Bulk Upload of SUGARBAG data to BetyDB

What do we have!!

We have three excel files containing data and one PDF document summarizing the SUGARBAG database.

Copy of SUGARBAG Data.xls

- ➔ List of Experiments Contains Experiment ID and Title of the Experiment. This information is also available in PDF booklet.
- ➔ Harvest Data contains information about measured variables, date of the sampling, sampling ID, and plot ID where sampling was done. Each measurement is associated with an experiment ID. There can be multiple plots for a single experiment ID and we can collect multiple samples (sample ID) from a single plot. Description of variable name (variables to be measured) is provided in the worksheet **"Crop Variable"** in **"Copy of SUGARBAG Fixed Information.xls".** Units of measurements are also provided.
- ➔ Weather Data contain climate data for Meteorological stations. Meteorological Station ID can be used to associate them with an experiment ID.
- ➔ Soil Layer Data contains measured values of soil properties from one specified depth to another specified depth. Each measurement is associated with an experiment ID and plot ID. Description of soil properties variables can be found in the worksheet **"Soil Variable"** in **"Copy of SUGARBAG Fixed Information.xls".** Units of measurements are also provided.

Copy of SUGARBAG Experiments.xls

- ➔ Experiment Summary contains experiment ID, Title(name), Description, Field ID, Beginning and Ending Date of Experiments, associated meteorological station ID, Design of Experiment, and Number of Replication associated with experiment design.
- ➔ Research per Experiment provided name of Researchers for each experiment ID.
- ➔ Experiment Design contains experiment ID, Treatment ID and number of replication for each treatment. It also tells us which plot is receiving what treatment. For each plot, information is provided regarding crop class, cultiver, Crop start date and time and implementation of different management practices (drying off, fumigation, fumigation and irrigation, harvest date and time, initiation date, lodging management, N rate, N rates, 2 yr N rates, and Ratoon date.
- ➔ Planting contains information about experiment ID, method of planting, notes (important) about preparation, date of planting, cultiver type, Depth of Planting and row span.
- ➔ Irrigation contains information about Experiment ID, Treatment ID receiving irrigation,, date, amount, and method of irrigation.
- ➔ Fertilization contains information about experiment ID, treatment ID (receiving fertilization), date of application, Fertilizer type, Amount of Fertilizer and units, Depth of placement, and method of application.
- ➔ Tillage contains experiment ID, treatment ID (receiving this management), date, depth of tillage, and method of tillage.
- ➔ *Notes suggest to delete every experiment that were not available in the design.*

Copy of SUGARBAG FIXED Information.xls

- ➔ Site contains information about site ID, site name, city, Region, Latitute, Longitude, and Elevation. However, no experiment ID.
- ➔ Fields contains information about individual blocks within a site, detailing soil type, depth, latitude, longitude, elevation, and slope.
- ➔ Soils contain description about soil type.
- ➔ Soil Layers contains information about hydraulic parameters of soils, bulk density, OM, depth wise for each soil type.
- ➔ List of Researcher contains Researcher ID, Name, Affiliated Institute and Location of the Institute.
- ➔ Weather Station contains weather station ID, Name (location??), Latitude, Longitude and Elevation of the station.
- ➔ Crop Variables contains measured crop variable name, description, and unit of measurement.
- ➔ Soil Variables contain information about soil variable name, description, and unit of measurements.
- ➔ Climate variable contain information about variable name, description and unit of measurements.
- ➔ Fertilizer contains information about Fertilizer name, percent of N, P, K, Ca, and S.
- ➔ Method contains summary of some of the terms(or method used) in this database.
- ➔ Factors contain information about Level of Each Factors. This is likely to play and important role in experiment Design.
- ➔ Read notes to gain/understand database

Database is also briefly summarized in the PDF booklet.

What do we want to obtain

We Want to format SUGARBAG data into a format of relational database that can be uploaded to the BetyDB. Templates to format data contains multiple Google spreadhseets. Here, I will describe the worksheets that need to be filled and how are they supposed to be filled using data from SUGARBAG.

Citations

- ➔ Enter citation ID in series (1,2,....). Each ID corresponds to each item in Experiment Summary Report. This is given in the PDF (page 8 to 22) and also in the **Copy of SUGARBAG Data.xls.** Remember to eliminate any experiment for which experiment design information is not provided.
  We could use experiment ID also but it makes more sense to use number in series starting with 1. These IDs will anyway be replaced when entered in BetyDB. It may be good idea to create two columns, one corresponding to ID in citations and another experiment ID in SUGARBAG.
- ➔ Enter Author for each citation ID. Author name for corresponding experiment ID can be obtained from **Copy of SUGARBAG Experiment.xls** (Researcher per experiment).
- ➔ Year= 2003 for all citations (year of publication of SUGARBAG). See front page of PDF booklet.
- ➔ Title= Experiment name. Experiment name corresponding to each experiment ID can be obtained from Experiment Summary in **CopyofSUGARBAGExperiment.xls**

- ➔ journal = Title of SUGARBAG database (Front Page of ODF Booklet). It's same for all the citations.
- ➔ You can leave remaining columns blanks.

Citations_sites

- ➔ Enter citations ID and corresponding sites ID. One citation can have multiple sites ID. Also one site can have multiple citations. By Site ID in betyDB, we mean Field ID of SUGARBAG. Field ID corresponding to each experiment ID is provided in Experiment Summary **(copy of SUGARBAG EXPERIMENT.xls).** They should be entered in Citations_sites. It'll be better to use numerical (1,2,3....) for sites ID also instead of using confusing texts such as MK bed13, ZN block 151 etc. (in SUGARBAG).

Sites

- ➔ Insert Site IDS (preferably 1,2,3....).
- ➔ Ignore usgsmuid.
- ➔ Enter information about state, country, lat, longitude from sites/Fields in copyogUGARBAG Fixed Information.xls
- ➔ Ignore gdd, and firstfrostkill
- ➔ include weather station ID in mat. Weather Station ID (MetStation) for each field is given in Experiment Summary in Copy of SUGARBAG Experiment.xls. It's best not to change weather station ID. So please do not use numerical to replace name of weather stations.
- ➔ Ignore masl
- ➔ In the soil column, insert soil Type from Fields in **Copy of SUGARBAGFixed Information.xls.** Do not change Soil Type (nomenclature.
- ➔ For some of the soil tyoes, layered soil properties are provided in Soil Layers in CopyofSUGARBAG Fixed Information.xls. Save this worksheet as Soil-layers. This will be an additional Table in betyD. As I'll need this information however these is no provision to entert his information currently. DO not change but retain all the 10 colums (including blank B column).
- ➔ Ignore remaining columns in the sites.