

## Metadata–Merged dataset (calibration cases and prediction sites)--342 sites

January 30, 2017

The FHX file is the computer file, containing fire-history data, provided by authors to the International Multiproxy Paleofire Database (IMPD), which is located here:

<https://www.ncdc.noaa.gov/data-access/paleoclimatology-data/datasets/fire-history>

The Fire History Analysis and Exploration System (FHAES) is software that reads the FHX file and allows computation of fire-history measures. It can be downloaded here:

<https://www.frames.gov/partner-sites/fhaes/download-fhaes>

### THE FIRST THREE COLUMNS DESCRIBE THE SAMPLE

C1 = **Author**, see the bibliographies in the paper

C2 = **Site**, as given in the FHX file from the IMPD or in the publication if there was no FHX file

C3 = **IMPDdataset**, the name assigned to the dataset in the IMPD

### THE NEXT SEVEN COLUMNS ARE ABOUT THE SAMPLE SIZE

C4 = **SampleArea(ha)**, as given in the FHX file or, if an FHX file not available, then as given in the publication. In many cases, neither source provided this information.

C5 = **ScarredTrees**, as given in the FHX file or, if an FHX file not available, then as given in the publication. In many cases, neither source provided this information.

C6 = **ScarredTrees/100ha** =  $C5/(C4/100)$

C7 = **TotalScars**, the total number of fire scars that were dated across all the wedges/cores and all the years on each wedge/core, based on the publication. This information was often not provided in the publication.

C8 = **MeanTotalScarredTrees**, calculated from the FHX file as the mean for the Total Samples column in the Summary Table available in FHAES after editing it to limit it to the Time Period (years) of the analysis. Missing data typically occur because the FHX file was a composite file, which cannot be used to produce a Summary Table.

C9 = **TotalEvents**, calculated from the FHX file as the sum of the Events column in the Summary Table available in FHAES after editing it to limit it to the Time Period (years) of the analysis. Missing data typically occur because the FHX file was a composite file, which cannot be used to produce a Summary Table.

C10 = **MeanScars/ScarredTree**, calculated as  $C9/C8$  if an FHX file and a Summary Table in FHAES were available, otherwise calculated as  $C7/C5$  from the publication, if the necessary information was reported. Missing data may occur because the FHX file was a composite file, which cannot be used to produce a Summary Table.

### THE NEXT THREE COLUMNS DEFINE THE SAMPLE PERIOD

C11 = **FIAalysisYears**, which extend from the first fire year in or after A.D. 1600 to the last fire year before or in A.D. 1900 in the case of the 252-site prediction dataset and over any period, from a first fire to a last fire, in the case of the 96-site calibration dataset.

C12 = **FIAuthorsYears**, which are the years analyzed by the authors, as reported in the publication. This information was not always provided.

C13 = **FIBegins**, which is the first year of the author(s)' analysis, as recorded in C12, or the first year of the analysis, as recorded in C11 if the author did not present the analysis period. The

purpose of this field is to allow calculation of the approximate age of the forest.

THE NEXT TWO COLUMNS CONTAIN THE ESTIMATOR AND PREDICTION

C14 = **WeibMeanITFI**, in years, which was calculated in FHAES for the FIAAnalysisYears (C11)

C15 = **Predicted FR/PMFI**, which was either (1) the reported value, in the case of the 96 calibration sites, or (2) calculated as  $1.216 * C14$ , based on the Weibull Mean ITFI regression model in Table 2 in the paper, in the case of the 252 prediction sites. C38 indicates whether the record is for a calibration or prediction site.

THE NEXT COLUMN DEFINES THE VEGETATION TYPE

C16 = **VegType**, which is either DryPine, for *Pinus ponderosa* and related pines of dry western forests, or DryMC for dry mixed-conifer forests, which are forests with ponderosa pine or other dry pines dominant but also with other trees (e.g., white fir, Douglas-fir, quaking aspen) present.

THE NEXT 8 COLUMNS ARE ABOUT TARGETTED SAMPLING

C17 = **Targeting**, which lists the page(s) in the publication where targeted sampling, if used, was explained.

C18 = **TargetBestInfoLongestRecord**, which is “Yes” if this form of targeting, which seeks to obtain the longest record of historical fires, was used, “No” if the author explained that this was not used, and is “No explanation” if this form of targeting was not mentioned or specifically explained.

C19 = **TargetMultiScarred**, which is “Yes” if this form of targeting, which favors trees with multiple fire scars, was used, “No” if the author explained that this was not used, and “No explanation” if this form of targeting was not mentioned or specifically explained.

C20 = **TargetClusters**, which is “Yes” if this form of targeting, which favors sampling fire scars in clusters, was used, “No” if the author explained that this was not used, and “No explanation” if this form of targeting was not mentioned or specifically explained.

C21 = **TargetDeadWood**, which may be either “Many” or “Some” if this form of targeting, which favors fire scars on standing dead or down wood, was used, “No” if the author explained that this was not used, and “No explanation” if this form of targeting was not mentioned or specifically explained.

C22 = **TargetSpecies**, whether tree species were preferentially sampled, which may be “Ponderosa pine,” “No” if the author explained this form of targeting was not used, and “No explanation” if this form of targeting was not mentioned or specifically explained.

C23 = **TargetPlots**, which may be “Many scarred trees” or “Old forests” or “Ponderosa pine” if plots were targeted in locations with these attributes, “No explanation” if targeting was not used or was not explained, “No plots” if plot-sampling was not used, “Objective-grid” if a non-targeted method based on an objective grid was used, “Objective-random” if a non-targeted method based on random sampling was used, “Proximity/access” if this criterion was used, which is ambiguous regarding objectivity, and “Range of settings/gradient” if the plots were spread across settings or a gradient (e.g., elevation) but it was unclear whether the actual plot placement was objective.

C24 = **TargetStudyAreas**, which may be many explanations

THE NEXT 3 COLUMNS ARE ABOUT WHETHER FIRE SEVERITY WAS STUDIED

C25 = **FireSeverity**, which is “No” if not studied, “Yes” if studied, and “No explanation” if the publication does not explain whether fire severity was studied.

C26 = **FireSeveritySource**, which is “Age structure-plots” if fire severity was determined at the plot scale, “Age structure-plots/aerial photos” if both plots and aerial photos were used, “Age structure-pools only” if fire severity was determined from a pool of plot data, which is not reliable as it masks locally severe fires, “Inference” if fire severity was inferred, rather than specifically studied, “Inference/dating” if some dating accompanied the inference, “Not studied” if not studied, and “No explanation” if it was not possible to determine whether the author studied fire severity.

C27 = **FireSeverityComments**, which explain some aspects of the fire-severity determinations

THE NEXT 9 COLUMNS ARE ABOUT THE LOCATION OF THE SAMPLE

C28 = **State**, there are these: AZ, CA, CO, ID, MT, MX, NM, OR, SD, WA, WY, plus MX, BC

C29 = **StateCode**, there are these: 1 = AZ, 2 = CA, 3 = CO, 4 = ID, 5 = MT, 6 = MX (Mexico), 7 = NM, 8 = OR, 9 = SD, 10 = WA, 11 = WY, 12 = BC

C30 = **LatSource**, which is the latitude given in the FHX file or publication

C31 = **LongSource**, which is the longitude given in the FHX file or publication

C32 = **Latitude**, which is the latitude converted to decimal degrees

C33 = **Longitude**, which is the longitude converted to decimal degrees

C34 = **UTM\_E**, which is the UTM easting coordinate in meters, if given in the FHX file or publication

C35 = **UTM\_N**, which is the UTM northing coordinate in meters, if given in the FHX file or publication

C36 = **Datum**, which is the datum, if given in the FHX file or publication

THIS LAST COLUMN IS WHETHER THIS CASE IS A CALIBRATION OR PREDICTION

C37 = **CalPred**, Calibration or Prediction