Package ‘SoyURT’

January 20, 2025

Title USDA Northern Region Uniform Soybean Tests Dataset

Version 1.0.0

Author Matheus Dalsente Krause [aut, cre]
(<`https://orcid.org/0000-0003-2411-9287`>),
William Dale Beavis [aut] (<`https://orcid.org/0000-0003-0790-2771`>)

Maintainer Matheus Dalsente Krause <`krause.d.matheus@gmail.com`>

Description Data sets used by 'Krause et al. (2022)' <doi:10.1101/2022.04.11.487885>. It comprises phenotypic records obtained from the USDA Northern Region Uniform Soybean Tests from 1989 to 2019 for maturity groups II and III. In addition, soil and weather variables are provided for the 591 observed environments (combination of locations and years).

URL `https://github.com/mdkrause/soyurt`

NeedsCompilation no

Language en-US

License CC BY 4.0

Encoding UTF-8

Depends R (>= 3.5.0)

LazyData true

LazyDataCompression xz

RoxygenNote 7.2.0

Suggests spelling

Repository CRAN

Date/Publication 2022-06-13 06:30:03 UTC

Contents

---

pheno *Phenotype*

Description

Modeled data set by Krause et al. (2022) from the USDA Northern Region Uniform Soybean Tests. The data contains 4,257 experimental genotypes evaluated at 63 locations and 31 years resulting in 591 location-year combinations (environments) with 39,006 yield values belonging to matirity groups II and III from 1989 to 2019. Annual PDF reports from the Northern Region of the USDA Uniform Soybean Tests were obtained from `https://ars.usda.gov/mwa/lafayette/cppcru/ust`. The data retrieved from the published PDF files represent averages for seed yield for each genotype evaluated at each trial in location-year combinations. Seed yield was adjusted to 13% moisture and results were reported in bushels per acre (bu/ac). For more information about the trial field plot design and agronomic practices, please refer to the PDF files. The raw data can also be downloaded from Soybase: `https://soybase.org/ncsrp/queryportal/`.

Usage

```
pheno
```

Format

A data frame in tidy format with 39,006 observations on the following 13 variables:

`year` years, 31 levels (1989 - 2019)

`location` locations, 63 levels (observed locations in the historical series)

`latitude` latitude

`longitude` longitude

`altitude` altitude

`trial` name of the trial that originated the phenotypic record

`check` indicator variable for variety checks, 2 levels (yes or no)

`maturity_group` genotype's maturity group, 2 levels (II or III)

`G` genotype, 4,257 levels

`eBLUE` empirical best linear unbiased estimate of genotype means

`SE` standard error of genotype means on a location level

`average_planting_date` average planting date on a location level (MM/DD/YY)

`average_maturity_date` average maturity date on a location level in days after planting

Source

Krause, M. D., Dias, K. O. G., Singh, A. K., and Beavis. W. D. (2022). Using large soybean historical data to study genotype by environment variation and identify mega-environments with the integration of genetic and non-genetic factors. bioRxiv, doi: 10.1101/2022.04.11.487885

`soil` *Soil variables*

Description

Soil variables in a depth interval of 5 to 15 cm were obtained from Soilgrids (`https://soilgrids.org/`) for the 63 observed locations in the historical series analyzed by Krause et al. (2022). The R code used to download and process the soil data can be retrieved at `https://github.com/mdkrause/VarComp-ME/blob/main/soil_data.R`.

Usage

```
soil
```

Format

A data frame in tidy format with 504 observations on the following 5 variables:

`Feature` soil variables, 8 levels

`location` locations, 63 levels (observed locations in the historical series)

`Soil_Grid` mean values of the soil variables (Feature)

`LAT` location latitude

`LON` location longitude

Details

Levels of `Feature`:

bdod_5-15cm_mean: Bulk density of the fine earth fraction (cg/m^3)

cec_5-15cm_mean: Cation Exchange Capacity of the soil (mmol(c)/kg)

clay_5-15cm_mean: Proportion of clay particles (< 0.002 mm) in the fine earth fraction (g/kg)

nitrogen_5-15cm_mean: Total nitrogen (cg/kg)

phh2o_5-15cm_mean: Soil pH ($\mathrm{pH}\times 10$)

sand_5-15cm_mean: Proportion of sand particles (> 0.05 mm) in the fine earth fraction (g/kg)

silt_5-15cm_mean: Proportion of silt particles (≥ 0.002 mm and ≤ 0.05 mm) in the fine earth fraction (g/kg)

soc_5-15cm_mean: Soil organic carbon content in the fine earth fraction (dg/kg)

Source

Krause, M. D., Dias, K. O. G., Singh, A. K., and Beavis. W. D. (2022). Using large soybean historical data to study genotype by environment variation and identify mega-environments with the integration of genetic and non-genetic factors. bioRxiv, doi: 10.1101/2022.04.11.487885

`weather` *Weather variables*

Description

Weather variables obtained from NASA's Prediction of Worldwide Energy Resource (https://power.larc.nasa.gov/) for the 591 environments in the historical series analyzed by Krause et al. (2022).

Usage

`weather`

Format

A data frame in messy format with 504 observations on the following 5 variables:

`location` locations, 63 levels (observed locations in the historical series)

`LON` longitude

`LAT` latitude

`DOY` day of the year

`YYYYMMDD` calendar date in the format YYYY/MM/DD

`daysFromStart` days from average planting date

`T2M` daily average temperature at 2 meters

`T2M_MAX` daily maximum temperature at 2 meters

`T2M_MIN` daily minimum average temperature at 2 meters

`PRECTOT` rainfall precipitation

`WS2M` wind speed at 2 meters

`RH2M` relative humidity at 2 meters

`T2MDEW` dew point at 2 meters

`ALLSKY_SFC_LW_DWN` downward thermal infrared (longwave) radiative flux

`ALLSKY_SFC_SW_DWN` insolation incident on a horizontal surface

`n` duration of sunshine in hours

`VPD` the deficit of vapor pressure

`SPV` the slope of saturation vapor pressure curve

`ETP` evapotranspiration

`PETP` deficit of evapotranspiration

`GDD` growing degree-days

`FRUE` effect of temperature on radiation use efficiency

`T2M_RANGE` daily temperature range at 2 meters

`PTT` photothermal time (GDD × daylight in hours)

`PTR` photothermal ratio (GDD / daylight in hours)

Note

Comprehensive R Archive Network (CRAN) policy limits R package size to 5 Mb. In order to give the users new opportunities of data analysis, we provide weather data for all combinations of locations (63) and years (31), resulting in information for 1,953 environments. If an environment was not observed in a given year, weather data was retrieved with the average planting and maturity data based on the empirical data for that location. This data set can be downloaded here.

Source

Krause, M. D., Dias, K. O. G., Singh, A. K., and Beavis. W. D. (2022). Using large soybean historical data to study genotype by environment variation and identify mega-environments with the integration of genetic and non-genetic factors. bioRxiv, doi: 10.1101/2022.04.11.487885

Index