

Connecticut Blue Ribbon Permanent Forest Plot Network - Public Dataset Metadata

Primary data owner: Connecticut Agricultural Experiment Station (CAES) **Secondary data owner / compiler:** Forest Ecosystem Monitoring Cooperative (FEMC) Connecticut Blue Ribbon Permanent Forest Plot Harmonization project

1. Dataset summary

A harmonized, analysis-ready database of long-term forest inventory data from a network of permanent research plots in Connecticut, established and maintained by the Connecticut Agricultural Experiment Station (CAES) and cooperating researchers. The dataset unifies two historically separate compilations (the West and East plot systems) into a single consistent schema, preserving the original observations while standardizing tree identities, crown classifications, mortality histories, and species codes.

Plots	13 permanent plots
Trees	2,256 individual tree/stem records
Observations	~15,000 tree-year measurement records
Temporal coverage	1994–2023 (irregular, plot-specific remeasurement)
Spatial coverage	Connecticut, USA (see plotInfo.csv / plot_locations.sql)
Coordinate system	WGS84 (EPSG:4326), decimal degrees
Tier	Public release (cleaned; internal QA fields removed)

This **public tier** is derived from the corrected (post-QAQC) analysis database. It drops internal working fields (QA flags, internal identifiers, raw archived JSON, the silvicultural species grouping) while retaining every field of scientific value, including field notes and the harmonized status/provenance.

2. Files in this release

File	Description
plotInfo.csv	One row per plot: location, area, dimensions, description, notes
treeInfo.csv	One row per tree/stem: species, coordinates, ingrowth, mortality, harvest
treeData.csv	Long-format repeated measurements (one row per tree-year): diameter, crown class, status, etc.
key_crown_class.csv	Crown-class code dictionary (cc_std)
key_status.csv	Tree-status code dictionary
key_species.csv	Species codes present in the data → common/Latin name, type, group

File	Description
key_data_dictionary.csv	Column-by-column data dictionary for the three data tables
METADATA.md	This document

Relationships: [treeData.treeid](#) → [treeInfo.treeid](#) → [treeInfo.plot](#) → [plotInfo.plot](#). Codes in [treeData](#) ([cc_std](#), [status](#)) and species codes join to the corresponding [key_*.csv](#) lookup tables. A normalized SQLite build of the full database ([blue_ribbon.sqlite](#)) is available separately.

3. Dataset descriptions and fields

The three data tables form a relational hierarchy ([plotInfo](#) → [treeInfo](#) → [treeData](#)), accompanied by four lookup/reference tables. [key_data_dictionary.csv](#) is the machine-readable version of the field lists below.

3.1 [plotInfo.csv](#)

One row per plot. Site-level metadata: location, area, dimensions, and descriptive history for each of the 13 permanent plots. Primary key: [plot](#).

Field	Units	Description
dataset		Source compilation: West or East
plot		Plot identifier (primary key)
latitude	decimal degrees (WGS84)	Plot latitude (blank for SharonB-E)
longitude	decimal degrees (WGS84)	Plot longitude (blank for SharonB-E)
area_acres	acres	Measured plot area (confirmed for West plots; blank for the 5 East plots, pending)
dimensions		Plot dimensions where recorded
description		Narrative site / treatment history
notes		Free-text plot notes

3.2 [treeInfo.csv](#)

One row per tree/stem identity. Attributes that stay with an individual tree (or stem) through time. Primary key: [treeid](#); joins to [plotInfo](#) on [plot](#) and to [key_species.csv](#) on [species_code](#).

Field	Units	Description
dataset		Source compilation: West or East
treeid		Permanent tree/stem identity, Plot_Group_Tree (primary key)
plot		Plot identifier (joins plotInfo)

Field	Units	Description
group		Sub-plot / block group (West only)
tree		Tree tag number within plot
species_code		Species code (joins key_species.csv)
common_name		Species common name
latin_name		Species scientific name
species_type		Record type (Tree / Site)
xcoor	m	Within-plot stem X coordinate
ycoor	m	Within-plot stem Y coordinate
ingrow_yr	year	Year tree crossed the measurement threshold (recorded for a subset of plots)
mort_yr	year	Recorded mortality year, if any
mort_cause		Recorded mortality cause, if any
mort_year_best	year	Best-estimate death year (reconciled)
harvest_flag	boolean	True if the tree was ever recorded as harvested
harvest_year	year	Year of harvest where determinable
mgmt_status		Management status note

3.3 treeData.csv

One row per tree-year (long format). The repeated measurements of each tree across survey years.

Composite key: [treeid](#) + [year](#); joins to [treeInfo](#) on [treeid](#), to [key_crown_class.csv](#) on [cc_std](#), and to [key_status.csv](#) on [status](#). Many columns are sparse (recorded only at the initial survey, or only for the Sharon forest-health surveys).

Field	Units	Description
dataset		Source compilation: West or East
treeid		Tree/stem identity (joins treeInfo)
year	year	Survey year
species_code		Species code (joins key_species.csv)
common_name		Species common name
latin_name		Species scientific name
dbh_mm	mm	Diameter at breast height
dbh_in	inches	Diameter at breast height

Field	Units	Description
cc_raw		Crown class as originally recorded
cc_std		Standardized crown class (joins <code>key_crown_class.csv</code>)
cc_desc		Crown-class description
lcr_pct	%	Live crown ratio
dsh_mm	mm	Diameter at stump height (initial survey)
height_tot_ft	ft	Total height (initial survey)
height_saw_ft	ft	Merchantable sawlog height (initial survey)
height_pulp_ft	ft	Merchantable pulpwood height (initial survey)
height_bol_ft	ft	Bole height (initial survey)
height_blc_ft	ft	Height to base of live crown (initial survey)
grade		Log grade 1-4 (initial survey)
bfc_defect		Butt/form-class defect indicator (initial survey)
cfc_defect		Crown/form-class defect indicator (initial survey)
fol_transp_pct	%	Foliage transparency (Sharon, 2021)
defoliation_pct	%	Defoliation (Sharon, 2021/2022)
norm_live_crown		Normalised live-crown score (Sharon, 2022)
mgmt_status		Management status on a known harvest year
obs_type		What the row contains: DBH / BOTH / NONE
status		Harmonized tree status (joins <code>key_status.csv</code>)
status_detail		Provenance of the status assignment
notes		Free-text field observation + QA/QC annotations (prefixed <code>POST-QAQC:</code>)

3.4 Lookup / reference tables

`key_crown_class.csv` - crown-class code dictionary. Fields: `code`, `description`, `category`, `numeric_equivalent`, `status_implication`, `notes`.

`key_status.csv` - tree-status dictionary. Fields: `status`, `description`, `assignment_rule`.

`key_species.csv` - species present in the data. Fields: `species_code`, `common_name`, `latin_name`, `species_type`, `species_group`.

`key_data_dictionary.csv` - machine-readable field list for the three data tables. Fields: `table`, `column`, `description`, `units`.

3.5 Conventions

- **Original values are retained** alongside standardized ones (e.g. `cc_raw` next to `cc_std`; `dbh_in` next to `dbh_mm`).
 - **Diameter** is provided in both millimetres (`dbh_mm`) and inches (`dbh_in`).
 - **Status** (`Live` / `Dead` / `Unknown` / `Not yet recorded`) is a reconstructed field; `status_detail` documents how each value was derived.
 - **Multi-stem trees** are stored as separate stem records sharing a tree tag with a `.1` / `.2` / `.3` suffix (e.g. `pikes_784.1`, `pikes_784.2` are two stems of one tree) - consider this when counting individuals vs stems.
 - **notes** carries original field observations plus QA/QC annotations (prefixed `POST-QAQC:`).
-

4. Methods (brief)

Source spreadsheets were harmonized through a documented pipeline: tree-identity normalization and joining; diameter unit/outlier standardization; crown-class harmonization across numeric and alphabetic coding systems; transparent mortality reconstruction (a per-tree rule set producing a standardized status for every tree-year); and measurement integration across inventories with differing variable sets. A manual FEMC QA/QC pass then resolved flagged conflicts (post-mortem diameter conflicts, identity relabels, species assignments) and recorded interpretive determinations. Full methods are in the project `README.md` and `cc_codes_reference.md`; every manual change is itemized in `postqaqc_changelog.csv`.

5. Known limitations

- **Plot areas:** the 8 West plot areas are confirmed (acres); the 5 East plot areas are pending field measurement (next field season) and are blank. Per-area (per-acre) metrics are therefore computable for the West plots only.
 - **Coordinates:** SharonB–E have no recorded coordinates (SharonA's coordinate is taken from a corner of Block E, not a centroid). See `plotInfo.csv` notes.
 - **Survey gaps:** some plots/years were not fully resampled (e.g. RobertsPines from 2003 on; SharonE in 2018/2023). Affected rows are `status = Unknown` and carry an explanatory note rather than being inferred as mortality.
 - **Recruitment year (`ingrow_yr`)** is recorded only for a subset of plots.
 - A small number of items remain documented but unresolved; see row-level `notes` and the project issue log (`femc_qaqc_questions.csv`).
-

6. Ownership, citation, and use

- **Data owner (primary):** Connecticut Agricultural Experiment Station (CAES)
- **Data owner (secondary) / compiler:** Forest Ecosystem Monitoring Cooperative (FEMC), University of Vermont
- **Authors:** Soren Donisvitch (FEMC); Joseph Barsky (CAES)
- **Original data:** Connecticut Agricultural Experiment Station (CAES) and cooperating researchers (see project `publications/`).
- **Suggested citation:** Connecticut Agricultural Experiment Station & Forest Ecosystem Monitoring Cooperative (2026). *Connecticut Blue Ribbon Permanent Forest Plot Network - harmonized analysis-ready database*. Compiled by S. Donisvitch (FEMC) and J. Barsky (CAES). New Haven, CT / Burlington, VT.

Disclaimer: This dataset is provided "as is", without warranty of any kind. It is intended for research and forest-monitoring purposes. Inventory data are never perfectly clean; users should review the **notes** fields and known limitations above and confirm results before using them in management or decision-making contexts.

Metadata last updated: 2026.