We would like to thank the referee for concise and constructive feedback. We will address these issues in the revised manuscript, as according to our responses to each comment below. Our response will be highlighted in blue below the referee's comments in black.

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.

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 in the text:

> "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:

> "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. 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, specifically:

> "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 zero-flow 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:

> "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:

> "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):

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

> 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.

In the text:

> "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."

And regarding the misalignment:

> "The outputs indicate a clear shift of placement in each of the probes over time, here denoted as x_1 and x_2 in each pine (Fig. 2). The eight probes, two per tree, all deviated from the ideal of 0.6 cm. Probe x_2 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:

> "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 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:

> "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 probes in the tree throughout the measurement period avoids this problem and enables the study to focus on the intrinsic factors affecting the sap flow pattern."

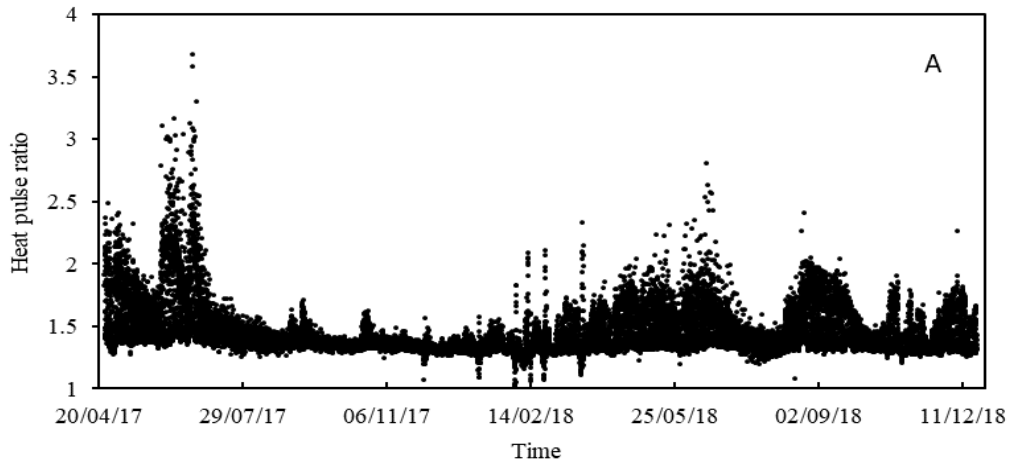
Also, do the authors think these results found on conifers are universal for all types of species? I would have expected a short discussion to clarify to the reader why these findings could be of general value to the application of the method.

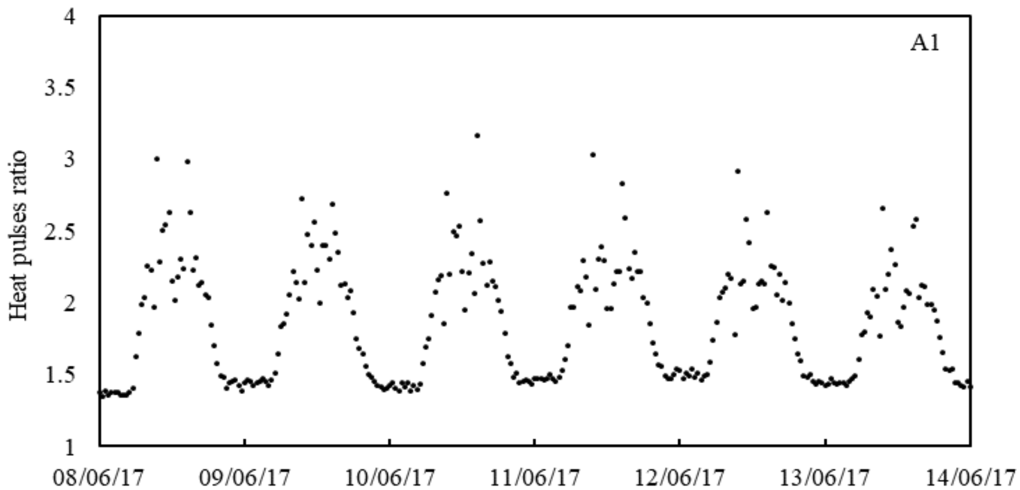
We agree that it's an important addition to the discussion and we have therefore included it in the text:

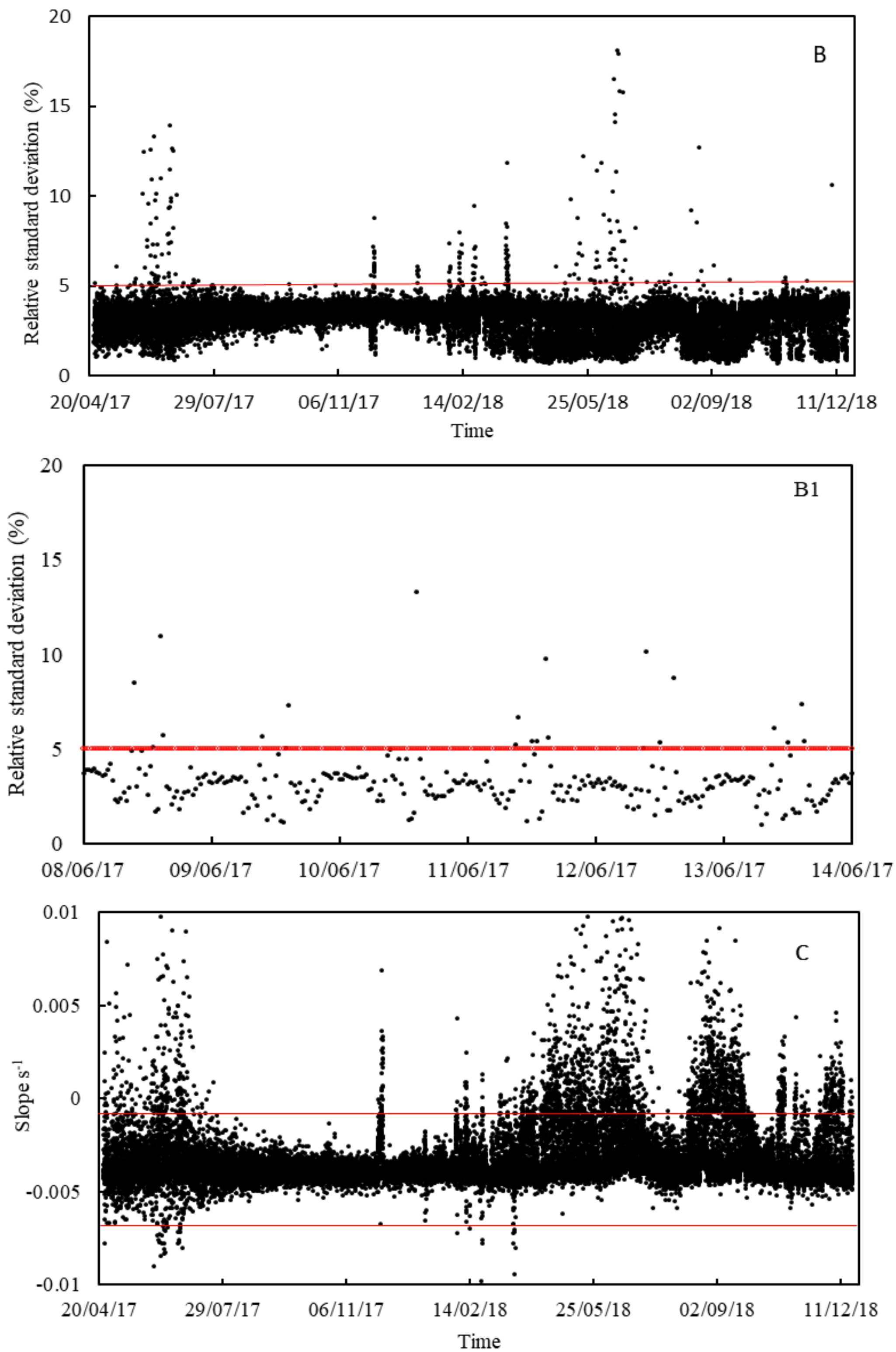
> "The time-dependent correction method could be useful for any species already tested with the HPR, where misalignment of probes can create a source of error and sensors are installed over a longer

period of time. Specifically, where the movement of the wood might cause further displacement of the sensor."

Corresponding figures below:







B
Relative standard deviation (%)
20
15
10
5
0
20/04/17
29/07/17
06/11/17
14/02/18
25/05/18
02/09/18
11/12/18
Time
B1
Relative standard deviation (%)
08/06/17
09/06/17
10/06/17
11/06/17
12/06/17
13/06/17
14/06/17
C
Slope s-1
0.01
0.005
0
-0.005
-0.01
20/04/17
29/07/17
06/11/17
14/02/18
25/05/18
02/09/18
11/12/18
Time

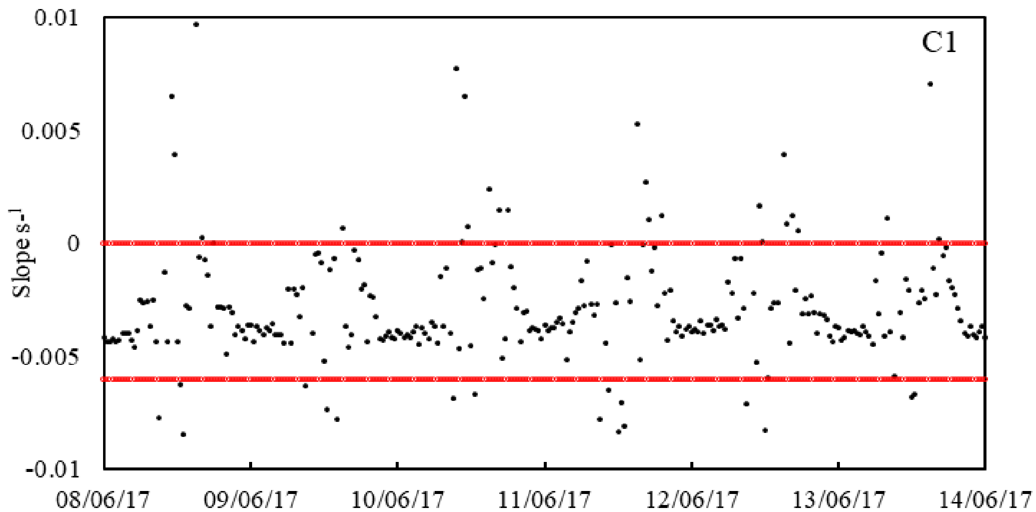
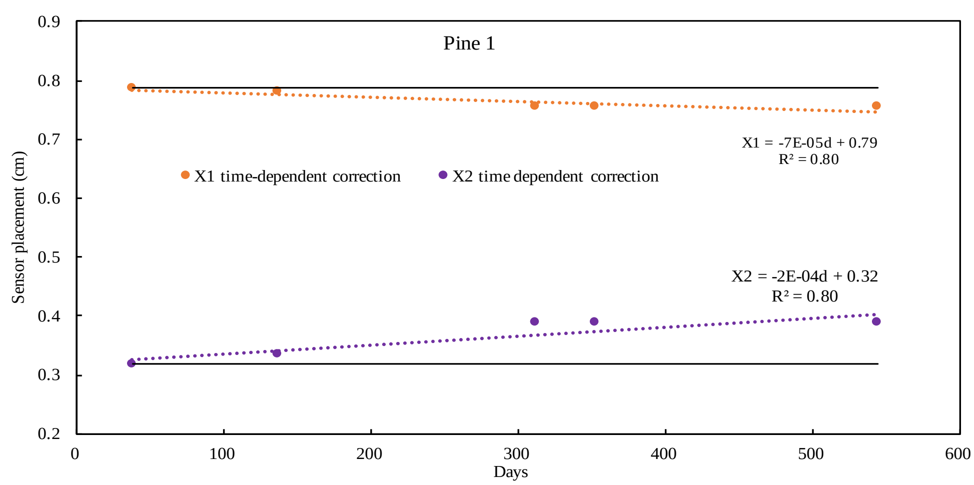


Figure 1. (A) Heat pulse ratios (HPR) throughout the measurement period in 30-minute intervals in pine number 1. Each HPR is an average of 41 instantaneous ratios corresponding to the temperature difference in two thermocouples at 0.6 cm up-and downstream from a heater probe at 0.2 cm depth. (A1) Zoomed in panel of HPR data for one week of measurements. (B) Relative standard deviation (%) for each HPR in pine number 1 for the whole measurement period. Red line indicates threshold used for the quality control were all HPR with 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 |slope – median (slope)| < 0.003 s^{-1} were removed from the data analysis. (C1) Zoomed in panel of slope data for one week of measurements.



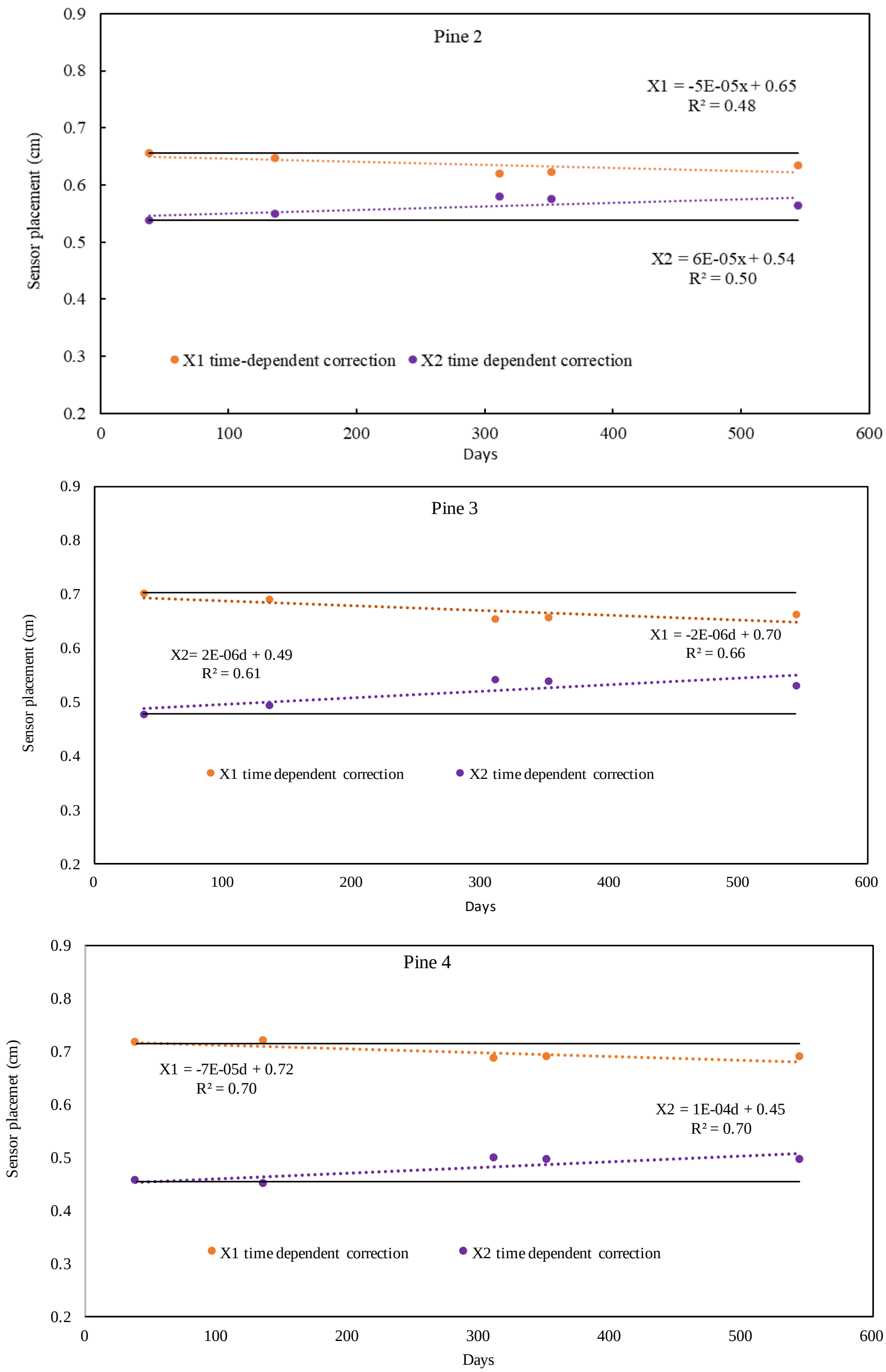
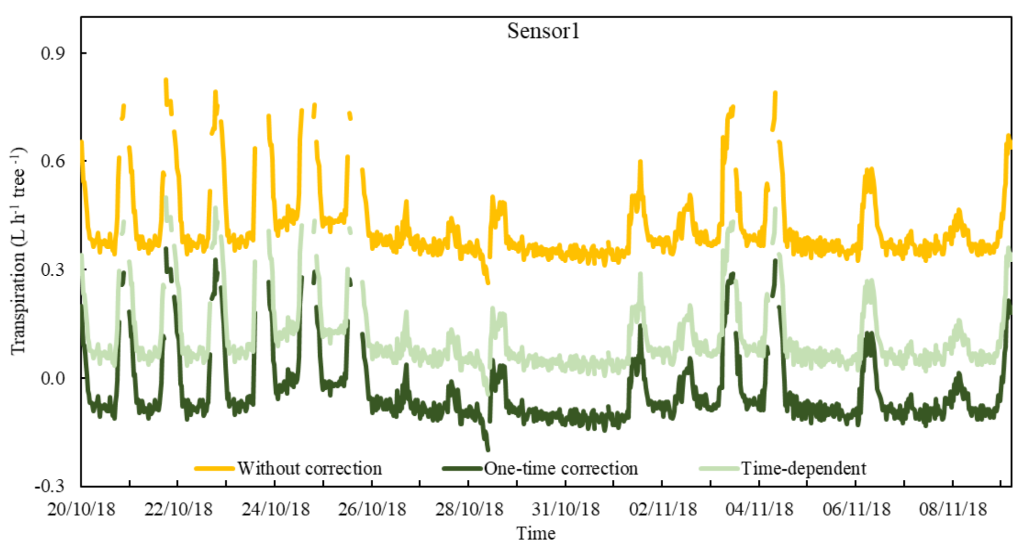


Figure 2. The placements of the probes are shown as distance from the heater (cm). Probe placement was calculated once (solid lines) for the whole study period and compared to probe placement calculated varies times (solid circles with dotted lines) throughout the study period. Each point represents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

Table 1. Seasonal averages of sap velocity for 4 different pines. All sap flow values are expressed in $cm^3\ cm^2\ h^{-1}$. Sap flow corrected for with time-dependent misalignment calculations are compared with sap 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.

Pine	Correction method	Sp-17	Su-17	Fa-17	Wi-17	Sp-18	Su-18	Fa-18
1	One-time correction	0.98±1.2	0.23±0.8	-0.31±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±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
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
	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
4	One-time correction	0.79±1.1	0.56±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



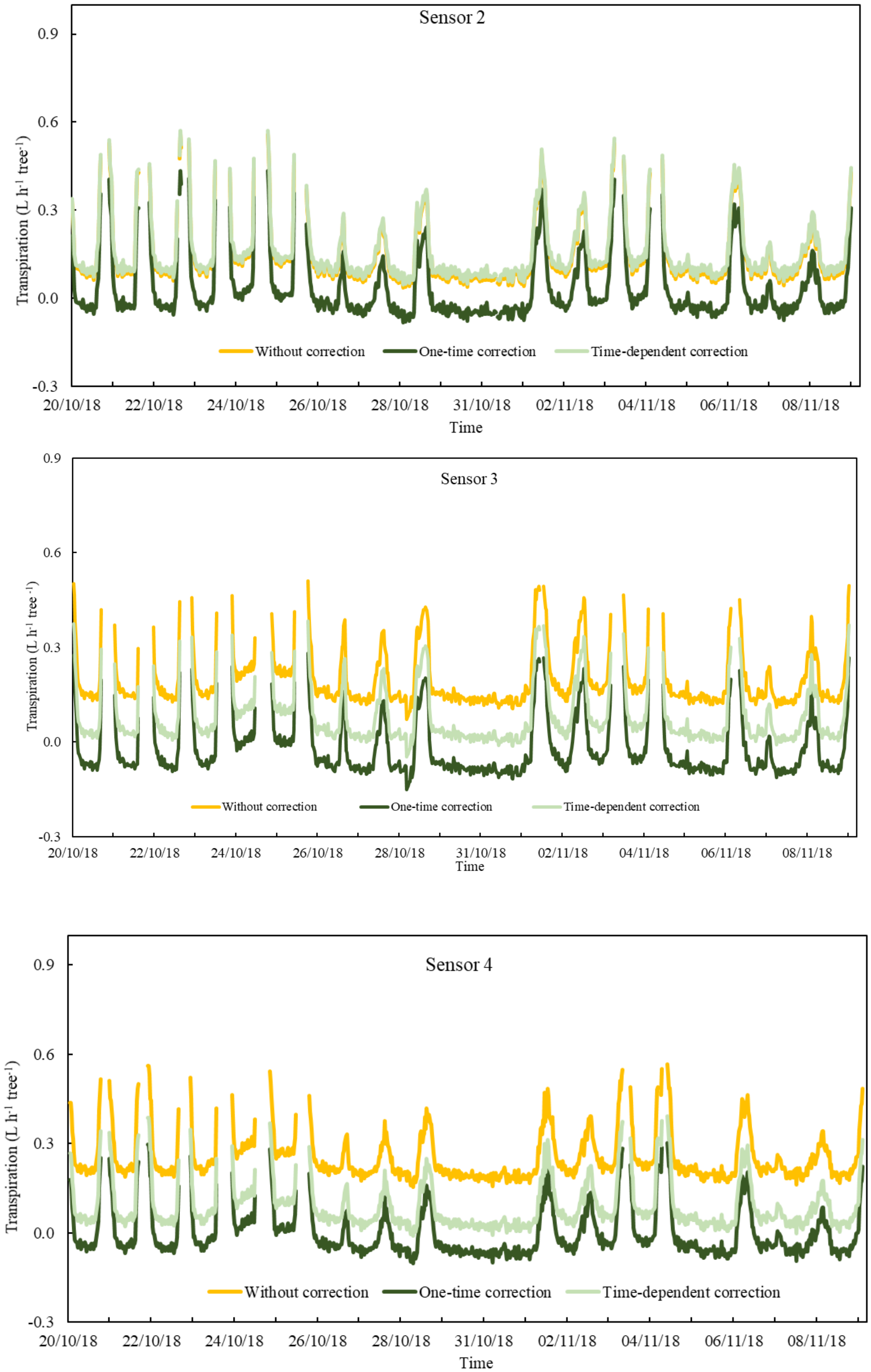
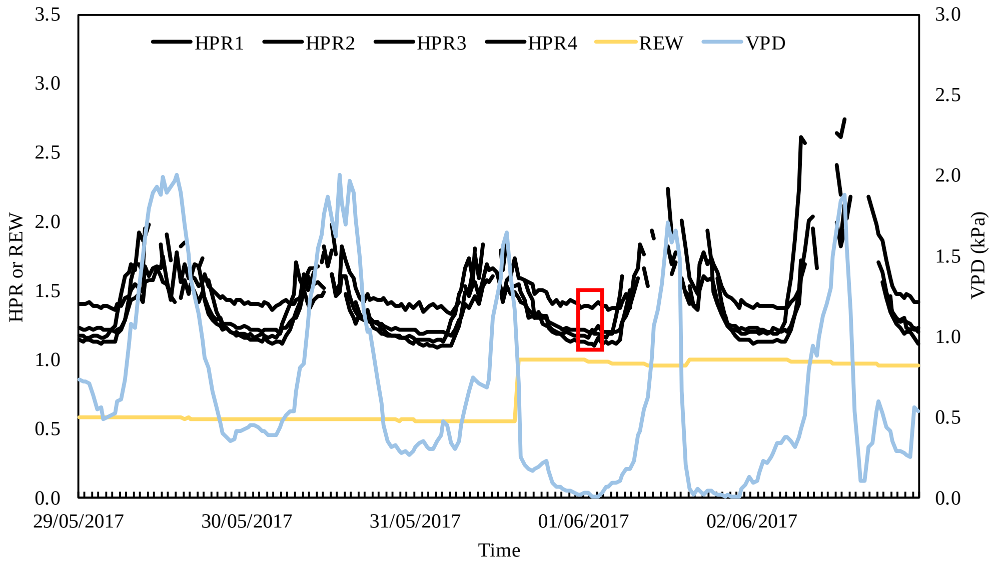
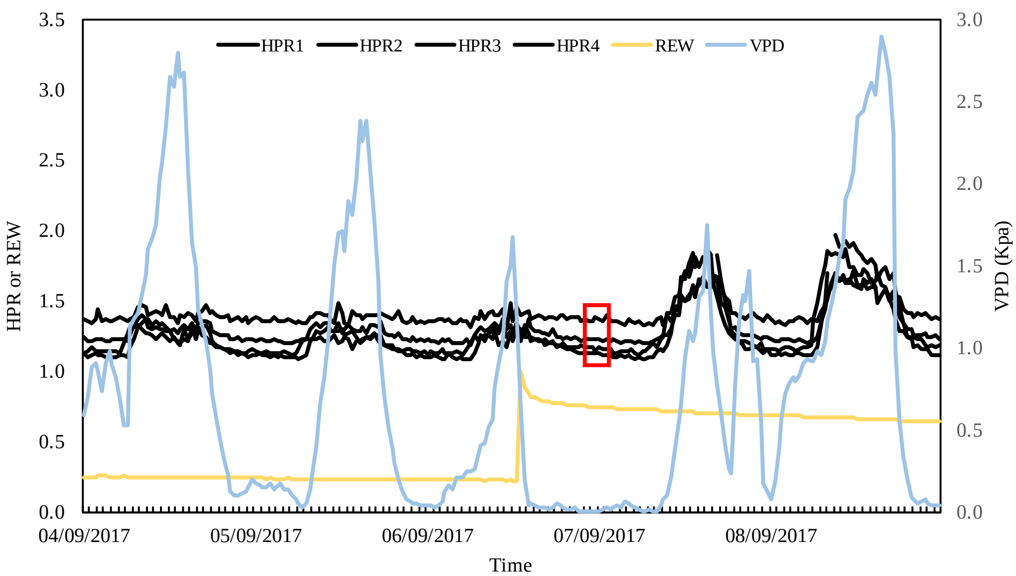


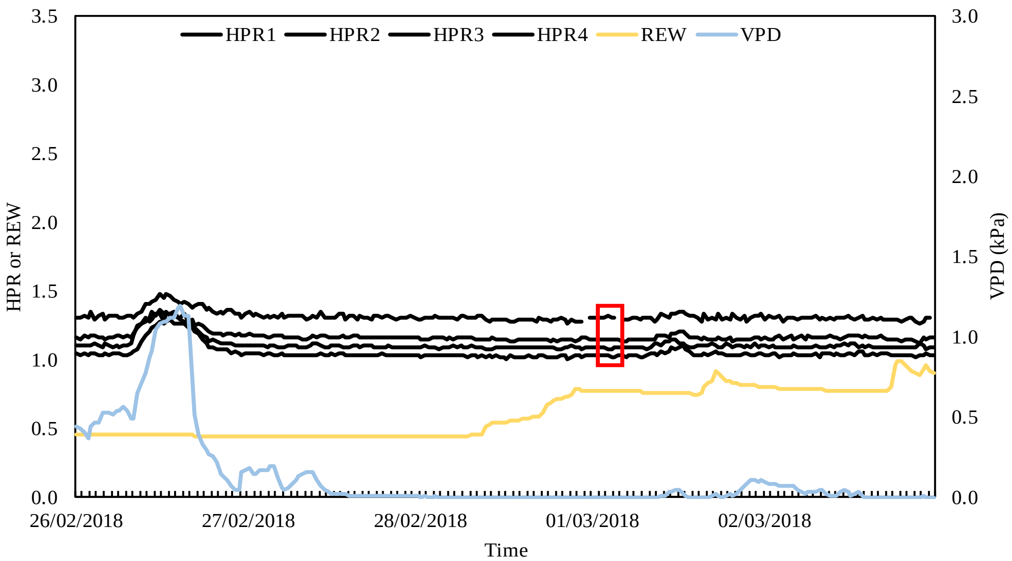
Figure 6. Timeline of transpiration estimates (L h^{-1} tree $^{-1}$) for each sensor and each tree during 3 weeks at the end of the measuring period. Three estimations are shown; without any correction for misalignment of the probes (yellow line), one-

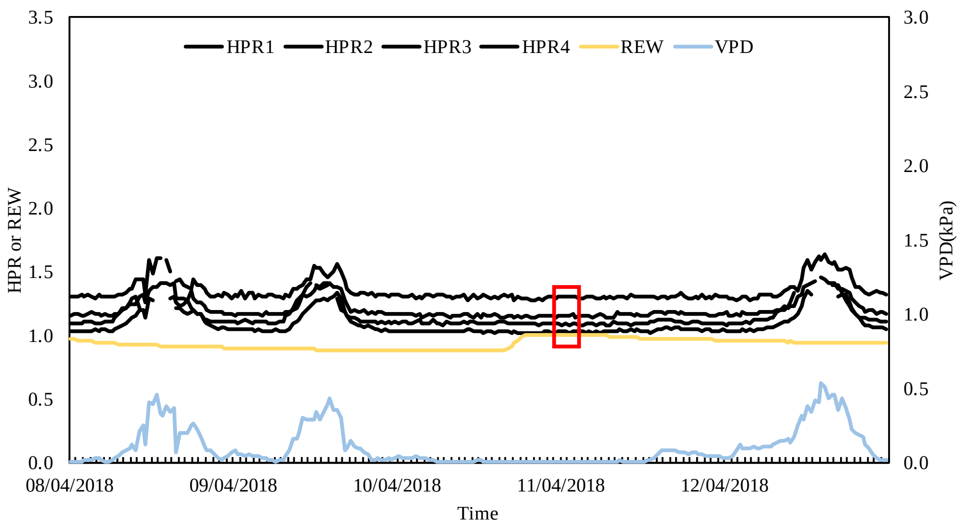
time correction for misalignment of the probes (dark green), and time-dependent correction for misalignment of the probes (light green). The timeline represents diurnal measurements taken at a 30-minute interval.

Supplementary material:






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


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

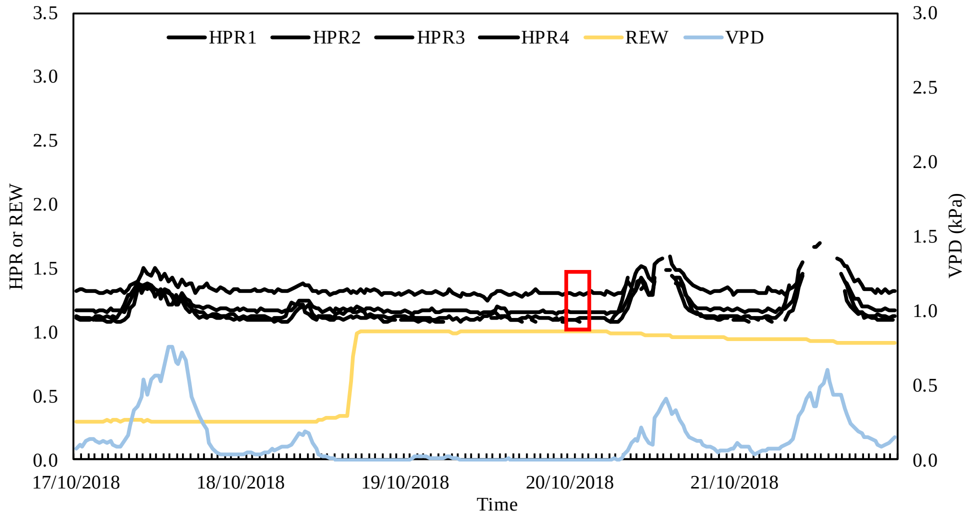


Figure. ***S2.*** *Vapour pressure deficit (VPD), relative extractable water (REW) and heat pulse ratio (HPR) for all trees at five different events assuming zero flow conditions. Squares indicate the HPR readings used for zero flow estimations, between 22:00 and 03:00 solar hours. Each panel includes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

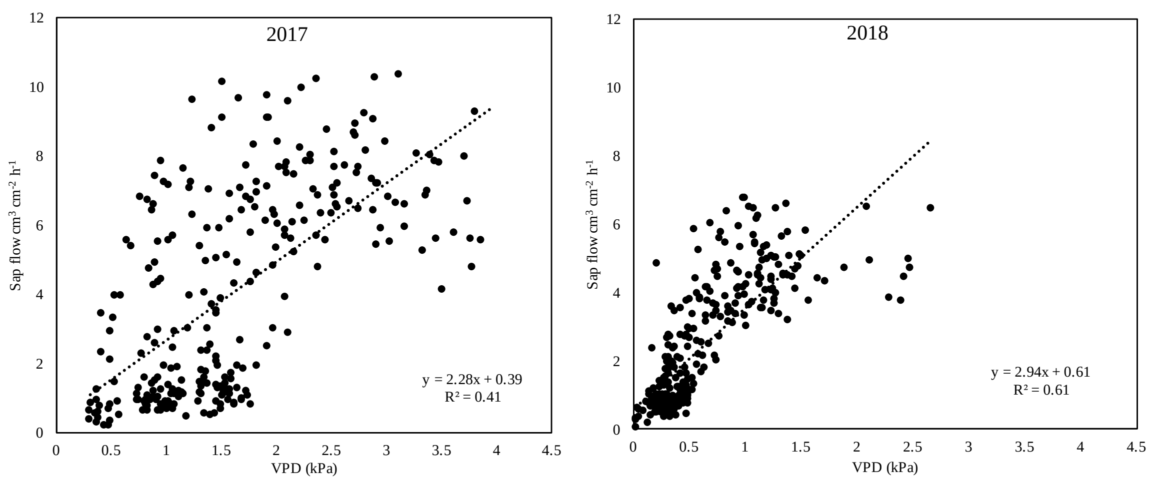


Figure S3. The observed relationship between averaged sap flow for four *Pinus halepensis* and vapour pressure deficit (VPD), between June 9 and June 15, for 2017 and 2018. Each point represents measurements every 30-minute.