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

Reviewer #2 comments:

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 <u>prior to publication</u> 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. <u>Using only empirical relationships between VIs and ET to retrieve ET is a novel idea</u>. 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. <u>Such relationships were not found before</u> 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 <u>after</u> 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 21-23(P01):

"After establishing empirical relationships, PaVI-E was used to retrieve ET_{Annual} at 250 m spatial resolution for the Eastern Mediterranean 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 17-19(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

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 23-24(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**. We also deleted the reference Helman and Lensky (2015) here and elsewhere because it has not been accepted for publication yet.

Lines 3-13(P06):

"Perennial and annual vegetation in Mediterranean regions have distinct phenology contributing differently to the VIs signal (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. 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)."

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(P08):

"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 18-19(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, 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)-3(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. The R for the annual ET-NDVI regression was higher (R = 0.88) than that for the ET-EVI regression (R = 0.79). 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 (Table 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). R for multiple regressions using annual data were 0.89 and 0.79 (ET against LST with NDVI or EVI, respectively) compared to 0.87 and 0.82 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 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. S4B). In AN, R from TG and 16-day ET–VIs regressions were not significantly different (p>0.1, Table 4 and Fig.S4A)."

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 **10(P12)** and added "a" to "such" in line **14(P12)**. We changed to "biases" in line **21(P12)** and replaced the slash by a parenthesis in line **24(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 **2(P13)** and to "ET-VI relationships" in line **20(P13)**. We deleted the word "Yet" in line **21(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 28(P13)-4(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. The advantage of PaVI-E is in its simplicity and spatial resolution (250 m) 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: