ASSESSING THE USE AND THE APPROACH OF SAP FLOW CALIBRATION IN TREES

by

MACKENZIE DIX

(Under the Direction of Doug P. Aubrey)

ABSTRACT

Sap flow calibration is crucial for accurately estimating transpiration in trees; however, we do not understand how different calibration approaches affect sap flow estimates of tree transpiration and most publications rarely apply calibration coefficients. To examine the prevalence of calibration, a review of the past 11 years of tree sap flow publications was conducted. Additionally, we performed an experiment comparing two calibration approaches that use either positive or negative force to move water through stems. Through this research, we confirmed that despite the repeated emphasis on the importance of calibration, it is far from common within the tree sap flow community. Furthermore, we determined that differences in sap flow approaches have significant effects on the calculation of transpiration. Sap flow studies should perform calibration and report results to improve our measurements and understanding of transpiration.

INDEX WORDS: ecohydrology, hydrologic cycle, sap flux, thermal dissipation, transpiration, trees, water use
ASSESSING THE USE AND THE APPROACH OF SAP FLOW CALIBRATION IN TREES

by

MACKENZIE DIX

BA, Franklin and Marshall College, 2014

A Thesis Submitted to the Graduate Faculty of The University of Georgia in Partial Fulfillment of the Requirements for the Degree

MASTERS OF SCIENCE

ATHENS, GEORGIA

2019
ASSESSING THE USE AND THE APPROACH OF SAP FLOW CALIBRATION IN TREES

by

MACKENZIE DIX

Major Professor: Doug P. Aubrey
Committee: Dan M. Johnson
C. Rhett Jackson

Electronic Version Approved:

Suzanne Barbour
Dean of the Graduate School
The University of Georgia
August 2019
ACKNOWLEDGEMENTS

This project would not have been possible without the support and help throughout from my advisor Doug Aubrey. I would also like to thank Seth Younger, who provided me with R code and helped me process my sap flow sensor data. In addition, Chris Lewis was instrumental in helping me set up calibration experiments before sunrise.
# TABLE OF CONTENTS

<table>
<thead>
<tr>
<th>Section</th>
<th>Page</th>
</tr>
</thead>
<tbody>
<tr>
<td>ACKNOWLEDGEMENTS</td>
<td>iv</td>
</tr>
<tr>
<td>LIST OF TABLES</td>
<td>vii</td>
</tr>
<tr>
<td>LIST OF FIGURES</td>
<td>viii</td>
</tr>
<tr>
<td>CHAPTER</td>
<td></td>
</tr>
<tr>
<td>1 INTRODUCTION AND LITERATURE REVIEW</td>
<td>1</td>
</tr>
<tr>
<td>2 RECALIBRATING THE SAP FLOW COMMUNITY</td>
<td>4</td>
</tr>
<tr>
<td>Abstract</td>
<td>5</td>
</tr>
<tr>
<td>2.1 Introduction</td>
<td>5</td>
</tr>
<tr>
<td>2.2 Methods &amp; Results</td>
<td>7</td>
</tr>
<tr>
<td>2.3 Discussion</td>
<td>8</td>
</tr>
<tr>
<td>2.4 Conclusion</td>
<td>11</td>
</tr>
<tr>
<td>References</td>
<td>12</td>
</tr>
<tr>
<td>3 SAP FLOW CALIBRATION APPROACHES SAP FLOW ESTIMATES</td>
<td>19</td>
</tr>
<tr>
<td>Abstract</td>
<td>20</td>
</tr>
<tr>
<td>3.1 Introduction</td>
<td>20</td>
</tr>
<tr>
<td>3.2 Methods</td>
<td>24</td>
</tr>
<tr>
<td>3.3 Results</td>
<td>30</td>
</tr>
<tr>
<td>3.4 Discussion</td>
<td>32</td>
</tr>
<tr>
<td>3.5 Conclusion</td>
<td>37</td>
</tr>
</tbody>
</table>
References.................................................................................................................. 37

4 CONCLUSION........................................................................................................... 52

REFERENCES ............................................................................................................. 54

APPENDICES

APPENDIX ................................................................................................................... 64

Table A1 ....................................................................................................................... 65
Table A2 ....................................................................................................................... 66
Figure A1 ..................................................................................................................... 67
Figure A2 ..................................................................................................................... 68
LIST OF TABLES

Table 1: Table 2.1 ..............................................................................................................16
Table 2: Table 3.1 ...........................................................................................................44
# LIST OF FIGURES

<table>
<thead>
<tr>
<th>Figure 1: Figure 2.1</th>
<th>Page</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>17</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Figure 2: Figure 2.2</th>
<th>Page</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>18</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Figure 3: Figure 3.1</th>
<th>Page</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>45</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Figure 4: Figure 3.2</th>
<th>Page</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>46</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Figure 5: Figure 3.3</th>
<th>Page</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>47</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Figure 6: Figure 3.4</th>
<th>Page</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>48</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Figure 7: Figure 3.5</th>
<th>Page</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>49</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Figure 8: Figure 3.6</th>
<th>Page</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>50</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Figure 9: Figure 3.7</th>
<th>Page</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>51</td>
</tr>
</tbody>
</table>
CHAPTER 1
INTRODUCTION AND LITERATURE REVIEW

Transpiration accounts for 60-80% of total evapotranspiration making it the dominant source of water lost to the atmosphere from the earth’s terrestrial surface (Jasechko et al., 2013; Schlesinger and Jasechko, 2014). It has been measured to create ecosystem water budgets and to determine watering regimes (Kostner et al., 1996; Krauss et al., 2015; Maier et al., 2017). This makes it an important measurement to help understand ecosystem functions and to manage agricultural and forest resources. Freshwater stores are being depleted globally while the rising population increases demand (Bernacchi and VanLoocke, 2015; Rodell et al., 2018). To successfully manage for this increased need for freshwater, we need to be able to accurately account for all major water inputs and outputs in ecological systems, which includes transpiration. With plant water use primarily coming from forests, where plant productivity is highest, accurately accounting for transpiration will be particularly important (Huston and Wolverton, 2009; Keenan et al., 2015; Köhl et al., 2015). Sap flow has been the most common technique for estimating transpiration because of its relative low cost and ease of installation (Poyatos et al., 2016).

Sap flow is the most common approach to quantifying transpiration, therefore it is important to understand how it is applied in scientific studies. Sap flow sensors have been used in many different capacities including to determine transpiration in agricultural fields such as orchards (Cammalleri et al., 2013) and vineyards (Braun and Schmid, 1999), to assess hydraulic redistribution (Nadezhdina et al., 2010) and to create watershed-scale water budgets (Ford et al., 2007). Early methods to estimate sap flow used dyes as tracers in stems (James and Baker, 1933;
Kramer, 1940). Unfortunately, these methods often required cutting stems, resulting in destroyed samples. One of the first methods to use heat as a tracer, calculated the amount of time it took for a heat pulse to reach a specific location up the stem (Huber 1932, Vandegehuchte and Steppe, 2013). More recently, heat pulse methods including heat ratio method (HR; Burgess et al., 2001), compensation heat pulse (CHP; Marshall, 1958) and Tmax (Green et al., 2003) have been employed to measure the full range of sap movement including transpiration and reverse sap flow (Berry and Smith, 2014; Gotsch et al., 2015; Simonin et al., 2009). There are also methods that require continuous heating: thermal dissipation probes (TD; Granier, 1985), heat-field deformation probes (HFD; Nadezhdina, 2009), and heat balance (HB; Cermak et al., 1973). Each method has various advantages and limitations that can influence how sap flow is applied including difficulty of construction, applicability to the research questions and expense (Vandegehuchte and Steppe, 2013). Thermal dissipation probes are applied twice as often as other sap flow sensors due to their low cost and ease of construction (Peters et al., 2018).

There has been some dispute over the need to perform calibration, as the original publications for many sap flow methods either stated that parameters were universal (Granier, 1985) or that sap flow methods do not require a comparison to a measured water use (Cermak et al., 1973; Marshall, 1958; Nadezhdina et al., 1998). As such, calibrations are currently the exception rather than the norm. Gutierrez and Santiago (2006) was one of the first studies to document the underestimate of TD probes using the original calibration coefficients (Granier, 1985). Although some studies have supported the use of the original parameters (Braun and Schmid, 1999; Lu et al., 2002; Macinnis-Ng et al., 2017; McCulloh et al., 2007; Nadezhdina et al., 2008), others have indicated substantial deviation (Bush et al., 2010; Fuchs et al., 2017; Steppe et al., 2010; Sun et al., 2012). Two studies published in 2010 clearly indicate the need for
calibration of multiple sap flow methods and across xylem anatomies (Bush et al., 2010; Steppe et al., 2010); however, it is not clear that researchers are performing and applying calibrations.

Furthermore, there are multiple calibration approaches applied to determine species-specific coefficients. One approach is the intact canopy potometer that involves placing all above-ground foliage and branches in a reservoir to track water use and relies on transpiration to draw water through a stem using negative pressure (Sun et al., 2012). Another approach is a gravimetric method that involves excising a stem section without branches or leaves and forces water through it using gravity, a positive pressure (Steppe et al., 2010). These calibration approaches use fundamentally different forces to move water through stems that may affect calculating water use in intact trees. However, there is a lack of knowledge as to how each calibration approach varies and how calibration coefficients derived from the different approaches affect sap flow estimates in intact trees.

The objectives of this research were to determine 1) the prevalence of sap flow calibration in trees from 2008-2018, and 2) how sap flow calibration approaches affect estimates of water use in intact trees. To determine the prevalence and application of calibration in tree sap flow literature, a literature search was conducted using Web of Science. To examine the effect of calibration approaches on sap flow estimates, gravimetric and potometer calibrations were performed and unique coefficients were generated. Coefficients from different calibration approaches were applied to sap flow data from intact trees growing in a plantation to determine the impact of calibration approaches on transpiration estimates. Through these two research projects, we have been able to 1) examine sap flow calibration and its application within the sap flow community and to 2) provide recommendations of how to move forward with sap flow methodologies.
CHAPTER 2

RECALIBRATING THE SAP FLOW COMMUNITY

Dix, Mackenzie J., Daniel M. Johnson, Doug P. Aubrey. To be submitted to Agricultural and Forest Meteorology.
Abstract. Transpiration dominates the flux of water from terrestrial ecosystems back to the atmosphere—a trend that increases with tree density across terrestrial biomes. Sap flow measurement is the most widely used method to quantify transpiration, and calibrations can greatly improve the accuracy of transpiration estimates. Here we review the tree sap flow literature to determine the prevalence and application of calibration across sap flow methods. Over the past decade, 6.8% of tree sap flow publications documented 1) calibrations, 2) an application of calibration coefficients from previous calibrations, or 3) an application of correction factors using calibration data. Moreover, a number of these studies focused on calibration instead of estimating accurate sap flow in intact trees, suggesting that the actual percentage of publications in which researchers are incorporating sap flow calibration is even smaller. These data indicate that the sap flow community has failed to adopt calibrations as a best practice. To improve our quantification and understanding of plant water use moving forward, researchers must adopt calibration as a best practice. Without the adoption of calibration within the sap flow community, our understanding of transpiration using this technology will be based on inaccurate measurements.

2.1 Introduction

With freshwater availability decreasing in many biomes, accurately accounting for how water is cycling through ecological systems is becoming increasingly important on a local, regional, national, and global levels with potential policy implications across scales (Bernacchi and VanLoocke, 2015; Rodell et al., 2018). Plants play a particularly important role in the global water cycle because transpiration (T) is the dominant source of water returning to the atmosphere from terrestrial systems (Jasechko et al., 2013; Lawrence et al., 2007). Transpiration contributes 60 – 80 % to terrestrial evapotranspiration (ET; Jasechko et al., 2013; Schlesinger and Jasechko,
2014; Wei et al., 2017). While occupying only 31% of the terrestrial surface, forests account for over 50% of global ET. Indeed, the relative importance of T increases with increasing tree density across terrestrial biomes (Figure 2.1; Chapin et al., 2011; Huston and Wolverton, 2009; Keenan et al., 2015; Schlesinger and Jasechko, 2014). This makes understanding tree water use and how it is quantified crucial to understanding the global water cycle.

Sap flow is the most common technique for estimating T because of its relative low cost and ease of installation (Peters et al., 2018; Poyatos et al., 2016). Heat-based sap flow methods measure differences in temperature as a proxy for measuring T directly and have been applied for nearly a century (Huber, 1932). Sap flow can generally be categorized into heat pulse and constant heating approaches, of which there are four common methods: heat pulse velocity, HPV (Marshall, 1958); heat balance, HB (Cermak et al., 1973; Sakuratani, 1981); thermal dissipation, TD (Granier, 1985); and heat field deformation, HFD (Nadezhdina, 1998). Each method has various advantages and disadvantages including ease of construction, cost, as well as size and type of stem on which sensors can be used (Vandegehuchte and Steppe, 2012). The application of TD probes is implemented twice as often as any other sap flow methods (Peters et al., 2018). Lu et al. (2004) was the first to suggest that calibration should be performed on TD probes whenever the design deviated from that of the original (Granier, 1985). In 2006, Gutierrez and Santiago demonstrated that TD probes are consistently underestimating sap flow rates.

Furthermore, the accuracy of TD probes, as well as other sap flow methods, was brought under wide scrutiny in 2010 (Steppe et al., 2010). Additional publications have since demonstrated the benefits of calibrating (Rubilar et al., 2017; Ma et al., 2017, Hultine et al., 2010). Calibration efforts have also explored the accuracy of sap flow estimates as a function of xylem anatomy (Bush et al., 2010; Sun et al., 2012). The magnitude and direction of error are highly variable,
across and within previous reports, ranging from underestimating actual water use by a factor of three (Rubilar et al., 2017) to overestimating actual water use by 55% (Sun et al., 2012). Recent publications have reaffirmed the inaccuracies of multiple methods and continue to call for calibration (Flo et al., 2019; Fuchs et al., 2017; Peters et al., 2018). Despite these numerous warnings and recommendations (Table 2.1), the performance and application of calibrations seemingly appear to be the exception rather than the rule in reports of tree sap flow.

Here we seek to determine how the sap flow community has responded to recommendations made over the past decade. Specifically, we conducted a literature search to quantify the number of publications that have been published since these calibration issues have been identified to determine if sap flow calibrations have been adopted as part of a best practice for sap flow methodologies.

2.2 Methods & Results

We performed a literature search in Web of Science using the search terms “‘Sap fl*’ OR ‘Sapfl*’”, AND ‘Tree*’” across an 11-year period from 2008 through 2018 to determine the total number of publications that reported on tree sap flow. Calibration was defined as comparing a reference water use to estimated water use from sap flow sensors. Reference water use was measured by lysimeters, gravimetrically, or by water added to a potometer. We also considered publications that described applying calibrated coefficients from previous studies or described applying a correction equation derived from calibration. Publications that documented using other sap flow sensors as a reference water use were not included as calibration publications in our exercise. We began our search in 2008 because two publications definitively demonstrated the importance of calibrations across multiple sap flow sensors and xylem anatomies in 2010 (Bush et al., 2010; Steppe et al., 2010), and we were hoping to identify how their seminal
publications were incorporated as best practices by the sap flow community. The search identified 1,610 publications during that 11-year period with 1,054 reporting on sap flow in trees. Publications that did not include sap flow sensors in trees (i.e., monocots, tree-like monocots, vines or lianas) were not included in this analysis.

Across the 11-year period, we observed a mean (± s.e.) of 95.8 ± 5.6 tree sap flow publications per year with a minimum and maximum of 65 and 130, respectively (Figure 2.2). The number of tree sap flow publications and the number of calibrations increased through the observation period. The rate of increase for tree sap flow publications over that period was 4.4 publications per year. Only 71 of 1,054 publications documented a calibration, use calibration coefficients from previous studies, or applied a correction after applying a calibration experiment. We observed a mean of 6.4 ± 0.8 publications documenting calibrations per year with a minimum and maximum of 2 and 12, respectively. There were 0.7 publications documenting calibrations per year. This translated to a mean of 6.5 ± 0.8% of tree sap flow publications documenting calibrations per year. The minimum observed percentage of tree sap flow publication documenting calibration in a year was 2.3 % in 2013 while the highest percentage observed was 11.1 % in 2017. Of the 71 publications that document calibration, 26 were published to emphasize the importance of calibration as opposed to correcting experimental sap flow. This indicated that only 4.4 % of the tree sap flow publications over this period applied calibration to improve estimates of sap flow in intact trees.

2.3 Discussion

We observed less than 5% of tree sap flow publications published over the past 11 years had documented calibration to validate original coefficients or generate and apply new coefficients. Hultine et al. (2010), Ma et al. (2017), and Rubilar et al. (2017) clearly
demonstrated that species-specific calibrations increased the accuracy of estimated water use. Steppe et al. (2010) and Fuchs et al. (2017) showed the need for calibration across many sap flow methods. Despite the ever-increasing evidence (Table 2.1) that species- and site-specific calibrations improve the accuracy of sap flow estimates and should be performed, it is far from being incorporated as a best practice. In fact, there seems to be a distinct lack of acknowledgment that calibrations are important to increase accuracy.

There is not only a lack of calibration among the sap flow community, but little to no reference of the seminal calibration papers previously published. Only 58, of the 802 sap flow papers published since 2010, have documented calibration. From July 2010 to 2018, Steppe et al. (2010) had been cited 140 times, and from December 2010 to 2018, Bush et al. (2010) had been cited 69 times, according to Web of Science. In other words, these seminal papers clearly demonstrating the importance of sap flow calibration have been cited by less than 20% of the tree sap flow publications. Furthermore, these papers are cited for purposes beyond calibration recommendations, which suggests their impact in terms of calibration is even less. This is also true for Lu et al. (2004), which is the first paper to our knowledge, to mention a need for calibration. It only has 194 citations from June 2004 to 2018, according to Web of Science. While Lu et al. (2004) is only in regard to TD, there have been over 450 publications since 2004, according to Peters et al. (2018), using TD. We would expect higher levels of citations for these important calibration publications, even if these publications were cited only to explain why calibration was unnecessary within the scope of various studies. The lack of acknowledgment in the form of citations is reflective of the sap flow community’s unwillingness to acknowledge the proverbial elephant in the room, calibration.
The improved accuracy in sap flow estimates that is gained through calibration represents an inconvenient truth for the sap flow community because the calibration process is resource intensive. Calibration is expensive in terms of plant material, how much time is needed to be devoted to it, and equipment. All calibrations require felling trees. In some cases, when trees are limited because the study involves a rare or endangered species (e.g. Zhao et al. 2009), calibration may not be feasible. There may be further restrictions placed on felling trees if they are located in restricted areas (e.g. national parks, or plantations). The equipment costs cannot be ignored when considering many sap flow sensors can only be used once. The cost of time is always a factor when considering protocols, as time is a limiting factor when making any decision. With all this being understood, if the purpose of a study is to accurately quantify sap flow estimates, without calibration, there is no way to know if estimates are accurate or not. It has been nearly two decades since Lu et al. (2004) identified a need for calibration in TD and a decade since Steppe et al. (2010) identified errors in estimated sap flow for TD, HFD and CHP, yet the sap flow community has failed to embrace calibration as a best practice. The ultimate reason for a lack of calibration is that as a community, we are not holding each other accountable to produce the highest quality data possible. While journal editors and reviewers need to hold authors accountable for applying calibration when estimating absolute water fluxes, the onus falls on researchers to perform calibrations prior to submitting manuscripts for publication. As researchers and reviewers are one in the same, this should initiate a positive feedback that establishes calibration as a best practice for publications reporting sap flow.

Through examining the tree sap flow literature, we have been able to identify aspects of calibration that may be important, but are unable to offer strict recommendations moving forward, as calibration literature is limited. The sap flow community may need to consider
applications of sap flow that do not require calibration. This may be possible, for example, when sap flow is being compared in trees of the same species of a similar age and size on similar sites. Sap flow methods in these cases would provide another tool to be able to address relative water use or qualitative data, as sap flow without calibration is highly correlated with plant water use (Flo et al., 2019). We do not, however, recommend comparing the same species in extremely different environments without applying calibration, as it has been suggested that site- and species-specific calibrations may be important (Steppe et al., 2012). Pooling data to determine correction factors based on site, sensor, or wood properties has been suggested in a published synthesis of 290 calibrations representing 55 studies (Flo et al., 2019) and in two studies that examined six species each (Bush et al., 2010; Fuchs et al., 2017). The suggestion to apply a pooled correction factor continues to perpetuate calibration avoidance. As a result, pooled corrections further impede the progress we could be making in understanding plant water use if calibrations were universally applied. Furthermore, there are some fundamental differences in calibration approaches, including driving force and calibration material, but there is not enough data to determine fully how these difference in calibrations affect water use estimates (Flo et al., 2019; Fuchs et al., 2017). While we are making some suggestions about when calibration is important and it may be less important, we cannot recommend sap flow best practices until calibration is more thoroughly researched and standardized. Only then can we assess the best calibration approaches and their affect on sap flow estimates.

2.4 Conclusion

We have shown that, despite a number of publications that demonstrated the importance of calibrating sap flow sensors, only a small portion of studies actually perform calibrations. Failure to adopt calibration as a best practice in the sap flow impedes our understanding of terrestrial
water cycling across stand, landscape, and global scales. Specifically, we sacrifice the accuracy in constructing water budgets and these inaccuracies propagate through synthesis and modeling efforts. The need to calibrate is an inconvenient truth and our literature review indicates that our community has failed to adequately acknowledge and address the issue. Failure to adopt calibration as a best practice in the sap flow community may leave a lasting impact on our ability to progress a variety of disciplines that are routinely implementing these measurements. The development of best practices for sap flow that include calibration is needed to progress the disciplines implementing sap flow as a common technique for quantifying forest transpiration. We encourage researchers to place additional effort into acquiring the highest-quality sap flow data possible through calibration and for journal editors and reviewers to hold the community to higher standards that will improve the quality of inference we strive to make through our research endeavors.

References


Hultine, K.R. et al., 2010. Sap flux-scaled transpiration by tamarisk (Tamarix spp.) before, during and after episodic defoliation by the saltcedar leaf beetle (Diorhabda carinulata). Agricultural and Forest Meteorology, 150(11): 1467-1475.


Table 2.1 Recommendations to calibrate taken directly from publications.

<table>
<thead>
<tr>
<th>Publication</th>
<th>Recommendations</th>
</tr>
</thead>
<tbody>
<tr>
<td>Lu et al. (2004)</td>
<td>“It is evident that recalibration should be undertaken on any sensor probes that have a different geometry or heating power to the original [TD(^1)] design.”</td>
</tr>
<tr>
<td>Bush et al. (2010)</td>
<td>“Our results suggest that the original calibration of Granier is not universally applicable to all species and xylem types and that previous estimates of absolute rates of water use for ring-porous species obtained using the original calibration coefficients may be associated with substantial error.”</td>
</tr>
<tr>
<td>Steppe et al. (2010)</td>
<td>“We conclude that a species-specific calibration is necessary when using any of these techniques [TD, HFD(^2), CHP(^3)] to insure that accurate estimates of sap flux density are obtained, at least until a physical basis for an error correction can be proposed.”</td>
</tr>
<tr>
<td>Sun et al. (2012)</td>
<td>“We conclude that species specific calibrations can substantially increase the accuracy of the thermal dissipation technique.”</td>
</tr>
<tr>
<td>Steppe et al. (2012)</td>
<td>“Recalibration for each new tree species on which [TD, HFD, HPV(^4)] are used is recommended, at least until a physically based error correction protocol is established or new sap flux density calculation approaches emerge.”</td>
</tr>
<tr>
<td>Fuchs et al. (2017)</td>
<td>“We conclude that (i) TD and HFD sensors require species-specific calibration to measure sap flux with high accuracy, (ii) the original Granier equation cannot be used for TD probes with deviating design, and (iii), at low to medium flow rates, the highest accuracy can be achieved with HRM(^5) sensors.”</td>
</tr>
<tr>
<td>Peters et al. (2018)</td>
<td>“Development of calibration curves is thus important for obtaining more accurate absolute [sap flux density] estimates.”</td>
</tr>
<tr>
<td>Flo et al. (2019)</td>
<td>“…all sap flow methods showed high precision, allowing potential correction of the measurements when a study-specific calibration is performed.”</td>
</tr>
</tbody>
</table>

\(^1\)TD: Thermal dissipation\(^2\)HFD: Heat field deformation\(^3\)CHP: Compensation heat pulse \(^4\)HPV: Heat pulse velocity \(^5\)HRM: Heat ratio method
Figure 2.1 Relationship between the proportion of transpiration to evapotranspiration (T/ET; %) from Schlesinger and Jasechko (2014) and global stem density (stem/ha) from Crowther et al. (2015), $y = 1.94x + 5.15$ ($R^2 = 0.48$).
Figure 2.2 (A) Total number of tree sap flow publications (black, solid line), the number of publications that document calibration or use calibrated parameters (black, dashed line), and the percent of publications that calibrate or use calibration (blue, dotted line) from 2008-2018. (B) Total number of publications published form 2008-2018 with percent of publications calibrated for each method including compensation heat pulse (CHP), heat balance (HB), heat field deformation (HFD), heat pulse (HP), heat ratio (HR), Sapflow+ (Sap+), Thermal Dissipation (TD), and Tmax (Tmax).
CHAPTER 3

SAP FLOW CALIBRATION APPROACH INFLUENCE SAP FLOW ESTIMATES

\[\text{2 Dix, Mackenzie J. and Doug P. Aubrey. To be submitted to Agricultural and Forest Meteorology.}\]
Abstract. Calibrating thermal dissipation sap flow sensors has become increasingly important to accurately estimate plant water use, but there is still uncertainty as to how calibration approaches influence sap flow estimates. Here we compare the two most common calibration approaches, gravimetric and potometer, using thermal dissipation probes. The gravimetric approach uses an excised stem segment devoid of leaves or branches and pushes water through the stem using gravity, a positive force. The potometer approach requires severing a tree stem at its base and placing it upright with the exposed stem in a reservoir where it is allowed to pull water through the stem via transpiration, a negative force. Gravimetric and potometer calibration was conducted using six *Eucalyptus benthamii* trees, 4.7-6.3 cm diameter at breast height. Estimates of conducting sapwood area were 11% larger under gravimetric compared to potometer calibration. The calibration approaches generated coefficients that differed from each other and from the original coefficients. Sap flow estimated derived from gravimetric coefficients differed from estimates derived from potometer coefficients by more than 200% when applied to intact trees, implying that conducting area is not the only factor affecting sap flow estimates. Two-piece calibration curves, which fit two separate curves to calibration data based on sap flux index (K) values observed in intact trees, diminished the magnitude of the difference in estimated sap flow derived from gravimetric and potometer coefficients. Future research should explore physical and physiological reasons why these calibration approaches differ and how these approaches are related to the transpiration of intact trees to develop a standard calibration approach to be used across tree sap flow studies.

3.1 Introduction

Plant water use has been measured to create ecosystem water budgets and to determine watering regimes making it an important measurement for understanding ecosystem functions
and in managing agriculture and forestry (Ford et al., 2005; Maier et al., 2017). Quantifying individual plant water use by measuring sap flow is a common approach to estimating transpiration and has substantially increased our understanding of species-specific physiological responses to changes in the environment (Allen et al., 2016; Smith and Allen, 1996). There are multiple sap flow methods employed including heat pulse (Green et al., 2003), heat balance (Cermak et al., 1992), heat-field deformation (Nadezhdina et al., 2012), and thermal dissipation (Granier, 1985). Each sap flow method has advantages and disadvantages, but thermal dissipation (TD) is the most commonly used method (Peters et al., 2018; Poyatos et al., 2016; Vandegehuchte and Steppe, 2013). Although some studies have supported Granier’s original coefficients (Braun and Schmid, 1999; Catovsky et al., 2002; Lu et al., 2002; McCulloh et al., 2007), others have indicated substantial deviation (Fuchs et al., 2017; Steppe et al., 2010; Sun et al., 2012), which has indicated a need to perform species- and site-specific calibrations.

Despite numerous recommendations, sap flow calibrations are relatively rare in the literature and standardized calibration approaches remain undeveloped. Flo et al. (2019) synthesized past calibration studies across calibration materials (i.e., whole trees, stem segments, and trees without roots) and concluded that stem segments produce higher accuracy when sap flow estimates are compared to reference water use. However, the higher accuracy of these measurements is likely the result of using high precision equipment in a highly controlled environment (Fuchs et al., 2017). One potentially critical aspect of sap flow calibration that has not previously been considered is how these calibration experiments are generating the flow of water through stems, by either applying a positive or negative.

The type of force used to move water through tree stems may influence the coefficients derived from calibrations. One approach for calibrating sap flow sensors has been to take excised
stems and push water through them using a gravitational force (Steppe et al., 2012). This approach was used by Granier (1985; 1996) to determine the original coefficients from experiments on five species. Positive pressure may force water through parts of a stem that do not usually conduct water; however, it is currently unknown whether this happens. An increase in the area of the stem’s conducting tissue will increase sap flow velocity and result in inflated coefficients. An overestimate of sap flow coefficients will cause an underestimate in sap flow calculations on experimental trees, which will have decreased sap flow velocities. Another common calibration approach is an intact canopy potometer (Gutiérrez and Santiago, 2006; Lu and Chacko, 1998; Maier et al., 2017; Roberts, 1977; Sun et al., 2012). Briefly, intact canopy potometer calibration requires felling entire trees, placing all aboveground biomass in a reservoir to track water use over time. An intact canopy potometer uses transpiration, a negative force, to pull water through the conductive tissue of the stem. Errors from this approach could occur from embolisms when trees are cut down which causes a decrease in conductive tissue. Decreased conductive tissue will result in decreased sap flow rates during calibration and an underestimate of coefficients. When coefficients are applied to experimental trees, which will have higher sap flow velocities, sap flow may be overestimated. Fuchs et al. (2017) considers that studies that have applied negative pressure, via a vacuum, have concluded that measured sap flow underestimates actual sap flow (e.g., Bush et al., 2010; Hultine et al., 2009; Taneda and Sperry, 2008). Steppe et al. (2010) concluded that heat field deformation, compensation heat pulse, and thermal dissipation sap flow methods underestimate reference water use when applying positive pressure. Sun et al. (2012) determined that calibrations relying on transpiration, a negative pressure, resulted in both over- and underestimates of cumulative water use across six species representing tracheid, diffuse-porous and ring-porous xylem anatomies. Taken together, these
results suggest that a potentially major influence on the coefficients derived from calibration coefficients is the approach used to move water through tree stems; however, we are unaware of previous side-by-side comparison calibration imposing different forces to move water.

Regardless of the physical force used to move water through tree stems, calibrations have maintained the non-linear power function to relate observed sap flux density ($F_d$) with the estimated dimensionless sap flow index ($K$) and calculate $\alpha$ and $\beta$ parameter coefficients, $F_d = \alpha K^\beta$, thereby, constructing a single set of coefficients to apply across the entire range of $K$. Bush et al. (2010) demonstrated with calibration data and Holtta et al. (2015) demonstrated through modeling that the relationship between $F_d$ and $K$ is sensitive to the range of $K$. These observations provide good evidence that $K$ values during calibration should be similar to those observed in experimental trees. Unfortunately, studies that perform calibration and also install sensors in trees in the field rarely explicitly examine this relationship (e.g., Hultine et al., 2010; Steppe et al., 2010). However, some studies have determined calibrations underestimate sap flow at low flow rates with decreased error as flow rates increase (Gutierrez and Santiago, 2006; Liu et al., 2008; Sun et al., 2012). Calibration removes the constraint of the non-linear power function and can provide an opportunity to generate coefficients based on the range of $K$ observed in the field, potentially resulting in improved estimates. However, we are unaware of any calibration studies that have explored this potential.

Here, we performed side-by-side calibrations of TD probes using intact canopy potometer and stem-segment gravimetric approaches, hereafter referred to as potometer and gravimetric respectively. We hypothesized that coefficients would differ between calibration approaches as a function of conducting sapwood area estimates made under positive or negative pressure. We predicted that the original TD coefficients (Granier, 1985) would more closely estimate the
observed sap flow under the gravimetric approach compared to the potometer approach and that coefficients generated under one approach would poorly estimate observed sap flow from the other approach. We further hypothesized that the different coefficients, when applied to intact trees, would yield different estimates of sap flow and that these differences would be proportional to the differences in conducting sapwood area observed under different types of pressure. Following the expected differences in conducting sapwood area, we predicted that, when applied to intact trees, coefficients derived from the gravimetric calibration approach would result in lower sap flow estimates than those derived from the potometer calibration approach. Finally, we hypothesized that calibration coefficients generated using the range of $F_d$ and $K$ observed in intact trees would yield different sap flow estimates than those generated across the full range of calibration $F_d$ and $K$.

### 3.2 Methods and Materials

#### 3.2.1 Plant material and experimental site description

Calibration experiments were performed throughout the summers of 2017 and 2018. The six eucalypts (*Eucalyptus benthamii*) used in each of the calibration experiments were planted in the spring of 2014 at the Savannah River Site (SRS), near Aiken, South Carolina, USA. The SRS is in the coastal plain region in the Southeastern United States and is characterized by sandy soils. The mean annual temperature is 20 °C and mean annual rainfall is 1219 mm.

#### 3.2.2 Sensor construction and theory

TD probe construction was based on the method developed by Granier (1985). Each sensor consists of two, 20 mm-long needles, a reference needle and a heater needle. The heater needle uses a constantan heating wire (TFCC-005, Omega Engineering, Stamford, USA) wound around the outside of the needle to supply heat to the length of the needle. Both reference and
heater needles contain a copper-constantan thermocouple (TFCP-005/TFCI-005, Omega Engineering, Stamford, USA) at their center, 10 mm, used to measure the temperature difference between needles. A low viscosity epoxy was applied to protect wire junctions. Sensors were inserted into tree stems using a predrilled, aluminum template that ensured parallel placement and 10 cm spacing between reference and heater needles. A leather punch was used to remove bark and expose sapwood. To be certain sensors were in direct contact with sapwood, drill bits for reference (1.5 mm) and heater needles (2.4 mm), were used with a battery-powered drill to install sensors. The upper needle was heated with a constant voltage of 1.8 v while the lower needle remaining unheated (Granier, 1985; Lu et al., 2004). As a result, when sap flow increases, the temperature difference between the two needles decreased asymptotically (Clearwater et al., 1999).

TD is based on the theory that for a system in thermal equilibrium, at a constant sap flux density, it can be assumed that input of heat, through an electrical current flowing through a resistor, is equal to the heat dissipated through convection and conduction (Cabibel and Do, 1991; Granier, 1985; Lu et al., 2004). Granier (1985) in his original paper discussing TD probe construction and calibration determined that the equation to represent this relationship is:

\[ F_d = \alpha K^\beta \]

where \( F_d \) (g m\(^{-2}\) second\(^{-1}\)) is equal to the sap flux density and \( K \) is the dimensionless flow index represented by the equation:

\[ K = \frac{\Delta T_{max} - \Delta T}{\Delta T} \]

where \( \Delta T_{max} \) is the maximum temperature difference between the two needles, when there is no sap flow, and \( \Delta T \) is the temperature difference at a given time. By integrating sapwood area (SA:m\(^2\)) at the point where probes are inserted, sap flow, \( F_s \) (g s\(^{-1}\)), can be calculated:
3.2.3 Gravimetric calibration

Gravimetric calibration was performed in a greenhouse at the University of Georgia’s Savannah River Ecology Laboratory (SREL) during the summer of 2018 at the SRS. Six, ~1 m long segments were removed from tree stems after felling. Stem segments were submerged in water during transport back to SREL. Once back at SREL, stem segments were either resubmerged underwater in a cooler at ~4°C for storage or used immediately for calibration. Stem segments were attached to a Mylar cylinder at least 100 cm high on stem segments to maintain the natural sap flow patterns. Bark was not be removed from eucalypts because it was determined through staining sapwood during calibration that water was not flowing through the bark. Mylar tubes were attached using silicon, large hose clamps, and expanding foam and allowed to dry for at least 18 hours. To prevent stem segment dehydration, stem segments ends were kept moistened throughout set up. Stem segments were secured to a Marriotte-based verification system that allowed for constant pressure heads, based on the method used in Steppe et al. (2010). Each stem segment was equipped with two TD probes inserted on opposite sides of the stem segments and connected to a multiplexor (AM16/32B multiplexer, Campbell Scientific Inc., Logan, UT, USA) that was connected to a battery-powered datalogger (CR1000 Campbell Scientific Inc., Logan, UT, USA). Water flowing through stem segments was measured with a scale to the 0.001 kg (T51XW, Ohaus, Parsippany, NJ, USA), also connected to the datalogger. Differential voltage and mass of water was collected every 5 seconds and averaged every 1 minute and stored in the datalogger. Measurements were taken at constant head heights for at least 45 minutes or until constant differential voltages were recorded for at least 15 minutes at each head height of 2, 5, 10, 20, 40, 60, 80, and 100 cm. After stem flow was measured at

\[ F_s = \alpha R^\beta \times SA \]
various head heights, TD probes remained connected to a datalogger overnight for zero flow ($\Delta T_{max}$) measurements. Blue food dye was added during calibration to stain conducting sapwood.

3.2.4 Potometer calibration

Potometer calibration was performed with six eucalypts in summer 2017. Trees were cut pre-dawn when transpiration was assumed to be negligible and transported to the experimental location nearby on the SRS. Stems ends were then submerged underwater and cut again to remove 10 cm of the stem that could have embolized during transport. Stems were transferred to 3 L reservoirs filled with 40 mM KCl solution. Reservoirs were placed on concrete platforms and stems were secured to living trees directly next to them to ensure they would remain upright throughout the potometer calibration. Each stem was equipped with two TD probes inserted on opposite sides of each stem and connected to a multiplexor (AM16/32B multiplexer, Campbell Scientific Inc., Logan, UT, USA) that were connected to a battery-powered datalogger (CR1000, Campbell Scientific Inc., Logan, UT, USA). Differential voltage from TD probes was collected every 15 seconds and averaged every minute for TD probes. Water uptake was collected by refilling reservoirs with KCl solutions to a predefined line and tracking how much solution was added. Water uptake was collected every 30 minutes over a two-day period during daylight hours. Throughout day two, red food dye was added to the reservoir to allow for the visualization of conductive sapwood. On the second night, stem tops were cut to remove foliage and zero-flow was determined over that night.

3.2.5 Sapwood area

Sapwood area estimates were calculated based on five stems that first underwent potometer calibration and then subsequently underwent gravimetric calibration in 2017. We used
a dual staining method that allowed us to visualize conducting sapwood area after potometer calibration and again after gravimetric calibration. Specifically, we used red food dye for staining during potometer calibration and blue food dye for staining during gravimetric calibration. To determine if conducting sapwood areas differed, a paired t-test was performed using R 3.4.3 (2019, R foundation for statistical computing, Vienna, Austria).

3.2.6 Calculating calibration curves

$K$ was calculated from equation 2, while $F_d$ was calculated from observed water use measured either gravimetrically or potometrically in each calibration. To determine calibration coefficients, $F_d$ was plotted as a function of $K$, and a power function was fit to describe the relationship. The calibration curves were calculated using R 3.4.3 (2019, R foundation for statistical computing, Vienna, Austria). Outliers were identified through studentized residual $>|2|$ and were removed from the nonlinear regressions. Studentized residuals were calculated with JMP® (Version 14.1. SAS Institute Inc., Cary, NC, 1989-2019).

3.2.7 Evaluating calibration coefficients

Observed $F_s$ was calculated from the water use measured either gravimetrically or potometrically in each calibration. Estimated $F_s$ was calculated following equation 3. Sapwood area was estimated from the conducting area stained during calibration experiments. A linear regression was performed to determine the relationship between estimated and observed $F_s$ (Appendix Table A1). Linear regressions were calculated using R 3.4.3 (2019, R foundation for statistical computing, Vienna, Austria).

3.2.8 Testing calibration coefficients

Calibration coefficients were applied to $K$ values in intact trees (N=5) to calculate estimated $F_s$. Cumulative water use was calculated from $F_s$ values over a 31-day period, May
Cumulative water use on day 31 was used to compare estimates from Granier, gravimetric and potometer coefficients. An analysis of variance (ANOVA) was used to determine if estimated sap flow differed significantly between calibration coefficients. The ANOVA treated calibration method as a fixed factor and individual tree as a random factor. A Tukey HSD test ($\alpha=0.05$) was used to determine significant mean differences. Statistical analyses were calculated using R 3.4.3 (2019, R foundation for statistical computing, Vienna, Austria).

3.2.9 Two-piece calibration curves

Two-piece calibration curves were generated based on the range of $K$ observed in intact trees. The $K$ threshold values were identified where 98%, 95% and 90% of the $K$ values were observed in intact trees. Three, two-piece calibration curves were generated for each calibration approach (i.e., one for each 98%; 95%; and 90% $K$ threshold value) to compare with the calibration curves (i.e., potometer and gravimetric calibrations curves; $K$ 100%). To generate two-piece calibration curves, calibration data was first separated into two pieces, one piece below and one piece above each of the threshold value described above. The two pieces were then fit with two separate power functions: one piece fit to $K$ less than the $K$ threshold value, and a second piece fit to $K$ greater than the threshold value. These two-piece calibration curves were then applied to intact tree data to estimate cumulative water use for May 2018. The ANOVA model was used to compare cumulative water use across calibration approaches and two-piece calibration curves. The ANOVA treated the series of $K$ thresholds for each calibration approach, as a fixed factor and individual tree as a random factor. A Tukey HSD test ($\alpha=0.05$) was used to determine significant mean differences. Statistical analyses were calculated using R 3.4.3 (2019, R foundation for statistical computing, Vienna, Austria).
3.3 Results

3.3.1 Conducting sapwood area

Calibration approach produced different estimates of conducting sapwood area (Appendix Figure A1; t₀.₉₅,₈=2.71; P = 0.0013). Conducting sapwood area of the same stem segment was 11% higher under gravimetric (mean ± s.e.; 17.5 ± 0.7 cm²) than under potometer calibration (15.5 ± 0.7 cm²).

3.3.2 Evaluating calibration curves

Gravimetric and potometer calibration approaches yielded different α and β coefficients (Figure 3.1), which also differed from the original TD coefficients (Table 3.1; Granier, 1985). Gravimetric and potometer coefficients estimated a lower \( F_d \) than Granier’s coefficients for \( K < 0.1 \) and \( K < 0.45 \), respectively. The range of \( K \) values observed in each calibration differed with maximum \( K \) values of 1.262 in gravimetric and 0.824 in potometer calibration. Maximum observed \( F_d \) values in gravimetric calibrations were 213.07 g m⁻² sec⁻¹ while maximum \( F_d \) values for potometer calibration were 481.68 g m⁻² sec⁻¹, representing a 2.3-fold difference in the range for observed \( F_d \). Because the range of \( K \) values was smaller while observed \( F_d \) was much larger in potometer calibration, the calibration curve for potometer calibration was much steeper than the calibration curve in the gravimetric calibration.

Calibration coefficients under one calibration approach were a poor predictor of sap flow under the other approach (Figure 3.2). For example, Granier and gravimetric coefficients underestimated potometer observed water use by 53.8% and 67.5%, respectively (Figure 3.2D, E; Figure 3.3B). Similarly, the potometer coefficients overestimated gravimetric observed water use by 235.9% (Figure 3.2 C; Figure 3.3). Granier coefficients overestimated gravimetric observed water use at lower flow rates (Figure 3.2A).
3.3.3 Testing calibration coefficients

Cumulative sap flow estimated in intact trees differed significantly among Granier, gravimetric and potometer coefficients (Figure 3.4; $F_{2,8}=3.26$, $P=0.004$). Granier coefficients predicted the largest cumulative sap flow in intact trees with potometer having predicted the second largest and gravimetric having predicted the smallest. Granier coefficients predicted mean cumulative sap flow that was 3.02 times larger than gravimetric coefficient prediction and 1.41 times larger than potometer coefficient predictions. Potometer coefficients predicted mean cumulative sap flow 2.15 times greater than gravimetric coefficients.

3.3.3 Two-piece calibration curves

The majority of $K$ values calculated in intact trees were small relative to $K$ values calculated during calibrations (Figure 3.5). Intact trees had 98% of $K$ values falling below 0.5, 95% falling below 0.38, and 90% falling below 0.315. These thresholds $K$ values observed in the field were used to calculated two-step calibration curves (Figure 3.4; Appendix Table A2). The two-piece calibration curves produced curves similar to Granier’s coefficients for $K$ values below threshold values; and curves similar to gravimetric and potometer (i.e., $K$ 100%) for $K$ greater than threshold values (Figure 3.6). Regardless of the calibration approach, two-piece calibration curves increased estimated water use in intact trees (Figure 3.7). Two-piece calibration curves produced estimates of cumulative water use in intact trees more similar to Granier’s coefficients than to gravimetric and potometer coefficients (i.e., $K$ 100%). The two-piece calibration curves increased cumulative sap flow estimates by 42.2, 44.2, and 34.3% relative to potometer coefficients (i.e., $K$ 100%) for 90, 95, and 98% $K$, respectively. Likewise, the two-piece calibration curves increased cumulative sap flow estimates by 191.8, 195.4, and 153.8 % relative to gravimetric coefficients (i.e., $K$ 100%) at 90, 95, and 98% $K$, respectively.
3.4 Discussion

Our study demonstrated that TD calibration approaches using different forces to move water through stems generated different coefficients that yielded significantly different cumulative sap flow estimates when applied to intact trees. Potometer coefficients yielded sap flow estimates that were approximately twice as high as those from gravimetric coefficients, and the Granier coefficients yielded sap flow estimates 1.5-times higher than those from potometer coefficients. Differences in sap flow estimates between the potometer and gravimetric calibration approaches (>200%) were not proportional to differences in sapwood area (11%), suggesting that additional factors influence the relationship between thermal dissipation and water movement in tree stems. Moreover, the two-piece calibration curves for both gravimetric and potometer calibration minimized differences between cumulative water use estimates in intact to trees to within 27%.

3.4.1 Negative versus positive driving force

Our side-by-side comparison of calibration employing positive and negative pressure demonstrated that the method used to move water through stems impacts resulting coefficients, the estimates of water use, and ultimately our scientific inference. Our results are similar to Macinnis-Ng et al. (2016) who confirmed Granier’s (1985) coefficients during gravimetric calibration, as we found that gravimetric and Granier coefficients yielded similar estimates of water use during calibration. In addition, the gravimetric calibration is the same approach that Granier (1985; 1987) used to calculate the original coefficients, indicating gravimetric pressure may result in coefficients similar to Granier’s. Fuchs et al. (2017) indicated that studies that validate Granier’s coefficients have also used positive pressure whereas studies that have concluded that sap flow is largely underestimated using Granier’s coefficients have used
subatmospheric, negative pressure. This presented dichotomy may be an oversimplification of how driving forces have effected calibration coefficients. Steppe et al. (2010), while using positive pressure, found Granier’s coefficients to underestimate sap flow rates in *Fagus grandifolia* by 60%. In addition, Bush et al. (2010) determined that Granier’s coefficients fit data well for *Populus fremontii* and *Tilia cordata* while using subatmospheric pressures. Sun et al. (2012) also observed both over- and under-estimates of sap flow ranging from -34% in *Populus deltoides* to 55% in *Pinus taeda*, while using negative pressure. Our results, in addition to these previous studies, indicate that accurately estimating sap flow may be more complicated than applying positive or negative forces. Specifically, there are likely other physiological or environmental forces affecting calibration and as a result derived coefficient.

3.4.2 Two-piece calibration curves

Our two-piece calibration curves diminished the differences in cumulative sap flow estimated among the different calibration approaches. For example, the difference between estimated water use in intact trees with gravimetric and potometer coefficients decreased from 200% to less than 30%. The two-piece calibration curves estimated sap flow of intact tree to within 2-27% of the estimated calculated from Granier’s (1985) coefficient estimates. Furthermore, we estimated water use more similar to Granier’s coefficients with potometer two-piece calibration curves than with gravimetric two-piece calibration curves. This result is likely driven by curves generated below $K$ value thresholds, as potometer two-piece calibration curves below $K$ value thresholds are similar to Granier’s original calibration curve. This may indicate that studies that have confirmed Granier’s original coefficients may be performing calibrations that produce a smaller range of $K$ and $F_d$ values. However, we cannot confirm this, as studies that document calibration rarely publish raw calibration data (e.g., Caterina et al., 2014).
The reason the two-piece calibration curves diminish differences in sap flow estimate is that the power function fit to the calibration data (K 100%) was sensitive to the larger values in the calibration, which caused an underestimate of $F_d$ at lower $K$ values for both potometer and gravimetric calibrations. Fuchs et al. (2017) drew attention to the fact that a power function is not flexible enough to describe data observed in *Acer pseudoplatanus* at low flows. Calibration with *Ochroma lagopus* and *Hyeronima alchorneoides* trees demonstrated that Granier’s coefficients underestimated sap flow by 52% and 28%, respectively, for each species when sap flow rates low while underestimation decreased to 14% for *H alchorneoides* and 2% for *O lagopus* at high flow (Gutierrez and Santiago 2006). Calibrations performed by Sun et al. (2012) determined that calibrated coefficients underestimated sap flow at $K$ values below 0.5 for *Ulmus americana*. These previous studies with the data presented here, confirm that the power function is unable to predict $F_d$ at the full range of $K$ during calibration. Furthermore, there was a large difference in the $K$ values we observe in our intact trees and in our calibration experiment. For instance, even with an intact canopy in the potometer the majority of $K$ values were $>0.5$ while the majority of $K$ values observed in intact trees were $<0.5$. The reason for this difference may be substantially decreased resistance from rooting systems during calibration, as Flo et al. (2019) points out there is a non-linear relationship between observed sap flow and estimated sap flow derived from calibration data using trees with intact rooting systems. Our calibration experiment is the first to explicitly relate the range of $K$ observed in independent intact tree with the range of $K$ observed during calibration. While several studies deploy sap flow in intact trees and perform potometer or gravimetric calibrations, we are unaware of studies that make this same comparison (e.g. Fassio et al., 2009; Himeno et al., 2017). We recommend that future studies aim to not only
capture a similar range of $K$ during calibration as that observed in intact trees but also calculate calibration curves that provide more meaningful fit to calibration data.

3.4.3 Predicting sap flow in intact trees

Although our results demonstrated different inferences from different approaches, we are still unable to determine how these estimates relate to intact trees. There have been some calibrations performed with lysimeters (e.g., Tfwala et al., 2018) that are able to measure water loss from transpiration of whole trees, but these studies are rare. The reason for this is likely because lysimeters are extremely expensive and are limited to smaller stems that can grow within a reservoir. Lysimeters provide a system to measure transpiration on a fine enough scale to relate it to sap flow measurements (Espadafor et al., 2015). Only with the comparison between sap flow and transpiration will we be able to apply calibration approaches to intact tree stems with a reasonable degree of confidence. Ultimately, a lysimeter approach would be prohibitive for nearly all sap flow studies, so it is critical to assess our available calibration approaches in relation to the calibration of an intact tree standard calibration system.

Despite not knowing which calibration approach produces better sap flow estimates in intact trees, there are situations when one calibration approach may be preferred over another. Gravimetric calibration allows for a larger stem to be used compared to the potometer calibration that is limited by the size of a tree that could reasonably be carried and set up right, although some of the size limitations in the potometer can be ameliorated with the use of heavy equipment. Gravimetric calibration requires expensive equipment (e.g., a precision scale, and a laboratory or indoor space) which provides the advantage of precise measurements in a controlled environment (Flo et al., 2019; Fuchs et al., 2017). For the potometer, beyond the equipment required to use TD, the equipment costs are much lower, as all that is needed is
equipment to fell trees, a reservoir for water, and a way to measure water added to the reservoir. The potometer required tracking water use for one or more days, but multiple trees can usually be measured simultaneously, making this approach reasonable to execute over a short period of time. Gravimetric calibration, with limited equipment, allows for one stem run at a time and each stem can take up to a day depending on how many head heights are used in the calibration. These are the known limitations, but there are multiple factors we know affect sap flow estimates in intact trees and lack an understanding of how they impact calibration.

While there have been several studies that indicate both the environmental effects and the physiological responses of a tree during experimental periods have substantial effects on sap flow estimates, these same environmental and physiological effects have not been examined during calibration. For example, installation of sap flow sensors often results a wounding response and correction for various sap flow methods have been introduced (Burgess et al., 2001; Green et al., 2003; Swanson and Whitfield, 1981; Vandegehuchte and Steppe, 2012). Maranon-Jimenez et al. (2018) determined that wounding from sensors installation causes decreased detection in sap flow for sensors installed in *Fagus sylvatica* and *Quercus petraea* throughout the growing season. However, this study only covers a single growing season when sap flow sensors are sometimes installed for multiple years (e.g., Ward et al., 2013; Wieser et al., 2018). Other factors include accurate zero flow rates (Oishi et al., 2008), accurate measurements of conducting area (Paudel et al., 2013) and natural thermal gradient effects (Vandegehuchte and Steppe, 2013). Furthermore, there are aspects of calibration itself that may change outcomes that have not been considered; for instance, water temperature throughout calibration, season of calibration, differences between early and late season wood, and the voltage applied to sensors.
These are all variables that need further investigation to understand how sap flow calibration relates to sap flow in intact trees.

3.5 Conclusion

Our results indicate that calibration approaches that rely on positive and negative forces to move water through tree stems yield inconsistent estimates of sap flow and that these differences were not explained by the area of conducting sapwood. We assumed that the potometer approach provided a more physiologically relevant framework for calibration because it uses the atmospheric driving force of transpiration (i.e., VPD) and maintains water movement consistent with our theoretical understanding of tension-cohesion (Tyree and Zimmermann, 2002). Regardless of this assumption and its basis in physiological theory, we could not independently quantify actual transpiration in intact stems and can only conclude that the three sets of calibration coefficients resulted in sap flow estimates that varied by a factor of three. However, by considering the range of sap flow measured in intact tree stems, we generated additional two-piece calibration curve approach to estimate sap flow. This two-piece calibration curve provided a new approach to relate ranges of $K$ in intact tree to ranges of $K$ in calibration. The two-step approach also minimized the difference between calibration approaches when coefficients were applied to $K$ values in intact trees. Future studies need to compare transpiration in intact tree with sap flow measurement to determine how calibration relates to measured transpiration and which calibration approach provides the best sap flow estimates.

Reference


Hultine, K.R. et al., 2010. Sap flux-scaled transpiration by tamarisk (Tamarix spp.) before, during and after episodic defoliation by the saltcedar leaf beetle (Diorhabda carinulata). Agricultural and Forest Meteorology, 150(11): 1467-1475.


Table 3.1 Summary of calibration coefficients and mean cumulative sap flow estimates (s.e.) for calibration trees.

<table>
<thead>
<tr>
<th></th>
<th>Gravimetric</th>
<th>Potometer</th>
<th>Granier</th>
</tr>
</thead>
<tbody>
<tr>
<td>$\alpha$</td>
<td>122.586</td>
<td>598.345</td>
<td>118.99</td>
</tr>
<tr>
<td>$\beta$</td>
<td>2.016</td>
<td>2.952</td>
<td>1.231</td>
</tr>
<tr>
<td>Gravimetric calibration cumulative $F_s$ (l)</td>
<td>1.38 (0.5)</td>
<td>5.08 (2.0)</td>
<td>1.88 (0.5)</td>
</tr>
<tr>
<td>Potometer calibration cumulative $F_s$ (l)</td>
<td>7.20 (1.3)</td>
<td>23.38 (5.1)</td>
<td>10.21 (1.6)</td>
</tr>
</tbody>
</table>
Figure 3.1 A comparison of calibrated coefficients for the gravimetric calibration (A) and potometer calibration (B). Dashed lines are fitted to non-linear regression, $y = ax^b$, to calculate coefficients where $F_d$, sap flux density, is the response variable and $K$, the sap flow index, is the explanatory variable. Solid lines are calibration curves from Granier (1985), $y = 119.6x^{1.231}$. Gravimetric data was averages during the last 15 minutes of each head height, when $F_d$ and $K$ were most stable. Not all data is shown for potometer calibration in order to easily compare calibrations (Appendix Figure A2).
Figure 3.2 Comparing observed water use from gravimetric calibration (A-C) and potometer calibration (D-F) with estimated water use derived from Granier’s coefficients (A, D), gravimetric coefficients (B, E) and potometer coefficients (C, F). The solid, gray line represents a 1 to 1 comparison between observed water use and estimated water use. Dotted lines represent a linear regression fit to $y = mx + b$ (Appendix Table A1).
Figure 3.3 Cumulative $F_s$ (±s.e) with (A) gravimetric and (B) potometer calibration trees.

Observed cumulative $F_s$ was measured by a scale during gravimetric calibration or by the amount of water added to reservoirs during potometer calibration. Granier, gravimetric and potometer cumulative $F_s$ calculated with Granier, gravimetric and potometer coefficients, respectively.
Figure 3.4 (A) Cumulative $F_s$ calculated for intact trees (N=5) over the month of May 2018 from Granier (1985), gravimetric and potometer coefficients. (B) Cumulative water use for the entire month with different letters above bars indicating significant differences using Tukey HSD test ($\alpha=0.05$).
**Figure 3.5** Frequency of $K$ in intact *Eucalyptus benthamii* trees (N=5) for May 2018. The $K$ threshold values are identified by vertical line with 98% $K < 0.5$ (grey, dotted), 95% $K < 0.38$ (green, dashed) and 90% $K < 0.32$ (purple, solid).
Figure 3.6 A comparison of calibration curves generated from Granier coefficients, potometer coefficients (K 100%; A-B) and gravimetric coefficients (K 100%; C-D) with two-piece calibration curves (K 98%, K 95%, and K 90%). Two-piece calibration curves were generated for calibration data above and below $K$ threshold values for each set of calibration data.
Figure 3.7 Estimated cumulative $F_s$ for intact trees using Granier coefficients, potometer coefficients (K 100%) and gravimetric coefficients (K 100%) with two-piece calibration curves (K 98%, K 95%, and K 90%). Two-piece calibration curves were generated from gravimetric and potometer calibration data (Figure 3.6). Different letters above bars indicate significant differences using Tukey HSD test ($\alpha=0.05$).
CHAPTER 4
CONCLUSION

This research performed a literature review of tree sap flow publications that indicated calibration is rarely applied to improve the accuracy of sap flow estimates. We have also presented the first side by side comparison of tree sap flow calibration approaches that use different driving forces to move water through stems. Our calibration experiments allowed us to compare how calibration affects estimates of water use in intact trees.

The results of our literature review indicate that 93.2% of publications using sap flow in trees are not calibrating. Not only is there a distinct lack of calibration throughout the literature in the last 11 years, but there is also a lack of acknowledgment that calibration has been demonstrated to be crucial to improve the accuracy of sap flow estimates. With sap flow as the most common method for measuring transpiration in trees, this means we as a community may failing to progress our understanding of transpiration, a critical ecosystem process. Although it is important to recognize that some publications confirm the application of sap flow methods without calibration, there are ample publications that determine sap flow estimates calculated from sap flow methods without calibration deviate substantially from measure water use. To see any change in the prevalence of calibration in sap flow methods, it will take the mutual agreement between researcher, journal editors and reviews that performing calibrations should be a best practice applied when using sap flow methods in trees.
With the comparison of calibration approaches, we determined not only that different calibration approaches result in different calibration coefficients, but also when coefficients are applied to intact trees, we get significantly different sap flow estimates. With two-piece calibration, we have decreased difference of estimated sap flow between calibration approach. Because we calibrated and compared calibration data to intact trees, we were able to use calibration data in a more meaningful way that allowed for more accurate estimates of tree water use. However, we were unable to quantify water use in intact trees and therefore unable to identify a calibration approach that can estimates water use most closely to water use in intact trees. Future studies need to compare transpiration in intact tree with sap flow measurement to determine how calibration relates to measured transpiration and which calibration approach provide the best sap flow estimates.

With the widespread application of sap flow, we need to be sure that it is being applied correctly. Calibration is the clear path forward, however reluctant the sap flow community is to embrace it. Without calibrations, there will be a lack of progress within the sap flow community. If the purpose of using sap flow is to estimate water use and determine how it is related to species, stand and global water cycling, we need to ensure sap flow estimates are accurate and the way to do that is to compare sap flow from sensors with observed water moving through stems.
REFERENCES


Hultine, K.R. et al., 2010. Sap flux-scaled transpiration by tamarisk (Tamarix spp.) before, during and after episodic defoliation by the saltcedar leaf beetle (Diorhabda carinulata). Agricultural and Forest Meteorology, 150(11): 1467-1475.


Table A1 Linear models and respective R² value for estimated $F_s (x)$ derived from Granier, gravimetric and potometer coefficients and observed $F_s (y)$ generated from potometer or gravimetric calibration.

<table>
<thead>
<tr>
<th>Calibration</th>
<th>Parameters</th>
<th>$y = mx + b$</th>
<th>R²</th>
</tr>
</thead>
<tbody>
<tr>
<td>Gravimetric</td>
<td>Granier</td>
<td>$y = 0.880x + 0.083$</td>
<td>0.83</td>
</tr>
<tr>
<td>Gravimetric</td>
<td>Gravimetric</td>
<td>$y = 0.998x - 0.020$</td>
<td>0.91</td>
</tr>
<tr>
<td>Gravimetric</td>
<td>Potometer</td>
<td>$y = 0.975x + 0.140$</td>
<td>0.79</td>
</tr>
<tr>
<td>Potometer</td>
<td>Granier</td>
<td>$y = 0.210x + 0.166$</td>
<td>0.56</td>
</tr>
<tr>
<td>Potometer</td>
<td>Gravimetric</td>
<td>$y = 0.207x + 0.078$</td>
<td>0.68</td>
</tr>
<tr>
<td>Potometer</td>
<td>Potometer</td>
<td>$y = 0.848x + 0.138$</td>
<td>0.72</td>
</tr>
</tbody>
</table>
Table A2 Power function equations for each two-piece calibration curve fit to 90%, 95% and 98% of $K$ in intact trees. Percent $K$ is based on the percent of $K$ values observed in intact trees. $K$ is the threshold value for each respective percent $K$ value.

<table>
<thead>
<tr>
<th>Calibration</th>
<th>Percent $K$ (%)</th>
<th>$K$</th>
<th>Power function $&lt; K$</th>
<th>$R^2$</th>
<th>Power function $&gt; K$</th>
<th>$R^2$</th>
</tr>
</thead>
<tbody>
<tr>
<td>Gravimetric</td>
<td>90</td>
<td>0.32</td>
<td>$F_d = 30.81K^{0.492}$</td>
<td>0.97</td>
<td>$F_d = 122.99K^{2.128}$</td>
<td>0.91</td>
</tr>
<tr>
<td>Gravimetric</td>
<td>95</td>
<td>0.38</td>
<td>$F_d = 31.73K^{0.509}$</td>
<td>0.96</td>
<td>$F_d = 122.99K^{2.144}$</td>
<td>0.91</td>
</tr>
<tr>
<td>Gravimetric</td>
<td>98</td>
<td>0.50</td>
<td>$F_d = 55.88K^{0.884}$</td>
<td>0.89</td>
<td>$F_d = 122.52K^{2.128}$</td>
<td>0.93</td>
</tr>
<tr>
<td>Potometer</td>
<td>90</td>
<td>0.32</td>
<td>$F_d = 143.91K^{1.288}$</td>
<td>0.89</td>
<td>$F_d = 602.45K^{2.971}$</td>
<td>0.59</td>
</tr>
<tr>
<td>Potometer</td>
<td>95</td>
<td>0.38</td>
<td>$F_d = 97.32K^{1.069}$</td>
<td>0.76</td>
<td>$F_d = 603.52K^{2.975}$</td>
<td>0.57</td>
</tr>
<tr>
<td>Potometer</td>
<td>98</td>
<td>0.50</td>
<td>$F_d = 524.35K^{2.605}$</td>
<td>0.72</td>
<td>$F_d = 618.18K^{3.046}$</td>
<td>0.53</td>
</tr>
</tbody>
</table>
**Figure A1.** Conducting sapwood area (±1 s.e.; n=9) for gravimetric and potometer calibrations.

A paired t-test was used to determine there was a significant difference between conducting area observed under gravimetric and potometer calibrations ($t_{0.95, 8} = 2.71; P=0.013$).
Figure A2. Potometer calibration curve with all potometer calibration data presented. Dashed lines are fitted to non-linear regression, $y = ax^\beta$, to calculate parameters where $F_d$, sap flux density, is the response variable and $K$, the sap flow index, is the explanatory variable. Solid line is the calibration curves from Granier (1985), $y = 119.6x^{1.231}$. 

\[ y = 598.345x^{0.952} \]
\[ R^2 = 0.654 \]