CanoPyHydro

Leveraging LiDAR to map water availability in tree canopies.

Vegetation coverage has a marked effect on the spatiotemporal distribution of terrestrial rainfall, marking the initial step in terrestrial rainfall-to-runoff pathways.

Growing interest from hydrologists and environmental scientists has led to numerous attempts to characterize these flows. However, these efforts have largely been correlative and regression-based, lacking clear frameworks to guide meaningful inferences (Van Stan et al., 2020).

CanoPyHydro empowers researchers to derive mechanistic inferences into the drivers underlying variability in canopy rainfall drainage fluxes and has garnered interest for its versatility across related use cases. By integrating precipitation partitioning data with increasingly available terrestrial lidar scans (TLS), CanoPyHydro offers a tailored environment to explore canopy water distribution, enhancing the precision and depth of hydrological analyses.

Summary

The main inputs to CanoPyHydro are Quantitative Structural Models (QSMs), which distill TLS point clouds of trees into topologically ordered cylinders representing branch structures. CanoPyHydro's functionality is divided into two groups: utilities for QSM ingestion and exploration, and utilities for predicting precipitation partitioning.

CanoPyHydro introduces a novel approach by treating tree canopies as watersheds to reveal the tributary-like flows within branch networks. It precisely distinguishes which flows reach the trunk (thus becoming stemflow) and delineates areas where water drips to the forest floor (as throughfall).

For tree-data explorers, CanoPyHydro's spatial utilities stand out. The use of Alpha Shapes enables novel interpretations of canopy coverage area, while detailed *intra*-canopy cross-sections offer precise insights into the vertical shading within the canopy and protection from wind/rain (in the horizontal direction).

Statement of Need

Net rainfall (throughfall + stemflow) reaching the surface beneath plant canopies influences all subsequent terrestrial hydrological processes, contributing to runoff (Savenije, 2004), recharging subsurface water (Friesen, 2020), or returning to the atmosphere via evaporation (Coenders-Gerrits et al., 2020). However, there is substantial spatiotemporal variability in net rainfall amount, timing, and

Figure 1: canoPyHydro logo

distribution, complicating reliable assessments of terrestrial water balances (Van Stan et al., 2020). The costly, labor-intensive techniques required to observe throughfall and stemflow (e.g., Voss et al., 2016; Zimmermann and Zimmermann, 2014) challenge current approaches to modeling and managing terrestrial water interactions (Gutmann, 2020).

Attempts to correlate whole canopy characteristics with stemflow measurements have produced inconclusive results (Sadeghi et al., 2020), despite advances in tree scanning and structural modeling tools (see references in Wischmeyer et al., 2024). A definitive method for accurately delineating these flux origins, crucial for understanding rainfall distribution, remains elusive.

CanoPyHydro addresses this gap with an innovative, bottom-up approach to estimating precipitation redistribution, supplementing QSMs generated using existing tools (Hackenberg et al., 2021) with complementary, graph-based models. CanoPyHydro's algorithm traverses these graph models to precisely delineate stemflow and throughfall drip-point drainage areas. Leveraging detailed canopy structural data from terrestrial LiDAR scans (TLS) to map out precise water pathways, CanoPyHydro transforms how rainfall drainage pathways are predicted and analyzed, offering configuration options to compare rainfall distribution under varying environmental conditions.

CanoPyHydro supports the application of model outputs with a robust suite of analytical tools suitable for various use cases. Its user-friendly filtering capabilities allow users to isolate branch subnetworks based on specified criteria (branches with a radius $>$ 10cm, branches with a branch order of 0 within 100cm of the ground, etc.). These filters may be used in tandem with integrated visualization functions to further enable the exploration and analysis of tree structures and hydrological processes.

By bridging the gap between advanced canopy scanning technologies and the need for precise hydrological insights, CanoPyHydro empowers researchers and environmental managers, enhancing their understanding of water flows in forested ecosystems and supporting more informed conservation and sustainability practices.

Functionality

QSMs

Quantitative Structural Models are 3D representations of tree branching structures using cylinders of varying radii, orientations, and spatial locations. These models effectively reduce point cloud data while preserving high-level structural information. The QSMs used in the creation of CanoPyHydro were generated by processing TLS point cloud data through the SimpleForest program (Hackenberg et al., 2021), resulting in `.csv` files.

(Left to right) A point cloud rendering of a tree, followed by a SimpleForest-

Point Cloud and QSM

Figure 2: Point Cloud and QSM

generated Quantitative Structural Model (QSM) that approximates the tree's branch structure using cylinders, and finally, a CanoPyHydro visualization of the same QSM, highlighting the model's features with color to distinguish different hydrological contributions such as stemflow and throughfall areas.

The below code demonstrates two different ways that CanoPyHydro can read in QSMs:

````
# A CylinderCollection object can be initialized directly
myCollection = CylinderCollection()

# Using the details of a QSM model stored in example_tree.csv
# to create a CylinderCollection object
myCollection.from_csv('example_tree.csv')

# Alternatively, the 'Forester' class can be used
myForester = Forester("data/test/")
print(f"Files available: {list(map(str, myForester.file_names))}")

## ... Read in single QSMs
myForester.qsm_to_collection("example_tree.csv")
print(len(myForester.cylinder_collections))


## Flow Identification

CanoPyHydro's hydrological estimates classify QSM cylinders as contributing either to stemfl

To identify such areas, a user-defined 'drip cut-off angle' is applied, which assumes water

<div align="center">
  <div class="container">
    ![]("../imgs/canhydro_prezi_algo_fig.png")500" width="600" alt="Flow ID Algorithm"/>
    <figcaption>The above diagram shows a minimal example of a QSM to demonstrate the core c
  </div>
</div>

The algorithm assigns an ID to each of the identified flows, with 'stemflow' always receivin

The below code demonstrates how the above is done in practice. Details regarding the various

```{python}
````
````

```python
# Initializing a CylinderCollection object
myCollection = CylinderCollection()
myCollection.from_csv('example_tree.csv')

# Setting a cut-off angle (in radians)
cut_off_angle = -0.166

# Initializing the graph based model
myCollection.initialize_digraph_from(in_flow_grade_lim=cut_off_angle)

# Running the above described algorithm
myCollection.find_flow_components()

# Printing the results of the algorithm

## Keys are equal to the cylinder ids of the cylinders in our collection
cyls = myCollection.cylinders
print(cyl_to_drip)
```

Flow Quantification

After the flows in a canopy's watershed have been identified, common statistics regarding these flows can be calculated though the use of the 'calculate_flows' function. In this process flows are characterized based off of the aggregate characteristics of the cylinders that contribute intercepted water to them. In this way, flows are discussed as having:

- A number of cylinders
- A projected area, volume and surface area
  - each being the sum of the same for their contained cyliners
- A surface area to volume ratio
- A sum of the angles of their cylinders - This is available to facilitate the calculation of average flow angle for one or many flows Most importantly, each non-stem flow also has a unique drip point and drip point location, representing a point in the canopy at which one would expect water to drip to the ground. Utilizing the above metrics, users can glean important information regarding a tree's rainfall-drainage watersheds. For example, the below graphic uses the projected area data for a tree's flows, along with canoPyHydro's visualization capabilities, to make the location and relative abundance of moisture beneath two tree canopies.

```
![]("../imgs/canhydro_drip_map_tale_of_2_trees.png")400" width="800" alt="Tale of Two Trees
<figcaption>Two trees with differing hydrologic characteristics, with drip points indicated
```

Visualization

```
{python}     myCollection = CylinderCollection()     myCollection.from_csv('example_tree.csv
myCollection.project_cylinders('XY')     myCollection.initialize_digraph_from()
myCollection.find_flow_components()     myCollection.calculate_flows()
myCollection.draw('XY', highlight_lambda=lambda:is_stem, save
= True, file_name_ext="docs_ex")     myCollection.draw('XZ',
highlight_lambda=lambda:is_stem, save = True, file_name_ext="docs_ex")
```

```
![]("../imgs/example_tree_XY_docs_ex.png")400" width="400" alt="Stem Flow Highlight XY"/>
```

```
![](../imgs/example_tree_XZ_docs_ex.png")400" width="400" alt="Stem Flow Highlight XZ"/>
```

Here we see an example of the visualization capabilities of canoPyHydro. The above images show the same tree from two different angles, with the stemflow contributing cylinders highlighted in blue

Metrics

Though a variety of metrics are available through this package, the majority are straight forward, summations of cylinder characteristics. Details regarding these metrics and more are available in the metrics definitions in this repository's documentation directory. However, custom functions are available for calculating a few more complicated metrics, which will be highlighted in this section

Shading Fraction

The internal shading of the canopy, as well as the ground beneath it, impacts the energy balance, surface temperature, and wind exposure. In turn, these environmental conditions influence moisture availability via related processes like evaporation. To support data exploration, CanoPyHydro includes robust tools for calculating these shade patterns.

In the calculation of canopy coverage area, we utilize Alpha Shapes rather than a circular region. In some popular Python packages, Alpha Shapes are referred to as 'hulls,' with CanoPyHydro using the tightly fitted 'concave hull' variant. This approach provides a more precise, often lower, estimate of canopy coverage compared to a circular approximation.

Future Direction

- We hope to widen the use cases for our tool by integrating additional real world data (i.e wind speed and direction, rain intensity and average angle, etc.).
- By integrating python libraries for spacial analysis (scipy-spacial, open3d) into canoPyHydro, we hope to allow for the projection of cylinders at an arbitrary angle. This will lead directly into supporting the afformentioned integration of weather data.

- Improve the efficiency of the flow finding algorithm and the flow caluclation algorithm. This will allow for the processing of larger QSMs and the use of more complex models (i.e. tessellated meshes).
  - Under the branch improve-find-flows-efficiency, you can see the current work being done to meet this goal. Early results so as much as a 200x increase in the speed of the algorithm as a result of:
    * migrating the the use of rust based graph models, using the rustworkx library
    * refactoring the current find flow algorithm as a graph traversal algorithm to enable parallel processing ## Future Direction As we continue to develop CanoPyHydro, several key advancements are planned to broaden its functionality and enhance its utility across a range of research applications. These improvements will focus on integrating real-world environmental data, advancing spatial analysis, optimizing computational efficiency, and expanding the scope of canopy hydrology modeling. Below are the primary directions for future development:
- Integrating Real-World Environmental Data We aim to broaden CanoPyHydro's application by incorporating additional real-world environmental data, such as wind speed and direction, rain intensity, and rainfall angle. These integrations will enhance the precision of canopy water distribution modeling by factoring in dynamic environmental conditions, allowing researchers to simulate and predict hydrological processes under various weather scenarios. This will be particularly useful for studying the effects of storm events, seasonal changes, and long-term climate impacts on canopy water redistribution.
- Advancing Spatial Analysis with Python Libraries By incorporating advanced Python libraries for spatial analysis, such as scipy-spatial and open3d, we aim to enable the projection of cylinders at arbitrary angles. This advancement will allow CanoPyHydro to support the aforementioned integration of weather data, improving the accuracy of stemflow and throughfall predictions. Additionally, this enhancement will facilitate more complex spatial queries, such as identifying microclimates within a canopy or evaluating wind-driven rain impacts on specific tree branches.
- Optimizing Computational Efficiency A critical goal is to improve the efficiency of the flow-finding and flow-calculation algorithms, enabling CanoPyHydro to process larger QSMs and handle more complex tree models, such as tessellated meshes. These improvements will expand the tool's scalability, making it more suitable for large-scale studies involving dense forest ecosystems or intricate tree structures. Under the branch improve-find-flows-efficiency, you can find current work toward this goal, which has shown promising results, with up to a 200x increase in algorithm speed. This improvement is attributed to migrating to Rust-based graph models using the rustworkx library and refactoring the flow-finding algorithm as a

graph traversal process, enabling parallel processing. These developments will significantly reduce computational time and resource requirements, allowing users to run complex models on standard computing setups.

- Expanding Canopy Coverage Analysis In the future, we aim to integrate additional methods for calculating canopy coverage beyond Alpha Shapes. By offering options for alternative geometric representations, such as 3D meshes or voxel-based models, CanoPyHydro will enable more flexible and accurate analyses of canopy structure and its influence on rainfall distribution. This expansion will help researchers tailor their analyses to specific forest types, vegetation structures, or ecological research questions.
- Supporting Multi-Scale Hydrological Modeling We plan to extend CanoPyHydro's capabilities to support multi-scale hydrological modeling, allowing for the simultaneous analysis of both individual trees and larger forest stands. This will facilitate more comprehensive ecosystem studies, enabling users to model how water flows through entire landscapes, from canopy to forest floor, while considering the collective behavior of multiple trees. This functionality will be particularly beneficial for watershed management, forest hydrology, and conservation planning at broader spatial scales.
- Customizable User Interfaces for Specialized Applications To make CanoPyHydro accessible to a wider range of users, we envision developing customizable user interfaces tailored to specific research applications. These could include simplified interfaces for educators and students, as well as advanced options for scientists requiring detailed control over modeling parameters and data inputs. Customization will also include the integration of automated workflows, enabling users to conduct common analyses with minimal manual intervention.

Through these future developments, CanoPyHydro will continue to evolve as a powerful and versatile tool for exploring tree hydrology, driving new insights into how tree canopies interact with environmental conditions and contribute to water redistribution in forest ecosystems.

Publications and Acknowledgements:

CanoPyHydro was developed in the process of authoring A LiDAR-driven pruning algorithm to delineate canopy drainage areas of stemflow and throughfall drip points., which has been accepted for publication by the *‘British Ecological Society’s’* ‘*Methods in Ecology and Evolution*’. Said paper, and the code within this repository, represents a collaboration between non-academic data professional Collin Wischmeyer, ecohydroloy scholar Professor John Van Stan with notable contributions from industry geo-scientist Travis Swanson. Likewise, this tool could not exist without the data collected and the ideas put forward by several graduate students working in Cleveland State University's ‘Wet Plant Lab’.

This paper's titular project and related efforts were made possible, in part, by the financial support of US-NSF DEB-2213623.

References

Coenders-Gerrits, A.M.J., Schilperoort, B., Jiménez-Rodríguez, C., 2020. Evaporative Processes on Vegetation: An Inside Look. Precipitation Partitioning by Vegetation: A Global Synthesis. https://doi.org/10.1007/978-3-030-29702-2_3

Friesen, J., 2020. Flow Pathways of Throughfall and Stemflow Through the Subsurface. Precipitation Partitioning by Vegetation: A Global Synthesis. https://doi.org/10.1007/978-3-030-29702-2_13

Cavelier, J., Jaramillo, M., Solis, D., de León, D., 1997. Water balance and nutrient inputs in bulk precipitation in tropical montane cloud forest in Panama. J Hydrol (Amst) 193, 83–96.

Dunkerley, D., 2020. A review of the effects of throughfall and stemflow on soil properties and soil erosion. Precipitation Partitioning by Vegetation: A Global Synthesis. https://doi.org/10.1007/978-3-030-29702-2_12

Hackenberg, J., Calders, K., Demol, M., Raumonen, P., Piboule, A., Disney, M., 2021. SimpleForest - a comprehensive tool for 3d reconstruction of trees from forest plot point clouds. bioRxiv. bioRxiv.

Sadeghi, S.M.M., Gordon, A.G., Van Stan, J.T., 2020. A Global Synthesis of Throughfall and Stemflow Hydrometeorology. Precipitation Partitioning by Vegetation: A Global Synthesis. https://doi.org/10.1007/978-3-030-29702-2_4

Savenije, H.H.G., 2004. The importance of interception and why we should delete the term evapotranspiration from our vocabulary. Hydrol Process 18, 1507–1511. https://doi.org/10.1002/hyp.5563

Van Stan, J.T., Hildebrandt, A., Friesen, J., Metzger, J.C., Yankine, S.A., 2020. Spatial variablity and temporal stability of local net precipitation patterns. Precipitation Partitioning by Vegetation: A Global Synthesis. https://doi.org/10.1007/978-3-030-29702-2_6

Voss, S., Zimmermann, B., Zimmermann, A., 2016. Detecting spatial structures in throughfall data: The effect of extent, sample size, sampling design, and variogram estimation method. J Hydrol (Amst) 540, 527–537. https://doi.org/10.1016/j.jhydrol.2016.06.042

Wischmeyer, C., Swanson, T., Mueller, K., Lewis, N., Bastock, J., Van Stan, I.J.T., 2024. A LiDAR-driven pruning algorithm to delineate canopy drainage areas of stemflow and throughfall drip points. Methods Ecol Evol In press. https://doi.org/10.2139/ssrn.4600550

Zimmermann, A., Zimmermann, B., 2014. Requirements for throughfall monitoring: the roles of temporal scale and canopy complexity. Agric For Meteorol

189, 125–139.