

About the dataset

This PF database is a preferential flow database that is generated by using NEON soil moisture and precipitation.

Metadata for PF base file

File format: XXXX_PF_database_00Y where xxxx is the NEON site name and y is the soil profile number ranging from 1 to 5.

Event data fields

- **stormStartTime** [UTC] : Precipitation event starts.
- **stormEndTime** [UTC] : Precipitation event ends.
- **stormStartValue** [mm/10min] : The precipitation at the time step, in which precipitation events starts.
- **stormEndValue** [mm/10min] : The precipitation at the time step, in which precipitation events ends.
- **stormSum** [mm] : Sum of precipitation of the precipitation event.
- **stormPeakIntensity** [mm/10min] : The max precipitation for a given precipitation event.
- **stormPeakTime** [UTC] : The time, in which *stormPeakIntensity* is observed.
- **stormDuration** [hours] : The duration of the precipitation events.
- **smReponseType_50X** ☐ : If the sensor responded to a precipitation event. X represent the sensor position from 1 - 8 in most of the case.
- **smOnsetTime_50X** [UTC] : Soil moisture onset time.
- **smPeakTime_50X** [UTC] : Soil moisture peak time.
- **smAtOnset_50X** [v/v] : Soil moisture value at the onset.
- **smAtPeak_50X** [v/v] : Soil moisture value at the peak.
- **smBeforePrecip_50X** [v/v] : Soil moisture value right before precipitation start.
- **flowTypes** ☐ : Type of flows. *notApplicable* = less than 2 sensors are working. *nonSequentialFlow* = at least 2 sensors responded to the precipitation event and the shallower sensor responses later than the deeper sensor. *SequentialFlow* = at least 2 sensors responded to the precipitation event and the shallower sensor responded earlier than the deeper sensor.
- **flowPosition** ☐ : record which two sensors are involved in preferential flow. [501Xvs50Y_PF_50Y] = nonsequential flow while comparing two sensors. The deeper sensor of each combo is recorded.
- **velocity_computed_50X** [cm/day]: The log scale computed flow velocity.

- **num_velocity_50X_median** [cm/day] : The median of matrix flow velocity by a model in log scale.
- **num_velocity_50X_95th** [cm/day] : The 95th percentile of matrix flow velocity by a model in log scale.
- **num_velocity_50X_99th** [cm/day] : The 99th percentile of matrix flow velocity by a model in log scale.
- **PF_velocity_metric_50X** ☐ : Compared the **velocity_computed_50X** to **num_velocity_50X_99th**. PF is identified (assigned as **True**) is **velocity_computed_50X** is greater than **num_velocity_50X_99th**.
- **num_velocity_name_50X** ☐ : The depth name where the **velocity_computed_50X** compares to.
- **sensor_nake_name_50X** [-]:The nake name of the sensor format as **soil plot name_sensor position** i.e., 001_501 represent the top sensor of the first soil plot/profile at a site.

Other data fields

- **total_sand_50X** [%]: The weighted sand percentage based on soil information for megapits and distributed pits.
- **total_silt_50X** [%]: The weighted silt percentage based on soil information for megapits and distributed pits.
- **total_clay_50X** [%]: The weighted clay percentage based on soil information for megapits and distributed pits.
- **total_porosity_50X** [-]:The weighted porosity based on soil information for megapits and distributed pits.
- **total_eff_porosity_50X** [-]:The weighted effective porosity based on soil information for megapits and distributed pits.
- **total_root_density_50X** [g/cm^3]:The weighted root density based on root sampling for megapits and distributed pits.
- **latitude** [-]: site latitude.
- **longitude** [-]: site longitude.
- **aridity** [-]: The aridity index at the site as PET/P.
- **MAT** [$^{\circ}\text{C}$]: Mean annual temperature of the site.
- **MAP** [mm]: Mean annual precipitation of the site.
- **field_utm_northing** [-]: northness of the site.
- **field_utm_easting** [-]: eastness of the site.
- **field_soil_subgroup**[-]: the soil types.

Authors

- Bonan Li - data collection, methology, coding
- Matthias Sprenger - discussion and advisory
- Tianfang Xu - discussion
- John Nimmo - discussion
- Hoori Ajami - discussion
- Ryoko Araki - discussion
- Octavia Crompton - discussion
- Daniel Gimenez - provided soil matrix flow simulation, discussion
- Jannis Groh - discussion
- Daniel Hirmas - methology, provided soil data, and discussion
- Nitin Singh - discussion
- Briana Wyatt - discussion
- Inge Wiekenkamp - discussion
- Pamela Sullivan1 - discussion and advisory

Online Repository link

The database is avaiable at <https://oregonstate.app.box.com/folder/247104527194>