2), 13–39, doi:10.1016/j.agrformet.2003.11.006, 2004.

Baret, F., Guyot, G. and Major, D. J.: Crop biomass evaluation using radiometric measurements, Photogrammetria, 43(5), 241–256, doi:10.1016/0031-8663(89)90001-X, 1989.

Casals, P., Gimeno, C., Carrara, A., Lopez-Sangil, L. and Sanz, M.: Soil CO2 efflux and extractable organic carbon fractions under simulated precipitation events in a Mediterranean Dehesa, Soil Biology and Biochemistry, 41(9), 1915–1922, doi:10.1016/j.soilbio.2009.06.015, 2009.

Chamizo, S., Cantón, Y., Miralles, I. and Domingo, F.: Biological soil crust development affects physicochemical characteristics of soil surface in semiarid ecosystems, Soil Biology and Biochemistry, 49(0), 96–105, doi:10.1016/j.soilbio.2012.02.017, 2012.

Chen, X., Su, Z., Ma, Y., Liu, S., Yu, Q. and Xu, Z.: Development of a 10-year (2001–2010) 0.1° data set of land-surface energy balance for mainland China, Atmos. Chem. Phys., 14(23), 13097–13117, doi:10.5194/acp-14-13097-2014, 2014.

Colaizzi, P. D., Kustas, W. P., Anderson, M. C., Agam, N., Tolk, J. A., Evett, S. R., Howell, T. A., Gowda, P. H. and O'Shaughnessy, S. A.: Two-source energy balance model estimates of evapotranspiration using component and composite surface temperatures, Advances in Water Resources, 50(0), 134–151, doi:10.1016/j.advwatres.2012.06.004, 2012.

Conradt, T., Wechsung, F. and Bronstert, A.: Three perceptions of the evapotranspiration landscape: comparing spatial patterns from a distributed hydrological model, remotely sensed surface temperatures, and sub-basin water balances, Hydrology and Earth System Sciences, 17(7), 2947–2966, doi:10.5194/hess-17-2947-2013, 2013.

Craine, J. M., Nippert, J. B., Elmore, A. J., Skibbe, A. M., Hutchinson, S. L. and Brunsell, N. A.: Timing of climate variability and grassland productivity, Proceedings of the National Academy of Sciences, 109(9), 3401–3405, 2012.

Duchemin, B., Hadria, R., Erraki, S., Boulet, G., Maisongrande, P., Chehbouni, A., Escadafal, R., Ezzahar, J., Hoedjes, J. C. B., Kharrou, M. H., Khabba, S., Mougenot, B., Olioso, A., Rodriguez, J.-C. and Simonneaux, V.: Monitoring wheat phenology and irrigation in Central Morocco: On the use of relationships between evapotranspiration, crops coefficients, leaf area index and remotely-sensed vegetation indices, Agricultural Water Management, 79(1), 1–27, doi:10.1016/j.agwat.2005.02.013, 2006.

Elhag, M., Psilovikos, A., Manakos, I. and Perakis, K.: Application of the SEBS water balance model in estimating daily evapotranspiration and evaporative fraction from remote sensing data over the Nile delta, Water Resour Manage, 25(11), 2731–2742, doi:10.1007/s11269-011-9835-9, 2011.

Glenn, E., Nagler, P. and Huete, A.: Vegetation index methods for estimating evapotranspiration by remote sensing, Surv Geophys, 31(6), 531–555, doi:10.1007/s10712-010-9102-2, 2010.

Glenn, E. P., Neale, C. M. U., Hunsaker, D. J. and Nagler, P. L.: Vegetation index-based crop coefficients to estimate evapotranspiration by remote sensing in agricultural and natural ecosystems, Hydrol. Process., 25(26), 4050–4062, doi:10.1002/hyp.8392, 2011.

Helman, D., Mussery, A., Lensky, I. M. and Leu, S.: Detecting changes in biomass productivity in a different land management regimes in drylands using satellite-derived vegetation index, Soil Use Manage, 30(1), 32–39, doi:10.1111/sum.12099, 2014a.

Helman, D., Lensky, I. M., Mussery, A. and Leu, S.: Rehabilitating degraded drylands by creating woodland islets: Assessing long-term effects on aboveground productivity and soil fertility, Agricultural and Forest Meteorology, 195–196(0), 52–60, doi:10.1016/j.agrformet.2014.05.003, 2014b.

Helman, D., Lensky, I. M., Tessler, N. and Osem, Y.: A phenology-based method for monitoring woody and herbaceous vegetation in Mediterranean forests from NDVI time series, Remote Sensing, (accepted), 2015.

Hollinger, S. E., Bernacchi, C. J. and Meyers, T. P.: Carbon budget of mature no-till ecosystem in North Central Region of the United States, Agricultural and Forest Meteorology, 130(1–2), 59–69, doi:10.1016/j.agrformet.2005.01.005, 2005.

Huete, A., Didan, K., Miura, T., Rodriguez, E. P., Gao, X. and Ferreira, L. G.: Overview of the radiometric and biophysical performance of the MODIS vegetation indices, Remote sensing of environment, 83(1), 195–213, 2002.

Hu, G., Jia, L. and Menenti, M.: Comparison of MOD16 and LSA-SAF MSG evapotranspiration products over Europe for 2011, Remote Sensing of Environment, 156(0), 510–526, doi:10.1016/j.rse.2014.10.017, 2015.

Jung, M., Reichstein, M. and Bondeau, A.: Towards global empirical upscaling of FLUXNET eddy covariance observations: validation of a model tree ensemble approach using a biosphere model, Biogeosciences, 6(10), 2001–2013, doi:10.5194/bg-6-2001-2009, 2009.

Jung, M., Reichstein, M., Ciais, P., Seneviratne, S. I., Sheffield, J., Goulden, M. L., Bonan, G., Cescatti, A., Chen, J., de Jeu, R., Dolman, A. J., Eugster, W., Gerten, D., Gianelle, D., Gobron, N., Heinke, J., Kimball, J., Law, B. E., Montagnani, L., Mu, Q., Mueller, B., Oleson, K., Papale, D., Richardson, A. D., Roupsard, O., Running, S., Tomelleri, E., Viovy, N., Weber, U., Williams, C., Wood, E., Zaehle, S. and Zhang, K.: Recent decline in the global land evapotranspiration trend due to limited moisture supply, Nature, 467(7318), 951–954, doi:10.1038/nature09396, 2010.

Kalma, J., McVicar, T. and McCabe, M.: Estimating land surface evaporation: A review of methods using remotely sensed surface temperature data, Surv Geophys, 29(4-5), 421–469, doi:10.1007/s10712-008-9037-z, 2008.

Karnieli, A.: Natural vegetation phenology assessment by ground spectral measurements in two semi-arid environments, Int J Biometeorol, 47(4), 179–187, doi:10.1007/s00484-003-0169-z, 2003.

Kutsch, W. L., Aubinet, M., Buchmann, N., Smith, P., Osborne, B., Eugster, W., Wattenbach, M., Schrumpf, M., Schulze, E. D., Tomelleri, E., Ceschia, E., Bernhofer, C., Béziat, P., Carrara, A., Di Tommasi, P., Grünwald, T., Jones, M., Magliulo, V., Marloie, O., Moureaux, C., Olioso, A., Sanz, M. J., Saunders, M., Søgaard, H. and Ziegler, W.: The net biome production of full crop rotations in Europe, Agriculture, Ecosystems & Environment, 139(3), 336–345, doi:10.1016/j.agee.2010.07.016, 2010.

Lu, H., Raupach, M. R., McVicar, T. R. and Barrett, D. J.: Decomposition of vegetation cover into woody and herbaceous components using AVHRR NDVI time series, Remote Sensing of Environment, 86(1), 1–18, doi:10.1016/S0034-4257(03)00054-3, 2003.

Maeda, E. E., Wiberg, D. A. and Pellikka, P. K. E.: Estimating reference evapotranspiration using remote sensing and empirical models in a region with limited ground data availability in Kenya, Applied Geography, 31(1), 251–258, doi:10.1016/j.apgeog.2010.05.011, 2011.

Maselli, F., Papale, D., Chiesi, M., Matteucci, G., Angeli, L., Raschi, A. and Seufert, G.: Operational monitoring of daily evapotranspiration by the combination of MODIS NDVI and ground meteorological data: Application and evaluation in Central Italy, Remote Sensing of Environment, 152(0), 279–290, doi:10.1016/j.rse.2014.06.021, 2014.

Maseyk, K. S., Lin, T., Rotenberg, E., Grünzweig, J. M., Schwartz, A. and Yakir, D.: Physiology–phenology interactions in a productive semi-arid pine forest, New Phytologist, 178(3), 603–616, 2008.

Ma, Y., Zhu, Z., Zhong, L., Wang, B., Han, C., Wang, Z., Wang, Y., Lu, L., Amatya, P. M., Ma, W. and Hu, Z.: Combining MODIS, AVHRR and in situ data for evapotranspiration estimation over heterogeneous landscape of the Tibetan Plateau, Atmos. Chem. Phys., 14(3), 1507–1515, doi:10.5194/acp-14-1507-2014, 2014.

Moffat, A. M., Papale, D., Reichstein, M., Hollinger, D. Y., Richardson, A. D., Barr, A. G., Beckstein, C., Braswell, B. H., Churkina, G. and Desai, A. R.: Comprehensive comparison of gap-filling techniques for eddy covariance net carbon fluxes, Agricultural and Forest Meteorology, 147(3), 209–232, 2007.

Nagler, P. L., Brown, T., Hultine, K. R., van Riper III, C., Bean, D. W., Dennison, P. E., Murray, R. S. and Glenn, E. P.: Regional scale impacts of Tamarix leaf beetles (Diorhabda carinulata) on the water availability of western U.S. rivers as determined by multi-scale remote sensing methods, Remote Sensing of Environment, 118(0), 227–240, doi:10.1016/j.rse.2011.11.011, 2012.

Oki, T. and Kanae, S.: Global Hydrological Cycles and World Water Resources, Science, 313(5790), 1068–1072, 2006.

Reichstein, M., Falge, E., Baldocchi, D., Papale, D., Aubinet, M., Berbigier, P., Bernhofer, C., Buchmann, N., Gilmanov, T., Granier, A., Grünwald, T., Havránková, K., Ilvesniemi, H., Janous, D., Knohl, A., Laurila, T., Lohila, A., Loustau, D., Matteucci, G., Meyers, T., Miglietta, F., Ourcival, J.-M., Pumpanen, J., Rambal, S., Rotenberg, E., Sanz, M., Tenhunen, J., Seufert, G., Vaccari, F., Vesala, T., Yakir, D. and Valentini, R.: On the separation of net ecosystem exchange into assimilation and ecosystem respiration: review and improved algorithm, Global Change Biology, 11(9), 1424–1439, doi:10.1111/j.1365-2486.2005.001002.x, 2005.

Reichstein, M., Ciais, P., Papale, D., Valentini, R., Running, S., Viovy, N., Cramer, W., Granier, A., Ogee, J., Allard, V., Aubinet, M., Bernhofer, C., Buchmann, N., Carrara, A., Grunwald, T., Heimann, M., Heinesch, B., Knohl, A., Kutsch, W., Loustau, D., Manca, G., Matteucci, G., Miglietta, F., Ourcival, J. M., Pilegaard, K., Pumpanen, J., Rambal, S., Schaphoff, S., Seufert, G., Soussana, J.-F., Sanz, M.-J., Vesala, T. and Zhao, M.: Reduction of ecosystem productivity and respiration during the European summer 2003 climate anomaly: a joint flux tower, remote sensing and modelling analysis, Global Change Biology, 13(3), 634–651, doi:10.1111/j.1365-2486.2006.01224.x, 2007.

Roderick, M. L., Noble, I. R. and Cridland, S. W.: Estimating woody and herbaceous vegetation cover from time series satellite observations, Global Ecology and Biogeography, 8(6), 501–508, doi:10.1046/j.1365-2699.1999.00153.x, 1999.

Rotem-Mindali, O., Michael, Y., Helman, D. and Lensky, I. M.: The role of local land-use on the urban heat island effect of Tel Aviv as assessed from satellite remote sensing, Applied Geography, 56(0), 145–153, doi:10.1016/j.apgeog.2014.11.023, 2015.

Rouse, J. W., Haas, R. H. and Schell, J. A.: Monitoring the vernal advancement and retrogradation (greenwave effect) of natural vegetation, Texas A and M University, College Station., 1974.

Scott, R. L., Hamerlynck, E. P., Jenerette, G. D., Moran, M. S. and Barron-Gafford, G. A.: Carbon dioxide exchange in a semidesert grassland through drought-induced vegetation change, J. Geophys. Res., 115(G3), G03026, doi:10.1029/2010JG001348, 2010.

Senay, G. B., Leake, S., Nagler, P. L., Artan, G., Dickinson, J., Cordova, J. T. and Glenn, E. P.: Estimating basin scale evapotranspiration (ET) by water balance and remote sensing methods, Hydrol. Process., 25(26), 4037–4049, doi:10.1002/hyp.8379, 2011.

Shi, Q. and Liang, S.: Surface-sensible and latent heat fluxes over the Tibetan Plateau from ground measurements, reanalysis, and satellite data, Atmos. Chem. Phys., 14(11), 5659–5677, doi:10.5194/acp-14-5659-2014, 2014.

Sims, D. A., Rahman, A. F., Cordova, V. D., El-Masri, B. Z., Baldocchi, D. D., Flanagan, L. B., Goldstein, A. H., Hollinger, D. Y., Misson, L., Monson, R. K., Oechel, W. C., Schmid, H. P., Wofsy, S. C. and Xu, L.: On the use of MODIS EVI to assess gross primary productivity of North American ecosystems, J. Geophys. Res., 111(G4), G04015, doi:10.1029/2006JG000162, 2006.

Sims, D. A., Rahman, A. F., Cordova, V. D., El-Masri, B. Z., Baldocchi, D. D., Bolstad, P. V., Flanagan, L. B., Goldstein, A. H., Hollinger, D. Y., Misson, L., Monson, R. K., Oechel,

W. C., Schmid, H. P., Wofsy, S. C. and Xu, L.: A new model of gross primary productivity for North American ecosystems based solely on the enhanced vegetation index and land surface temperature from MODIS, Remote Sensing of Environment, 112(4), 1633–1646, doi:10.1016/j.rse.2007.08.004, 2008.

Skiba, U., Drewer, J., Tang, Y. S., van Dijk, N., Helfter, C., Nemitz, E., Famulari, D., Cape, J. N., Jones, S. K., Twigg, M., Pihlatie, M., Vesala, T., Larsen, K. S., Carter, M. S., Ambus, P., Ibrom, A., Beier, C., Hensen, A., Frumau, A., Erisman, J. W., Brüggemann, N., Gasche, R., Butterbach-Bahl, K., Neftel, A., Spirig, C., Horvath, L., Freibauer, A., Cellier, P., Laville, P., Loubet, B., Magliulo, E., Bertolini, T., Seufert, G., Andersson, M., Manca, G., Laurila, T., Aurela, M., Lohila, A., Zechmeister-Boltenstern, S., Kitzler, B., Schaufler, G., Siemens, J., Kindler, R., Flechard, C. and Sutton, M. A.: Biosphere–atmosphere exchange of reactive nitrogen and greenhouse gases at the NitroEurope core flux measurement sites: Measurement strategy and first data sets, Agriculture, Ecosystems & Environment, 133(3–4), 139–149, doi:10.1016/j.agee.2009.05.018, 2009.

Sprintsin, M., Karnieli, A., Berliner, P., Rotenberg, E., Yakir, D. and Cohen, S.: Evaluating the performance of the MODIS Leaf Area Index (LAI) product over a Mediterranean dryland planted forest, International Journal of Remote Sensing, 30(19), 5061–5069, doi:10.1080/01431160903032885, 2009a.

Sprintsin, M., Karnieli, A., Sprintsin, S., Cohen, S. and Berliner, P.: Relationships between stand density and canopy structure in a dryland forest as estimated by ground-based measurements and multi-spectral spaceborne images, Journal of Arid Environments, 73(10), 955–962, doi:10.1016/j.jaridenv.2009.04.011, 2009b.

Suyker, A. E. and Verma, S. B.: Interannual water vapor and energy exchange in an irrigated maize-based agroecosystem, Agricultural and Forest Meteorology, 148(3), 417–427, doi:10.1016/j.agrformet.2007.10.005, 2008.

Tillman, F. D., Callegary, J. B., Nagler, P. L. and Glenn, E. P.: A simple method for estimating basin-scale groundwater discharge by vegetation in the basin and range province of Arizona using remote sensing information and geographic information systems, Journal of Arid Environments, 82(0), 44–52, doi:10.1016/j.jaridenv.2012.02.010, 2012.

Trambauer, P., Dutra, E., Maskey, S., Werner, M., Pappenberger, F., van Beek, L. P. H. and Uhlenbrook, S.: Comparison of different evaporation estimates over the African continent, Hydrology and Earth System Sciences, 18(1), 193–212, doi:10.5194/hess-18-193-2014, 2014.

Tsarouchi, G. M., Buytaert, W. and Mijic, A.: Coupling a land-surface model with a crop growth model to improve ET flux estimations in the Upper Ganges basin, India, Hydrol. Earth Syst. Sci., 18(10), 4223–4238, doi:10.5194/hess-18-4223-2014, 2014.

Wilson, T. B. and Meyers, T. P.: Determining vegetation indices from solar and photosynthetically active radiation fluxes, Agricultural and Forest Meteorology, 144(3–4), 160–179, doi:10.1016/j.agrformet.2007.04.001, 2007.

Zemp, D. C., Schleussner, C.-F., Barbosa, H. M. J., van der Ent, R. J., Donges, J. F., Heinke, J., Sampaio, G. and Rammig, A.: On the importance of cascading moisture recycling in South America, Atmos. Chem. Phys., 14(23), 13337–13359, doi:10.5194/acp-14-13337-2014, 2014.

Ziv, B., Saaroni, H., Pargament, R., Harpaz, T. and Alpert, P.: Trends in rainfall regime over Israel, 1975–2010, and their relationship to large-scale variability, Reg Environ Change, 14(5), 1751–1764, doi:10.1007/s10113-013-0414-x, 2014.

Table 1. Description of the 16 selected FLUXNET sites. Horizontal line divides between the six FLUXNET sites in PA systems (Top) and the nine FLUXNET sites in AN systems (Bottom). Plant functional types (PFT) include CSH: closed shrublands, WDL: woodland, SAV: savannah, ENF: evergreen needle-leaved forest, WSA: woody savannah, CRO: croplands, and GRA: grasslands. Mean annual precipitation (*P*) is in mm yr^{-1} for the years in which ET data was used (Period).

Site ID	Lat	Lon	PFT	Main species	*P*	Period	Reference
ES-Amo	36.83	-2.25	OSH	Dwarf shrubs	200	2009–11	Chamizo et al. (2012)
IL-Yat	31.35	35.05	WDL	*Pinus halepensis*	300	2003–09	Maseyk et al. (2008)
ES-LMa	39.94	-5.77	SAV	*Quercus ilex*	660	2004–09	Casals et al. (2009)
ES-ES	39.35	-0.32	ENF	*Pinus halepensis*	580	2001–06	Reichstein et al. (2007)
FR-Lbr	44.72	-0.77	WSA	*Pinus pinaster*	825	2004–08	Reichstein et al. (2007)
US-Blo	38.90	-120.63	ENF	*Pinus ponderosa*	1350	2001–06	Sims et al. (2006)
ES-ES2	39.28	-0.32	CRO	Rice	620	2005–08	Kutsch et al. (2010)
IT-Cas	45.07	8.72	CRO	Rice	960	2007–10	Skiba et al. (2009)
US-Bo1	40.01	-88.29	CRO	Corn–soybeans	795	2001–06	Hollinger et al. (2005)
US-Ne1	41.17	-96.48	CRO	Maize	590	2002–04	Suyker and Verma (2008)
US-Ne2	41.16	-96.47	CRO	Maize–soybean	590	2002–04	Suyker and Verma (2008)
US-Ne3	41.18	-96.44	CRO	Maize–soybean	590	2002–05	Suyker and Verma (2008)
US-Var	38.41	-120.95	GRA	C3 grass & herbs	465	2003–09	Baldocchi et al. (2004)
US-Kon	39.08	-96.56	GRA	C4 grasses	660	2007–12	Craine et al. (2012)
US-Wkg	31.74	-109.94	GRA	C4 grasses	190	2005–07	Scott et al. (2010)
US-Goo	34.25	-89.87	GRA	C4 grasses	1300	2003–06	Wilson and Meyers (2007)

Table 2. Water balances from six catchments along the north to south rainfall gradient in the Eastern Mediterranean (Fig. S2). Catchments area is in 10^3 ha. Precipitation (*P*), discharge (*Q*) and calculated ET as $P - Q$, are all in mm yr^{-1}. Fluxes were averaged over the years 2000 – 2013. MA-N, MA-CS and MA-S stand for the northern, central-southern and southern parts of the Mountain Aquifer of Israel, respectively, as defined by the Hydrological Service of Israel (HSI).

Name	Area	*P*	*Q*	ET
Kziv	13	799	284	515
HaShofet	1.2	654	183	471
MA-N	59	615	193	422
MA-CS	93	592	202	390
MA-S	28	619	257	362
Mamashit	6	130	28	102

Table 3. Correlations coefficients (*R*) of 16-day ET averages and MODIS-derived NDVI, EVI and daytime LST (LST, °C); and of annual ET and NDVI, EVI and LST from the 6 FLUXNET sites in PA systems (perennials and annuals vegetation systems, i.e. forests, woodlands, savannah and shrublands). Statistically significant correlations at $p<0.05$ were indicated by * while ** indicates $p = 0.06$ and *** $p = 0.07$.

	NDVI		EVI		LST	
Site ID	16-day	Annual	16-day	Annual	16-day	Annual
ES-Amo	0.63^{*}	0.89	0.62^{*}	0.71	-0.51^{*}	-0.33
IL-Yat	0.62^{*}	0.88^{*}	0.70^{*}	0.89^{*}	-0.36^{*}	-0.84^{*}
ES-LMa	0.17^{**}	0.93^{*}	0.28^{*}	0.80^{**}	-0.22^{*}	-0.93^{*}
ES-ES	0.41^{*}	0.91^{*}	0.30^{*}	0.94^{*}	-0.62^{*}	-0.32
FR-Lbr	0.36^{*}	0.85^{***}	0.68^{*}	0.93^{*}	-0.65^{*}	-0.63
US-Blo	0.17^{**}	0.92^{*}	0.46^{*}	0.66	-0.87^{*}	-0.59

Table 4. Correlations coefficients (*R*) of three VIs-based ET models using MODIS-derived NDVI, EVI and daytime LST (LST, °C). Results are for models using 16-day/annual data in AN (annual vegetation systems i.e., croplands and grasslands), and PA (perennials and annuals vegetation systems i.e., forests, savannah and shrublands) systems. All *R* were significant at $p<0.05$ except for the 16-day ET-LST simple regression in PA. Mean annual NDVI and EVI were regressed against annual ET using linear and exponential functions.

		AN		PA	
Model type	Variables used	16-day	Annual	16-day	Annual
Simple regression	NDVI (linear)	0.86	0.88	0.51	0.94
	NDVI (expo)	–	0.87	–	0.94
	EVI (linear)	0.86	0.79	0.61	0.95
	EVI (expo)	–	0.82	–	0.94
	LST	-0.42	-0.64	0.00^{ns}	-0.89
Multiple regression	NDVI, LST	0.87	0.89	0.71	0.94
	EVI, LST	0.87	0.79	0.73	0.96
Modified TG	NDVI, LST_{scaled}	0.87	–	0.78	–
	EVI, LST_{scaled}	0.87	–	0.80	–

Table 5. The mean absolute error (MAE) for Table 4. The 16-day MAE is in mm d^{-1}, while annual MAE is in mm y^{-1}. In parenthesis is the mean relative error (MAE/mean) in %.

		AN		PA	
Model type	Variables used	16-day	Annual	16-day	Annual
Simple regression	NDVI (linear)	0.51(32)	66(12)	0.65(47)	52(11)
	NDVI (expo)	–	83(15)	–	58(12)
	EVI (linear)	0.52(33)	79(14)	0.59(43)	53(11)
	EVI (expo)	–	90(16)	–	63(13)
	LST	0.94(60)	119(21)	0.78(57)^{ns}	74(15)
Multiple regression	NDVI, LST	0.51(32)	63(11)	0.57(41)	52(11)
	EVI, LST	0.51(33)	79(14)	0.54(40)	49(10)
Modified TG	NDVI, LST_{scaled}	0.48(30)	–	0.47(34)	–
	EVI, LST_{scaled}	0.50(32)	–	0.45(33)	–

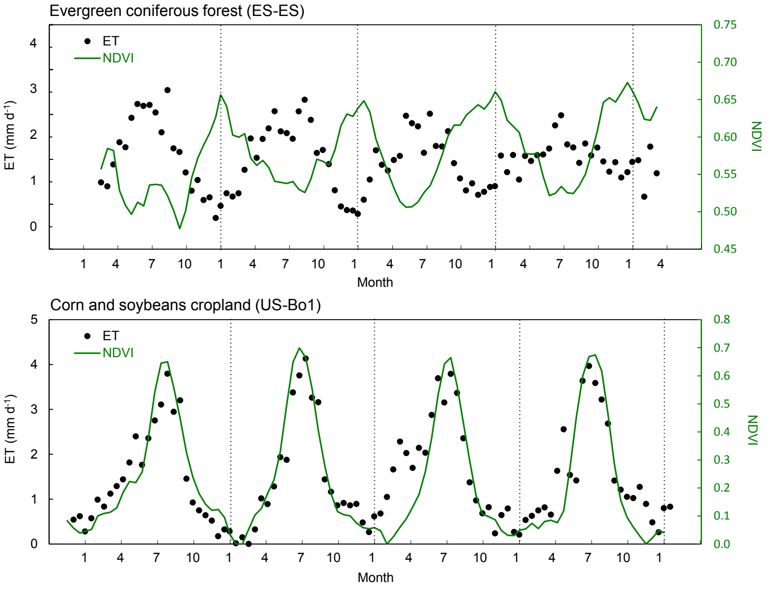


Figure 1. Sixteen-day ET averages and MODIS-derived NDVI at two vegetation systems: (Top) PA, i.e. comprising perennial and annual vegetation (evergreen coniferous forest), and (Bottom) AN, i.e. annual vegetation alone (corn and soybean cropland). Note: In the cropland site (Bottom) is the NDVI during the growing season after the annual minimum NDVI was subtracted.

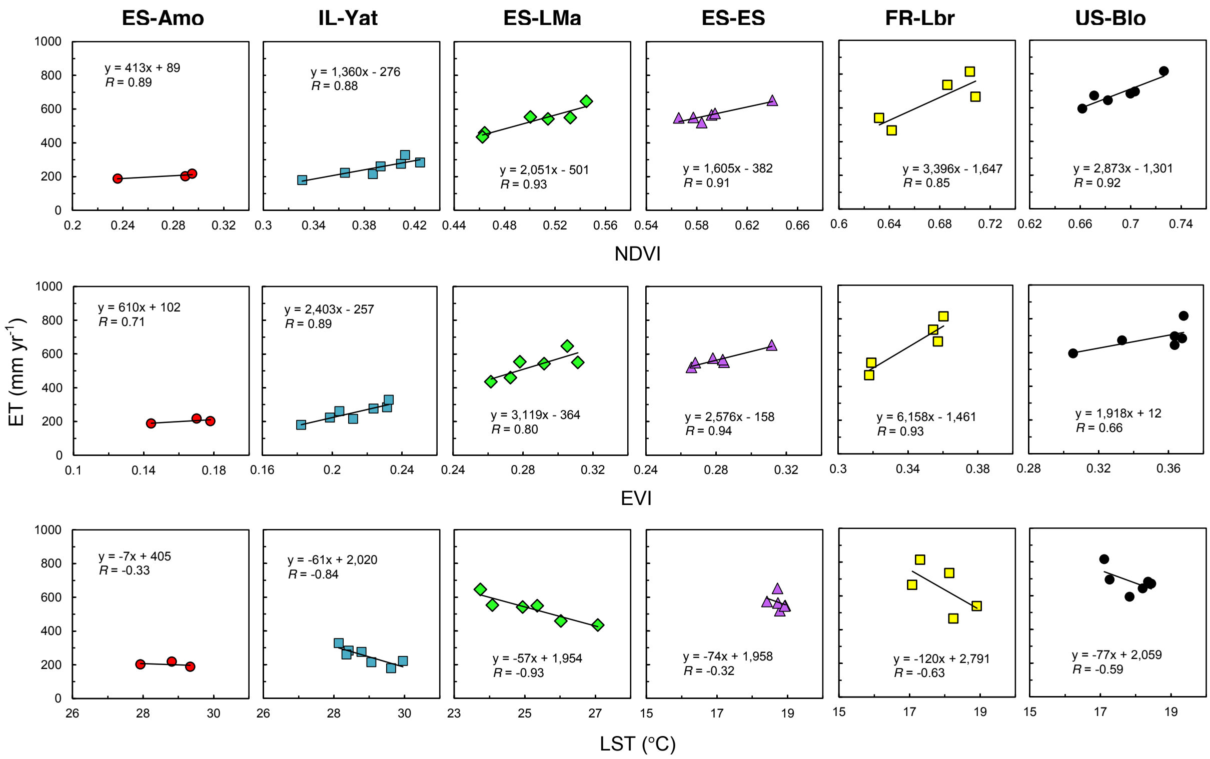


Figure 2. Relationships between annual ET (mm yr^{-1}) from eddy covariance towers and mean annual MODIS-derived NDVI, EVI and daytime LST (LST, °C) in PA sites (perennials and annuals vegetation systems, i.e. forests, woodlands, savannah and shrublands).

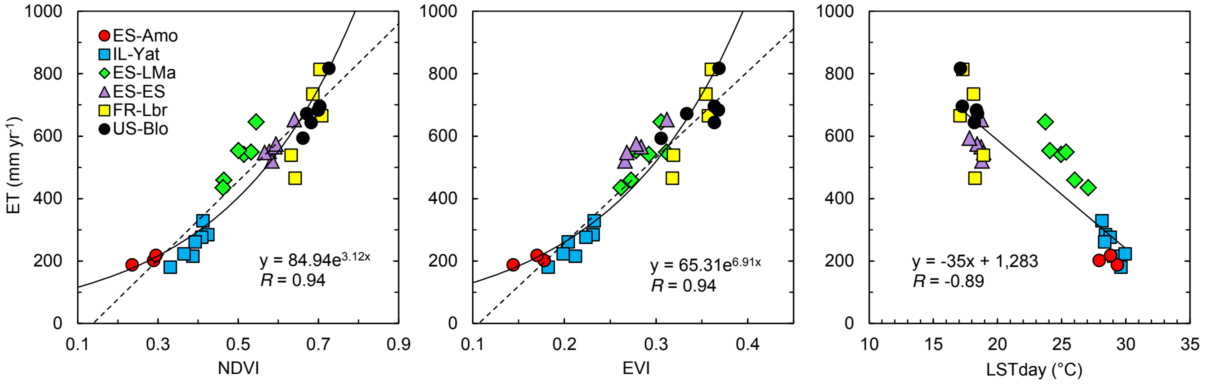


Figure 3. Same as Fig. 2 but for all PA sites together. The linear (dashed line) and exponential (solid line) functions are presented in ET-VIs with the *R* for the exponential function.

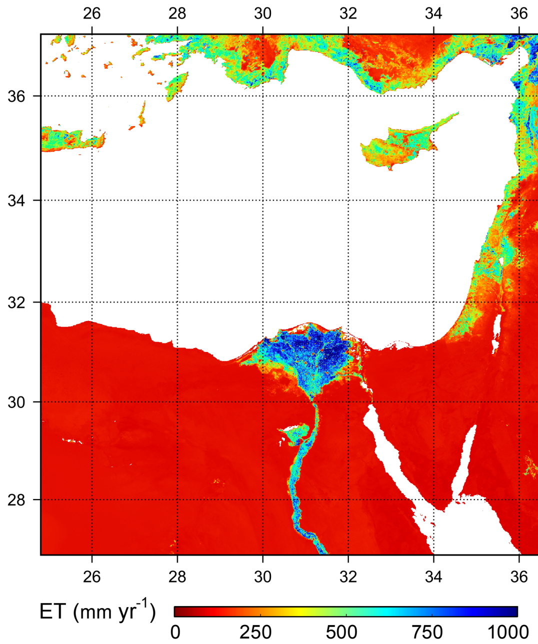


Figure 4. Mean annual ET from PaVI-E for the Eastern Mediterranean (2000–2014).

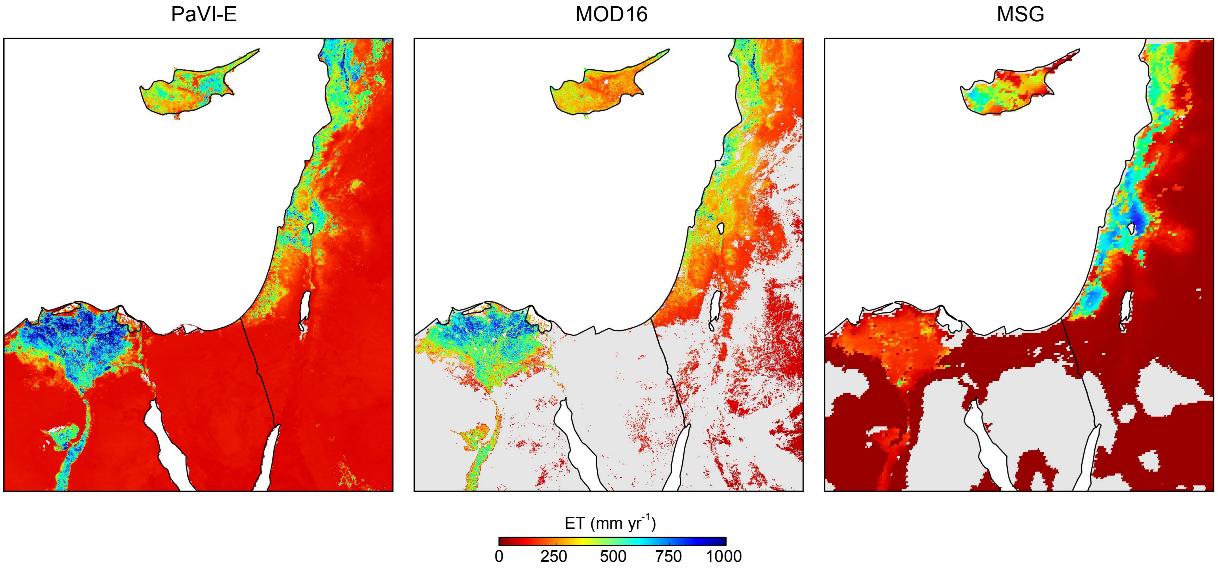


Figure 5. Maps of the total annual ET for the Eastern Mediterranean from PaVI-E, MODIS (MOD16) and MSG (LSA-SAF MSG ETa) for 2011. Grey colour indicates pixels with no data in MOD16 and MSG products.

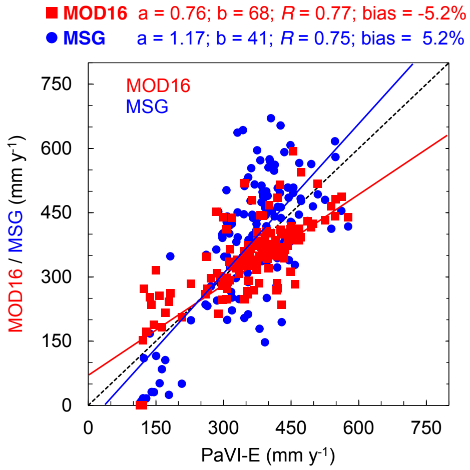


Figure 6. Total annual ET at 148 Eastern Mediterranean basins (Fig. S1) from MODIS (MOD16) and MSG (LSA-SAF MSG ETa) vs. PaVI-E. The slope (a), intersection (b), Pearson's (*R*) and relative bias (bias/mean) are also presented for each one of the linear regressions. Dashed line indicates the 1:1 ratio.

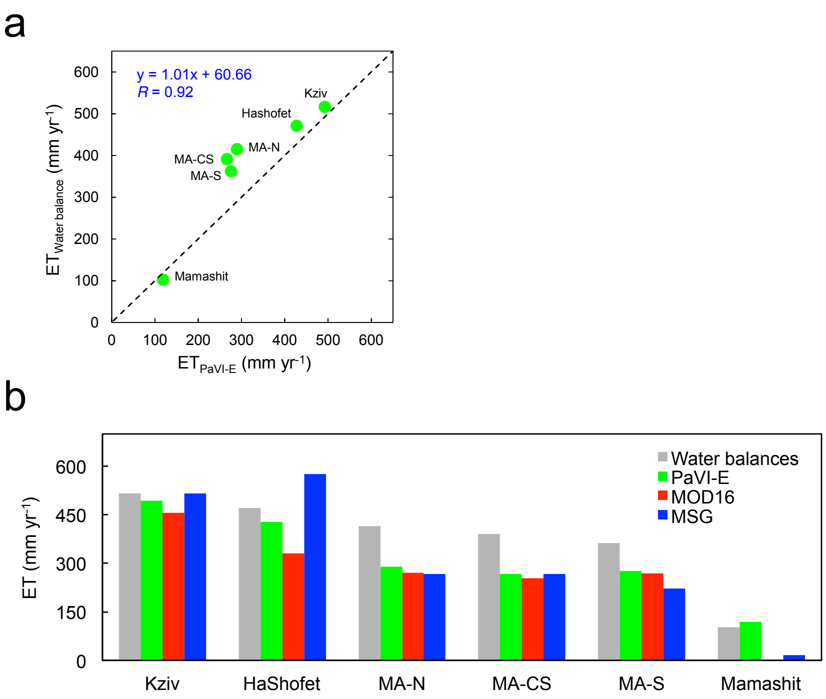


Figure 7. (a) Scatter plot of the mean total annual ET (2000-2013) retrieved from PaVI-E and calculated from water balances at six water catchments along the EM north – south rainfall gradient (Fig. S2). (b) Comparison between mean annual ET estimates from PaVI-E, MOD16, MSG and water balances in those six water catchments. MA-N, MA-CS and MA-S stand for the northern, central-southern, and southern parts of the Mountain Aquifer of Israel, respectively.

Supplement of

Annual evapotranspiration retrieved from satellites' vegetation indices for the Eastern Mediterranean at 250 m spatial resolution

D. Helman et al.

Correspondence to: D. Helman (davidhelman.biu@gmail.com)

Figure S1. Main basins (>10 km^2) in the Eastern Mediterranean, used to compare annual ET estimates from PaVI-E, MODIS (MOD16) and MSG (LSA-SAF MSG ETa) (see Fig. 5 in main text).

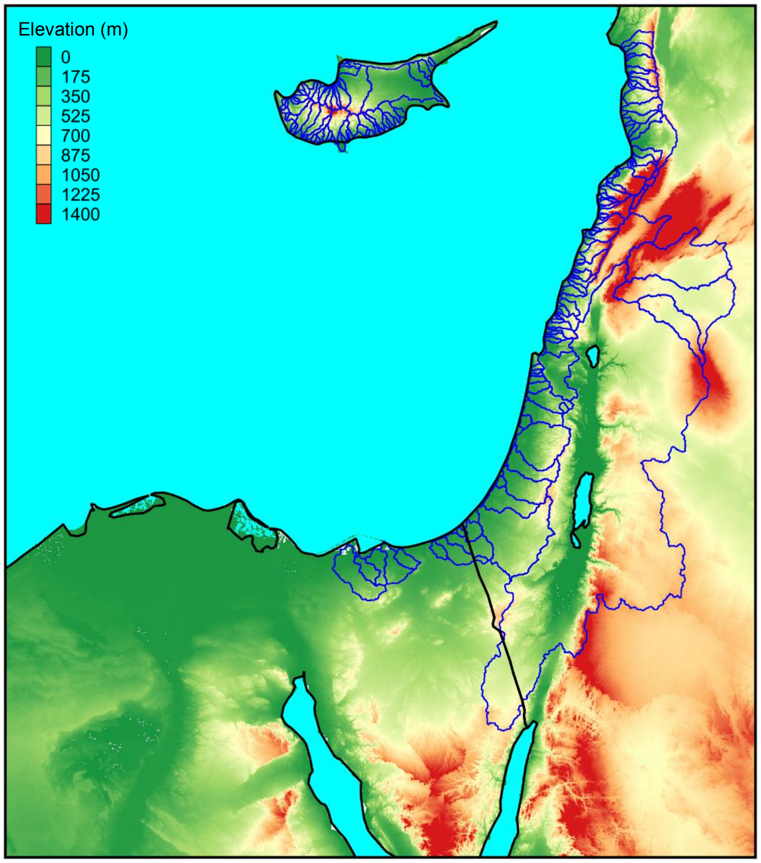


Figure S2. Location of six water catchments used to evaluate PaVI-E along the south-north rainfall gradient in the Eastern Mediterranean. ET from water catchments balances were calculated as $ET = P - Q$, for 2000-2013 (see in main text). MA-N, MA-CS, and MA-S are for north, centre-south and southern mountain aquifers, respectively. Contours of mean annual rainfall amount (isohyet) are also shown.

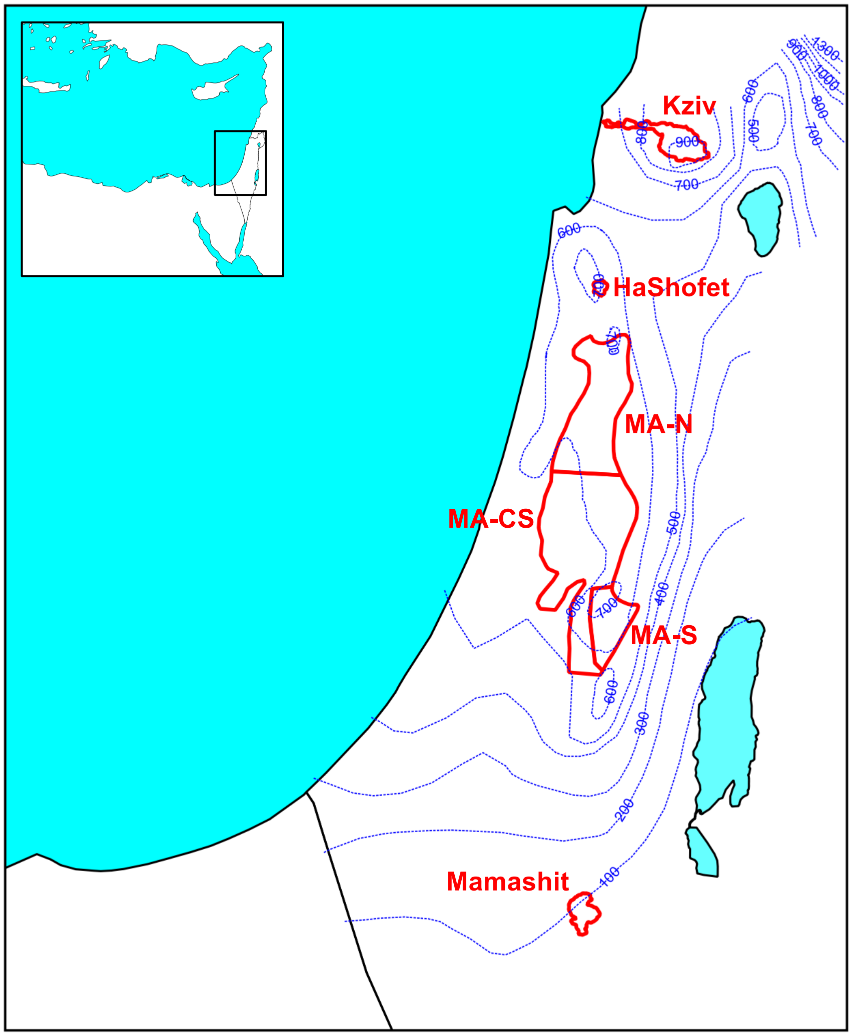


Figure S3. Seasonal and interannual time series of the eddy covariance ET at the PA sites. The vertical dashed line indicates the beginning of the year (from January 1).

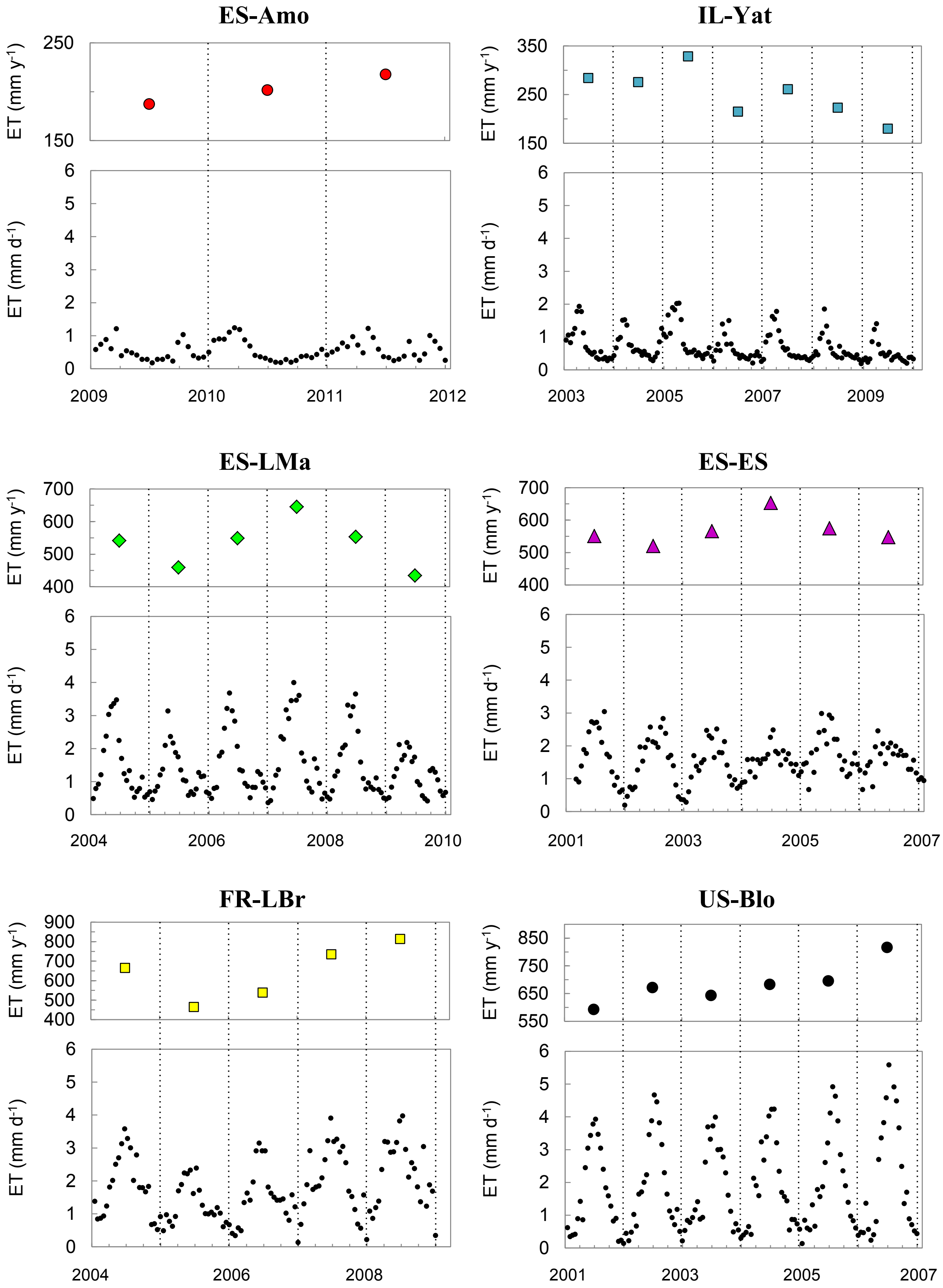
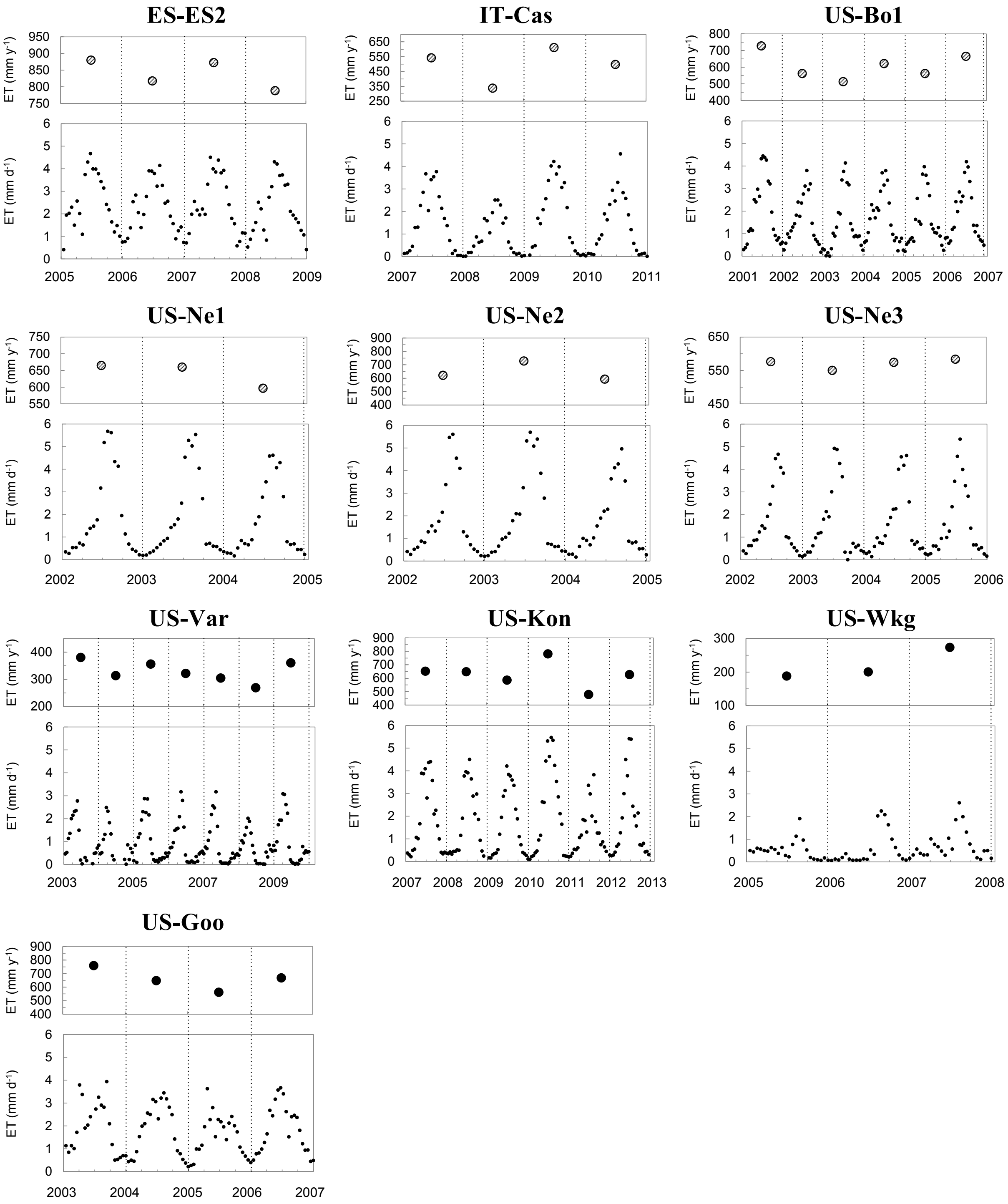


Figure S4. Seasonal and interannual time series of the eddy covariance ET at the AN sites (**CRO**: ES-ES2, IT-Cas, US-Bo1, US-Ne1, US-Ne2 and US-Ne3; **GRA**: US-Var, US-Kon, US-Wkg and US-Goo). The vertical dashed line indicates the beginning of the year (from January 1).



ES-ES2

IT-Cas

US-Bo1

US-Ne1

US-Ne2

US-Ne3

US-Var

US-Kon

US-Wkg

US-Goo

Figure S5. Mask map used to apply PaVI-E model for annual (AN) and annual + perennial (PA) vegetation systems in the Eastern Mediterranean. Pixels identified as AN are shown in red, while remaining land cover is classified as PA (see explanation in main text - Section 3.3). Water bodies are shown in cyan colour.

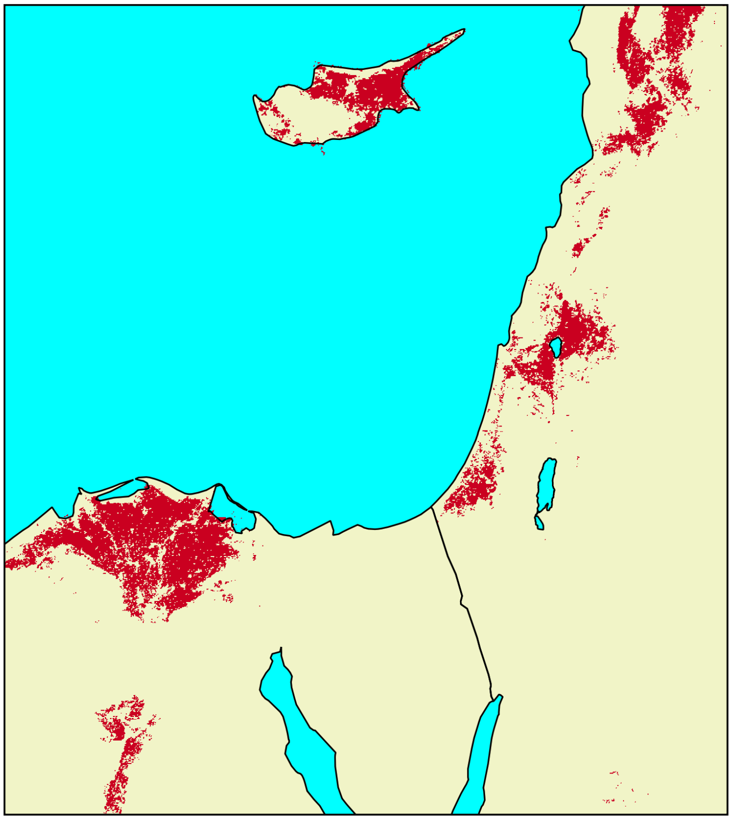


Figure S6. Scatter plots of the 16-day eddy covariance ET vs. MODISs' vegetation indices (NDVI and EVI) and the modified TG model (Sims et al., 2008), for **(A)** annuals vegetation systems (AN), and **(B)** perennials + annuals vegetation systems (PA).

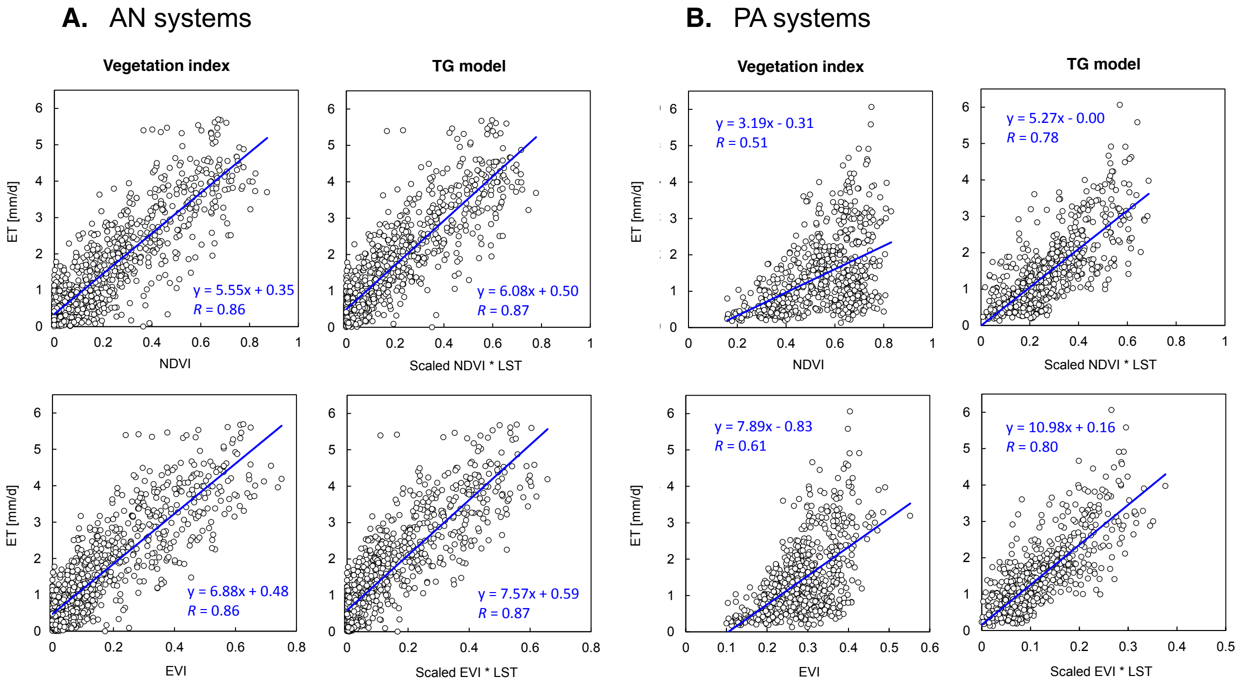
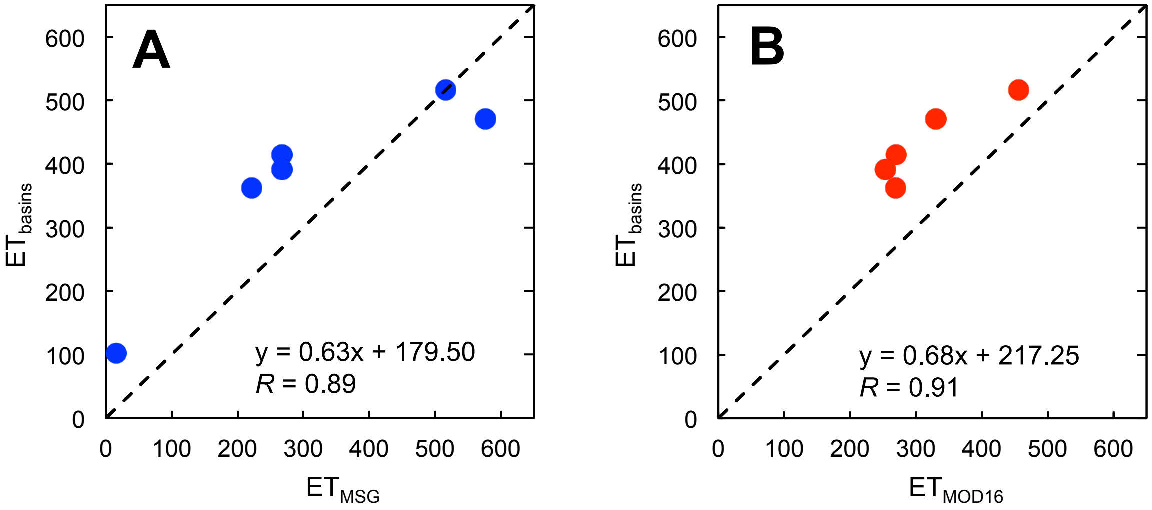


Figure S7. Scatter plot of the mean annual ET calculated from water balances against ET retrieved from (**A**) MSG and (**B**) MODIS ET products, at six water catchments along the EM north – south rainfall gradient. There were no ET values from MODIS for the area of Mamashit catchment.



Reference

Sims, D. A., Rahman, A. F., Cordova, V. D., El-Masri, B. Z., Baldocchi, D. D., Bolstad, P. V., Flanagan, L. B., Goldstein, A. H., Hollinger, D. Y., Misson, L., Monson, R. K., Oechel, W. C., Schmid, H. P., Wofsy, S. C. and Xu, L.: A new model of gross primary productivity for North American ecosystems based solely on the enhanced vegetation index and land surface temperature from MODIS, Remote Sensing of Environment, 112(4), 1633–1646, doi:10.1016/j.rse.2007.08.004, 2008.