Dear Theresa Blume, thank you for handling our manuscript.

Below is a point-by-point response to the comments. Referees' comments are in black with our response in blue below each comment. The marked-up manuscript is found below the text.

Best regards,

Elisabeth K. Larsen with co-authors

Referee number 1.

The main issue I see with regards to the manuscript is that the authors could spend more attention on the variability of misalignment issues between installed trees and present in a clear way how they define zero flow conditions. Additionally, the implications of the correction could be more concretely quantified by showing time-series of daily water use, with and without the proposed correction. Finally, within the discussion there is space for further elaborating on the other issues related to the installment of these type of sensors. Clearly circumferential variability, wounding and other biases should be further investigated in the future.

We have gone through and corrected for the issues the referee points out. Because this is a technical note and we didn't want to increase the figures-to-text ratio too much, some of the additional figures are added in the supplementary materials. Figures are here shown below the text.

Specific comments:

Line 44-46: Please provide sources describing each of the method. Now readers cannot read related literature to explain the method.

We've included references in the text for each method mentioned, Line 40:

> "There are a range of different approaches to sap flow measurements, and methods vary between heat dissipation (HD), (Granier, 1987; Lu et al., 2004), steam heat balance (SHB), trunk segment heat balance (THB), (Smith and Allen, 1996), and heat pulse velocity (HPV), (Marshall, 1958). However, they are all based on tracing heat within the xylem (Burgess et al., 2001; Davis et al., 2012; Forster, 2017)."

Line 56-59: See also: Steppe et al. 2010 A comparison of sap flux density using thermal dissipation, heat pulse velocity. This is a relevant study which addresses the offset of sap flow methods from gravimetric measurements.

Thank you for providing the reference, even if the paper is comparing the overall heat pulse velocity technique rather than the specific heat ratio method, we found it very relevant and decided to include the reference, line 50:

> "By accounting for these sources of error and additionally estimating the stem moisture content and radial variability, the heat ratio method (HRM) has been evaluated the heat pulse velocity method (HPV) with highest accuracy, although with a tendency of underestimating transpiration values (Forster, 2017), an error that is shown to increase with higher sap flow values (Steppe et al., 2010)."

The paper also made us aware of an error of concept; using the term "sap velocity" and the unit cm h -1 is not precise, and we have therefore changed it throughout the paper to "sap flow" per unit of sap wood (cm^3 cm^{-2} h^{-1}).

Figure 1: It would be good if there would be a zoom in panel where you can see more detail on the patterns. The current point cloud does not give the reader a full idea on the diurnal quality of the data.

As suggested by the referee, we have included a zoomed in panel in Fig.1, representing one week of data for the HRP, slope and RSD, marked A1, B1 and C1. We agree that it enhance the understanding of the filtration as it gives the reader the diurnal pattern of the data and therefore demonstrates the range of values that has been filtered out.

Line 220: Please help the reader to understand what RSD is. I went back to the methods to check, yet this part of the text should be understandable on its own.

We included a definition for the RSD also in this part of the text so the reader doesn't need to go back in the text to check, line 230:

> "Therefore, the quality of the measurements was indicated by calculating the RSD, the relative standard deviation divided by the sample mean, and the slope versus time, for each HPR."

Line 237-238: It would be good to understand when and why there is a limited amount of periods with zeroflow conditions. Additionally, it is not clear how these period where exactly defined. It would help to include an appendix figures which details these periods and the underlying environmental conditions.

As suggested by the referee, a definition of zero-flow conditions was included. The definition now references five additional figures in the appendix, showing the relative extractable water and VPD along with the heat pulse ratios during the estimated zero-flow events, line 250:

> "Zero-flow conditions were assumed at night (22:00-03:00h solar time) during days when relative extractable water, REW > 0.75, and vapor pressure deficit (VPD) was close to zero (Fig. S2). Multiple readings were used to produce an average of each event. These conditions were limited to five occasions during our study period."

An additional paragraph was included in the method section due to the inclusion of environmental data, line 95-100:

> "**Environmental conditions**
>
> Air relative humidity (%) and air temperature (ºC) were registered every 30-minute (U23 Pro V2, Onset Computer Corporation, USA). Precipitation was registered using a rain gauge with 0.2 mm resolution (RG3-M, Onset Compute, USA). Three soil moisture probes were inserted at 20-25 cm depth to register soil water content, SWC, (S-SND.M005, Onset Computer Corporation, USA), using a datalogger for logging specifications (HOBO Micro Station, USA). To assume periods of zero-flow events, relative extractable water (REW) was calculated using the method of Bréda et al. (1995):
>
> $$REW = \frac{(\theta_t - \theta_{min})}{(\theta_{max-}\ \theta_{min})} \qquad [1]$$
>
> Were θ_t is the registered SWC, θ_{min} is the minimum SWC observed during the measurement period, and θ_{max} is the SWC at field capacity. When values of REW surpassed 1 they were converted to 1 according to Granier et al. (2000)."

Additionally, now the displacement is provided for one tree and the average of all trees and sensors. Yet, it would be good to see whether there are differences between the sensors themselves. Would the authors be able to provide the change presented in Figure 2 for each sensor?

On request from the referee, we decided to include the change described in Figure 2 for each sensor. We also included a better definition of the sensor to clarify that each sensor consists of three needles, two probes and one heater. When referring to misalignment we refer to both probes in a sensor.

> Line 85:

> "Each sap flow sensor consists of three needles: one heater and two thermocouples. We will refer to the thermocouples as probes, and when using the term "sensors" we refer to both probes and the heater."

Regarding the misalignment, line 250:

> "The outputs indicate a clear shift of placement in each of the probes over time, here denoted as x1 and x2 in each pine (Fig. 2). The eight probes, two per tree, all deviated from the ideal of 0.6 cm. Probe x2 in pine number 1 had the highest inaccuracy with an initial value close to 0.3 cm".

> Line 263-265: It would be interesting to see temporally what the difference are in daily water use (L d-1). This will clarify if the offset due to misalignment is progressively getting worse or whether, in these species, the impact is not that bad. Additionally, when presenting these numbers, it is critical that the standard deviation is also provided for these estimates.

The authors originally didn't include the difference of daily water use (L d^{-1}) because this introduces the errors related to upscaling to whole tree transpiration, whereas we wanted to focus on the correction of the point measurements. However, we chose to include three weeks of data towards the end of the measurement period, also including the transpiration values without any misalignment correction (Fig. 6). This also highlighted a very small difference between the "non-correction" method and "time-dependent-correction" when the misalignment was small or converged towards the ideal distance of 0.6 cm between the probes. In the original manuscript, Fig. 5 represents how the difference between the two methods changes in time and is meant to give an indication of how the misalignment estimation is getting progressively worse if the misalignment is measured only once at the beginning of a measurement campaign. Line 266-268 in the original manuscript shows the overall difference between using the correction during the entire study period of 20 months. Standard deviation was provided for the estimates in table 1.

> Line 296-299: Indeed, there could be a reduction in the amplitude due to wounding effects or other changes within the stem. Did the authors analyse whether they would see a reduction in the amplitude over time? It would be important to make this test as the data is available.

We have taken the suggestion of the referee into consideration and included the test results in the paper and the figures in the supplementary material, line 320-325:

> "Due to the variation in rainfall between the two years, SWC differed significantly on similar calendar dates with the exception of six days in June (9 – 15). When comparing the relationship of sap flow versus VPD on these days, there was an increase in the slope the second year: 2.3 for 2017 and 2.9 for 2018. This, we attributed to an overall weaker correlation between VPD and sap flow in 2017; R 2 = 0.4* in 2017 versus R 2 = 0.6* in 2018, due to higher values of VPD in 2017 (S3). Within the normal variability of the measured data we concluded that the sap flow values had not decreased in the second year, when compared to the first year, under similar environmental conditions."

Line 316-318: This is indeed a valid point, yet I would propose that the authors would elaborate on the fact that reinstalling sensors along the stem will introduce change due to circumferential variability in the stem. This could be critical when generating continuous series of sap flow over the long term.

We appreciate the referee pointing this out, as it is an important point to make which we now have included in the text, line 350:

> "However, as pointed out by Moore et al. (2010), re-installing the sensors might create a shift in the data due to spatial variation within the tree. Leaving the same sensors in the tree throughout the study period avoids this problem and enables the study to focus on the intrinsic factors affecting the sap flow rates"

Also, do the authors think these results found on conifers are universal for all types of species? I would have expected a short discussion to clarify to the reader why these findings could be of general value to the application of the method.

We agree that it's an important addition to the discussion and we have therefore included it in the text, line 390 - 340:

> "The time-dependent correction method could be useful with any species already tested with the HRM, where misalignment of probes can create a source of error and sensors are installed over a longer period. Specifically, where the movement of the wood might cause further displacement of the sensor."

Response to referee number 2:

Overview: Larsen et al. in this Technical note address two important issues of sap flow measurements with the heat pulse method, namely (i) data filtering/quality control of the raw heat pulse records, and (ii) errors due to misalignment of the sap flow probes. The authors suggest some statistical thresholds/filters to be applied in the raw heat pulse ratios for data cleaning and present a time-dependent correction to account for probe misalignment. Moreover, they demonstrate the importance of such uncertainties for robust transpiration estimates, by presenting sap velocity and transpiration estimates with and without applying the proposed correction. I find the study topical and of interest for the scientific community working on transpiration estimates with the heat pulse sap flow method. However, I feel that the manuscript needs major revisions to better present the motivation and rationale of the study, the developed methods, and the broader implications of the obtained results.

To better present the motivation of the study we have made some changes in the text, line 55-60:

> "Previous studies have suggested additional solutions for probe misalignment (Ren et al., 2017), for determining thermal diffusivity (Vandegehuchte and Steppe, 2012), and correcting for heterogeneous heat capacity in wood (Becker and Edwards, 1999). However, there is no recent recommendation for how long newly deployed sap flow sensors can be employed. Some studies have shown how sensor probes inserted into the xylem can dampen the signal due to blocking or destruction of vessels (Moore et al., 2010; Wiedemann et al., 2016). One way to account for changes over time has been to reinstall sensors throughout the study period (Moore et al. 2010), however there is little information to be found on the exact interval for which this needs to happen (Vandegehuchte and Steppe, 2013), and for continuous measurements this will interrupt the dataset (Moore et al. 2010). Therefore, we aim to find a

> quality indication that can ensure that the readings don't deteriorate over time, or if they do, that it would be detectable. Attention should be given to check the accuracy of the heat pulse ratio itself, in which the rest of the methodology is built on. In addition, to allow for sensors to be employed over longer periods it's necessary to develop a dynamic probe misalignment correction method due to observed change in probe position over time."

The text needs significant editorial improvements to eliminate vague/unclear wording and grammatical errors. Moreover, several parts of the text (including the abstract) need to be revised/rephrased/rewritten to improve the clarity of the text and better communicate the design of the study, results, discussion and conclusions. I have highlighted few specific points below (see Specific comments), yet several other cases exist throughout the manuscript.

We have corrected specific comments and rewritten phrases in the abstract, methods, discussion and conclusion. We hope this has led to more clarity and improved the communication of our study. Specifically we rewrote the conclusion, line 355-365.

The methods need to be revised and clarified. In some parts, there are inconsistencies and it is hard to follow. Sometimes the authors refer to V as sap velocity (L130) and other times as heat pulse velocity (L113, L135). Please clarify and use consistently the terms/variables/abbreviations throughout the manuscript. Also, the selected thresholds (L153-161) for the raw data filtering need to be better justified, since at the moment seem quite arbitrary or could be interpreted as case-specific. Also, the data from all eight sensors (or averages across trees, since two sensors per tree were deployed) should be presented, either in the main text, or in the supplementary material. Apart from Fig 5, all figures illustrate data from a single sensor. In addition, more details should be provided in the methods session on how the positions of sensor misalignment were estimated in Fig 2 (and the misalignment for all eight sensors would be interesting to be illustrated, too).

Inconsistencies highlighted by the referee has been addressed and corrected for. Further inconsistencies or vague formulations has been checked throughout the manuscript. We argue that the selected threshold needs to be case-specific because it depends on both wound width and sap velocity. However, our suggested threshold is within the magnitude of the threshold observed by Burgess et al. (2001). We have elaborated this justification in the text, line 165-170:

> "The magnitude of the threshold chosen for the slope was taken from the modelled output of instantaneous ratios performed by Burgess et al. (2001), were low sap flow velocities (5 cm h-1) combined with a small wound width (0.17 cm) were shown to display a slope of 0.001. Due to our low sap flow velocities (< 15 cm h^{-1}) and small probe diameter (0.12 cm), we expected the slope to be as close to 0.001 as possible. The specific threshold of 0.003 was decided upon inspection of the natural variability of the measurements and can be modified according to needle size and magnitude of the sap velocities. According to Burgess et al. (2001), higher values of slope (0.01) can be expected with greater wound width and higher velocities."

In relation to the referees' request for details regarding how the misalignment was estimated in figure 2, we included the exact equation in the method section as equation 6. We have also decided to include the misalignment from all 8 probes (Fig. 2).

I feel that the hydrological community and the readership of this Journal, would appreciate also some figures with the up-scaled transpiration estimates and the resulting biases do to probe misalignment, complementing the existing figures with the sap velocities and the results presented in L265-268.

We decided to combine the answer for this request with number the request for comparison between sap flow with and without corrections. A comparison is now included as figure 6 expressed as transpiration estimates (L tree -1 h-1) for each pine over a three-weeks period at the end of the study.

The suggested time-depended correction accounts for two effects: probe misalignment and wounding effects. The current experimental design does not allow to disentangle the two. Therefore, the text should be revised so it is clear that the proposed correction addresses issues related to both wounding and probe misalignment.

The text was revised accordingly, line 145-150.

I suggest to include a comparison between sap velocities/transpiration estimates averaged throughout the study period/growing season as calculated with (i) no wounding correction, (ii) traditional (no time depended) corrections, and (iii) the presented time-depended correction. This would better emphasise/illustrate the advantages of this Technical note.

We appreciate the suggestion and agree that this would highlight the advantage of our study. We decided to include three weeks' worth of data towards the end of the study period to illustrate the biggest differences. This also highlighted a very small difference between the "non-correction" method and "time-dependent correction" when the misalignment was small, or converged towards the ideal distance of 0.6 cm between the probes, as Figure 6 in text.

Specific comments: Abstract: the study location, tree species, number of instrumented trees, study period should be clearly stated in the abstract. L16: 'Whole-plant transpiration' reads redundant, just 'Transpiration' should be enough here. L17: and Hydrology. L18: 'wide application range' and L19: 'ready automation': unclear what you mean here. Please consider revising/rephrasing. Similar for 'data readings', I guess what you mean here is the sap flow sensors can provide long-term measurements of sap flow in tree stems with high temporal resolution (e.g., minutes, hours etc.). L19: 'Several different': reads redundant. 'Several methods' or 'Different methods' should be enough. Le20: how the methods were adjusted to different climatic conditions? Unclear statement. Maybe 'tested' instead of 'adjusted'? L21: 'in the method', unclear to which method you are refereeing to, here. Please rephrase/revise. L21-22: if you focus only on the heat pulse method, then that is probably fine, but if you are referring to sap flow methods in general, then additional sources of uncertainty should be listed here, e.g., Granier's empirical parameters, zero-flow conditions.

We rephrased the abstract according to the specific comments proposed by the referee.

We have further corrected for all the specific comments mentioned by the referee. In addition, we have gone through the manuscript to make sure of the consistency of terms and expressions.

L90-100: mentioned that you deployed eight sensors in total, two per tree. I found this information further below in the text, but this has to be very clear from the methods session.

We apologise for not using consistent terminology when referring to the sensors. There is one sap flow sensor per tree. Each sap flow sensor consists of two probes and one heater. When calculating the misalignment, we refer to each probe, of which there are eight. We have now declared a definition in material and methods, 85:

> "Each sap flow sensor consists of three needles: one heater and two thermocouples. We will refer to the thermocouples as probes, and when using the term "sensors" we refer to both probes and the heater"

L94-97: mention the specific depth where the thermocouples are located, and thus the heat velocity is measured. I found this information mentioned in a figure caption (L229) but has to be included in the methods description.

The depth of which the thermocouples are located is described in material and methods in the original paper (L94-97):

> "The sensors were drilled into the uphill side of each tree trunk. Since P. halepensis has a higher sap flow average near the cambium with the flow steadily declining nearer to the heartwood (Cohen et al., 2008), sensors were installed at 20 mm depth below the cambium for average sap velocity rates, as estimated by Manrique-Alba (2017)."

The figure caption (L229) refers to the vertical distance between the heater and each of the thermocouples. This information is also included in materials and methods in the original manuscript (L98). However, we understand that this information can be interpreted as the depth, and we have therefore added to the caption:

> "Figure 1. (A) Heat pulse ratio (HPR) throughout the measurement period in 30-minute intervals in tree number 1. Each HPR is an average of 41 instantaneous ratios corresponding to the temperature difference in two thermocouples at 0.6 cm up-and downstream from a heater probe at 0.2 cm depth:"

L138: you are referring to the raw heat velocities here I assume and not to sap flow measurements. Here and throughout the text clarify and use carefully and consistently terms such as heat velocity, sap velocity, and sap flux density.

We included a clarification in the specific phrase, line 140:

> "On these premises, we have built a methodology utilising a quality check of systematic sap flow measurements by means of a statistical analysis performed on the instantaneous heat pulse ratio acquired between 60 and 100 seconds after the release of a heat pulse."

To me be more precise, we have gone through the whole text and decided to go away from the term sap velocity (cm h-1) and use the term sap flow (cm^3 cm^{-2} h^{-1}), denoting the sap volume flowing per square centimetre of sapwood per hour. This also makes it clearer to distinguish from heat pulse velocity.

Additional changes:

- After adding more figures as suggested by both the referees, we decided that figure number 3 in the original manuscript was superfluous. We realised that the figure, showing heat pulse velocities from the whole study period, gives an unclear picture of the actual baseline due to the large amount of data points. The zoomed-in panels of the same data (Figure 4 in the original manuscript), gives the same information for one week at the time.
- Another paragraph was added in the result section to complement the graph showing the transpiration values, and the differences between the values in the new manuscript (line 295-100).

Technical note on long-term probe misalignment and proposed quality control using the heat pulse method for transpiration estimations

Elisabeth K. Larsen[1,3], Jose Luis Palau[1], Jose Antonio Valiente[1], Esteban Chirino[2] and Juan Bellot[3,4].

[1]Mediterranean Centre for Environmental Studies (Foundation CEAM), Charles R. Darwin 14, Parque Tecnològico, 46980 Paterna, Valencia, Spain

[2]Faculty of Agricultural Sciences, Lay University Eloy Alfaro of Manabí, Ave. Circunvalación, Vía San Mateo, Manta 130802, Ecuador

[3]Department of Ecology, University of Alicante, Apdo. 99, E-03080 Alicante, Spain

[4]IMEM Ramón Margalef, Department of Ecology, Faculty of Sciences, University of Alicante, 03080 Alicante, Spain

Correspondence to: Elisabeth K. Larsen (eli@ceam.es)

1 Abstract

~~Whole-plant T~~transpiration is a crucial component in the hydrological cycle and a key parameter in many disciplines like agriculture, forestry,~~ and~~ ecology and hydrology. Sap flow measurements are one of the most widely used methods to estimate whole-plant transpiration in woody species due to its ~~wide application range~~applicability in different environments and in a variety of species, as well as ~~and being easily programmed to obtain~~having the capacity of ~~its ready automation for~~ continuous high temporal resolution

~~data readings~~measurements. Several~~Several different~~ methods have been developed and tested~~adjusted~~ under~~to~~ different climatic conditions and wood properties. However, the scientific literature also identifies several sources of error ~~in the method that~~when using sap flow measurements that needs to be accounted for; misalignment of the probes, wound to the xylem, thermal diffusivity and stem water content. This study aims to integrate probe misalignment as a function of time to improve ~~readings~~ measurements during long-term studies ~~measurements~~ (> 3 months). Heat ratio method (HRM) sensors were installed in four *Pinus halepensis* during 20 months, in a coastal valley in South-Eastern Spain (39º57'45" N 1º8'31" W) in a Mediterranean climate. We conclude that even when geometrical misalignment~~s~~ errors are small, the introduced corrections can generate~~imply~~ an important shift in sap flow estimations. Additionally, we propose a new set of statistical ~~record~~ information to be recorded during the measurement period, which can be ~~to~~ used as a quality control of the heat ratio readings obtained from the sensors. Relative ~~S~~standard deviation and slope against time of the averaged heat pulse ratio, was used as quality indicators to conclude that no general time limit can be decided for the longevity of the sensors, but should rather be determined from individual performance over time.

Formatted: Font: Italic

2 Introduction

Plant transpiration is a key process in the hydrological cycle, and in forest ecosystems it is often the largest component of total evapotranspiration (Jasechko et al., 2013). Accurate estimations of transpiration are still difficult to obtain, and field assessments of transpiration estimations are therefore crucial in hydrological planning as well as in forestry, ecophysiological research and climate forecasting. Sap flow measurements are~~is~~ one of the most widely used techniques to estimate ~~whole plant~~ transpiration in woody species as it is readily automated for continuous readings and ~~is~~ not limited to single leaf measurements (Forster, 2017). Although some sap will go to stem and leaf storage, it is estimated that 99% is lost through transpiration, and sap flow measurements can therefore be used to directly estimate transpiration values (Forster, 2017). ~~In addition, sap flow sensors estimate plant transpiration rates regardless of the orographic complexity and atmospheric conditions of stability or stratification, which can hinder the direct measurement of transpiration flows using in forest environments.~~

There are a range of different approaches to sap flow measurements, and methods vary between heat dissipation (HD) (Granier, 1987; Lu et al., 2004), ste~~a~~m heat balance (SHB), trunk segment heat balance (THB) (Smith and Allen, 1996), and heat pulse velocity (HPV) (Marshall, 1958). However, they are all based on tracing heat within the xylem (Burgess et al., 2001; Davis et al., 2012; Forster, 2017). Marshall (1958) developed a theoretical method to determine sap flow from ~~the~~ thermal diffusion and dissipation theory of heat pulses in heterogeneous material. His theory relies on calculating the heat ratios measured in two parallel thermocouples aligned vertically and symmetrically with respect to a line heater along the direction of the xylem. Burgess et al. (2001) developed an improved HPV technique, termed the heat ratio method (HRM), based on the Marshall (1958) methodology, which is sensitive to the direction of sap flow and is capable of measuring low and reverse rates (Burgess et al., 2001)~~, which is not possible using energy balance methods like eddy-covariance and which often is the case for nightly values (Burgess et al., 2001; Novick et al., 2009).~~ . The HRM is therefore more appropriate to use in water-deficit environments were flow rates are low. Burgess et al. (2001) developed two steps of corrections for sap flow calculations by considering probe misalignment and wounding (caused by the implementation of the sensors). By accounting for these sources of error and additionally estimating the stem moisture content and radial variability, the ~~heat ratio method (~~HRM~~)~~ has been evaluated as the ~~heat pulse velocity method (~~HPV-method~~)~~ with the highest accuracy, although with a tendency of underestimating transpiration values (Forster, 2017), an error that is shown to increase with higher sap flow values (Steppe et al., 2010).

Previous studies have suggested additional solutions for probe misalignment (Ren et al., 2017),~~ or~~ for determining thermal diffusivity (Vandegehuchte and Steppe, 2012), and correcting for heterogeneous heat capacity in wood (Becker and Edwards, 1999)~~. However, we suggest attention should be given to check the accuracy of the heat pulse ratio itself, in which the rest of the methodology is built on~~. However, we could not find any recent recommendation for how long the sensors could be employed. Some studies has shown how, when sensor probes are inserted into the xylem, it can dampen the signal due blocking or destruction of vessels (Moore et al., 2010; Wiedemann et al., 2016).~~,~~ ~~In addition, we found it necessary to develop a dynamic probe misalignment correction method.~~ One way to account for changes over time has been to reinstall sensors throughout the study period (Moore et al. 2010), however there is little information to be found on the exact interval of which this needs to happen (Vandegehuchte and Steppe, 2013), and for continuous measurements this will interrupt the dataset (Moore et al. 2010). Therefore, we wanted to find a quality indication that would ensure that the readings didn't deteriorate over time, or if they did, that it would be detectable.~~,~~ Attention should be given to check the accuracy of the heat pulse ratio itself, in which the rest of the methodology is built on. In addition, we found it necessary to develop a dynamic probe misalignment correction method due to observed change in probe placement over time. The adjustment proposed here, which is built on the calculations of Burgess et al. (2001), is necessary when monitoring transpiration continuously for more than 3 months because, wood properties, the heterogeneity of the xylem, and plant tissue growth, might further misplace the sensor after its implementation (Barrett et al., 1995). Burgess et al. (2001) corrected for linear probe misalignment *in situ*. This correction must be effectuated during zero flow. Actual probe placement is therefore suggested to be calculated one time, usually at the end of the experiment when the root system can be severed to enforce zero flow, to correct for the misalignment (Burgess et al., 2001). However, this solution is not suitable for long-term measurements, as the misplacement will change over time (Ren et al., 2017), or when intrusive methods are not an option. Thus, the objective of this research is firstly, to develop a statistical filtering~~ration~~ method to ensure the quality and consistency of the measurements over time, and~~,~~ secondly, to implement a modified version of the probe-placement ~~misalignment~~ calculation for sap flow series longer than 3 months.

This methodological paper is structured in two parts: the first one, dealing with the statistical analysis of long-term time series of averaged heat pulse ratios to ensure the quality of data and their stability over time; the second part, proposing an adaptation of the method developed by Burgess et al. (2001) to propose~~contemplate~~ a dynamic probe misalignment correction for the ~~heat ratio method (~~HRM~~)~~. The aim is to obtain a more precise calculation of transpiration by parameterising the probe misalignment as a function of time to correct for the effect of tree growth.

3 Materials and method

3.1 Field site

This study was carried out in the Turia river basin, Eastern Spain (39º57'45" N 1º8'31" W), in a Mediterranean climate. Average annual rainfall is 475 mm, average annual maximum and minimum temperature is 15.5 ºC and 4.4 ºC respectively. Sap flow sensors were installed in four pine trees (*Pinus halepensis* Mill.) according to the heat ratio method (HRM, Burgess et al., 2001). Each sap flow sensor consists of three needles; one heater and two thermocouples. We will refer to the thermocouples as probes, and when using the term "sensors" we refer to both probes and the heater. ~~Three needles, one heater and two~~The sensors ~~thermocouples (0.13 cm x 4 cm),~~ were drilled into the uphill side of each tree trunk. Since *P. halepensis* has a higher sap velocity average near the cambium with the velocity steady declining nearer to the heartwood (Cohen et al., 2008), sensors were installed at 20 mm depth below the cambium for average sap velocity rates, as estimated by Manrique-Alba (2017). A metal plate was used as guide to ensure~~assure a~~ 0.6 cm spacing between the drilling holes. The ~~*P. halepensis*~~ selected pines had an average ~~mean~~ diameter of 24.5 cm at breast height (150cm). Continuous measurements were obtained from April 2017 to December 2018.

3.2 Environmental conditions

Air r~~R~~elative humidity (%) and air temperature (ºCº) were registered every 30-minute (U23 Pro V2, Onset Computer Corporation, USA). Precipitation was registered using a rain gauge with 0.2 mm resolution (RGR-M, Onset Computer, USA). Three soil moisture probes were inserted at 20-25 cm depth to register SWC (S-SND.M005, Onset Computer Corporation, USA), using a datalogger for logging specifications (HOBO Micro Station, USA). To assume periods of zero-flow, relative extractable water (REW) was calculated using the method of Bréda et al. (1995).

$$REW = \frac{(\theta_t - \theta_{min})}{(\theta_{max} - \theta_{min})} \qquad [1]$$

Were θ_t is the registered SWC, θ_{min} is the minimum SWC observed during the measurement period, and θ_{max} is the SWC at field capacity. When values of REW surpassed 1 they were converted to 1 according to Granier et al. (2000).

Formatted: Font: (Default) Times New Roman, Font color: Auto

Formatted: Indent: Left: 0 cm

~~3.2~~3.3 Construction of the sensors

The thermocouples were made after Davis et al. (2012), with a type E junction of chromium and constantan. ~~The E type has a higher accuracy, stronger signal and more stability than the type K (Davis et al., 2012).~~ The wires of the thermocouples were soldered together at temperatures not surpassing 200°C, ~~and~~ placed 2~~0 mm~~ ~~cm~~ inside a micropipettes ~~a glass tube~~ (10µL~~0.1 cm x 4 cm~~), and then into a needle (0.12~~3~~ cm x 4 cm, Sterican, Braun). The heater was made from~~of~~ a constantan wire of 20 cm coiled around a 7 cm long ~~aluminium~~aluminum wire~~, before~~ placed inside the same type of~~a~~ needle as the thermocouples ~~of 4 cm~~. The wire-ends were then soldered on to an electrical cable. Another resistance of 4.95~~10~~ ohms were soldered onto to the cable to get a total resistance of 14.95~~20~~ ohms. The heater was connected to a 12 V battery and delivered 7.0 W of power. All three sensors were connected to a CR800 datalogger (Campbell Scientific Inc., USA).

~~3.3~~3.4 Quality control of heat pulse ratios

Marshal (1958) parameterised the heat pulse velocity (*V*) in the HRM as a function of time following a heat pulse, and the instantaneous ratio of the increase in~~of~~ temperature (from the temperature prior to the release of the thermal pulse) at the downstream and upstream, v_1 and v_2 respectively, from a line heater:

$$V = \frac{4Kt\ln\left(\frac{v_1}{v_2}\right) - (x_2^2 + y_2^2) + (x_1^2 + y_1^2)}{2t(x_1 - x_2)} \qquad [2]$$

where *K* is the thermal diffusivity, *t* is the time passed from the release of thermal pulse in seconds, (x_1, y_1) and (x_2, y_2) are the relative positions of the thermocouples to the line heater (considering x-axis along the xylem and y-axis the perpendicular direction both to the xylem and to the heater line), and v_1,v_2 represents the temperature increases following the heat release~~, in the downstream and upstream thermocouple respectively.~~.

If probes are installed symmetrically above and below the heater line, $x = x_1 = -x_2$ and $y_1 = -y_2$, ~~the~~ equation 2~~1~~ simplifies into a function that is not dependent on time:

$$V = \phi \ln\left(\frac{v_1}{v_2}\right)$$

$$\phi = \frac{K}{x} \qquad [3\cancel{2}]$$

where $\boldsymbol{\phi}$ is a priori, a constant only depending on the placement of the probes and on the thermal diffusivity of both the xylem and the material used in the sensors.

The “perfect symmetry” assumption renders that the heat pulse ratio~~HPR~~ remains constant with time if heat pulse~~sap flow~~ velocity (V), thermal diffusivity (K) and probe positions (in both, x and y directions) have negligible variations during the time following each heat pulse (Marshall, 1958). ~~However,~~ Burgess et al. (2001) further demonstrated~~scribe~~ how empirical results initially differed from the ideal approach described by equation 3~~2~~ due to blocking of xylem vessels and misplacement of sensors.~~,~~ However, the study concludes that ~~although they~~the ratio~~HPR~~ converge asymptotically at least 60 seconds after the heat pulse release and, for at least, 40 seconds more (until 100 seconds after the heat pulse release), which is when the ratio~~HPR~~ should be measured. Our study argues that a~~A~~ visual inspection of heat pulse velocities (V in equations 2~~1~~ and 3~~2~~) does not necessarily give enough information to decide if measured values are a good representation of the sap flow.~~velocity or not.~~ The method does not consider that random HPR can arise, which due to the sensitivity of the method~~asurement~~, are likely to occur in practice. On these premises, we have built a methodology ~~utilising a~~ to quality check ~~of systematic~~ sap flow measurements systematically by ~~means of~~ introducing a statistical analysis performed on the instantaneous heat pulse ratio~~s values~~ acquired between 60 and 100 seconds after the heat pulse release. Hereafter, we will denote the averaged instantaneous heat pulse ratio between 60 and 100 seconds as HPRi. The quality check consisted of establishing ~~T~~threshold values ~~were established~~ for ~~the~~ relative standard deviation (RSD), statistically defined as the standard deviation divided by the mean, and the slope of the ~~time evolution of~~ HPRi.~~instantaneous heat pulse ratios~~. The statistical information obtained would account for any deterioration of the measurement. Burgess et al. (2001) proposed two separately methods to correct for wound and misalignment. The methods assume that errors arising from the wound inflicted by a sensor probe can be estimated using an empirical factor, whereas a misalignment of the probe needs to be calculated in situ. We propose a development of the misalignment correction method, while arguing that a statistical check of the HPR would detect a deterioration of the signal caused by a worsening of the wound. This would lead to a smaller sample mean and hence a higher RSD and was therefore chosen as a quality-check parameter along with the slope, which was a parameter proposed by Burgess et al. (2001).

~~3.3.1~~3.4.1 Logging specifications

The proposed analysis to ensure the reliability of long time series of sap flow measurements require the storage of statistics that are not usually recorded (as it is considered unnecessary). The datalogger ~~to be used~~ must have a minimum performance of the storage capacity (memory), and the processing speed of the algorithm implemented (especially in the routines related to the statistical calculations to be performed). In this study a CR800 (Campbell Scientific, USA) was used. A flow chart of sampling and data log (Fig. S1) was specifically designed, programmed and implemented in the datalogger to enable the calculus that are presented in the next sections of this paper.

Formatted: Font: 10 pt

Formatted: Font: 8 pt

Formatted: Font: 10 pt

~~The RSD is statistically defined as the standard deviation divided by the mean. Therefore, we~~We selected RSD (%) and Slope (s^{-1}) of the instantaneous heat pulse ratio versus time, calculated for each of the periods from 60 to 100 seconds ~~and used for the quality control~~to filter out random ratios (Fig.1A and B). All HPR with relative standard deviation > 5% and a |slope – median (slope)| < 0.003 s^{-1} were removed. The 5 % threshold was chosen to ensure that 95 % of the dataset was considered in the data processing. The slope median was taken from all slope values obtained during the measurement period. The magnitude of the~~is~~ ~~slope~~ threshold ~~was~~ chosen for the slope was taken from a modelled output of instantaneous ratios performed by Burgess et al. (2001), were low sap flow velocities (5 cm h^{-1}) combined with a small wound width (0.17 cm) were~~are~~ shown to display a slope of 0.001. Due to our low sap flow velocities (< 8 cm h^{-1}) and small probe diameter (0.12 cm), we expected the slope to be as close to 0.001 as possible. The specific threshold of 0.003 was decided upon inspection of the measurements and can be modified according to needle size and magnitude of the sap velocity measured~~the sensor~~. According to Burgess et al. (2001), higher values of slope (0.01) can be expected with wound width and higher velocities~~y~~.

Formatted: Superscript

~~3.4~~3.5 Correction for long-term probe misplacement

Under the assumption of “perfect symmetry”, in periods when V = 0 cm h^{-1}, the ratio of the increase of temperature at the downstream and upstream from a line heater would be equal to 1~~one~~ ($v_1 = v_2$ in equation 2). However, this is not always the case with field measurements due to misalignment of the probe, damage to xylem vessels when inserting the probes, and the different thermal properties of ~~or heterogeneous~~ wood and needle ~~properties~~ (Fig. 1A, Burges et al., 2001). The same study ~~Burgess et al. (2001)~~ showed how probe placement can be estimated *in situ* with the HRM. Further, it~~He~~ proposed a methodology for probe misalignment which builds on the assumption that errors arising from inaccurate probe spacing can be treated one-dimensionally. Th~~H~~is approach assume that the total effect of probe misalignment (in both axis directions), observed in the ratio of the increase in~~of~~ temperature, can be parameterised calculating an “effective” probe misalignment in only one direction (x direction, parallel to the xylem). Thus, without the assumption of "perfect symmetry", in periods when V = 0 cm h^{-1}, equation 2~~1~~ takes a more simplified form:

$$4Kt\ln\left(\frac{v_1}{v_2}\right) = (x_2^2 + y_2^2) - (x_1^2 + y_1^2) \quad [4]$$

Which becomes equation (5) if $y_1 = -y_2$

$$4Kt\ln\left(\frac{v_1}{v_2}\right) = x_2^2 - x_1^2$$

[5]

As it is unknown if the misalignment is in x_1 or in x_2, the calculation is repeated twice, assuming first that x_1 is correct when calculating actual x_2 placement, and vice versa. Each probe placement is calculated as:

$$x_2 = \sqrt{(4kt\ \ln(\frac{v_1}{v_2}) + x_1^2} \qquad [6]$$

Formatted: Centered

Two different heat pulse velocities, V_1 and V_2, are then derived (using equation 2~~1~~ but with the assumption $y_1 = -y_2$) for the~~both misalignments~~ x_1 and x_2 obtained; and the final V provided as their average.

Zero ~~sap~~ flow conditions can either be imposed artificially by severing the root or stem (Burgess et al., 2001), or ~~assumed~~ found when there is~~are~~ no biophysical driving force (Forster, 2007). That is, when atmospheric vapour pressure deficit (VPD) is close to zero, the soil is saturated, usually after substantial precipitation, and values are taken at predawn ~~to avoid any biophysical driving force~~ (Forster, 2017). Zero flow conditions were assumed at night (22:00-03:00h solar time) during days when relative extractable water, REW > 0.7, and vapor pressure deficit (VPD) was close to zero (Fig. S2). Multiple readings were used to produce an average of each event. These conditions were limited to five occasions during our study period. Saturated soil is a necessary criterion due to the possibility of reverse flow at night-time (Forster, 2014). If not considered, low HPR values representing reverse flows can be interpreted as zero flow. The non-intrusive~~latter~~ approach of zero flow allows the parameterisation of misalignment as a function of time (equation 5), if several calculations of x_1 and x_2 (equation 4) are performed. By applying equation 4 at varies times throughout the measurement period, it is possible to calculate a linear regression using the estimation of Burgess et al. (2001) for misalignment for each of the thermocouples:

$$x_i = m_i\, d + n_i \quad ; \quad i = 1, 2 \qquad [6]$$

where x_i are the relative positions of the thermocouples to the line heater in cm, d is the time along the measuring campaign in days, m_i is the slope of the regression for thermocouple i, and n_i is the interception coefficient.

By introducing equation 5 in equation 1 and allowing the simplification assuming $y_1 = -y_2$, two equations of the corrected heat pulse velocities are obtained as a function of the time in the measuring campaign (equation 6).

$$V_1 = \frac{4Kt\,ln\,\left(\frac{v_1}{v_2}\right) + (m_1\,d + n_1)^2\;-0.6^2}{2t(m_1 d + n_1 + 0.6)} \qquad [7]$$

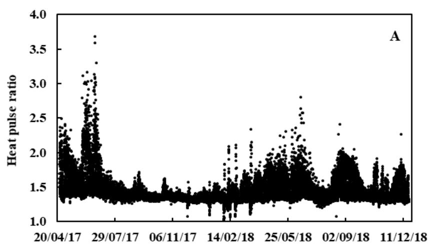
$$V_2 = \frac{4Kt\,ln\,\left(\frac{v_1}{v_2}\right) +\,0.6^2 - (m_2\,d + n_2)^2}{2t(0.6 - m_2 d - n_2)}$$

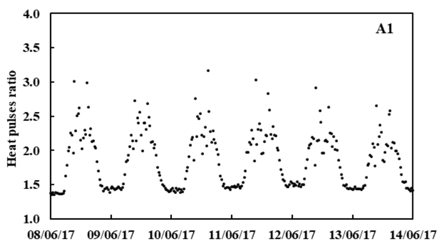
In accordance to what Burgess et al. (2001) proposes, our approach averages the two estimates obtained from equation 6 to obtain a corrected heat pulse velocity.

4 Results

4.1 Heat pulse ratios

The HPR obtained during the measurement period displayed~~s~~ a clear positive shift away from the theoretical ideal where the HPR would equal to one at zero flow (Fig. 1A). This gives an indication of the necessity of corrections, that being due to wound inflicted by the probe, misalignment, misestimation of thermal diffusivity-or stem water content. However, the HPR data by itself does not give an indication of the quality of each measurement, nor if it deteriorates over time. Therefore, the quality of the measurements was indicated by calculating the RSD, the relative standard deviation divided by the sample mean, -and the slope for each HPR (Fig. 1B and 1C). All HPR with RSD higher than 5% were eliminated. The data points eliminated corresponded to a 1 % of the total dataset. Because the HRM ~~method~~ is built on the theoretical assertion that the temperature in each of the thermocouples is steady with time, specifically between 60 and 100 seconds after the release of a heat pulse, the slope of the HPR should be close to zero. All HPR with |slope – median (slope)| < 0.003 s^{-1} were eliminated, which corresponded to a 12 % of the original dataset (Fig. 1C).





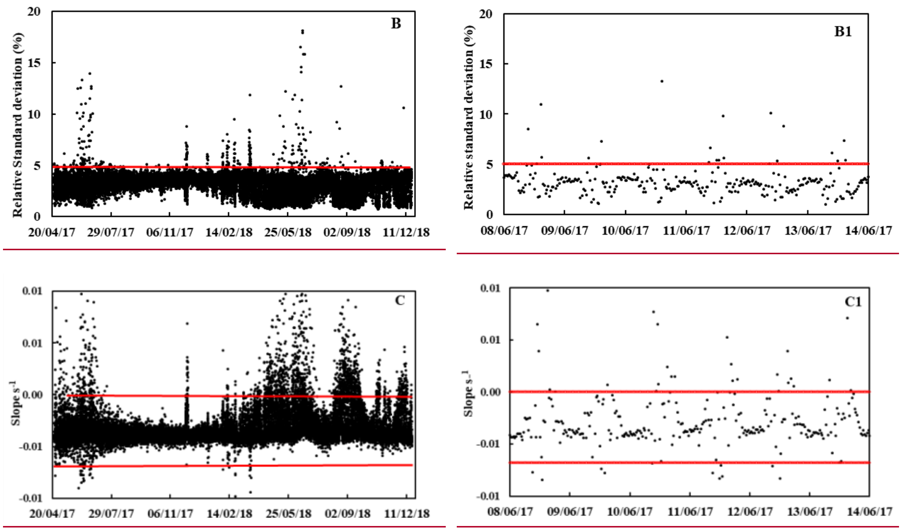


Figure 1. (A) Heat pulse ratios (HPR) throughout the measurement period in 30-minute intervals in tree number 1. Each HPR is an average of 41 instantaneous ratios corresponding to the temperature difference in two thermocouples at 0.6 cm up and down stream from a heater probe at 0.2 cm depth. (A1) Zoomed in panel of HPR data for one week of measurements. (B) Relative standard deviation (%) for each HPR in tree number 1 for the whole measurement period. Red line indicates threshold used for the quality control were all HPR relative standard deviation > 5% were removed from the data analysis. (B1) Zoomed in panel of RSD data for one week. (C) Slope (s^{-1}) for each HPR in tree number 1 for the whole measurement period. Red line indicates threshold used for the quality control for this particular sensor. HPR with $|\mathrm{slope} - \mathrm{median(slope)}| < 0.003\ s^{-1}$ were removed from the data analysis. (C1) Zoomed in panel of slope data for one week of measurements. Red squares in A,B,C indicate which week is represented in A1,B1,21.

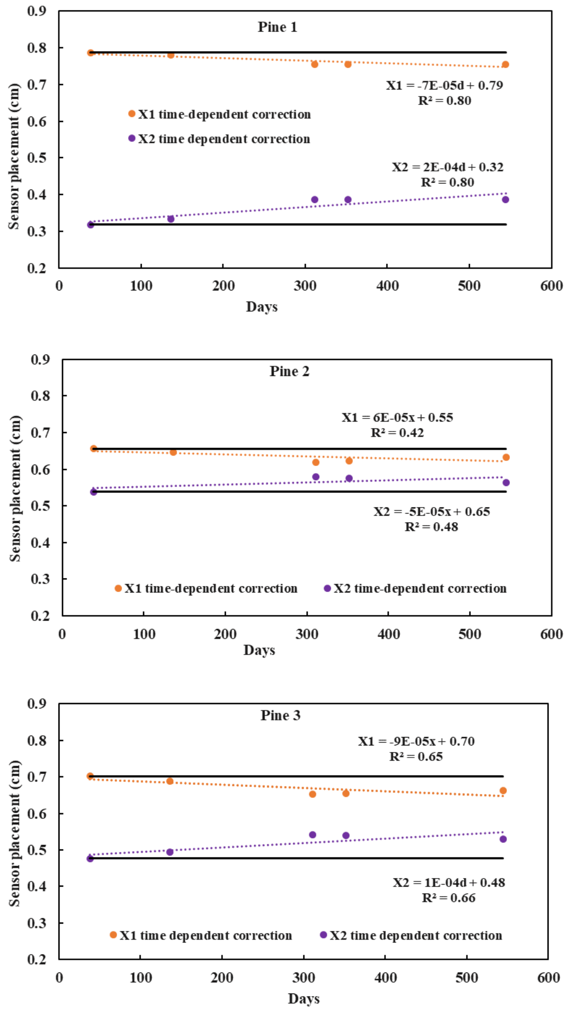
4.2 Heat pulse velocities

A linear regression was obtained from misalignment calculations for each sensor performed during zero-flow conditions. Zero-flow conditions were assumed at night (22:00-03:00h solar time) during days when relative extractable water, REW > 0.75, and vapor pressure deficit (VPD) was close to zero (Fig. S2). Multiple readings were used to produce an average of each event. These conditions were limited to five occasions during our study period.~~Because of the dry climate, only five events fulfilled these criteria:), during the measurement period of 20 months.~~ The outputs indicate a clear shift of placement in each of the sensors over time, here denoted as x_1 and x_2 in each tree~~, with a greater shift in x_2 than x_1~~ (Fig. 2). The eight probes~~-sensors~~, two per tree, all deviated from the ideal of 0.6 cm. Probe~~Sensor~~ X_2 in pine~~tree~~ number 1 had the highest inaccuracy with the initial value close to 0.3 cm. On average, the ~~eight~~ probes~~sensors (two per tree)~~ showed a averaged shift of 0.04 cm in placement after twenty months of measurement (Fig. 2). The equation obtained from the linear regression was then implemented in equation 6, and corrections to the heat pulse velocity data was done using an average of V_1 *and* V_2 (Fig. 3,4). Outputs were compared with one-time misalignment correction calculated in the beginning of the measurement period to demonstrate the evolution of the probe misalignment. The one-

Formatted: Font: Italic

Formatted: Font: Italic, Subscript

time correction demonstrated a steady decline in accuracy over time (Fig. 3, 4). The difference between the two correction methods showed significance after three months of employment (Fig. 3).



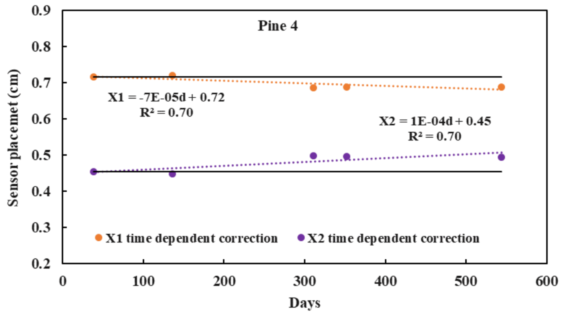
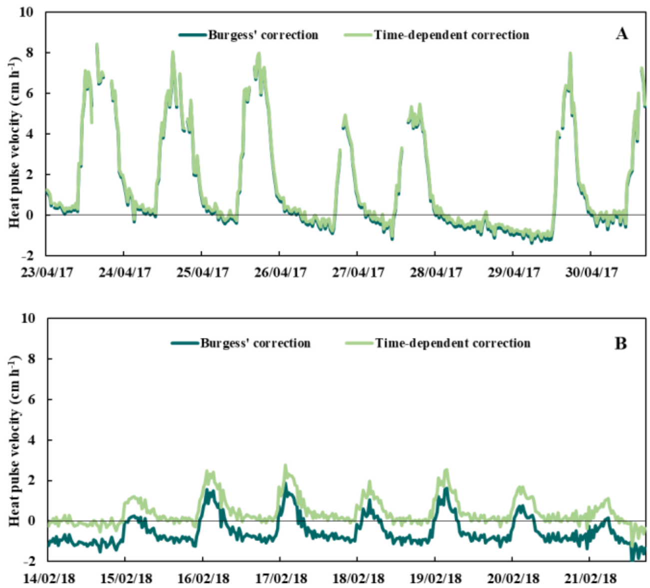


Figure 2. P~~The p~~lacement~~s~~ of the probes are shown as distance from the heater (cm). Probe placement was calculated once (solid lines) for the whole study period and compared to probe placement calculated varies times (solid circles with dotted lines) throughout the study period. Each point represents the probe position calculated during its respective zero-flow event

~~Sensor misalignment positions calculated once (solid lines) compared to probe misplacement over time in (solid circles with dotted lines). Each point represents the probe misalignment position calculated during its respective zero flow event.~~



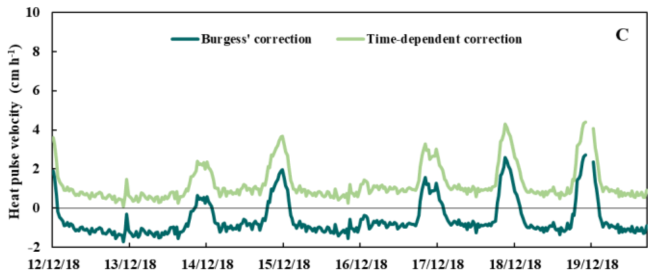


Figure 4. Heat pulse velocities during one week of measurements in the beginning (A), halfway (B) and at the end (C) of the measurement period. Dark green lines represent velocities with probe misalignment corrected for once at the beginning of the experiment. Light green lines represent velocities with the time- dependent probe misalignment corrections.

4.3 Sap flow

Heat pulse velocities (~~both corrections~~cm h^{-1}) were converted into sap flow (cm^3 cm-^{2} h^{-1})~~velocities~~ according to Burgess et al. (2001). Our data demonstrated that by not correcting for changes in probe misalignment under continuous measurement for more than 3 months, the errors corresponded to an averaged difference of 0.29 cm cm^3 cm^2 h^{-1}~~h^{-1}~~ in sap flow per quartile for the four trees. At the end of the 20-months period, this corresponded to an averaged difference of 0.53(±0.23) cm^3 cm^2 h^{-1} ~~cm h^{-1}~~ (Table 1, Fig. 5) In terms of transpiration values, this corresponded to a mean difference of 0.7 L tree $^{-1}$ day^{-1} (sapwood area = 170 cm^2), which would correspond to a difference of 542 L ha $^{-1}$ day $^{-1}$ (775 tree ha $^{-1}$) assuming 8 hours of daylight. It is relevant to note our conversion did not consider the differences in stem moisture content, which can also affect the output values (Vandegehuchte and Steppe, 2013). The outputs obtained should be considered as relative differences as the one-time correction was applied in the beginning of the experiment.

Table 1. Seasonal averages of sap velocity for 4 different pines~~trees~~. All sap flow values are expressed in cm^3 cm^2 h^{-1}~~cm h^{-1}~~. Sap ~~velocities~~ flow corrected for with time-dependent misalignment calculations are compared with sap ~~velocities~~flow corrected for once in the beginning of the measurement period. Averages were taken from daily values. Abbreviations Sp, Su, Fa, and Wi indicates Spring, Summer, Fall and Winter respectively, each with the corresponding year.

Pinee ~~number~~	Correction method	Sp-17	Su-17	Fa-17	Wi-17~~/18~~	Sp-18	Su-18	Fa-18
1	One-time correction	0.98±1.2	0.23±0.8	-0.31~~1~~±0.2	-0.37±0.5	0.12±0.7	0.05±0.5	-0.16±0.5
	Time-dependent correction	1.08±1.2	0.44±0.7	0.04±0.2	0.10±0.5	0.64±0.7	0.70±0.6	0.69~~7~~±0.5
2	One-time correction	1.33±2.1	0.49±1.2	-0.28±0.4	-0.25±0.5	0.06±0.6	0.04±0.5	-0.05±1.8
	Time-dependent correction	1.40±2.1	0.66±1.2	-0.02±0.4	0.09±0.5	0.36±0.7	0.42±0.5	0.45±2.0

Formatted: Font: 9 pt

Formatted Table

Formatted: Left

Formatted: Font: 9 pt, Pattern: Clear

Formatted: Font: 9 pt, Pattern: Clear

3	One-time correction	0.82±1.7	0.73±1.1	-0.02±0.5	-0.24±0.5	0.57±0.6	0.75±0.5	0.53±0.60 ~~0.53~~
	Time-dependent correction	0.84±1.7	0.79±1.1	0.09±0.5	-0.09±0.5	0.76±0.6	0.99±0.5	0.80±0.63 ~~0.80~~
4	One-time correction	0.79±1.1	0.5~~65~~±0.7	0.09±0.4	-0.05±0.4	0.29±0.6	0.18±0.5	0.05±0.5
	Time-dependent correction	0.82±1.1	0.66±0.7	0.29±0.4	0.23±0.4	0.61±0.6	0.56±0.5	0.55±0.5

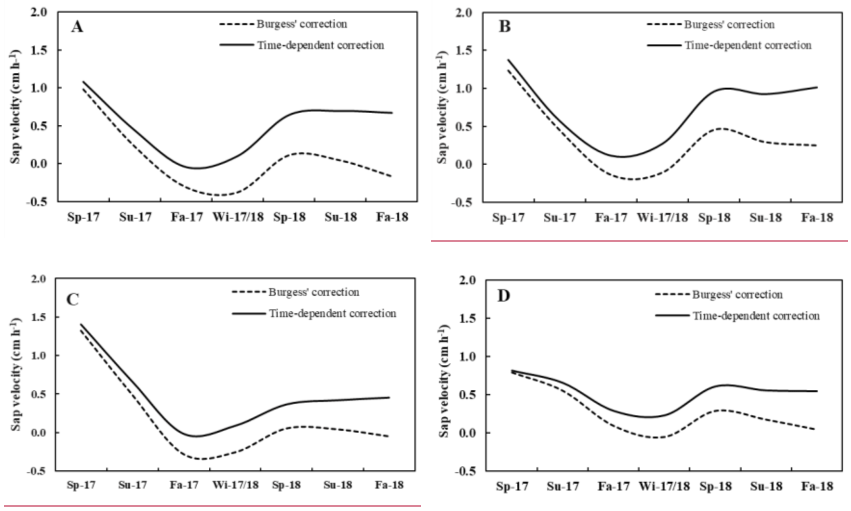


Figure 4. Seasonal averages of sap ~~velocities~~ flow (cm³ cm² h⁻¹~~cm h⁻¹~~) calculated using the two misalignment corrections throughout the 20-month measurement period. Dotted lines represent probe misalignment corrected for once. Solid lines represent sap ~~velocity~~ flow corrected for with the time-dependent probe misalignment method. Abbreviations Sp, Su, Fa and Wi indicates Spring, Summer, Fall and Winter respectively, each with the corresponding year. 1, 2, 3, 4 corresponds to different trees.

4.4 Transpiration

To demonstrate the difference in terms of transpiration, sap flow values ($cm^3\ cm^{-2}\ h^{-1}$) without misalignment correction (Tr_r), with one-time correction (Tr_{one}), and with time-dependent correction (Tr_{time}) were converted into transpiration values ($L\ tree^{-1}\ h^{-1}$) and compared against each other during a period of three weeks. Data from three of the pines (1, 3 and 4) were taken towards the end of the measurement period, to demonstrate the greatest differences (Fig. 5). However, the sensor placement in pine number two demonstrated a decreasing misplacement over the course of the measurement period (Fig. 2), and another measurement period was therefore chosen for this tree, for a better demonstration of the differences between the estimations. Note that the transpiration estimations without misalignment correction still went through the wound correction step as shown in Burgess et al. (2001). Pine number 1 displayed the biggest differences between the T_r and Tr_{time}, with an averaged daily difference of 1.5±0.002 ($L\ tree^{-1}$) during the three weeks, and an averaged daily difference of 0.3±0.003 ($L\ tree^{-1}$) between Tr_{one} and the Tr_{time}. Pine number 3 and 4 showed similar differences between

Formatted Table

Formatted: Font: 9 pt

Formatted: Font: 10 pt

Formatted: Heading 3, Left, Space After: 0 pt, Outline numbered + Level: 2 + Numbering Style: 1, 2, 3, … + Start at: 1 + Alignment: Left + Aligned at: 0.63 cm + Indent at: 1.27 cm, Border: Top: (No border), Bottom: (No border), Left: (No border), Right: (No border), Between : (No border)

Formatted: Font: 9 pt

the methods, with a daily average of 0.1±0.001 and 0.2±0.002 (L tree^{-1}) between Tr and T_{time} respectively, and both had an average daily difference of 0.1±0.001 (L tree^{-1}) between Tr_{one} and Tr_{time} (Fig. 5).

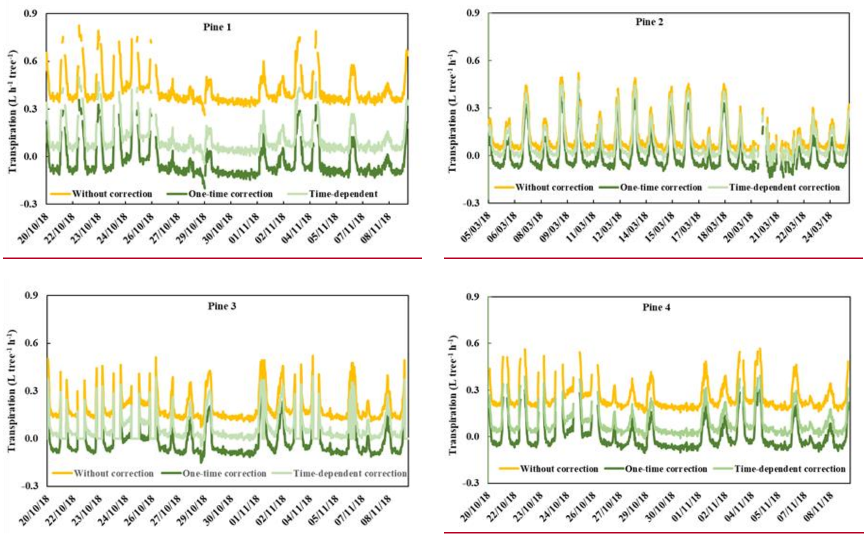


Figure 5. Three weeks of transpiration rates (L tree^{-1} h^{-1}). 'Without correction' represents rates without misalignment-correction, 'One-time correction' represents rates corrected for once in the beginning of the measurement period, and 'Time-dependent correction' represents rates being corrected for at varies days throughout the measurement period.

5 Discussion

5.1 ~~Filtration of~~Filtering heat pulse ratios

Because the HPR is an average of instantaneous temperature ratios it is difficult to ensure its accuracy only by visual inspection of the averaged output, and without further statistical information. We suggest RSD and slope to filter out random ratios, and to ensure the quality of the data. A clear tendency was seen of higher RSD and slope values during periods of higher flow rates when compared to the HPR data (Fig. 1A, 1B and 1C), indicating more noise and less trustworthy data during higher flow rates. The HRM is known for being limited at higher rates (> 45 cm h^{-1}, Forster, 2017), but our dataset showed no sap velocity values higher than 6 cm h^{-1}, and therefore was not initially considered as a limitation.

During twenty months of field measurements the thermocouples showed no visible sign of deterioration with time. However, it is still important to note that this does not consider the possible diminishing amplitude of the ratios over time, which can also lead to underestimations of actual flow (Barett et al., 1995; Green et al.,

Formatted

Formatted: Font color: Text 1

Formatted: Justified

2003; Forster, 2017). Therefore, this should be observed separately, comparing sap flow ratios using data obtained under similar climatic conditions (Moore et al., 2010). Comparing the data from 2017 to the data of 2018, SWC differed significantly on similar calendar dates with the exception of six days in June (9 – 15). When comparing the relationship of sap flow versus VPD on these days, there was an increase in the slope; 2.3 for 2017 and 2.9 in 2018. This, we attributed to an overall weaker correlation between VPD and sap flow in 2017; $R^2 = 0.4$ in 2017 versus $R^2 = 0.6$ in 2018, due to higher values of VPD the first year (S3). In conclusion, it was clear that the HPR readings had not decreased in the second year, when compared to the first year, under similar environmental conditions.

5.2 Long term variation in probe placement

When applying the HRM for longer than a few weeks it is relevant to quantify how $\boldsymbol{\phi}$ in equation (2) and the misalignment term in equation (4) evolve during the measuring period. The predicted variation of x_1 and x_2 is due to the growth of the tree and, on the other hand, periodical variations of K due to annual and seasonal variations of the physiological properties of the tree (Green et al., 2003; Vandegehuchte and Steppe, 2012; Ren et al., 2017). In the HRM the probe misalignment calculations can be corrected using the methodology proposed here, considering that as an average, each sensor displayed a 0.04 cm displacement within the tree after twenty months of measurements. The correction would also be more rigorous with more assumed zero flow events, which would be easier to obtain in humid environments with more periods of saturated soil.

The theory behind the HPV method were tested on conifers (Marshall, 1958), whereas Burgess et al. (2001) more generally refers to woody species when working with HPR. The time-dependent correction method could be useful for any species already tested with the HPR, where misalignment of probes can create a source of error and sensors are installed over a longer period. Specifically, where the movement of the wood might cause further displacement of the sensor.

By going back to the original assumption of "perfect symmetry", we investigated the original premises the method is built on. Even though Burgess et al. (2001) elaborated a correction method for sensor misalignment we saw that changes in sensor misplacement was detectable after each season. Therefore, multiple corrections should be carried out throughout the measurement period. The proposed modified method coincides with the one-time correction method (Burgess et al., 2001) for short time periods (< 3 months) but differs progressively over time. We found that this shift in placement is significant already after 3 months, and therefore dynamically misplacement calculations should be carried out or sensors should be reinstalled at this frequency. However, as pointed out by Moore et al. (2010), re-installing the sensors might~~each time a sensor is replaced it~~ create~~s~~ a shift in the data due to spatial variation within the tree. Leaving the same sensors in the

Formatted: Not Highlight

Formatted: Not Highlight

tree throughout the study period avoids this problem and enables the study to focus on the intrinsic factors affecting the sap flow rates.

Formatted: Not Highlight

2. Conclusion

Formatted: Font: (Default) Times New Roman

In conclusion, we found that high quality measurements with sap flow sensors can be ensured over longer periods (>3 months), if the HPR is assessed using the proposed filtering method, and the probe misalignment variability over time corrected for. In this study, we observed data over a 20- month period in Pinus halepensis, and saw no sign of detoriation in the second year compared to the first, when observing the values obtained for slope and relative standard deviation. However, when observing the alignment of each probe, there was a clear shift from the beginning to the end of the measurement period. This indicate that measurements can be obtained during a second season without the need of re-installing sap flow sensors, if the proposed time-dependent misalignment correction is incorporated in the data processing. This would increase the accuracy of point measurements, and consequently transpiration estimations. The different errors related to upscaling are beyond the scope of this paper, but significant differences were observed when comparing sap flow estimations with no correction, one-time correction, and time-dependent correction for probe misalignment. To avoid sensor reinstallation, this should therefore be considered."

Author contributions

JLP and ECH conceptualized the study. EKL made the sensors with the supervision of JB. EKL implemented the sensors and had the responsibility of the data processing and field site. JLP, EKL and JAV all worked on the data analysis. JAV wrote the script for the datalogger and had the technical responsibility. EKL and JLP prepared the manuscript with inputs from all the co-authors. JLP and JAV designed the flow chart to be programmed in any data logger system.

The authors declare that they have no conflict of interest

Acknowledgements

The CEAM Foundation is supported by the Generalitat Valenciana (Spain). This study is partially funded by the Spanish Ministry of Science, Innovation and Universities and the European Regional Development Fund CGL2015-67466-R (MICINN/FEDER) through the research project "VERSUS", by the project "IMAGINA" PROMETEU/2019/110 (Consellería de Cultura GVA, and by the research project "DESESTRES"–PROMETEOII/2014/038 (Consellería de Cultura GVA),. We thank the municipality of Aras de los Olmos for permitting field experiments on a public managed forest site. We thank J. Puértolas and H. Moutahir for their feedback and encouragement on the manuscript.

References

Barrett, D.J.; Hatton, T.J.; Ash, J.E.; Ball, M.C.: Evaluation of the heat pulse velocity technique for measurement of sap flow in rainforest and eucalypt forest species of south-eastern Australia. Plant Cell Environment, 18, 463–469. https://doi.org/10.1111/j.1365-3040.1995.tb00381.x, 1995

Becker, P., Edwards, W. R. N.: Corrected heat capacity of wood for sap flow calculations. Tree Physiology, 19, 767–768. doi:10.1093/treephys/ 19.11.767, 1999

Bréda, N., Granier, A., Aussenac, G. Effects of thinning on soil and tree water relations, transpiration and growth in an oak forest (Quercus petraea (Matt.) Liebl.). Tree Physiology. 15, 295-306, 1995.

Burgess, S. S. O., Adams M, Turner NC, Beverly CR, Ong CK, Khan AAH, Bleby TM,: An improved heat pulse method to measure low and reverse rates of sap flow in woody plants. Tree Physiology, 21, 589–598. doi:10.1093/treephys/21.9.589, 2001

Cohen, Y., Cohen, S., Cantuarias-aviles, T., Schiller, G.: Variations in the radial gradient of sap velocity in trunks of forest and fruit trees. Plant and soil, 305, 49-59. DOI:10.1007/s11104-007-9351-0, 2008

Davis, T.W., Kuo, C. M., Liang, X., Yu, P. S.,: Sap Flow Sensors: Construction, Quality Control and Comparison. Sensors, 12, 954-971; doi: 10.3390/s120100954, 2012

Forster, M. A., How reliable are heat pulse velocity methods for estimating tree transpiration? Forests 8(9), 350: doi:10.3390/f8090350, 2017

Granier, A. Evaluation of transpiration in a Douglas-fir stand by means of sap flow measurements. Tree Physiology, 3, 309-320, 1987.

Granier, A. A generic model of forest canopy conductance dependent on climate, soil water availability and leaf area index. Ann. For. Sci, 57, 755 – 765, 2000.

Green, S.; Clothier, B.; Jardine, B. Theory and practical application of heat pulse to measure sap flow. Agron. J. 95, 1371-1379. doi:10.2134/agronj2003.1371, 2003

Jasechko, S., Sharp, Z. D., Gibson J. J., Birks S.J., Yi, Y., Fawcett, P. J.: Terrestrial water fluxes dominated by transpiration. Nature, 496, 347-350. Doi: https://doi.org/10.1038/nature11983, 2013

Lu, P., Urban, L., Zhao P. Graniers Thermal Dissipation Probe (TDP) Method for Measuring Sap Flow in Trees: Theory and Practice. Acta Botanica Sinica, 46 (6), 631-646, 2004.

Manrique-Alba, A.: Ecohydrological relationships in pine forests in water-scarce environments. Ph.D. thesis, University of Alicante, Spain, 165 pp., 2017.

Marshall, D. C.: Measurement of sap flow in conifers by heat transport. Plant Physiology, 33, 385–396. doi:10.1104/pp.33.6.385, 1958

Moore, G. W., Bond, B. J., Jones, J. A., Meinzer, F. C.: Thermal-dissipation sap flow sensors may not yield consistent sap-flux estimates over multiple years. Trees – Structure and Function, 24, 165–174.

doi:10.1007/s00468-009-0390-4, 2010

~~Novick, K.A., Oren, R., Stoy, P. C., Siquiera M. B. S., Katul, G. G. Nocturnal evapotranspiration in Eddy-covariance records from three co-located ecosystems in the Southeastern U.S.: Implications for annual fluxes. Agricultural and Forest Meteorology, 149, 1491-1504. doi:10.1016/j.agrformet.2009.04.005, 2009~~

Ren, R.; Liu, G.; Wen, M.; Horton, R.; Li, B.; Si, B. The effects of probe misalignment on sap flow measurements and in situ probe spacing correction methods. Agric. For. Meteorol. 232, 176–185. https://doi.org/10.1016/j.agrformet.2016.08.009, 2017

Smith, D.M., Allen, S.J. Measurement of sap flow in plant stems. Journal of Experimental Botany, 305, 1833-1844. https://doi.org/10.1093/jxb/47.12.1833, 1996.

Formatted: Justified

Steppe, K., De Pauw, D.J.W, Doody, T.M., Teskey, R.O. A comparison of sap flux density using thermal dissipation, heat pulse velocity and heat field deformation methods. Agricultural and Forest Meteorology. 7-8, 1046-1056. doi: https://doi.org/10.1016/j.agrformet.2010.04.004, 2010

Vandegehuchte, M.W., Steppe, K.: Sapflow+: A four-needle heat-pulse sap flow sensor enabling nonempirical sap flux density and water content measurements. New Phytologist, 196, 306–317. https://doi.org/10.1111/j.1469-8137.2012.04237.x, 2012

Vandegehuchte, M. W., Steppe, K.: Sap-flux density measurement methods: working principles and applicability. Functional Plant Biology, 40, 213-223. http://dx.doi.org/10.1071/FP12233_CO, 2013

Wiedemann, A., Marañón-Jiménez, Rebmann, C., Herbst, M., Cuntz, M.: An empirical study of the wound effect on sap flux density measured with thermal dissipation probes. Tree Physiology, 36, 1471-1484. https://doi.org/10.1093/treephys/tpw071, 2016

Supplementary material

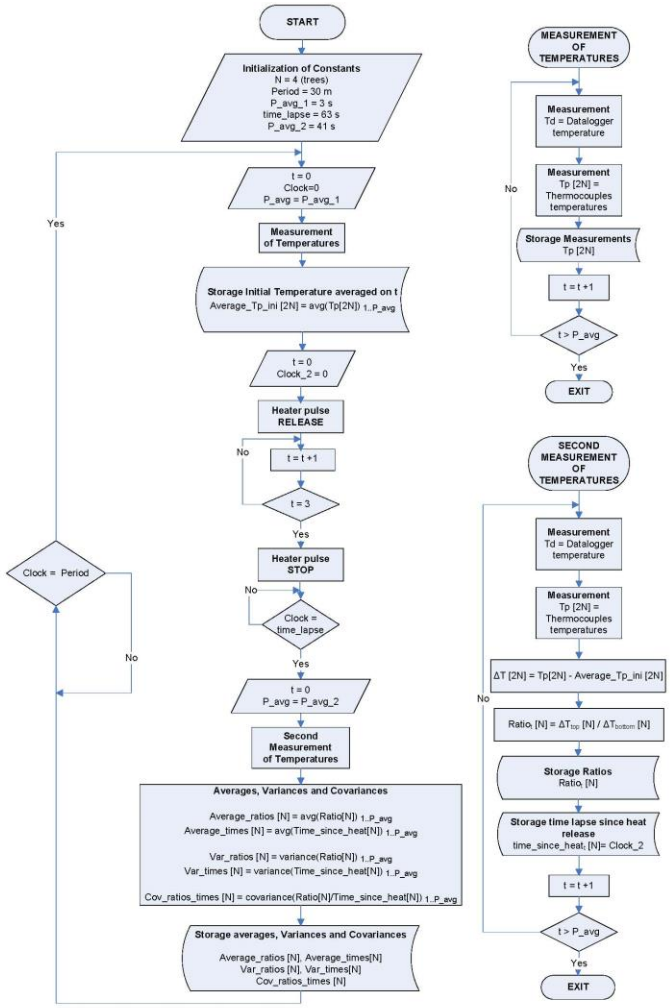


Figure. S1. Detailed flow chart of the procedure for data sampling. Each step represents a command in the script for the datalogger.

List of variables referenced in the flow chart:

N: Number of trees sampled (in our case, 4 trees).

Period: Time between two consecutive measurements or heat pulse releases (in our case, 30 minutes).

P_avg_1: Period to average the initial temperatures (before the heat release) by the thermocouple probes (in our case, 3 seconds).

P_avg_2: Period to average the final temperatures (after the heat release) by the thermocouple probes (in our case, 41 seconds).

Time_lapse: Elapsed time from the beginning of initial temperatures to the beginning of final temperatures corresponding to the measurement of the cooling ramp at the thermocouples (in our case this lapse is 63 s).

t: time counter, or dummy variable, to count the elapsed time inside each of the different loops of the programme (in seconds).

Clock: Marks the temporal sequence of the whole process, from the beginning of the process of the measurements (instant zero) to the end of the periodic process for all measurements (the maximum number allowed is determined by the constant Period, which in our case is 30 minutes). The processing speed of the data logger must have the capacity to perform all the calculations in a time (measured by "Clock") lower than the pre-set measurement period (determined by "Period").

Clock_2: Monitor the time elapsed between the release of the heat pulse and the instant when thermocouples are measured.

Td: Temperature at the data logger (this temperature is necessary to be known to correctly measure temperature using thermocouples).

Tp [2N]: Array of the temperatures measured at the 2N thermocouple probes (there are two thermocouples per tree).

Average_Tp_ini [2N]: Array of the averaged initial temperatures at the 2N thermocouple probes. Average is calculated using the P_avg_1 period, previous measurements to the beginning of the heat release.

ΔT [2N]: Array of the variations of the temperatures in the 2N thermocouples due to the heat release (variations with respect to their initial temperatures).

$\mathbf{Ratio_t}$ [N]: Ratio between the temperature variations in the thermocouple above the heater (ΔT_{top} [N]) and those below the heater (ΔT_{bottom} [N]). Sub index indicates the time corresponding to the ratio (in our case the ratios are measured at the 60s second after the beginning of the heat release and for the next 40 seconds).

$\mathbf{Time\_since\_heat_t}$ [N]: Storage of the time elapsed between the start of the heat pulse release and the instant when thermocouples are measured. Sub index indicates the time corresponding to each measurement, this sub index matches this variable with Ratio $_t$ [N].

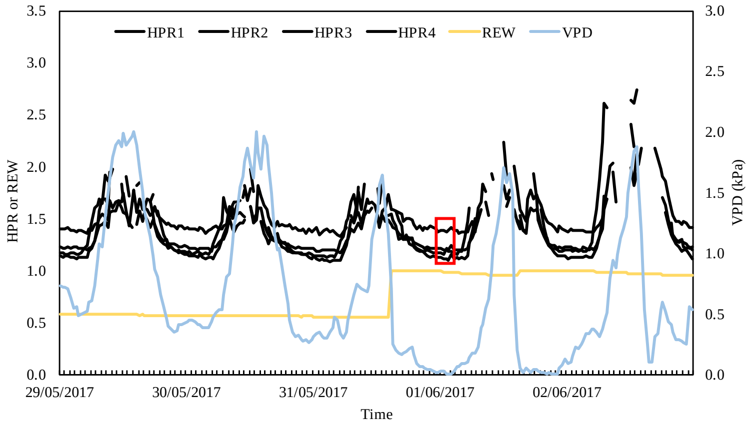
Average_ratios [N]: Array of the averaged ratios for each tree (N). Average is calculated for the period determined by P_avg_2; in our case for 41 seconds (from second 60 to second 100 after the heat release).

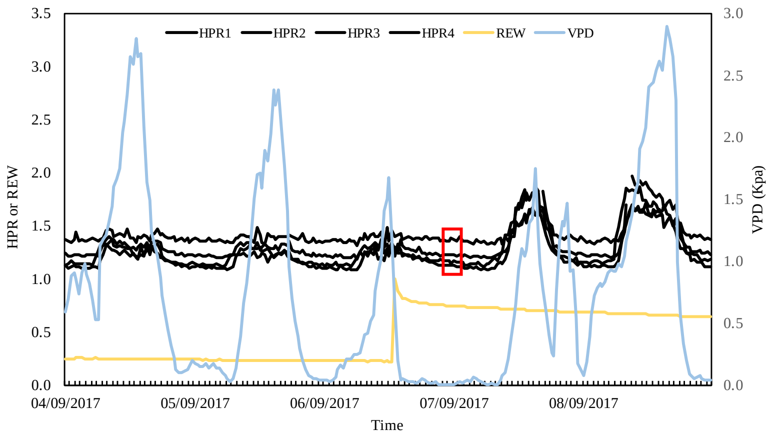
Average_times [N]: Array of the averaged time elapsed between the release of the heat pulse and the instant when thermocouples are measured. Average is calculated for the period determined by P_avg_2; in our case for 41 seconds, from second 60 to second 100 after the heat release). If everything works well, this value will be constant and, in our case, equal to 80 seconds (median measurement time, between 60 and 100 seconds).

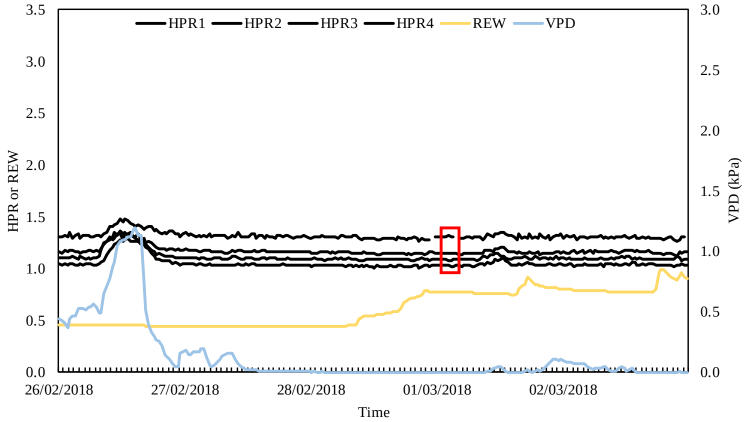
Var_ratios [N]: Array of the variances calculated for the ratios measured for the period determined by P_avg_2; in our case for 41 seconds (from second 60 to second 100 after the heat release).

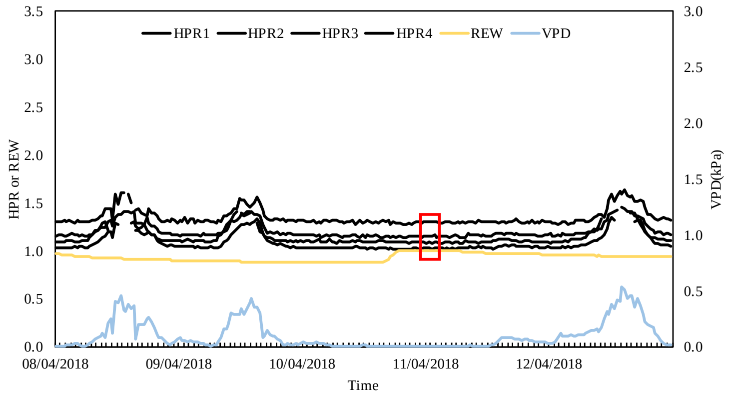
Var_times [N]: Array of the variances calculated for the times elapsed between the release of the heat pulse and the instant when thermocouples are measured. Variance is calculated for the period determined by P_avg_2; in our case for 41 seconds (from second 60 to second 100 after the heat release).

Cov_ratios_times [N]: Array of the covariances calculated with the ratios of temperature variations and the times elapsed between the releases of heat pulse and the instants when thermocouples are measured.






HPR1
HPR2
HPR3
HPR4
REW
VPD
HPR or REW
VPD (kPa)
3.5
3.0
2.5
2.0
1.5
1.0
0.5
0.0
26/02/2018
27/02/2018
28/02/2018
01/03/2018
02/03/2018
Time


HPR1
HPR2
HPR3
HPR4
REW
VPD
HPR or REW
VPD(kPa)
3.5
3.0
2.5
2.0
1.5
1.0
0.5
0.0
08/04/2018
09/04/2018
10/04/2018
11/04/2018
12/04/2018
Time

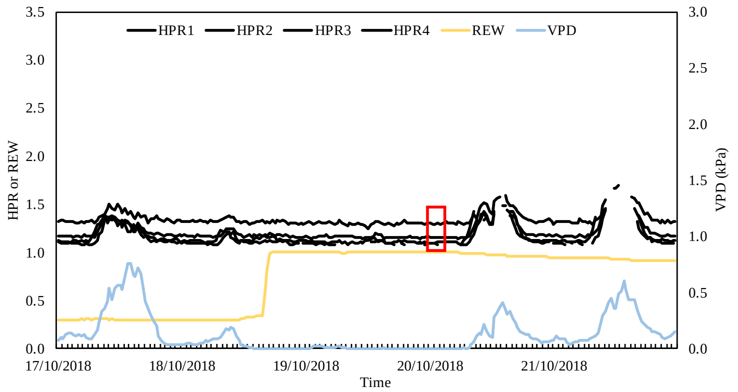


Figure. S2. Vapour pressure deficit (VPD), relative extractable water (REW) and heat pulse ratio (HPR) for all trees at five different events assuming zero flow conditions. Squares indicate the HPR readings used for zero flow estimations, between 22:00 and 03:00 solar hours. Each panel includes five days of data.

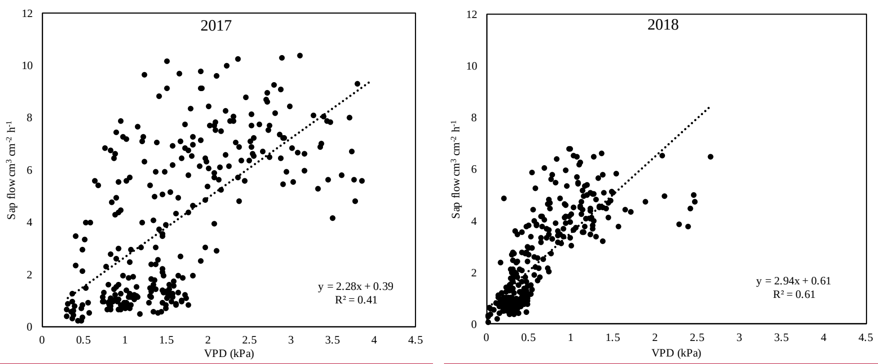


Figure S3. The observed relationship between averaged sap flow for four *Pinus halepensis* and vapour pressure deficit (VPD), between June 9 and June 15, for 2017 and 2018. Each point represents measurements every 30-minute.